HAL
open science

Visualisation d'information pour une décision informée en exploration d'espace de conception par shopping

Audrey Abi Akle

▶ **To cite this version:**

Audrey Abi Akle. Visualisation d'information pour une décision informée en exploration d'espace de conception par shopping. Autre. Ecole Centrale Paris, 2015. Français. NNT : 2015ECAP0039 . tel-01220661

HAL Id: tel-01220661

https://theses.hal.science/tel-01220661v1

Submitted on 26 Oct 2015

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.

CentraleSupélec

THÈSE

présentée par

Audrey ABI AKLE

pour l'obtention du

GRADE DE DOCTEUR

Spécialité : Génie Industriel

Laboratoire d'accueil : Laboratoire Génie Industriel

SUJET :

Visualisation d'information pour une décision informée en exploration d'espace de conception par shopping

Soutenue le : 10 juillet 2015

Jury composé de :

Carole BOUCHARD, Professeure des Arts et Métiers ParisTech	Présidente
Jean-François BOUJUT, Professeur de Grenoble INP	Rapporteur
Jean-François PETIOT, Professeur de l'Ecole Centrale de Nantes	Rapporteur
Mauricio CAMARGO, Professeur de l'INP Lorraine	Examinateur
Bernard YANNOU, Professeur de CentraleSupélec	Directeur
Stéphanie MINEL, Enseignant-Chercheure de l'ESTIA	Co-encadrante

2015ECAP0039

« J'aime bien la vie.
Ça remplit mes journées.»
Ronnie Shakes

Remerciements

Tout d'abord, mes remerciements s'adressent à Messieurs Jean-François PETIOT, Professeur à l'Ecole Centrale de Nantes, et Jean-François BOUJUT, Professeur à l'INP Grenoble, pour avoir accepté de juger ce travail en tant que rapporteurs.

Je remercie également Carole BOUCHARD, Professeure des Arts et métiers ParisTech, et Mauricio CAMARGO, Professeur de l'Université de Lorraine, pour avoir accepté de juger ce travail en tant qu'examinateurs.

Je tiens à remercier tout particulièrement le Conseil Général des Pyrénées Atlantiques qui a accepté de financer mes travaux et M. Jean-Roch GUIRESSE, Directeur de l'ESTIA, d'avoir mis les moyens à l'entreprise de cette thèse.

Je tiens à exprimer mes plus grands remerciements à mes directeurs de thèse Bernard YANNOU et Stéphanie MINEL. Merci à toi Bernard pour ta rigueur, ta grande patience et ton éternel enthousiasme. Le résultat de cette thèse ne pourrait pas être sans toi. Stéphanie, il est difficile d'exprimer en quelques lignes toute ma gratitude, la liste est longue. En bref, tout ceci ne serait pas arrivé sans toi, merci !

Je tiens à remercier toutes les personnes qui ont participé à cette aventure que ce soit de près ou de loin.

Merci à la PEPSS pour l'eye-tracker et de m'avoir toujours grand ouvert votre porte.

Merci aux membres du LGI pour m'avoir toujours très bien accueillie. Merci à Delphine MARTIN pour son aide.

Merci à tous les membres d'Ikasi. Je n'aurai pas rêvé de meilleur entourage pour faire ma thèse. Merci surtout à toi Marion pour ta relecture, Christophe pour ton aide dans la mise en place de l'expé 2 et à toi Mikele (le cerveau de l'équipe) pour ta bienveillance.

Merci aux doctorants de l'ESTIA. Plus particulièrement, Xavier pour ton swag (oh my buddha !), Kat pour toutes ses pauses clopes où tu m'as écouté et encouragé et Marion merci de m'avoir aidé à retrouver ma problématique de thèse (plusieurs fois…). Bref, les filles, merci pour tous ses afterworks !

Julien (fall's de son nom de scène) merci pour ton écoute, ton amitié, ta folie et tous ces vendredis (et merci pour les bières). Merci aussi à toi Max pour tous ces apéros et surtout ta patience (pas facile de supporter une doctorante qui doute).

Enfin, merci à ma famille pour leur soutien et leur amour. J'en profite pour souhaiter un joyeux anniversaire à ma filleule Lylou qui fête son 1 an.

Table des matières

Remerciements 3
Table des matières 4
Table des figures 7
Table des tableaux 11
INTRODUCTION GENERALE 13
Champ de notre recherche 13
Organisation du Manuscrit 16
1 PREMIERE PARTIE : Positionnement scientifique et état de l'art 18
1.1 OPTIMISATION MULTICRITERE EN CONCEPTION 19
1.1.1 Les différentes approches 21
1.1.2 Le Design by Shopping 24
1.1.3 L'exploration d'espace de conception 27
1.1.4 L'essentiel du Chapitre 1 36
1.2 DECISION INFORMEE 37
1.2.1 La décision 38
1.2.2 La décision informée : Etat de l'art 40
1.2.3 Les thématiques gravitant autour de la décision informée 45
1.2.4 Importance du terme « décision informée » 48
1.2.5 Notre définition de la décision informée 48
1.2.6 L'essentiel du chapitre 2 49
1.3 LES REPRESENTATIONS GRAPHIQUES 50
1.3.1 La visualisation d'information 51
1.3.2 Les représentations des ensembles multidimensionnels 53
1.3.3 Comparaison 65
1.3.4 Les tâches 68
1.3.5 Bilan 71
1.3.6 L'essentiel du chapitre 3 80
CONCLUSION DE LA PREMIERE PARTIE 81
2 DEUXIEME PARTIE : Problématique et démarche de recherche 83
2.1 Positionnement de nos travaux de recherche 84

2.1.1 Définition du périmètre ... 85
2.1.2 Définition des deux composantes de notre problématique ... 86
2.2 Les hypothèses de résolution ... 89
2.3 Notre plan d'actions ... 91
CONCLUSION DE LA DEUXIEME PARTIE ... 92
3 TROISIEME PARTIE : Expérimentations ... 94
3.1 INTRODUCTION DE L'EXPERIMENTATION N°1 ... 96
3.1.1 Description des graphiques et du jeu de données ... 96
3.1.2 Déroulement de l'expérimentation n°1 ... 98
3.1.3 Les mesures ... 102
3.1.4 Synthèse ... 105
3.2 INDICE DE QUALITE ... 106
3.2.1 Recherche des solutions non-dominées ... 106
3.2.2 Indice Basique ... 114
3.2.3 Notre indice de qualité ... 119
3.3 TRAITEMENT DES RESULTATS 1 ... 127
3.3.1 Le compromis ... 128
3.3.2 La recherche de compromis ... 128
3.3.3 Le ressenti ... 129
3.3.4 Conclusion pour la sélection en M.A.D.M. ... 132
3.4 INTRODUCTION DE L'EXPERIMENTATION N°2 ... 133
3.4.1 Présentation de la plateforme d'expérimentation ... 133
3.4.2 Déroulement de l'expérimentation n°2 ... 146
3.4.3 Synthèse ... 151
3.5 ANALYSE PRELIMINAIRE : GROUPES DE PARTICIPANTS ... 152
3.5.1 Connaissance en exploration d'espace de conception ... 153
3.5.2 Confiance du Participant ... 156
3.5.3 Conclusion pour les données des participants ... 157
3.6 EXPERIMENTATION 2.1 : PHASE SELECTION ... 159
3.6.1 Mesures liées à la phase de sélection ... 159
3.6.2 Traitement des résultats 2.1 ... 159
3.6.3 Conclusion pour la phase de sélection ... 162

3.7 EXPERIMENTATION 2.2 : PHASE OPTIMISATION 163
3.7.1 Mesures liées à la phase d'optimisation 163
3.7.2 Traitement des résultats 2.2 165
3.7.3 Conclusion pour la phase d'optimisation 168
3.8 EXPERIMENTATION 2.3 : PHASE DECOUVERTE 169
3.8.1 Mesures liées à la phase de découverte 169
3.8.2 Traitement des résultats 2.3 172
3.8.3 Conclusion pour la phase de découverte 176
3.9 PERFORMANCE DES GRAPHIQUES DANS L'EXPLORATION D'ESPACE DE CONCEPTION 177
3.10 EXPERIMENTATION 2.4 : DECISION INFORMEE 179
3.10.1 Mesures liées à la décision informée 179
3.10.2 Traitement des résultats 2.4 179
3.10.3 Conclusion sur la décision informée 182
CONCLUSION DE LA TROISIEME PARTIE 183
4 QUATRIEME PARTIE : Apports, limites et perspectives 186
4.1 Apports 187
4.1.1 Préconisations de visualisation d'information pour la prise de décision 187
4.1.2 Apports orientés méthodologie 188
4.1.3 Apports à destination de la communauté scientifique en génie industriel 188
4.2 Limites et perspectives 189
CONCLUSION GENERALE 194
Liste des publications 197
BIBLIOGRAPHIE 198
ANNEXES 207
Résumé / Abstract 242
Résumé 242
Abstract 243

Table des figures

Figure 1 : Positionnement de notre recherche en Design Science 14
Figure 2 : Positionnement de nos travaux 15
Figure 3 : Plan du manuscrit 16
Figure 4 : Illustration de la composition de l'espace de conception 20
Figure 5 : Illustration de la place du concepteur dans l'approche automatique 21
Figure 6 : Illustration de la place du concepteur dans l'approche semi-automatique 22
Figure 7 : Illustration de la place du concepteur dans l'approche par réduction de l'incertitude 22
Figure 8 : Illustration de la place du concepteur dans l'approche Design by Shopping 23
Figure 9 : Illustration de l'enchainement des activités Explorer, Optimiser et Sélectionner en Design by Shopping 25
Figure 10 : Illustration de l'approche typique de l'exploration d'espace de conception décrite par Simpson *et al*. (2008) 28
Figure 11 : Illustration de l'approche « work-centered » de Yan *et al*. (2012) 29
Figure 12 : SADT A0 Explorer un espace de conception 30
Figure 13 : SADT combinant A1 Découvrir, A2 Optimiser et A3 sélectionner 32
Figure 14 : Impressions d'écran de l'ARL Trade Space Visualizer: a) Glyph Plot, b) Parallel Coordinates, d) Scatter Matrix. 33
Figure 15 : Impression d'écran de l'outil VIDEO 34
Figure 16 : Impressions d'écran de l'outil LIVE: (a) scatter plot des variables de performance; (b) scatter plot des paramètres de conception (montrant des clusters de résultats); (c) règles de décision; (d) treemap de l'arbre de décision; (e) sliders ajustables des intervalles des variables. 35
Figure 17 : Processus de « Visual analytics » de Keim *et al*. (2008b) 43
Figure 18 : Les éléments clés de la décision informée 45
Figure 19 : Nombre de papier mentionnant les éléments clés par champs disciplinaires 47
Figure 20 : Nombre de fois que le terme « informed decision » est utilisé dans chaque papier 48
Figure 21 : Impression d'écran de « classification of visual data exploration techniques » dans (Keim, 2001) 51
Figure 22 : Impression d'écran de « The pixel-oriented circle segment technique » dans (Keim, 2001) 52
Figure 23 : Graphique en barres 53
Figure 24 : Exemple de scatterplot matrix issue de (Stump *et al*., 2009) 54
Figure 25 : Impression d'écran de l'amélioration du scatterplot matrix issu (Cui *et al*., 2006) 55
Figure 26 : Paramètres du visage de Chernoff (Chernoff, 1973) 56
Figure 27 : Impression d'écran de l'inconvénient des valeurs extrêmes (Flury & Riedwyl , 1981) 56

Figure 28 : Diagrammes étoile avec des aires 57
Figure 29 : Diagrammes étoile avec des segments 57
Figure 30 : Illustration du spider graph 58
Figure 31 : Illustration du Value Paths 59
Figure 32 : Construction du Parallel Coordinates Plot 59
Figure 33 : Illustration d'un Parallel Coordinate 60
Figure 34 : Impression d'écran illustrant des clusters séparés en x et en y issue de (Wegman, 1990) 60
Figure 35 : Impression d'écran illustrant des clusters séparés en x issue de (Wegman, 1990) 61
Figure 36 : Impression d'écran illustrant 3 clusters suivant une corrélation négative issue de (Wegman, 1990) 61
Figure 37 : Impression d'écran issu de (Ankerst et al., 2011) illustrant la construction du circle segment 62
Figure 38 : Exemple d'un circle segment issu de (Ankerst *et al.*, 2001) 62
Figure 39 : Exemple d'un Treemap issu (Shneiderman & Wattenberg, 2001) 63
Figure 40 : Illustration Treemap Vs. Structure arbre 64
Figure 41 : Positionnement sur le modèle de Keim (2001) 64
Figure 42 : Tableau de synthèse des différentes techniques de visualisation d'après Miettinen (2014) 65
Figure 43 : Tableau de comparaison préliminaire et subjective de Keim (2000) 67
Figure 44 : Exemple de visualisation d'un scatterplot matrix issu de (Stump *et al.*, 2009) 74
Figure 45 : Illustration des relations pour chaque graphique dans espace à 6 dimensions 76
Figure 46 : Fiche identité du Scatter Plot Matrix 77
Figure 47 : Fiche identité du Simple Scatter Plot 77
Figure 48 : Fiche identité du Parallel Coordinates Plot 78
Figure 49 : Fiche identité du Radar Chart 78
Figure 50 : Fiche identité du Combined Table 79
Figure 51 : Proposition du périmètre de nos travaux 86
Figure 52 : Illustration de la première composante de notre problématique 87
Figure 53 : Illustration de la deuxième composante de notre problématique 88
Figure 54 : Découpage de notre plan d'actions 91
Figure 55 : Illustration du découpage et positionnement des expérimentations 94
Figure 56 : Guide de lecture de la troisième Partie 95
Figure 57 : Description du Radar Chart 96
Figure 58 : Description du Combined Table 97
Figure 59 : Description du Parallel Coordinates Plot 98
Figure 60 : Illustration du déroulement de l'expérimentation n°1 99
Figure 61 : Synthèse des cinq objectifs des trois scénarios 102
Figure 62 : Synthèse des mesures de l'expérimentation n°1 105
Figure 63 : Exemple scénario TAILLE (en rose les solutions dites non-dominées au sens de Pareto) 108

Figure 64 : Carte des dominations pour le scénario Taille. Les ronds gris représentent les solutions dominées et la taille des ronds est proportionnelle au nombre de dominations qu'effectue la solution. 108
Figure 65 : Carte des dominations pour le scénario Puissance. Les ronds gris représentent les solutions dominées et la taille des ronds est proportionnelle au nombre de dominations qu'effectue la solution. 109
Figure 66 : Matrice de nuages de points des variables de performance du scénario TRAJET (en rose les solutions dites non-dominées au sens de Pareto et en vert la solution 34) 110
Figure 67 : Carte des dominations pour le scénario Trajet. Les ronds gris représentent les solutions dominées et la taille des ronds est proportionnelle au nombre de dominations qu'effectue la solution. 111
Figure 68 : Illustration de la « cascade » de dominations avec un cas bi-objectif (en rose et vert les solutions Pareto-Optimales) 112
Figure 69 : Exemple de solutions pour un problème bi-objectif où il faut minimiser g1 et g2 112
Figure 70 : Illustration des étapes pour l'identification du point Optimal 114
Figure 71 : Exemple de solution Idéale et Anti-idéale dans un problème où il faut minimiser g1 et g2 115
Figure 72 : Illustration du Point Optimal dans le cas d'un front de Pareto convexe. En rose le front de Pareto, en jaune la diagonale de boite et les indices dits basiques à coté de chaque point 116
Figure 73 : Illustration du Point Optimal dans le cas d'un front de Pareto concave. En rose le front de Pareto, en jaune la diagonale de boite et les indices dits basiques à coté de chaque point 116
Figure 74 : Illustration des indices que prennent les solutions dans les trois scénarios : a.) Indice i_j b.) Indice i 117
Figure 75 : Carte des indices pour le scénario PUISSANCE avec « vitesse » à maximiser et « poids » à minimiser 118
Figure 76 : Illustration de la construction de l'indice n°1 120
Figure 77 : Illustration de la construction de l'indice n°2 121
Figure 78 : Illustration de la construction de l'indice n°3 123
Figure 79 : Illustration des différences de classement des solutions entre les trois indices 123
Figure 80 : Illustration des classements des solutions Pareto-Optimales avec les trois indices pour le scénario TRAJET. 124
Figure 81 : Illustration des indices de qualité n°1 des quarante voitures pour le scénario TRAJET 125
Figure 82 : Illustration des indices de qualité n°2 des quarante voitures pour le scénario TRAJET 125
Figure 83 : Illustration des indices de qualité n°3 des quarante voitures pour le scénario TRAJET 125

Figure 84 : Meilleures et pires voitures définies par l'indice de qualité pour les trois scénarios 126
Figure 85 : Composition de l'interface – Exemple du Simple Scatter Plot 134
Figure 86 : Interface avec le graphique Scatter Plot Matrix (a.) et avec le Parallel Coordinates Plot (b.) 134
Figure 87 : Les trois menus 135
Figure 88 : Onglet « Design Problem Description » 135
Figure 89 : Onglet « Design Points Sampling » 136
Figure 90 : Illustration du générateur de points non-dominés au sens de Pareto 136
Figure 91 : Onglet « Range Constraints Control » 137
Figure 92 : Onglet « Preference Control » 138
Figure 93 : Illustration de la mise en lumière des optimums 138
Figure 94 : Onglet « Color Glyph Control » 139
Figure 95 : Onglet « Axis Control » 139
Figure 96 : Onglet « Mouse Control » 139
Figure 97 : Illustration de la fonction « selection » avec le SSP 140
Figure 98 : Illustration de la fonction « brush » avec le SPM 141
Figure 99 : Illustration de la fonction « drag & drop » avec le PCP 141
Figure 100 : Illustration du problème Truss 143
Figure 101 : Illustration du problème Gear 144
Figure 102 : Illustration du problème Disk 145
Figure 103 : Illustration du problème Vessel 146
Figure 104 : Illustration du déroulement de l'expérimentation n°2 147
Figure 105 : Illustration des trois étapes de la phase Formation 148
Figure 106 : Illustration des différentes étapes de la phase Test 149
Figure 107 : Exemple de la question n°8 du Q.C.M. 1 150
Figure 108 : Illustration des différentes évolutions de « connaissances » possibles par profil 151
Figure 109 : Moyennes des notes obtenues aux QCM et écart-type 155
Figure 110 : Moyennes et écart-types des résultats pour la qualité de la solution sélectionnée 161
Figure 111 : Illustrations des indicateurs (b et c) : « sens », (d et e) : différence de l'étendue de qualité et (f et g) : différence de la qualité moyenne 164
Figure 112 : Export de l'eye-tracking du participant n°11 pour le problème Disk 170
Figure 113 : Export de l'eye-tracking du participant n°9 pour le problème Vessel 171
Figure 114 : Export de l'eye-tracking du participant n°26 pour le problème Disk 171
Figure 115 : Moyennes et écart-types des résultats pour le nombre de découvertes réalisées 173
Figure 116 : Moyennes et écart-types des résultats pour le temps pour la première découverte 174
Figure 117 : Moyennes et écart-types des résultats pour le temps moyen par découverte ... 176

Figure 118 : SADT de l'exploration d'espace de conception avec nos recommandations 184
Figure 119 : Illustration des premiers apports de nos travaux 187
Figure 120 : Illustration de l'apport scientifique (a et b) et d'un apport terrain (c et d). M pour construire le Modèle, D pour Découvrir, O pour Optimiser et S pour Sélectionner. 190
Figure 121 : Illustration des perspectives. M pour construire le Modèle, WB pour la définition du *WellBeing*, D pour Découvrir, O pour Optimiser et S pour Sélectionner 191
Figure 122 : Variables visuelles de Bertin 1/3 208
Figure 123 : Variables visuelles de Bertin 2/3 209
Figure 124 : Variables visuelles de Bertin 3/3 209
Figure 125 : Exemple d'une réponse pour une décision informée 241
Figure 126 : Exemple d'une réponse pour une décision non-informée 241

Table des tableaux

Tableau 1 : Définitions de la décision, chronologie et thématiques 38
Tableau 2 : Comparaison des graphiques selon le nombre de critères et d'alternatives..... 71
Tableau 3 : Comparaison des graphiques selon les différentes « découvertes » 73
Tableau 4 : Contrebalancement des graphiques 100
Tableau 5 : Mesures et objectifs du questionnaire post-test 104
Tableau 6 : Nombre de solutions non-dominées pour les trois scénarios 113
Tableau 7 : Construction de l'indice n°1 pour les solutions dominées et non-dominées..... 119
Tableau 8 : Construction de l'indice n°2 pour les solutions dominées et non-dominées..... 120
Tableau 9 : Construction de l'indice n°3 pour les solutions dominées et non-dominées..... 122
Tableau 10 : Tests statistiques utilisés par indicateurs de l'expérimentation n°1 127
Tableau 11 : Synthèse des résultats pour la sélection en M.A.D.M. 132
Tableau 12 : Résultats des tests de Kruskal-Wallis pour les réponses aux trois questions de confiance (Vessel) 156
Tableau 13 : Résultats des tests de Kruskal-Wallis pour les réponses aux trois questions de confiance (Disk) 157
Tableau 14 : Tableau de contingence des variables groupes et standard 160
Tableau 15 : Tableau de contingence des variables groupes et « Pareto » 160
Tableau 16 : Synthèse des résultats pour la phase de Sélection 162
Tableau 17 : Tableau de contingence des variables groupes et erreur 166
Tableau 18 : Synthèse des résultats pour la phase d'optimisation 168
Tableau 19 : Synthèse des résultats obtenus pour la phase de Découverte 176
Tableau 20 : Synthèse des résultats de l'expérimentation n°2 177
Tableau 21 : Tableau de contingence des variables groupes et réponse à la question 181
Tableau 22 : Synthèse des résultats des expérimentations 183

Introduction Générale

INTRODUCTION GENERALE

Avec les progrès de l'informatique en puissance et en vitesse de calculs, les concepteurs peuvent aujourd'hui analyser et évaluer des milliers voire des millions d'alternatives de conception plus rapidement et à moindre coût (Wang & Shan, 2007). Ces avancées permettent une simulation accrue d'alternatives de conception et entrainent de nouvelles formes d'exploration de l'espace de conception. Cependant, comme l'ont souligné Simpson et al., (2008), cette profusion des données d'alternatives de conception peut conduire à une surcharge d'information et incommoder les concepteurs particulièrement lorsqu'il n'y a pas d'outils appropriés pour accompagner la visualisation de données multidimensionnelle.

Dans nos travaux, nous nous intéressons à la performance des représentations graphiques pour l'exploration d'espace de conception i.e. pour une méthode d'optimisation multicritère en conception. Lorsqu'on parle d'espace de conception, on fait référence au champ des possibles, i.e. à l'ensemble des solutions possibles à un problème de conception. L'espace de conception peut également être vu comme un ensemble de données multidimensionnel. Avec cette méthode d'exploration, le concepteur est engagé dans un paradigme dit de Design by Shopping. En ce sens, à l'aide de graphiques, il « parcourt » l'espace de conception pour gagner en intuition sur les compromis, sur les solutions faisables et infaisables. Le concepteur apprend de ces alternatives avant l'optimisation et la sélection d'une solution optimale. L'objectif de cette thèse est donc l'identification de visualisations de données appropriées et performantes pour l'exploration d'espace de conception.

Nous proposons d'inscrire nos travaux dans le domaine du génie industriel et en décrivons le positionnement dans la section suivante.

Champ de notre recherche

Pour Womack et Jones (2010), le génie industriel concerne la conception, l'amélioration et l'installation de systèmes intégrés mettant en jeu des hommes, des matériaux, des équipements et de l'énergie. Il s'appuie tant sur les connaissances spécialisées et les aptitudes dans le domaine des mathématiques, de la physique et des sciences sociales que sur les principes et méthodes d'analyse et d'**ingénierie de la conception**, ceci pour spécifier, prédire et évaluer les résultats de ce type de système.

Cette définition met en exergue l'aspect pluridisciplinaire du génie industriel. De plus, elle souligne l'utilisation des méthodes et des principes de l'ingénierie de conception (engineering design) pour spécifier, prédire et évaluer des résultats attendus. L'ingénierie de conception apparait comme une branche du génie industriel.

Plus précisément, l'ingénierie de conception fait appel à un vaste corpus de connaissances défini par Hubka et Eder (1987) par le terme *Design Science* c'est-à-dire « a scientific study of

the design activity in its context ». L'information comprise dans le *Design Science* est découpée en quatre groupes :

- La théorie des systèmes techniques
- La théorie de la conception et du processus de conception
- La connaissance appliquée issue des sciences humaines et sociales
- Les méthodologies de conception

Notre recherche se positionne à l'intersection des catégories de la théorie de la conception et du processus de conception (2), de la connaissance appliquée issue des sciences humaines et sociales (3) et des méthodologies de conception (4) (**Figure 1**).

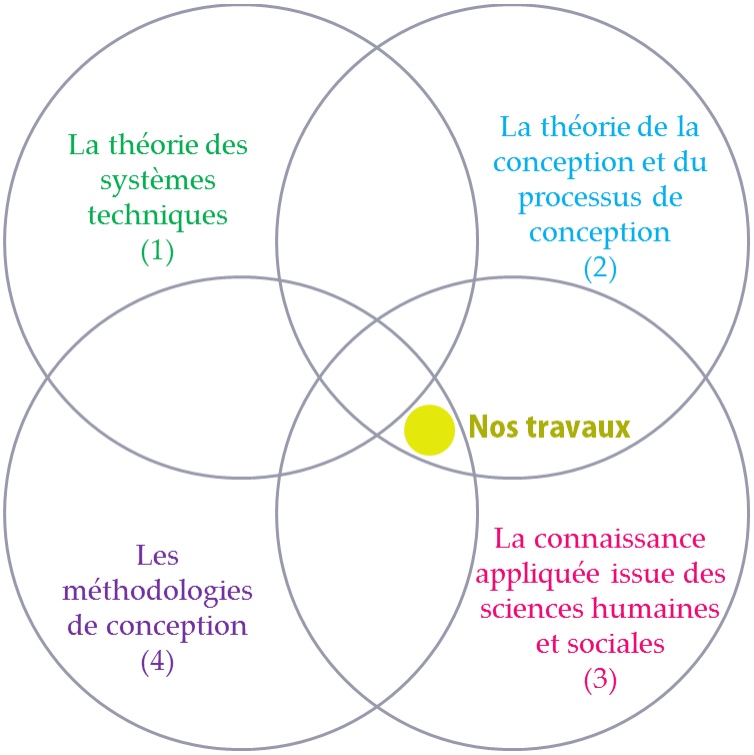


Figure 1 : Positionnement de notre recherche en Design Science

L'objet de notre recherche est l'utilisation de graphiques adaptés aux concepteurs pour l'exploration d'espace de conception. Ici, le concepteur est considéré comme un utilisateur. En ce sens, notre recherche fait donc appel à la connaissance issue des sciences humaines et sociales (3). Aussi, l'exploration peut être vue comme une méthodologie pour l'optimisation (4) mais aussi comme un processus de décision (2). Ainsi, nous avons pour objectif l'identification de graphiques qui sont performants pour la prise de décision dans ce cas particulier d'exploration d'espace de conception. Nos travaux sont donc intégrés dans la théorie du processus de conception (2). Enfin, nous cherchons des graphiques performants pour les différentes étapes (ou phases) de la méthode d'exploration. Le cadre de notre recherche s'intègre également dans les méthodologies de la conception (4).

Nous illustrons le cadre de notre recherche sur la **Figure 2.**

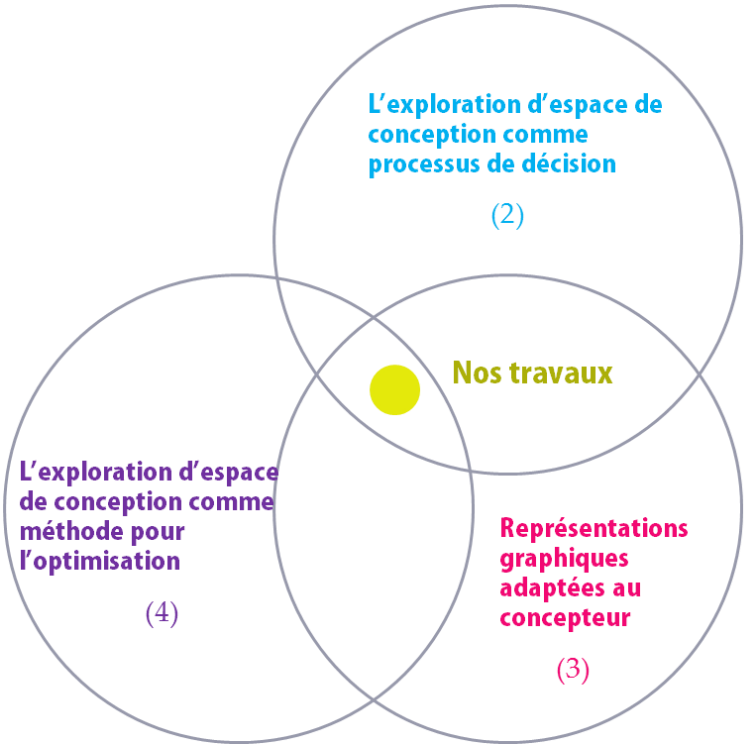


Figure 2 : Positionnement de nos travaux

Organisation du Manuscrit

Notre manuscrit est divisé en 4 parties distinctes. La **Figure 3** illustre le plan du manuscrit de thèse.

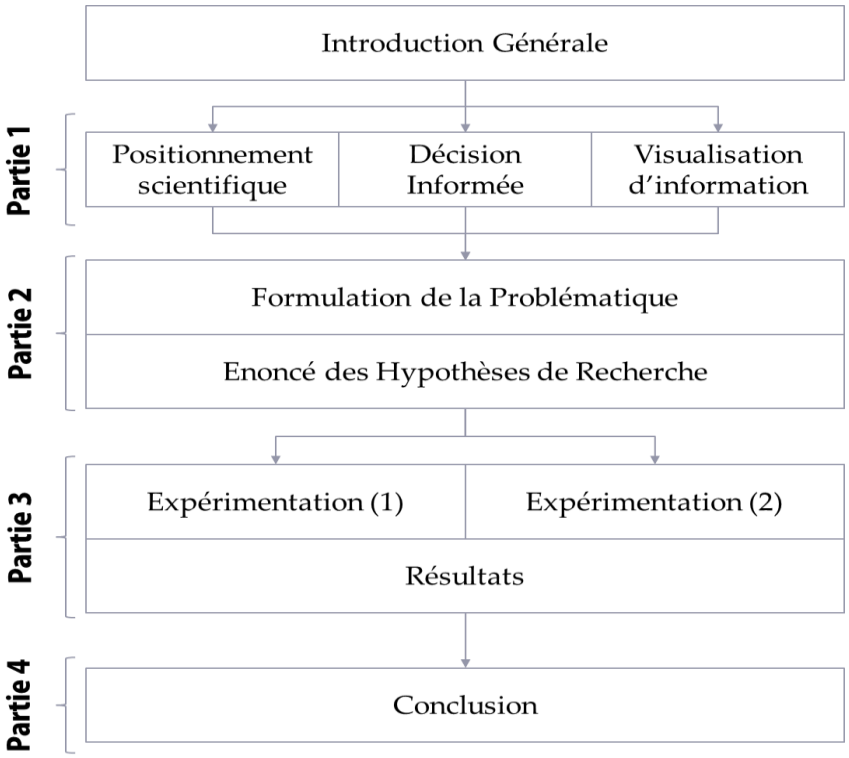


Figure 3 : Plan du manuscrit

La Partie 1, intitulée Positionnement Scientifique et Etat de l'Art, précise notre positionnement scientifique et présente notre état de l'art. La partie est scindée en trois sections. La première section situe nos travaux de recherche en exploration d'espace de conception. La deuxième section propose un état de l'art de la décision informée et sa définition. La troisième section propose un état de l'art des représentations graphiques pour des ensembles de données multidimensionnels.

La Partie 2, intitulée Problématique et Démarche de Recherche, présente notre problématique et la logique de construction de cette dernière, la définition des hypothèses de résolution et notre démarche de recherche.

La partie 3, intitulée Expérimentations, présente les expérimentations réalisées dans nos travaux pour répondre à nos hypothèses de recherche. Cette partie décrit le protocole expérimental et les mesurables, l'analyse statistiques des données et les résultats des expérimentations.

Enfin, la Partie 4, intitulée Apports, Limites et Perspectives, établit les apports de la thèse, les limites actuelles et les perspectives envisagées.

Le manuscrit s'achève avec une conclusion générale.

Première partie

Positionnement Scientifique et Etat de l'Art

1 PREMIERE PARTIE : Positionnement scientifique et état de l'art

Dans cette première partie, nous présentons notre positionnement scientifique et notre état de l'art. Elle est divisée en trois chapitres distincts.

Dans le premier chapitre, nous situons nos travaux de recherche dans le domaine de l'exploration d'espace de conception qui est une méthode pour l'optimisation multicritère en conception. Afin de décrire cette méthode, nous présentons le paradigme de Design by Shopping en conception, le processus d'activité et les outils existants pour appliquer cette méthode. Ce chapitre nous conduit à deux premiers constats et nous pousse à détailler le cas particulier de décision introduit par le Design by Shopping et les représentations graphiques comme support à la conception.

Dans le deuxième chapitre, nous dressons un état de l'art de ce que nous appelons, dans le cadre de nos travaux, la « décision informée ». Nous relevons, entre autre, des moyens et des prérequis nécessaires à la formulation de ce type de décision. Nous terminons ce chapitre en exposant notre définition de cet objet.

Le troisième chapitre propose un état de l'art des représentations graphiques pour des ensembles de données multidimensionnels. En considérant la méthode d'exploration d'espace de conception et la décision informée, nous réalisons un comparatif des graphiques et faisons émerger ceux qui sont potentiellement pertinents pour notre cas d'exploration d'espace de conception.

1.1 OPTIMISATION MULTICRITERE EN CONCEPTION

La conception est une activité de transformation provoquant un changement d'état d'un produit/service (Mistree et al., 1990) qui est perçue comme un processus[1], une séquence d'activités de conception nécessaires pour créer une ou plusieurs représentations d'un produit (Reymen, 2001).

Les processus de conception sont souvent composés des phases d'expression du besoin, de recherche de concepts, de recherche d'architecture et de conception détaillée étant elle-même constituée de la réalisation de calculs, de simulations, du dimensionnement, etc (Pahl & Beitz, 1996).

Le processus de conception est également définit comme un processus de décision (voir par exemple (Scaravetti, 2004)).

Dans les processus de conception ou de décision, l'objectif est de faire le meilleur choix parmi plusieurs autres possibles. En science, le « meilleur » fait référence à une décision qui maximise ou minimise plusieurs critères simultanément, on parle alors d'optimisation (Stadler, 1988).

La prise de décision multicritère (MCDM pour Multi Criteria Decision Making) fait référence à la prise de décision en présence de multiples critères habituellement en conflits (Zanakis *et al.*, 1998). Les problèmes de MCDM sont communément catégorisés comme continus ou discrets en fonction du type d'alternatives de solution, selon qu'elles sont en nombre fini (discret) ou en nombre infini (caractéristiques évoluant de manière continue). Yoon et Hang (1995) les classent comme :

- Prise de décision multi-attribut (MADM), avec un nombre discret et habituellement limité d'alternatives prédéfinies nécessitant des comparaisons inter et intra attributs – les attributs sont des critères - qui impliquent la réalisation de compromis par le décisionnaire.
- Prise de décision multi-objective (MODM), avec des valeurs de variables de décision qui sont à déterminer dans un domaine continu et un nombre infini d'alternatives (ou très grand) pour satisfaire les contraintes et les préférences du décisionnaire en optimisant au mieux tous les objectifs – ou encore performances -.

En ingénierie de la conception, les critères sont représentés par les vecteurs X (en entrée) et Y (en sortie) : X représente les paramètres de conception et forme l'espace des paramètres de

[1] D'après le Centre National de Ressource Textuelles et Lexicales (CNRTL), un processus est « une suite continue de faits, de phénomènes présentant une certaine unité ou une certaine régularité dans leur déroulement. Il s'agit d'un ensemble d'opérations successives, organisées en vue d'un résultat déterminé ».

conception ou l'espace de décision. Y représente les variables de performance et forme l'espace de performance ou l'espace objectif. Nous appelons l'espace de conception Z=[X Y] l'ensemble combiné des deux espaces de décision et de performance (voir **Figure 4**). Ainsi un point de conception (Design point) est une solution possible (i.e. faisable qui respecte les contraintes) dans l'espace de conception.

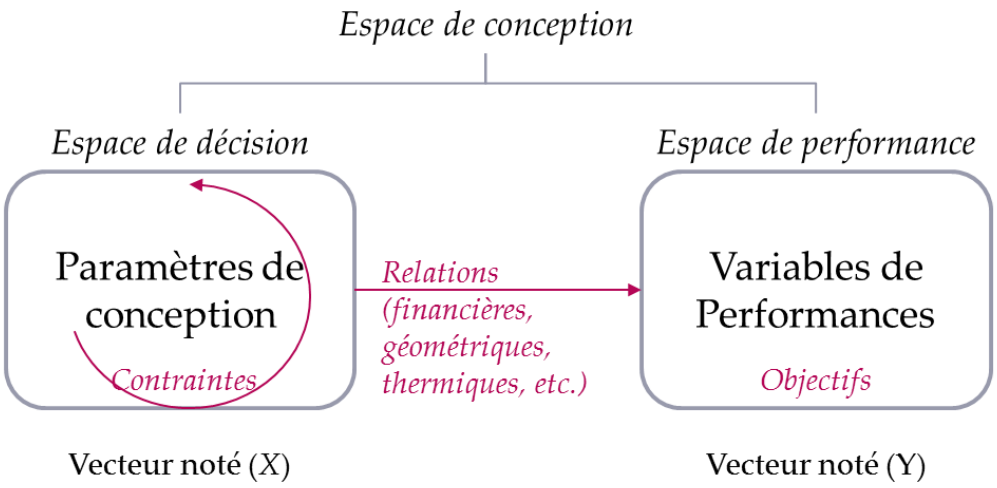


Figure 4 : Illustration de la composition de l'espace de conception

Pour passer d'un espace de conception composé d'une infinité de solutions possibles à une ou un ensemble de solutions optimales, il existe plusieurs approches et méthodes. Nous proposons de décrire les grandes familles de l'optimisation en conception et de nous positionner dans ce champ plus particulièrement en considérant la place du concepteur dans le processus.

1.1.1 Les différentes approches

L'optimisation en conception est un domaine répandu de la recherche. Les approches de l'optimisation en conception sont divisées en deux familles principales : L'optimisation d'une unique fonction objective et l'optimisation de multiples objectifs (Yannou *et al.*, 2013). Dans notre cas, nous nous intéressons à une situation d'optimisation multicritère en conception. Nous sommes donc concernés par la deuxième famille dans laquelle des solutions de conception sont calculées et/ou un compromis souhaité est trouvé automatiquement ou de manière interactive.

Notons que dans les **Figure 5**, **Figure 6**, **Figure 7** et **Figure 8**, les « bonhommes » en noir représentent les interventions du concepteur dans l'approche illustrée.

Dans l'approche automatique, un ensemble fini de « meilleures » solutions est directement proposé au concepteur (voir **Figure 5**). En suivant cette approche, le concepteur intervient au début du processus, au moment de définir le problème de conception, et à la fin pour sélectionner une des « meilleures » solutions. La **Figure 5** illustre la place du concepteur dans l'approche automatique. Nous pouvons qualifier cette approche de traditionnelle car elle est alimentée par plusieurs méthodes comme le *weighted-sum of preferences, indifference points* ou le *compensation of the weight* (Scott and Antonsson, 2000 ; 2005).

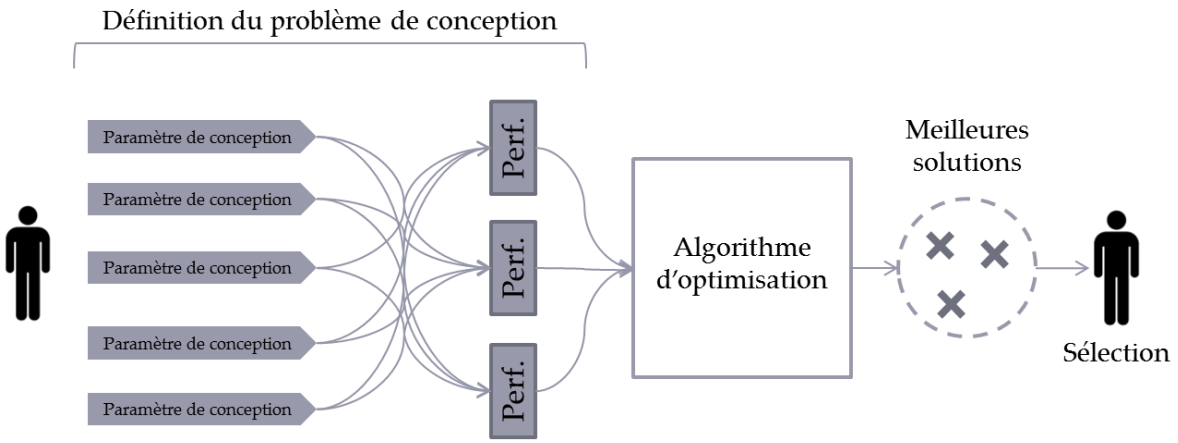


Figure 5 : Illustration de la place du concepteur dans l'approche automatique

Dans notre vision, nous considérons également l'approche «semi-automatique» dans la famille d'optimisation « automatique ». Bien que la spécification directe des poids d'importance puisse être un processus ad hoc dans l'approche semi-automatique et que le concepteur intervienne à la fin du processus, l'optimisation dans cette approche reste, pour nous, relativement opaque pour le concepteur. Nous illustrons la place du concepteur dans l'approche semi-automatique dans la **Figure 6**.

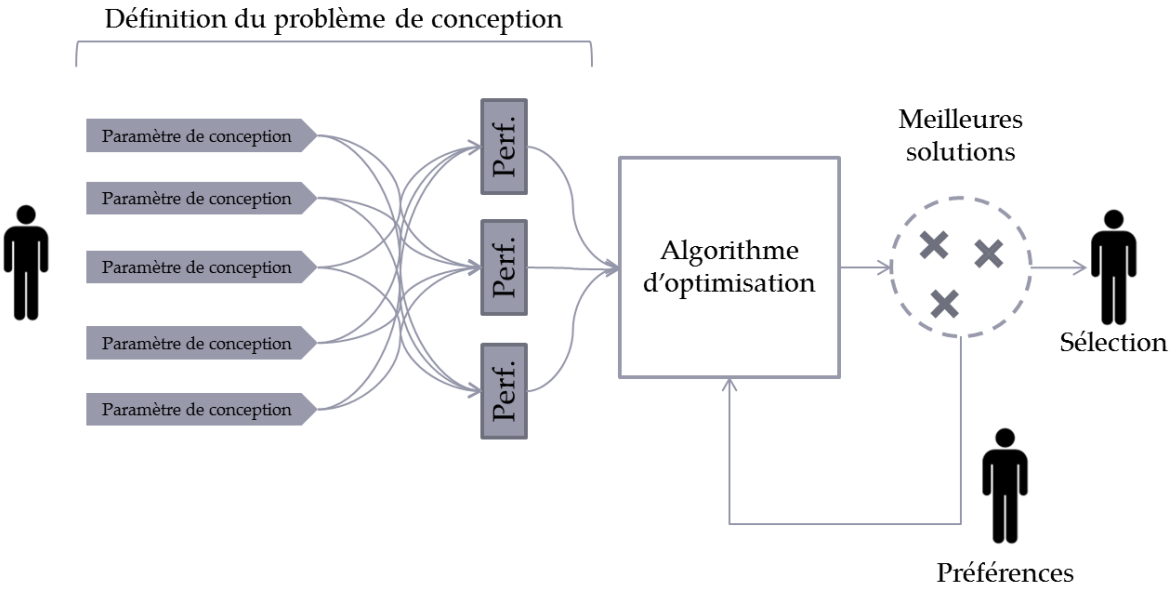


Figure 6 : Illustration de la place du concepteur dans l'approche semi-automatique

Dans l'approche interactive nous distinguons deux paradigmes particulièrement proches : la réduction d'incertitude et le Design by Shopping.

Dans le paradigme de réduction d'incertitude, le concepteur réduit de manière itérative les intervalles sur les variables. Cette réduction itérative lui permet de gagner en connaissance et, de cette manière, l'incertitude est réduite à travers le processus de convergence vers une solution optimale (Paredis *et al.*, 2007; Malak *et al.*, 2009). En phase préliminaire de conception, nous relevons l'utilisation de la programmation par contrainte comme une méthode pertinente pour la réduction d'incertitude (voir Yannou and Hamdi, 2004; Yannou *et al.*, 2009). Comme nous l'illustrons dans la **Figure 7** avec ce paradigme, le concepteur est intégré dans le processus et la sélection d'une solution optimale en est issue.

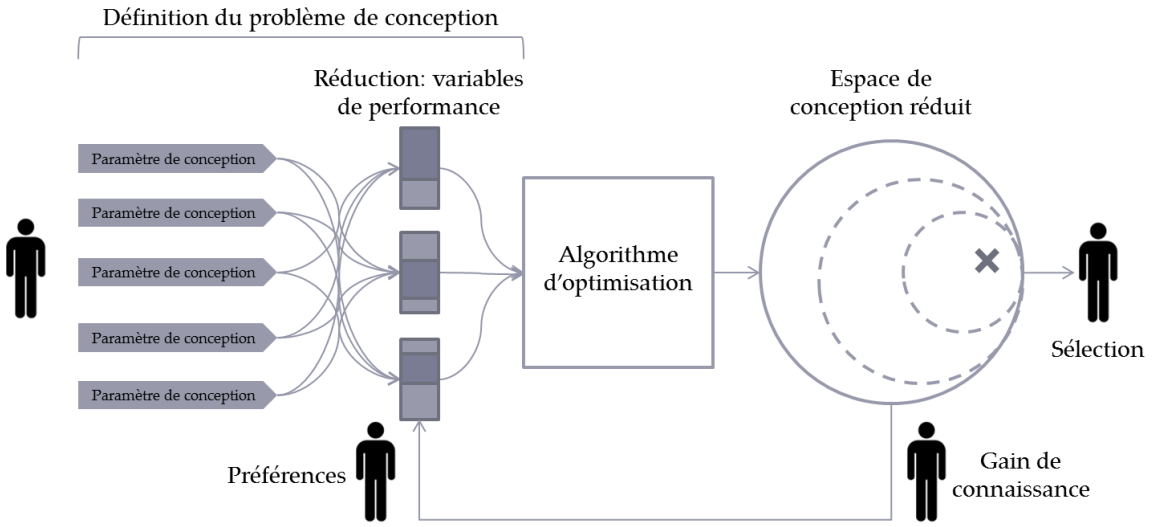


Figure 7 : Illustration de la place du concepteur dans l'approche par réduction de l'incertitude

En Design by Shopping, le concepteur explore les solutions possibles et les interactions entre les variables. La sélection d'un point de conception et, par conséquent, la réduction de l'espace de conception est réalisée par la découverte de connaissance, de révélations, de prises de conscience – on parle d'insights - au cours du processus (Zhang *et al.*, 2012). Avec ce paradigme le concepteur est dans le processus et directement dans la boucle de simulation (voir **Figure 8**).

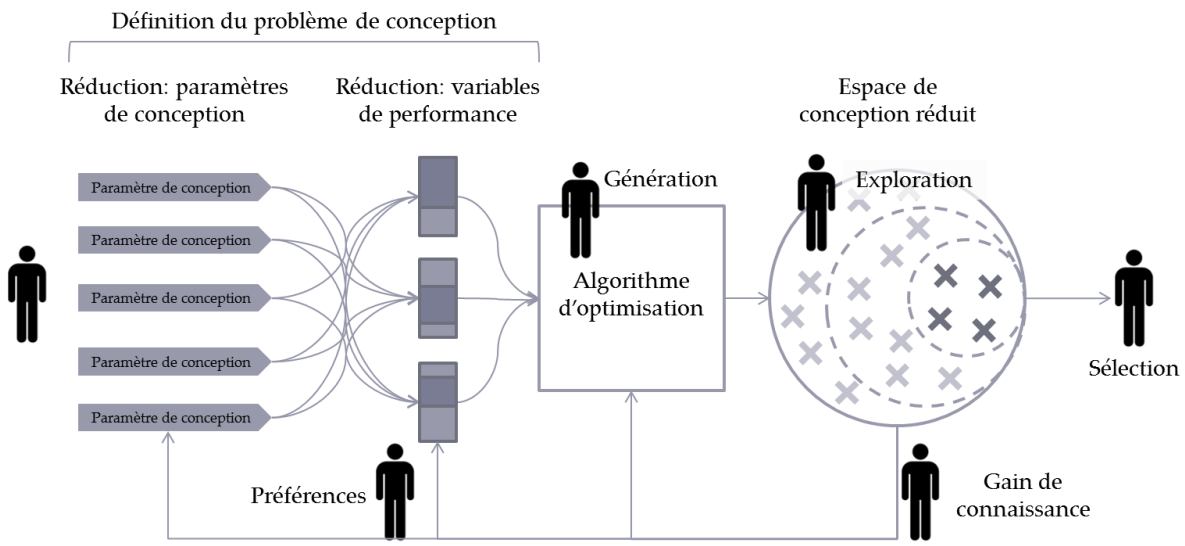


Figure 8 : Illustration de la place du concepteur dans l'approche Design by Shopping

Outre le processus de conception et les interventions du concepteur, nous pouvons différencier les deux paradigmes susmentionnés par la stratégie de conception (Fathianathan et Panchal, 2009). Dans la réduction d'incertitude, le processus est ascendant (Bottom-Up) ou ascendant étendu (Canbaz, 2013). Tandis qu'en Design by Shopping, le processus est dans une approche combinée : ascendante et descendante (Top-Down). En effet, le concepteur engagé dans une approche de Design by Shopping contrôle et interagit avec l'espace de performance et l'espace de décision de manière simultanée.

Nous positionnons nos travaux de recherche en Design by Shopping et proposons de décrire plus en détail ce paradigme dans la section suivante.

1.1.2 Le Design by Shopping

Balling (1999) a relevé que le processus de conception traditionnel, basé sur l'optimisation, qui est de (1) formuler le problème de conception, (2) obtenir un modèle d'analyse et (3) exécuter un algorithme d'optimisation, laisse les concepteurs insatisfaits de leurs résultats. Pour lui, cette insatisfaction vient du fait que le problème de conception n'est pas correctement formulé. Nous pouvons souligner qu'avec l'approche traditionnelle, le concepteur ne fait pas parti du processus. Il se trouve en entrée et en sortie. Comme nous l'avons illustré dans la section précédente, le processus de conception est relativement contraignant pour le concepteur. De plus, Balling indique que l'approche traditionnelle réduit la créativité du concepteur.

Comme le souligne Zhang et ses collègues (2012), dans de nombreux cas, les concepteurs (et les gens en général) ne savent pas ce qu'ils veulent vraiment jusqu'à ce qu'ils voient certains modèles :

"*The objectives and constraints used in optimization were not what the owners and stakeholders really wanted. In many cases, people do not know what they really want until they see some designs*." (cité dans Zhang *et al*., 2012)

Le processus de conception traditionnel est opaque pour les concepteurs. Ils ne peuvent pas « toucher » certains modèles et/ou se représenter le possible (i.e. des solutions faisables). Les concepteurs ont besoin d'apprendre de leurs alternatives, de faire le lien entre les paramètres de conception et les variables de performance avant de prendre une décision. Les concepteurs, comme les consommateurs, veulent faire leur « marché » (notre traduction de "*shopping*") pour gagner en intuition sur les compromis, sur les solutions faisables et infaisables.

Le Design by Shopping propose, dans un premier temps, de permettre aux concepteurs d'explorer l'espace de conception et, dans un deuxième temps, d'optimiser et de sélectionner une solution optimale dans l'ensemble des solutions possibles. Ainsi, le concepteur peut développer des attentes réalistes de ce qui est possible.

Explorer l'espace de conception est l'activité qui permet aux concepteurs d'accéder au « champ » des possibles et d'apprendre de leurs alternatives. Car l'exploration leur permet de « toucher » des modèles et de faire le lien entre les paramètres de conception et les variables de performance en découvrant et manipulant les solutions possibles.

L'optimisation est itérative et est réalisée à l'aide de l'exploration. Les flèches (1) et (2) sur la **Figure 9,** illustrent le passage d'une activité à l'autre, pouvant être répété. Cette activité permet aux concepteurs de réduire petit à petit le « champ » des possibles premièrement en écartant les solutions infaisables et celle n'étant pas de « bonnes » réponses. Les concepteurs

sont libres d'utiliser des algorithmes leur permettant d'avancer vers des solutions plus performantes. Le paradigme replace l'homme directement dans la boucle de simulation.

La sélection est également réalisée à l'aide de l'exploration mais à partir de l'optimisation (**Figure 9**).

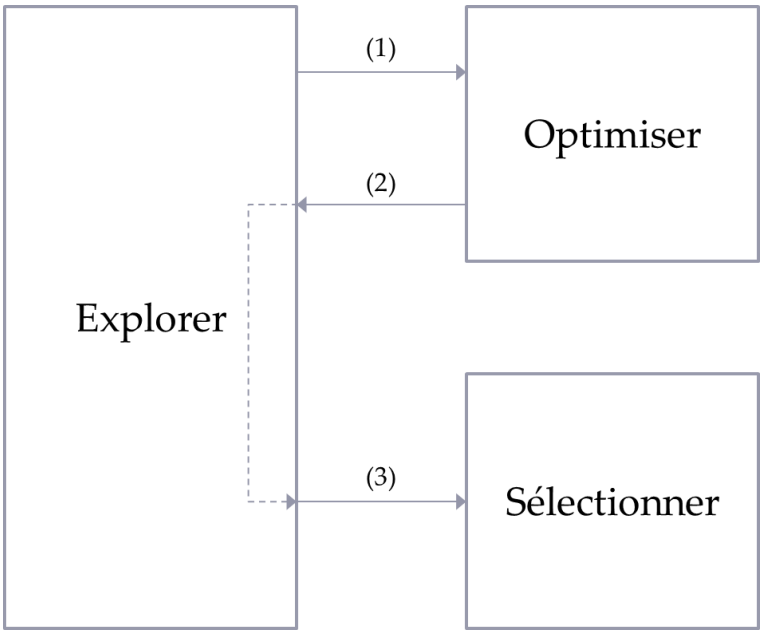


Figure 9 : Illustration de l'enchainement des activités Explorer, Optimiser et Sélectionner en Design by Shopping

La description de l'exploration en Design by Shopping révèle que ce paradigme implique que le concepteur réalise sa décision à partir de l'information acquise durant le processus de conception. C'est ce que nous appelons la décision « informée ». Nous proposons une définition de ce type de décision dans le **chapitre 2**.

Ce paradigme introduit plus de « contrôle » pour le concepteur car il lui permet de constituer ses préférences à partir de la visualisation de l'espace de conception avant de sélectionner une solution optimale :

"*allowing them to first form their preference based on the visualization of the entire design space, and then choose an optimal design*." (Rangavajhala, 2007)

Plus précisément, ce paradigme porte plus sur l'élimination des mauvaises réponses plutôt que de trouver la bonne réponse. Le Design by Shopping est défini comme une articulation a posteriori des préférences utilisées pour résoudre des problèmes d'optimisation multicritère.

Pour Feather et ses collègues, le Design by Shopping met l'accent sur révéler l'espace des options disponibles aux concepteurs sans présumer que tous les critères de sélection soient déjà fixés (Feather *et al.*, 2008). En effet, l'exploration de l'espace de conception et donc le

gain de connaissance sur le problème de conception, peuvent être réalisés même si tous les « aspects » du problème ne sont pas définis.

Aussi, Ferringer et ses collègues montrent que les concepteurs préfèrent la sélection d'une solution optimale lorsqu'ils sont dans une approche de Design by Shopping.

"*The ultimate reduction of the solution space to a single Decision-Maker preferred design proceeds using this tradeoff information in a Design by Shopping paradigm*." (Ferringer *et al*., 2009)

Rothrock & Narayanan (2011) relèvent des travaux de recherche ayant des résultats proches du Design by Shopping. Par exemple, nous trouvons dans le domaine de la psychologie cognitive que :

- Wilson & Schooler (1991) ont montré que les gens sont moins performants à certaines tâches de décision lorsqu'on leur demande d'analyser les raisons de leurs préférences ou d'évaluer tous les attributs de leurs choix.
- Shanteau (1992) a observé que lorsque les gens ne sont pas satisfaits avec les résultats d'un processus de prise de décision rationnelle, ils changent souvent leurs notations pour atteindre les résultats souhaités.

Ce paradigme, en replaçant le concepteur au cœur du processus de conception, apporte plus de flexibilité pour la conception et plus de satisfaction chez le concepteur.

Nous proposons, dans la section suivante, de décrire une méthode qui permet la conception en Design by Shopping.

1.1.3 L'exploration d'espace de conception

L'exploration d'espace de conception est une méthode qui « incarne » le paradigme de Design by Shopping (O'Hara *et al.*, 2007). Cette méthode est considérée comme un moyen prometteur pour la prise de décision car elle fournit un moyen visuel et intuitif pour formuler, ajuster et résoudre les problèmes d'optimisation multicritère en conception (Carlsen *et al*, 2008).

Cette méthode est utilisée pour l'optimisation en conception avec des projets complexes. Nous identifions des exemples d'utilisation dans des situations d'optimisation multi-objective :

- Dans le domaine de l'automobile avec la configuration d'un nouveau véhicule (Ferguson *et al.*, 2005 ; Donndelinger *et al.*, 2006)
- Dans le domaine de l'aéronautique avec la conception d'une aile (Simpson *et al.*, 2008)
- Dans le domaine de l'aérospatial avec l'optimisation de trajectoire d'un satellite (Stump *et al.*, march 2004 ; Jordan *et al.*, 2008)

Plus particulièrement dans le domaine de l'aérospatial, nous trouvons l'utilisation de cette méthode dans des situations d'optimisation multi-attribut notamment lors de la conception conceptuelle (conceptual design). Nous relevons tout particulièrement les travaux de Ross et ses collègues issus du Massachusetts Institute of Technology, pour l'architecture système en ingénierie concourante (Ross, 2003), pour la conception de systèmes spatiaux (Ross *et al.*, 2004 ; Ross et Hastings, 2005) et pour le gain de connaissance et de compréhension en conception (Ross *et al.*, 2010).

Pour présenter l'exploration d'espace de conception (ou Trade Space Exploration, Trade faisant référence au compromis que doit souvent réaliser le concepteur), nous proposons de décrire le processus et les grandes phases de l'exploration et les outils existants pour appliquer cette méthode.

1.1.3.1 Déroulement du processus

Simpson *et al*. (2008) décrivent l'approche typique de l'exploration d'espace de conception en trois phases principales : construire le modèle, exécuter les expérimentations et explorer (voir **Figure 10**). Ici, l'activité « explorer » englobe l'optimisation. Par ailleurs, la vision proposée par Simpson et ses collègues leur permet principalement de présenter et de situer l'outil ARL Trade Space Visualizer. Dans nos travaux, nous avons une vision centrée sur le concepteur en Design by Shopping. Nous cherchons donc une illustration sensiblement différente du processus.

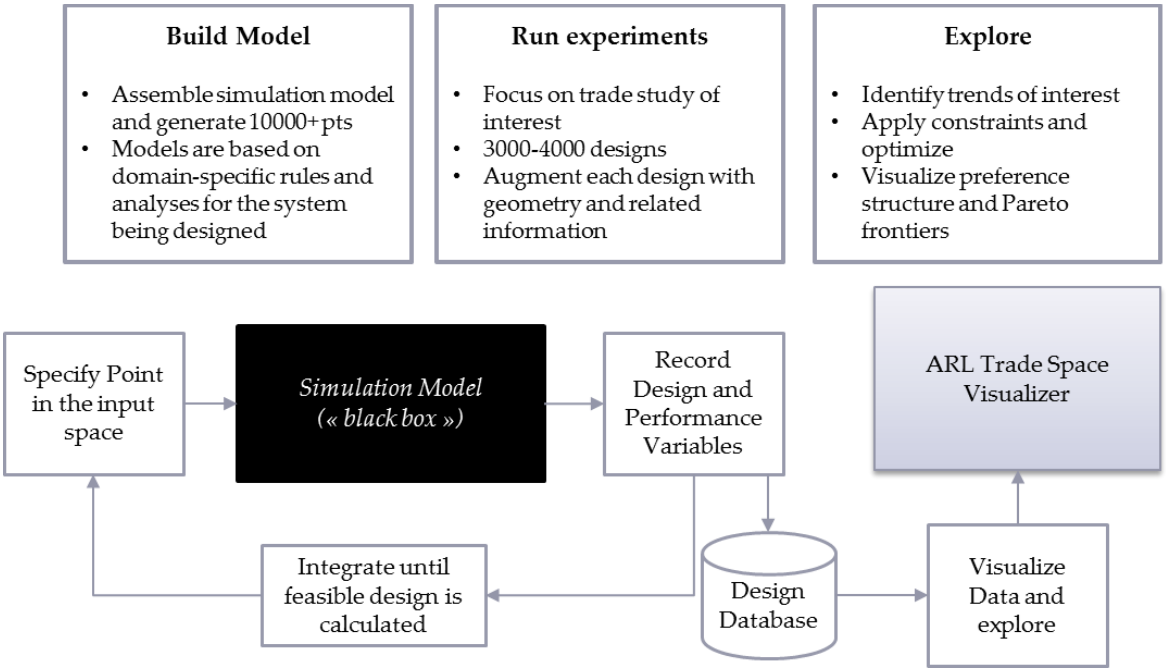


Figure 10 : Illustration de l'approche typique de l'exploration d'espace de conception décrite par Simpson *et al*. (2008)

Yan *et al*. (2012) proposent une approche « work-centered » (voir **Figure 11**) qui met l'emphase sur l'intégration d'algorithmes computationnels orientés données avec un processus de conception centré utilisateur. Dans leur vision, la visualisation est l'interface entre les données, les modèles et l'utilisateur. Leur approche s'inscrit dans l'analyse visuelle (Visual Analytics) dans laquelle ils ajoutent la manipulation d'algorithmes par interactions directes sur le graphique pour l'exploration d'espace de conception.

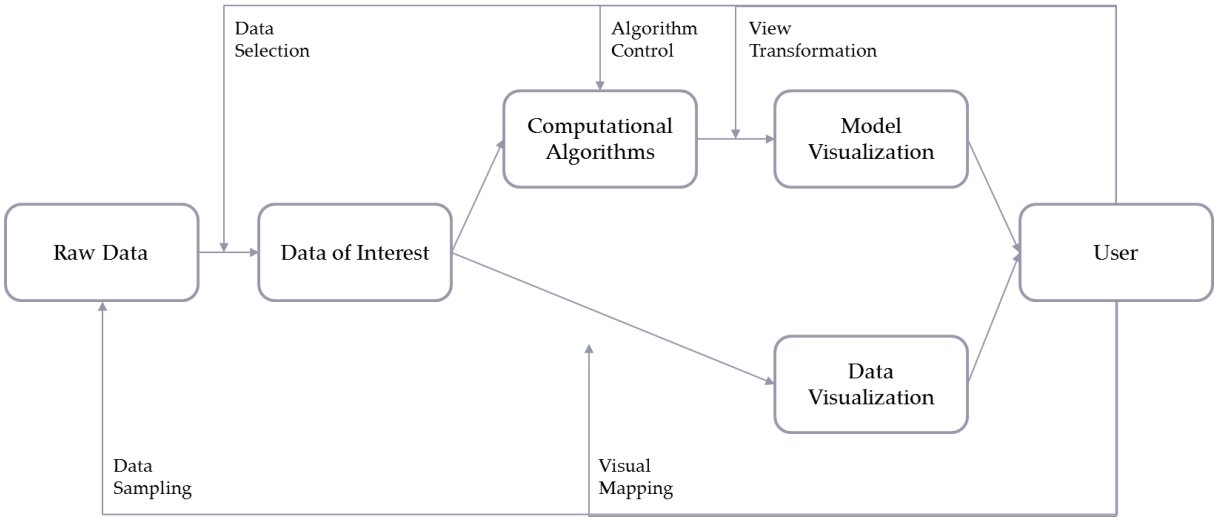


Figure 11 : Illustration de l'approche « work-centered » de Yan *et al.* (2012)

En s'appuyant sur les exemples et les deux approches d'exploration d'espace de conception précités, nous illustrons, sous forme d'une SADT (*Structured Analysis and Design Technique*), le processus d'exploration d'espace de conception en nous focalisant sur l'activité du concepteur (voir **Figure 12**). Trois phases principales sont identifiées : la découverte, l'optimisation et la sélection. La construction du modèle par le concepteur n'apparait pas dans notre représentation car elle est pour nous une activité à part entière en amont du processus d'exploration d'espace de conception (nous pouvons remarquer « modèle de données » en entrée de la SADT).

Plus précisément, l'exploration d'un espace de conception en Design by Shopping implique en variables d'entrée :

- Un niveau de connaissance du concepteur
- Un modèle de données
- Une question qui est le problème à résoudre.

Il y a plusieurs moyens de contrôle :

- L'échantillonnage qui permet de générer des solutions dans l'espace de conception
- Le graphique qui permet une représentation de l'espace et des points de conception
- Les outils d'optimisation qui permettent les calculs de performances
- Le contrôleur de contraintes qui permet de réduire l'espace de conception
- Les interacteurs qui permettent la transformation de la représentation de l'espace de conception

Enfin, l'exploration d'un espace de conception comporte des variables en sortie :

- Le « niveau de connaissance ++ » qui correspond aux connaissances acquises durant le processus
- La réponse qui correspond à la solution apportée par le concepteur

Dans notre cas nous utilisons (pour la suite de nos travaux) comme support la plateforme développée pour l'expérimentation.

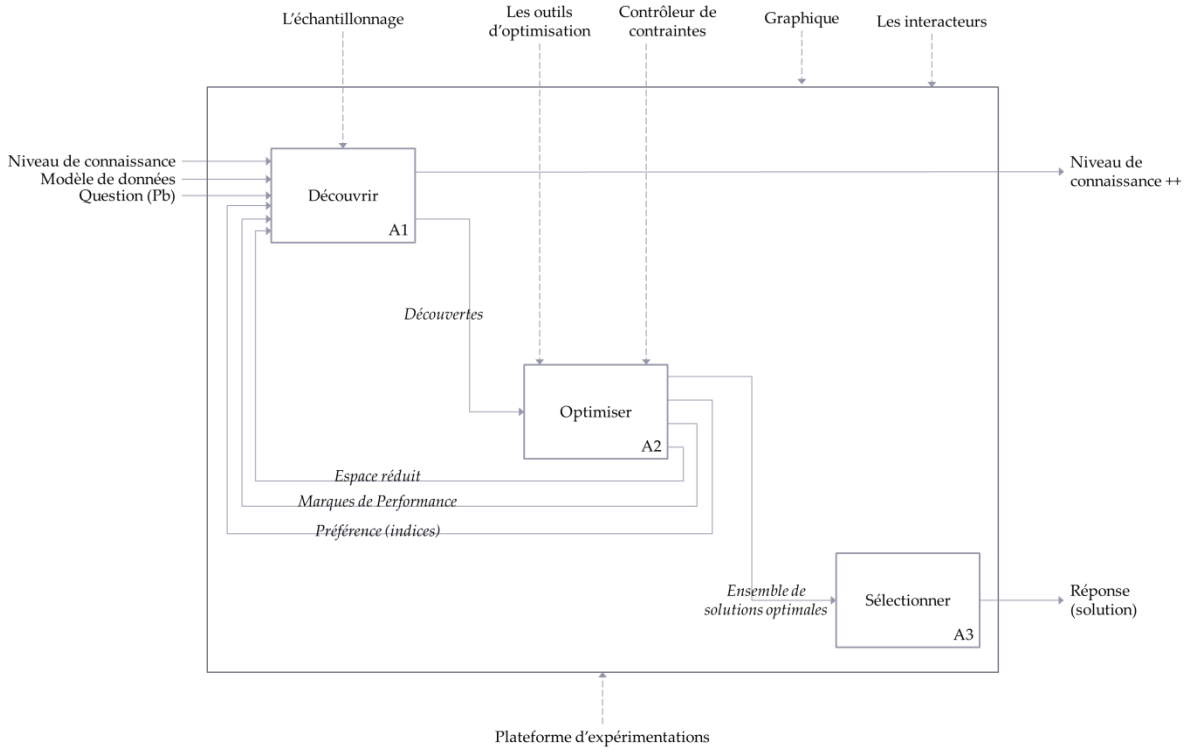


Figure 12 : SADT A0 Explorer un espace de conception

L'utilisation de la méthode d'exploration d'espace de conception avec l'ARL Trade Space Visualizer par un ingénieur automobile est décrite dans les travaux de Stump *et al*. (2009). La description de ce cas (configuration et conception d'un véhicule automobile) nous permet d'identifier des micro-situations auxquelles doit faire face le concepteur.

Nous illustrons dans la **Figure 13** les sous-activités repérées dans (Stump *et al*., 2009) et que nous appelons des micro-situations. En étendant donc nos trois phases du processus – découverte, optimisation, sélection -, nous trouvons ce qui suit.

« La découverte » rassemble plusieurs « micro-situations », à savoir la compréhension du problème en observant les interactions entre les variables d'un point de vue global, l'identification d'une zone de performance, et l'interaction entre les variables d'un point de vue local, c'est-à-dire centré sur les solutions dites optimales.

« L'optimisation » se compose également de « micro-situations » telles que la réduction de l'espace de conception vers une zone de performance en s'appuyant sur un vecteur de

préférence et/ou sur le sous-ensemble de points appelés Pareto-optimals et l'analyse de sensibilité des variables en vue de limiter l'espace de conception à un ensemble de solutions performantes.

Les micro-situations de la « sélection » sont la comparaison de solutions performantes afin d'identifier celles qui sont un bon compromis ; et la sélection d'une solution qui réponde à des critères et/ou préférences. Dans la première des deux micro-situations, il s'agit de comparer les solutions entre elles alors que dans la deuxième il s'agit de comparer une solution parmi les autres sur l'ensemble des critères.

En observant les micro-situations de la phase de sélection, il apparait que cette phase peut être traitée en situation MADM ou en situation MODM. En effet, cette phase peut être réalisée suite au deux premières i.e. dans un espace continu et un nombre quasi-infini de solutions (MODM). Elle peut également être réalisée par une personne supplémentaire (« supra-décisionnaire ») i.e. déconnectée des deux premières phases lorsque l'espace est défini performant avec un nombre limité de solutions qui nécessitent des comparaisons inter et intra attributs (MADM).

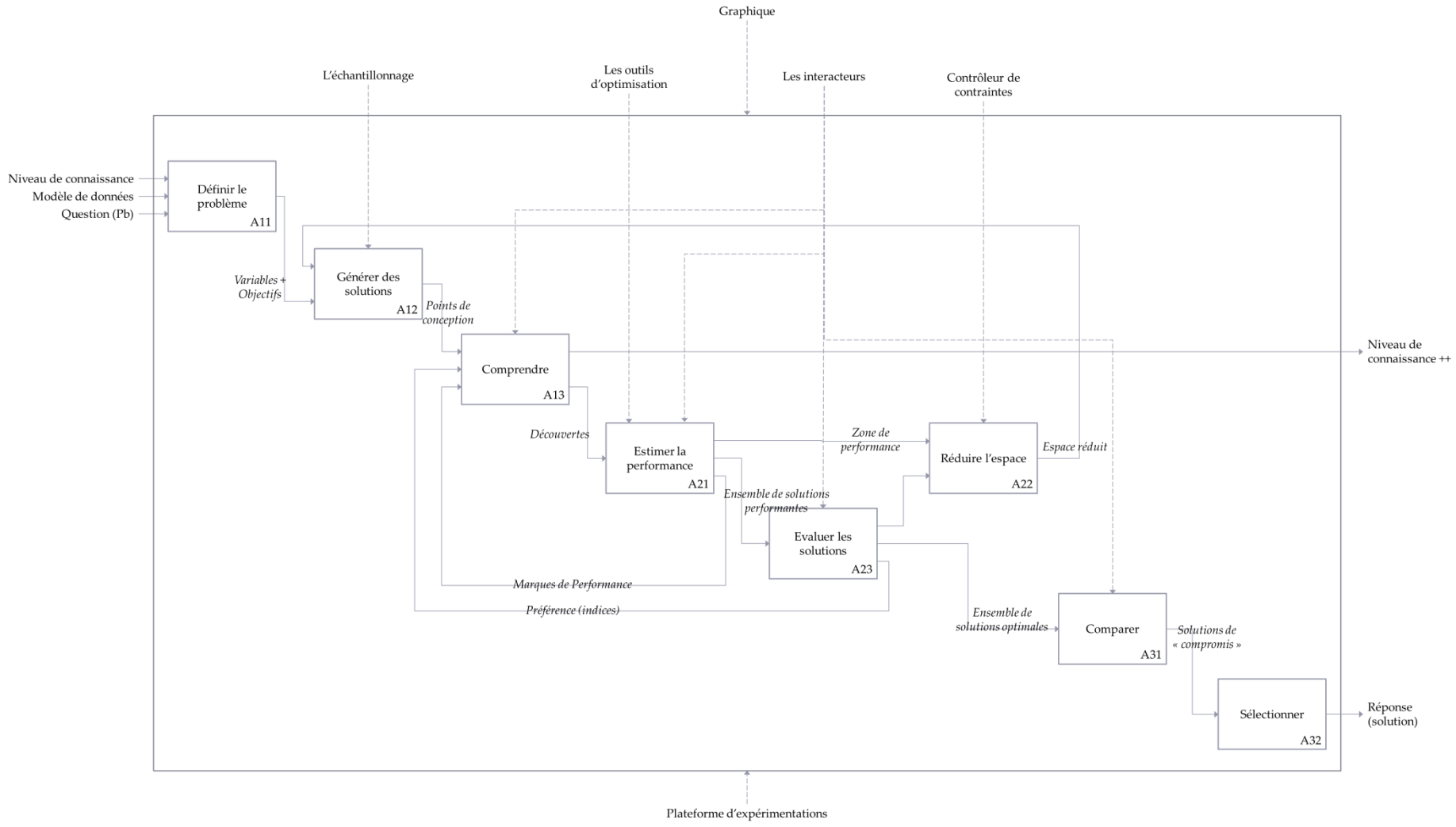


Figure 13 : SADT combinant A1 Découvrir, A2 Optimiser et A3 sélectionner

1.1.3.2 Les outils existants

Les supports graphiques sont efficaces pour l'optimisation en ingénierie (Barron *et al.*, 2004) et la conception conceptuelle (Yannou *et al.*, 2005).

Il a déjà été montré que les interfaces graphiques impactent la performance utilisateur en conception et lors du processus de recherche en conception (Ligetti *et al.*, 2003). Il existe des outils pour l'exploration de l'espace de conception avec différents graphiques, notons : l'ARL Trade Space Visualizer (Stump *et al.*, 2004), l'outil VIDEO (Kollat et Reed, 2007) et l'outil LIVE (Yan *et al.*, 2011, 2012).

L'ARL Trade Space Visualizer (Stump *et al.*, 2004) permet de visualiser sur le même graphique les paramètres de conception et les variables de performance avec Scatter Plot Matrix, Parallel Coordinates Plot et des graphiques à barres (**Figure 14**). Nous notons que les travaux de Wolf *et al.* (2009 ; 2011), qui comparent l'utilisation de l'outil par des experts et des novices en conception, montrent que les graphiques à barres ne sont, à priori, pas utilisés par les concepteurs.

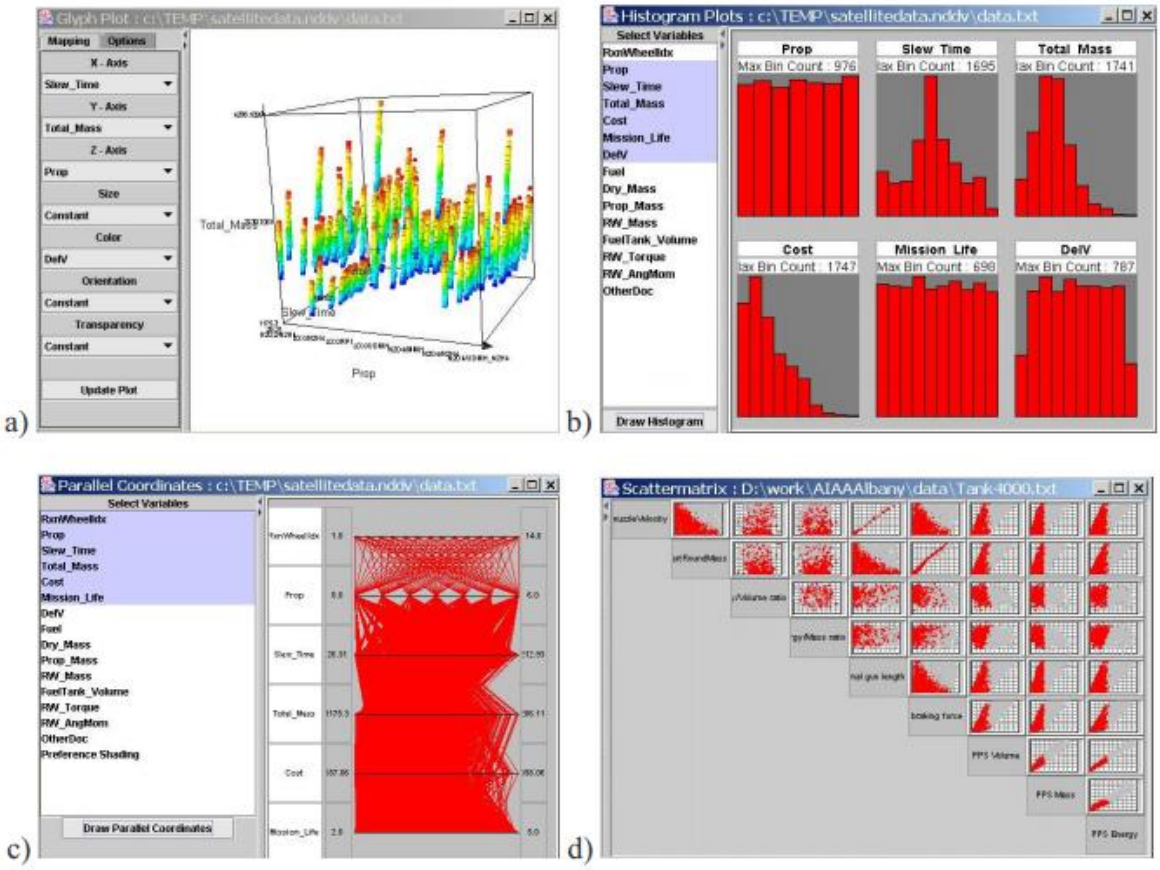


Figure 14 : Impressions d'écran de l'ARL Trade Space Visualizer: a) Glyph Plot, b) Parallel Coordinates, d) Scatter Matrix.

L'outil VIDEO de Kollat et Reed (2007) permet de visualiser séparément sur deux graphiques l'espace objectif (c'est à dire des variables de performance) et l'espace de décision (c'est à dire

des paramètres de conception) avec un 3-D Scatter Plot utilisant des glyphes (marqueurs) de couleur et de taille (**Figure 15**). VIDEO signifie *Visually Interactive Decision-making and Design using Evolutionary Multi-objective Optimization*.

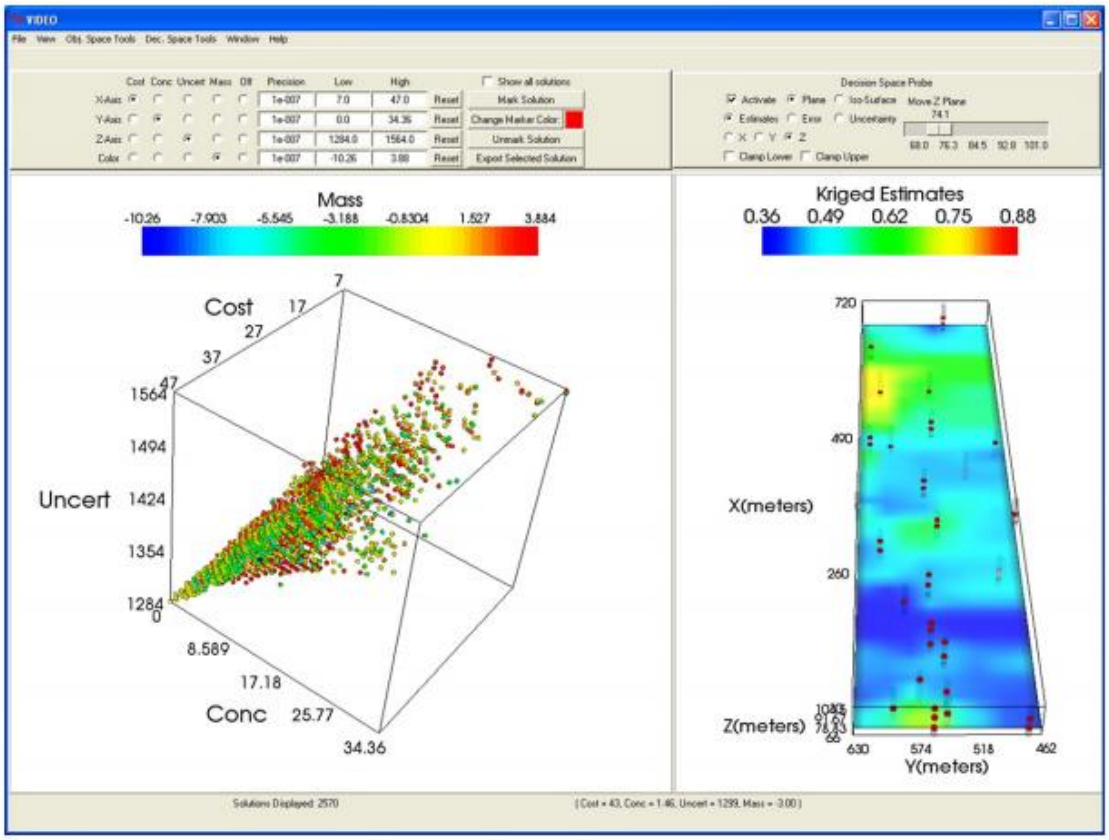


Figure 15 : Impression d'écran de l'outil VIDEO

L'outil LIVE (Yan *et al.*, 2011, 2012) permet de visualiser, dans une première partie, sur deux Scatter Plot les paramètres de conception (**Figure 16 b.**) et les variables de performance (**Figure 16 a.**), et dans une deuxième partie, un treemap représentant la classification d'une combinaison de variables d'entrée (**Figure 16 d.**). Sur les deux Scatter Plot il est possible d'ajouter des marqueurs de couleur et de forme pour la visualisation de clusters par exemple.

Un « cluster » est un groupement de solutions (ou d'alternatives) homogènes, en ce sens que les solutions partagent des caractéristiques communes qui, par exemple, correspondent à un critère.

Dans la **Figure 16 d.**, la structure hiérarchisée est présentée dans un treemap (Shneiderman, 1992) dans lequel chaque nœud à un attribut d'entrée avec une valeur de « splitting ». Chaque feuille de l'arbre (i.e. les nœuds sans enfants) spécifie la valeur de sortie (avec une échelle de couleur) comme une conséquence de la valeur particulière de l'entrée décrite par le chemin de la racine jusqu'à la feuille.

Figure 16 : Impressions d'écran de l'outil LIVE: (a) scatter plot des variables de performance; (b) scatter plot des paramètres de conception (montrant des clusters de résultats); (c) règles de décision; (d) treemap de l'arbre de décision; (e) sliders ajustables des intervalles des variables.

1.1.4 L'essentiel du Chapitre 1

Le chapitre 1 positionne nos travaux de recherche dans le domaine de l'exploration d'espace de conception. L'exploration d'espace de conception est une méthode d'optimisation multicritère de conception (multi-attribut ou multi-objectif) en Design by Shopping.

Le processus de conception, selon cette méthode, est décomposé en trois phases principales : la découverte, l'optimisation et la sélection. Les deux premières phases s'exécutent de concert jusqu'à la phase de sélection. Cette dernière montre une particularité car elle peut être réalisée sans nécessairement être précédée de la découverte et l'optimisation i.e. par une autre personne : « supra-décisionnaire ». Ainsi, nous faisons le constat que la phase de sélection peut être dans une situation multi-objectif (intégré dans le processus complet) ou dans une situation multi-attribut (indépendante des deux premières phases). Chacune des trois phases est composée de micro-situations propres auxquelles doit faire face le concepteur.

Nous avons vu que cette méthode d'exploration d'espace de conception s'applique à travers des outils graphiques notamment pour supporter le Design by Shopping (gain d'insights, génération d'hypothèses, etc.). Les outils identifiés présentent différents moyens de représenter l'espace de conception (différents graphiques).

Le Design by Shopping est un paradigme dans lequel le concepteur est replacé directement dans la boucle de simulation (i.e. une place importante dans le processus). Par ailleurs, ce paradigme implique que le concepteur aboutisse à une décision que nous proposons comme une décision informée.

1.2 DECISION INFORMEE

Le paradigme de Design by Shopping implique que le concepteur réalise sa décision à partir de l'information acquise durant le processus de conception. C'est ce que nous appelons la décision informée.

Il est difficile de trouver une définition claire de la « décision informée » dans la littérature scientifique. Cependant nous retrouvons l'emploi de ce terme dans certains articles. Nous proposons ainsi d'identifier les entrées / sorties et les mécanismes liés à la décision informée afin de construire notre propre définition.

Nous présentons la décision d'un point de vue global puis nous dressons un état de l'art de la décision informée. Nous poursuivons avec la troisième section de ce chapitre qui fait émerger les thématiques et les éléments clés qui soutiennent la décision informée. Enfin, nous soulignons l'importance du terme « décision informée » avant de conclure avec notre définition de la formulation d'une décision informée.

1.2.1 La décision

Au préalable, il nous parait opportun et bienvenu de définir la décision avant de s'attaquer au cas particulier d'une décision informée. Nous proposons, pour la première section de ce chapitre, de présenter plusieurs définitions de la décision issues de différents domaines de recherche (Voir **Tableau 1**). Les différentes définitions ont été choisies en considérant la « notoriété » des auteurs par rapport à ce sujet (la décision).

Tableau 1 : Définitions de la décision, chronologie et thématiques

Auteur	Année	Définition	Domaine
Scott Morton	1971	La décision est assimilée à un processus dans le temps avec des changements	Economie, gestion des entreprises, science de gestion
Le Moigne	1974	Décider, c'est "identifier et résoudre les problèmes que rencontre toute organisation"	SHS, génie industriel, informatique
Mintzberg	1979	Une décision, qu'elle soit individuelle ou résultant d'un travail de groupe, peut être définie comme l'engagement dans une action, c'est à dire, une intention explicite d'agir	Economie, gestion des entreprises, science de gestion
Lévine & Pomerol	1989	Une décision est une action qui est prise pour faire face à une difficulté ou répondre à une modification de l'environnement, c'est à dire, pour résoudre un problème qui se pose à l'individu ou à l'organisation.	Informatique
Roy & Bouyssou	1993	la décision est souvent présentée comme le fait d'un individu isolé (décideur) qui exerce librement un choix entre plusieurs possibilités d'actions à un moment donné dans le temps.	Economie, gestion des entreprises, science de gestion
Kast	1993	La décision est synonyme de "choix entre plusieurs solutions existantes, ayant chacune des conséquences différentes, le choix sera fait selon des critères précis de sélection".	Economie, gestion des entreprises, science de gestion
Schneider	1994	La décision est un choix entre plusieurs alternatives.	Economie, gestion des entreprises, science de gestion
Barabel	1996	La décision correspond à "l'ensemble des actions que le décideur effectue pour prendre une décision et la mettre en œuvre".	Economie, gestion des entreprises, science de gestion
Stal le Cardinal	2000	"Une décision est un processus qui conduit un acteur à répondre à une question posée"	Système industriel, génie industriel, conception, design
Berchet	2000	"Une décision est une sélection de l'éventualité dont on pense qu'elle permettra d'obtenir les	Système industriel, génie industriel,

		avantages optimaux et les inconvénients minimaux".	conception, design
Bellut	2002	"Décider, c'est choisir l'action à entreprendre, c'est-à-dire choisir une action parmi toutes celles qui sont possibles. Les actions sont aussi appelées des solutions ou des options".	Economie, gestion des entreprises, science de gestion
Longuevil le	2003	"La décision est un processus de transformation d'informations. Il conduit un acteur ou un groupe d'acteurs de l'organisation à répondre à une question donnant lieu à une action (par la réponse)".	Système industriel, génie industriel, conception, design
Dumont d'Ayot	2005	La prise de décision. Cela concerne un choix parmi plusieurs buts, plans, actions possibles, notamment lors de situations nouvelles (adaptation), lors de situations de danger (point de focus, décision en temps contraint), et en cas d'erreur (réparation, retour-arrière).	Système industriel, génie industriel, conception, design
Seguy-Garcia	2008	une décision a pour objectif de résoudre un problème, elle doit se concrétiser par une action, elle peut être individuelle ou collective et, parfois, distribuée, elle peut être construite à partir de solutions existantes partielles parmi lesquelles le décideur choisit une solution pouvant répondre au problème posé.	Système industriel, génie industriel, conception, design
Adla	2010	« nous appellerons prise de décision tout processus mental à la suite duquel tout individu, placé devant plusieurs alternatives, choisit l'une d'entre elles. »	Informatique

Au vu de ces différentes définitions, nous relevons que la décision implique :

- Un processus de transformation d'information
- Un choix selon des critères de sélection
- Un choix entre plusieurs alternatives
- Les avantages optimaux et les inconvénients minimaux
- Une réponse à une question posée (situation nouvelle, danger, un problème, etc.)

Ainsi, nous définissons la décision comme étant un processus de transformation d'information. Il s'agit d'une réponse à une question, de la sélection d'une alternative parmi plusieurs qui permettra d'obtenir les avantages optimaux et les inconvénients minimaux.

Notons également que la notion de décision ne se limite pas à la simple tâche de prise de décision mais bel et bien à l'ensemble des tâches qui aboutit à la prise de décision. Nous parlons alors de processus décisionnel.

1.2.2 La décision informée : Etat de l'art

Nous avons relevé seize articles utilisant le terme « décision informée » dans divers champs disciplinaires impliquant tout de même l'analyse de grands ensembles de données. Nous avons écarté le champ des soins et de la santé qui utilise ce terme dans un sens légèrement éloigné du notre (le consentement d'un patient à suivre un soin/traitement ou non).

1.2.2.1 Monitoring et supervision

Dans le champ du monitoring & supervision, Ireson (2009) mentionne qu'il est crucial de gérer la masse d'information afin d'obtenir la « situation consciente » et prendre une décision informée. Bass (2000), dans le même esprit, indique le besoin de la « situation consciente » pour la formulation d'une décision informée. Il ajoute la nécessité de fusionner les données en information et en connaissance pour que les opérateurs prennent des décisions informées. Riveiro et ses collègues indiquent également la « situation consciente » dans le cadre de prise de décision informée (Riveiro *et al*., 2008).

"*The management of this mass of information is crucial in aiding the decision-making process, ensuring, as far as possible, that* ***the responders have full situational awareness to make informed decisions.***" (Ireson, 2009)

"*fusing data into information and knowledge, so network operators can make* ***informed decisions*** " (Bass, 2000)

Il convient ici d'établir la différence entre la « donné » et l' « information » :

- Pour Mennis et ses collègues les données sont "*observational measurements that have been recorded in some way, whereas information is data that is generalized, ordered and contextualized in ways that give them meaning. Information thus is selective toward data, separating the important from the relatively unimportant*." (Mennis *et al*., 2000)
- Dans la même idée, Bertini et Lalanne décrivent les données comme " *a collection of facts usually collected by observations, measures or experiments. Data consist of numbers, words, or images. It is generally called abstract data in infovis, since it refers to data that has no inherent spatial structure enabling further mapping to any geometry*." (Bertini et Lalanne, 2009)

Ainsi, les données sont de simples faits, dépourvus de tout contexte. Si les données ne nous informent pas, ce n'est pas l'information.

1.2.2.2 Business et marketing

En business / marketing, Lurie et Mason (2007) évoquent, eux aussi, la gestion d'un grand ensemble de données et la disposition d'outils de visualisation qui peuvent entrainer une décision informée (Lurie & Mason, 2007). Glaser et Tolman (2008) relient la décision

informée à un processus d'analyse de grands ensembles de données, de « tracking » de performance et de détection des formes - patterns - et de tendances.

"*Business intelligence is the process of analyzing large amounts of corporate data, usually stored in large databases such as a data warehouse, tracking business performance, detecting patterns and trends, and helping enterprise business users make* ***more informed decisions***." (Glaser & Tolman, 2008)

1.2.2.3 Système d'information

Dans le champ des systèmes d'information, Wiederhold mentionne l'application d'une variété de connaissances en de l'information.

« ***The making of informed decisions*** *requires the application of a variety of knowledge to information* "(Wiederhold, 1992)

1.2.2.4 Construction

Dans le domaine de la construction, Petersen et Svendsen démontrent que la prise de décision de conception informée nécessite de gérer une grande quantité d'information sur les propriétés détaillées des options de conception et d'opérer des simulations de leur performance. Pour eux, le concepteur a besoin d'un large espace de conception et d'une vue d'ensemble des conséquences de variations des paramètres pour l'aider à acquérir une compréhension profonde des performances et ainsi prendre une décision de conception informée (Petersen & Svendsen, 2010). Russell et ses collègues considèrent que le « visual analytics » permet de meilleures décisions informées :

"*Visual analytics, the science of analytical reasoning facilitated by interactive visual interfaces, has the potential to improve the construction management process through the enhanced understanding of project status and reasons for it, better* ***informed decision*** *making, and improved communication amongst project participants.*" (Russell *et al.*, 2009)

1.2.2.5 Visualisation d'information

Dans le champ de la Visualisation d'information Keim *et al.* indiquent, en 2006, le fait qu'il est indispensable que l'humain soit inclus dans le processus.

"*For* ***informed decisions****, it is indispensable to include humans in the data analysis process to combine flexibility, creativity, and background knowledge with the enormous storage capacity and the computational power of today's computers.*" (Keim *et al.*, 2006)

En 2008, Keim *et al.* couvrent une surface plus large de l'interaction entre l'humain et le système de « visual analytics » qui, d'après eux, est le système permettant de prendre une décision informée.

"*Visual Analytics systems need to be easily used and interacted with by the analyst. The analyst needs to be able to fully focus on the task at hand, not on overly technical or complex user interfaces, which potentially distract. To this end, novel interaction techniques need to be developed which fully support the seamless, intuitive visual communication with the system. User feedback should be taken as intelligently as possible, requiring as little manual user input as possible, which guarantees the full support of the user in navigating and analyzing the data, memorizing insights*[2] *and* ***making informed decisions***." (Keim *et al*., 2008a)

Meyer *et al*. soulignent la nécessité d'un couplage étroit entre la cognition, l'interaction et le « visual analytics » pour que l'utilisateur puisse prendre une décision informée.

"*A tight coupling between cognition, interaction and visual analytics is necessary to enable the user to make* ***informed decisions***" (Meyer *et al*., 2010).

La cognition, pour eux, est l'action mentale ou le processus d'acquisition de connaissance et de compréhension à travers la pensée, l'expérience et les sens. Elle entrecoupe les trois niveaux hiérarchiques perception / décision / action.

En 2008, Keim et ses collègues proposent une modélisation du "visual analytics" (voir **Figure 17**) qui, entre autre, indique des représentations visuelles qui permettent la découverte de patterns robustes qui forment la connaissance utilisée dans la prise de décision informée.

"*Our work definition looks at visual data mining as the process of interaction and analytical reasoning with one or more visual representations of an abstract data that leads to the visual discovery of robust patterns in these data that form the information and knowledge utilized in* ***informed decision making****. The abstract data can be the original data set or/and some output of data mining algorithm(s).*" (Keim *et al*., 2008b)

Dans la **Figure 17** nous trouvons:

- *S* pour *Source* représente l'ensemble de données. *Dw* représente les fonctions de prétraitement des données.
- *I* pour *Insights* représente l'objectif du processus
- *V* pour *Visualization. Uv* représente l'exploration visuelle des données.
- *H* pour *Hypothesis. Uh* correspond au « data mining ».
- *U* pour *User* représente les interactions utilisateurs. *Ucv* sont les insights conclus à partir de la Visualisation et *Uch* sont ceux conclus à partir des hypothèses.
- Vh représente les hypothèses vérifiées avec la visualisation.
- Hv représente les hypothèses générées à partir de la visualisation.

[2] Tel que défini par Saraiya *et al*. « individual observation about the data by the participant, a unit of discovery » (Saraiya *et al*., 2005)

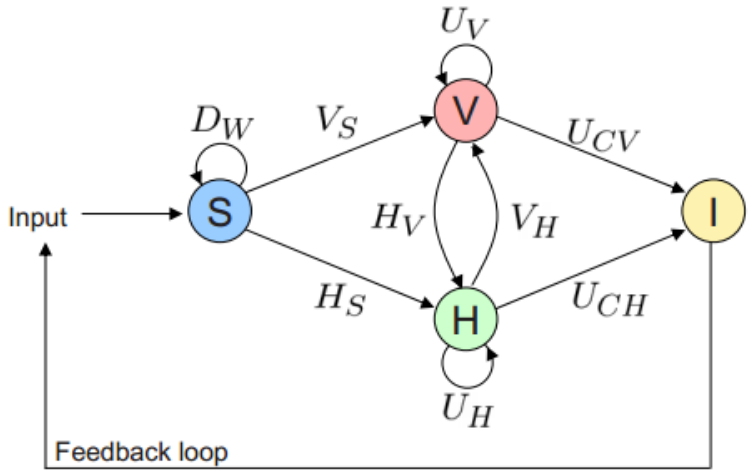


Figure 17 : Processus de « Visual analytics » de Keim *et al.* (2008b)

Riveiro *et al*. utilisent la modélisation de Keim *et al*. dans le cadre de détection de comportement maritime anormal. Ils soulignent que "***Human informed decision-making*** *can be seen as a complex information processing task*." (Riveiro *et al*., 2008)

1.2.2.6 Ingénierie de la conception

Enfin, nous retrouvons le terme de décision informée dans le champ de l'ingénierie de la conception.

Wood *et al*. indiquent, dans leur recherche, que lors des phases de conception préliminaire, la manipulation des données telles que les incertitudes permet aux concepteurs de prendre des décisions mieux informées.

"*We model and manipulate such uncertainties in a computer-assisted environment, under the hypothesis that doing so will allow the designer to make faster and* ***more-informed decisions***." (Wood *et al*., 1992)

Sullivan et ses collègues ont trouvé que pour prendre une décision informée, il y a un besoin de savoir comment des changements dans l'environnement peuvent affecter les paramètres de conception (sensibilités).

"*In particular, to make* ***informed decisions*** *about the choice of design rules and clustering of design parameters, we found we needed to know how changes in the environment would affect them.*" (Sulllivan *et al*., 2001)

Chandrasegaran *et al*. indiquent que « *An effective computer support tool that helps the designer make* ***better-informed decisions*** *requires efficient knowledge representation schemes*" (Chandrasegaran *et al*., 2013). Le processus de transformation des données en information et par la suite en connaissance est crucial. De plus, il est essentiel d'équiper les concepteurs

d'outils proposant des représentations du problème dès le début du processus de conception pour prendre des décisions mieux informées.

Pour finir, Mavris et ses collègues avancent que l'intégration du « visual analytics » dans le processus de conception apporte aux concepteurs la capacité de gagner la connaissance et les *insights* nécessaires pour prendre des décisions informées. Pour eux, la visualisation est essentielle pour faciliter la génération d'hypothèses et la formulation d'une décision informée. Ils ajoutent que : "*problem-solving activity requires that the data, knowledge, and insight necessary for the formulation of* ***informed decisions*** *be generated throughout the design process*."

L'un des points supplémentaires écrit dans leur papier est :"*to be able to provide the decision maker with the justified means of making an* ***informed decision***." (Mavris *et al*., 2010).

1.2.2.7 Synthèse

De cet état de l'art, nous retenons des éléments clés qui conduisent à la formulation d'une décision informée :

Le décisionnaire : Il est nécessaire que l'humain (pour nous le concepteur) soit inclus dans le processus.

L'activité : la décision informée nécessite au préalable une activité d'analyse et de traitement d'information (et/ou des données).

Les prérequis : la formulation d'une décision informée est réalisée suite à :

- La transformation de données en information et par la suite en connaissance
- Le gain de connaissance et d'insights tels que :
    - la détection de patterns et de tendances
    - La découverte de patterns robustes
    - Le tracking des performances
    - La sensibilité des paramètres (conséquence des variations)
- La compréhension du projet, du problème et des performances
- L'obtention d'une « situation consciente »

Les moyens : la formulation d'une décision informée est réalisée en s'appuyant sur des moyens :

- La visualisation d'information et de données pour l'analyse (visual analytics).
- La manipulation (interaction) : de données (ou *model*), des incertitudes et d'information à travers la navigation dans ces ensembles, la génération (et test) d'hypothèses et la simulation de performances.

1.2.3 Les thématiques gravitant autour de la décision informée

1.2.3.1 D'un point de vue général

Nous relevons dans la littérature que quatre thèmes (ou éléments clés) sont plus souvent cités en lien avec la décision informée : « *Knowledge and insights gain* », « *Visualization* », « *Analysis and treatment* » et « *Manipulation and management* » (voir **Figure 18**). Il semble que ces quatre éléments clés contribuent davantage à la formulation d'une décision informée.

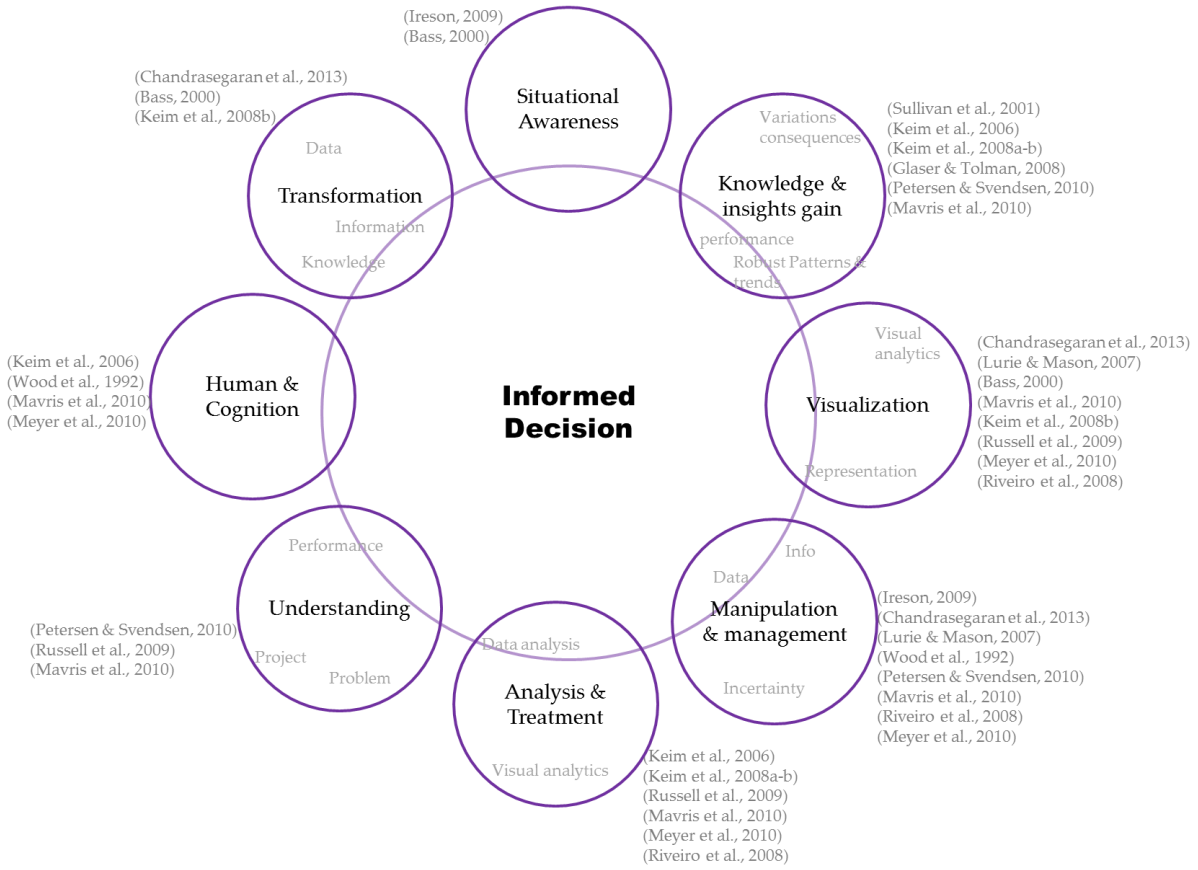


Figure 18 : Les éléments clés de la décision informée

1.2.3.2 Pour chaque champ disciplinaire

Des huit éléments clés identifiés ci-dessus (**Figure 18**), nous avons déjà relevé que certains d'entre eux semblent contribuer davantage à la définition d'une décision informée avec par exemple Keim qui se retrouve sur cinq éléments clés sur huit. Cependant, nous n'avons pas précisé la dominance des champs disciplinaires vis-à-vis de ces éléments.

La **Figure 19** illustre le nombre de papiers mentionnant les éléments clés par champs disciplinaires (champs déjà indiqués dans la deuxième section du chapitre).

Notons par exemple que « *situational awareness* » est exclusivement employé par le champ Supervizion & Monitoring. Nous décidons pour la suite de ne pas utiliser ce terme dans notre définition.

Nous avons vu précédemment les 4 éléments clés qui semblent contribuer les plus largement à la formulation d'une décision informée. Mais si on regarde où se trouve l' « engineering design » dans les éléments clés, nous relevons :

- *Knowledge & Insights gain*
- *Manipulation & Management*
- *Visualization*
- *Human & Cognition*

En comparaison, et bien qu'il semble très proche, nous relevons pour l'InfoVis (*Information Visualization*), une plus grande présence des éléments clés suivants :

- *Knowledge & Insights gain*
- *Visualization*
- *Human & Cognition*
- *Analysis & Treatment*

La différence s'explique par le besoin d'interagir avec les algorithmes et les modèles de données dans l'ingénierie de conception, spécificité déjà évoquée pour l'approche « work-centered » de Yan *et al.* (2012), **p.28**.

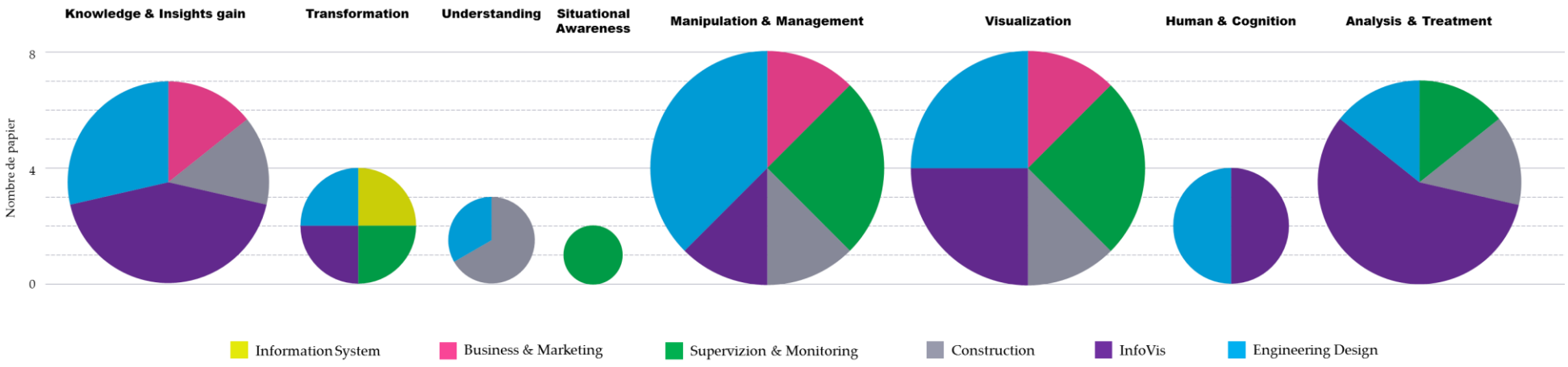


Figure 19 : Nombre de papier mentionnant les éléments clés par champs disciplinaires

1.2.4 Importance du terme « décision informée »

Afin de relever l'importance du terme « décision informée » pour les auteurs, nous relevons le nombre de fois qu'il est employé dans chaque papier (voir **Figure 20**).

Nous notons l'utilisation importante du terme pour Mavris *et al.* (2010) avec un nombre de 11, Petersen et Svendsen (2010) avec un nombre de 5 et Glaser et Tolman (2008) avec l'utilisation du terme à hauteur de 4 fois. Nous relevons également Keim qui l'utilise dans (Keim *et al.*, 2006), (Keim *et al.*, 2008a-b) et (Meyer *et al.*, 2010).

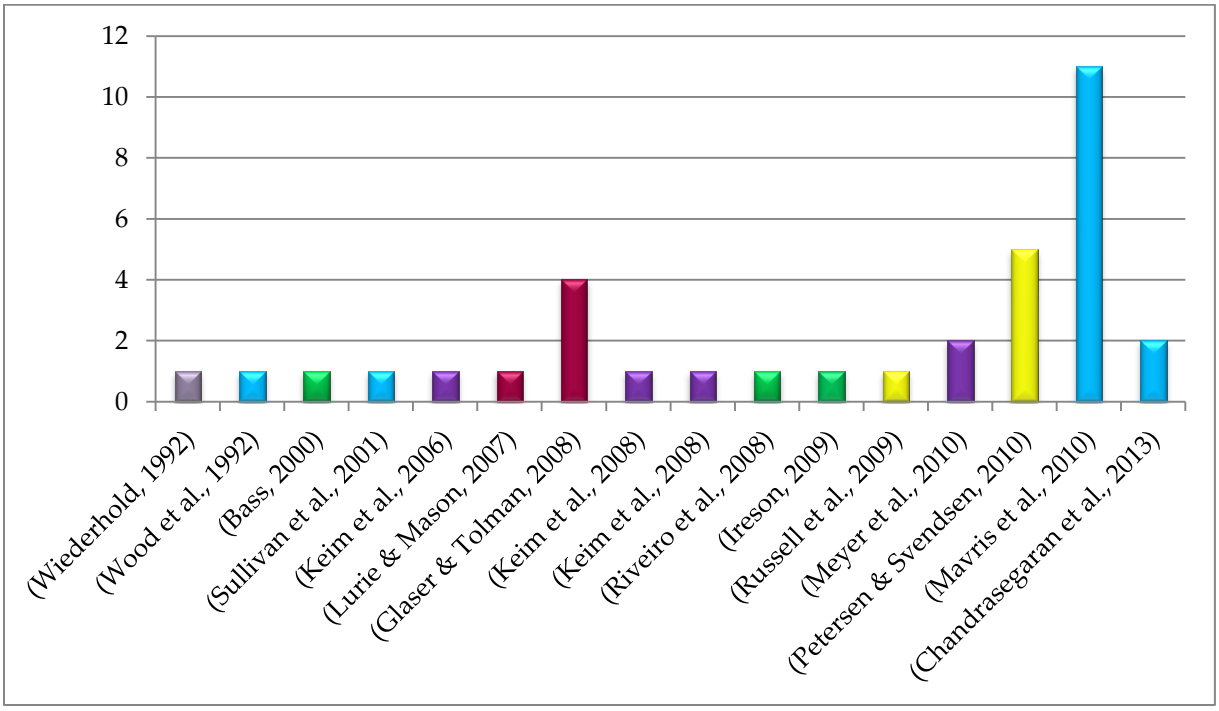


Figure 20 : Nombre de fois que le terme « informed decision » est utilisé dans chaque papier

Ainsi, nous donnons plus d'attention aux entrées/sorties et mécanismes cités par ces auteurs pour notre définition de la décision informée.

1.2.5 Notre définition de la décision informée

Afin de définir la formulation d'une décision informée, nous prenons donc en considération la définition de la décision énoncée dans la première section du chapitre en combinaison avec les éléments et les thématiques clés mis en lumière dans les sections précédentes.

Ainsi nous définissons la décision informée comme **une réponse à une question, l'élection d'une alternative ou un ensemble d'alternatives parmi plusieurs qui permettra d'obtenir les avantages optimaux et les inconvénients minimaux suite à un processus de traitement et d'analyse itératif et interactif durant lequel l'acteur gagne en compréhension, connaissance et « insights » grâce à la visualisation et la manipulation de grands ensembles et/ou modèle de données**.

1.2.6 L'essentiel du chapitre 2

Ce chapitre propose un état de l'art de la décision informée et permet d'en faire émerger notre propre définition.

Notre état de l'art, sans prétendre à l'exhaustivité, confirme la formulation d'une décision informé pour la conception en Design by Shopping notamment en soulignant la nécessité d'inclure l'humain dans le processus, l'activité d'analyse et de traitement de données et le besoin de transformation de données en information et par la suite en connaissance.

L'état de l'art nous conduit à dresser les 4 thématiques clés qui gouvernent la décision informée en ingénierie de la conception : *Knowledge & Insights gain*, *Manipulation & Management*, *Visualization*, *Human & Cognition*. Il est également souligné que ce type de décision se fait avec un sens relativement justifié par le décisionnaire.

Enfin, dans ce chapitre, nous faisons le constat que les moyens utiles à la décision informée sont la visualisation d'information et l'interaction (avec les données et/ou modèles).

1.3 LES REPRESENTATIONS GRAPHIQUES

Dans ce chapitre, nous présentons les caractéristiques des graphiques afin de déterminer ceux qui seraient utiles dans le processus d'exploration d'un espace de conception multicritère en Design by Shopping. Nous cherchons, par exemple, des graphiques qui puissent afficher une infinité ou du moins un très grand nombre d'alternatives (autrement dit des milliers de solutions possibles).

Nous proposons premièrement une introduction aux techniques pour l'exploration visuelle de données afin de nous positionner dans ce champ. Puis, nous présentons un état de l'art composé d'une partie décrivant des représentations de l'information pour des ensembles de données multidimensionnels, une deuxième partie présentant des comparatifs de ces représentations puis une dernière partie traitant des tâches réalisables avec ces représentations.

Pour clore ce chapitre, nous établissons un bilan qui nous permet de faire émerger les graphiques revêtant des caractéristiques utiles à notre contexte de conception multicritère.

1.3.1 La visualisation d'information

Comme synthétisé par Keim (2001), les techniques de visualisation d'information sont classifiées selon trois critères : les données qui sont à visualiser, la technique de visualisation et la méthode d'interaction et de distorsion (voir la **Figure 21**). Pour visualiser un type de données spécifique, il est possible de combiner les techniques de visualisation en conjonction avec les méthodes d'interaction et de distorsion.

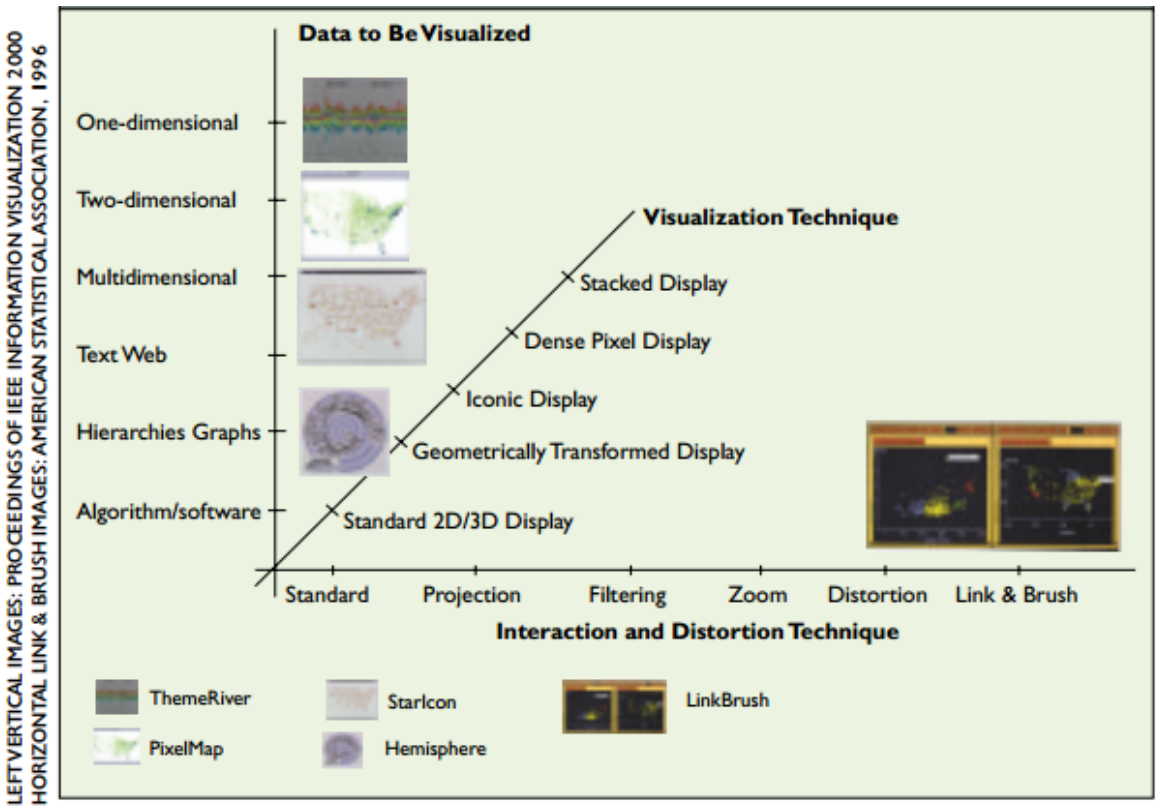


Figure 21 : Impression d'écran de « classification of visual data exploration techniques » dans (Keim, 2001)

Dans notre cas d'optimisation multicritère en conception, nous nous intéressons à des ensembles de données à N-dimensions (« *Multidimensional* » dans la **Figure 21**).

Cinq techniques de visualisation sont indiquées :

- « Standard 2D/3D Display » utilise les techniques standards de visualisation en deux dimensions et trois dimensions comme les nuages de points et les graphiques en barres.
- « Geometrically transformed displays » utilise des transformations et projections géométriques pour produire des visualisations utiles comme le Parallel Coordinate.
- « Icon-based displays » affiche chaque item de données comme un icône et chaque valeur de dimension comme une fonction de l'icône. (voir Chernoff's faces, Star-Diagram, Spider Graph, etc.)
- « Dense pixel displays » affiche chaque valeur de dimensions comme un pixel de couleur et les pixels appartenant à chaque dimension sont groupés en des zones

adjacentes. En organisant et colorant les pixels d'une manière appropriée, la visualisation résultante fournit des informations détaillées sur les clusters, et les dépendances (voir un exemple en **Figure 22**).

- « Stacked displays » affiche des données divisées d'une manière hiérarchique comme le Treemap.

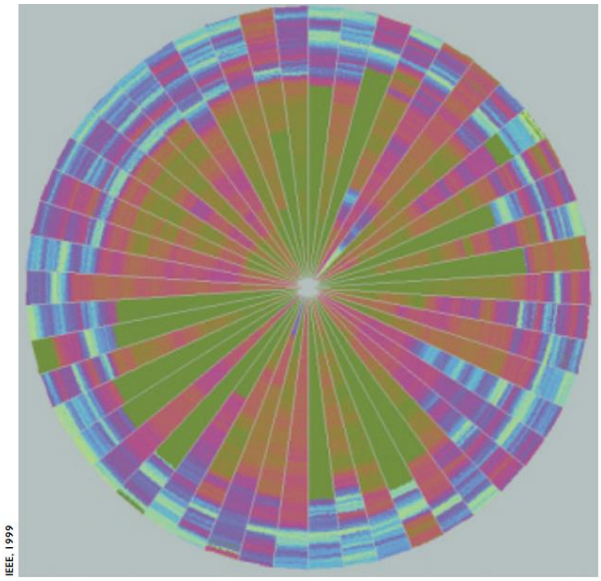


Figure 22 : Impression d'écran de « The pixel-oriented circle segment technique » dans (Keim, 2001)

Keim (2001) propose également des méthodes d'interactions et de distorsions :

- Méthode standard comme la sélection
- La projection qui est une technique utile pour un ensemble de données multidimensionnel (changement dynamique des projections).
- Le filtre qui permet une division interactive de l'ensemble de données soit en choisissant un sous-ensemble désiré (« *browsing* ») soit en spécifiant les propriétés du sous-ensemble désiré (« *querying* »).
- Le Zoom (spatial et sémantique)
- La distorsion qui est une technique permettant d'obtenir une partie de détail tout en gardant une vue d'ensemble (le *fish-eye* par exemple).
- Le lien interactif et brossage (*brush*) qui permet la sélection d'un ensemble d'items dans une vue et la mise en lumière de cette sélection dans d'autres vues.

Toutes les méthodes d'interactions susmentionnées sont réalisables avec tous types de représentations de données.

Comme nous l'avons énoncé plus haut, nous nous intéressons plus précisément à la représentation d'ensembles de données multidimensionnels car nos travaux portent sur l'optimisation multicritère en conception. Ainsi, nous proposons de détailler les représentations utiles à notre cas (le Scatterplot par exemple) dans la section suivante.

1.3.2 Les représentations des ensembles multidimensionnels

Nous présentons dans cette section une liste non-exhaustive des différentes représentations d'ensembles de données à N-dimensions.

1.3.2.1 Le graphique en barres

Le graphique en barres ou diagramme en barres ou histogramme est une représentation utilisant une technique dite « Standard 2D/3D Display ». Il y a plusieurs possibilités pour l'affichage des alternatives (i.e. points de conception) et des critères (i.e. paramètres de conception ou variables de performance). Un groupe de barres représentent les valeurs des alternatives d'un seul critère (voir **Figure 23**). L'inverse peut être réalisé, c'est-à-dire un groupe de barres représentant les valeurs des critères pour une seule alternative. Il est possible d'afficher les barres à l'horizontale ou à la verticale et même dans une projection 3D. Notons cependant l'emploi d'une seule échelle (si les variables n'ont pas la même échelle, le graphique devient difficilement réalisable). D'après Miettinen (2014), ce graphique est limité en nombre d'alternatives et de critères pour rester lisible.

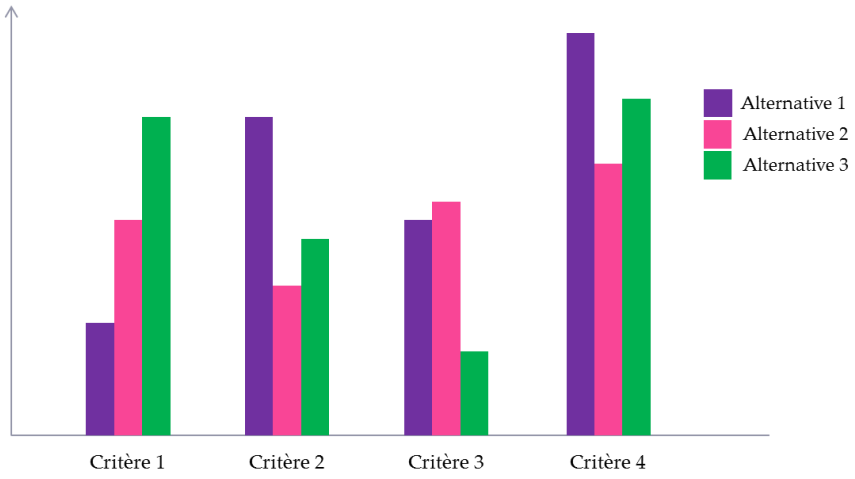


Figure 23 : Graphique en barres

Une autre utilisation répandue des graphiques en barres est la visualisation de la distribution des alternatives sur une variable.

1.3.2.2 Le Scatterplot (nuage de points)

1.3.2.2.1 Simple (x-y plot)

Un Scatterplot est la méthode conventionnelle pour visualiser la relation entre deux variables. Comme une structure visuelle, le Scatterplot utilise la position pour encoder les valeurs des deux variables et leur relation (voir l'encodage visuel de Bertin (1973) en annexe p.207). Il s'agit d'une projection des données (représentant les alternatives) dans un espace 2D. Siirtola (2007) considère que le Scatterplot est utile pour détecter facilement des patterns non-linéaires, des corrélations positives ou négatives entre les variables et les valeurs

aberrantes. Les Scatterplots sont des représentations Cartésiennes et ont donc une longue histoire. Il en résulte un fort développement d'intuitions[3] sur l'apparence de ce type de représentation (Wegman, 1990).

L'intuition gagnée à partir d'un Scatterplot est hautement dépendante du système d'axes choisis. En effet, il est aisé de passer à côté de structures de données réelles en examinant simplement des projections à dimensions plus basses (typiquement visualiser deux dimensions sur six). De plus, nous devons noter que ces « intuitions » sont à prendre avec précaution car elles « échouent » rapidement avec l'augmentation des dimensions. Par exemple, une illustration simple est l'intersection de deux plans dans un espace 4D. L'intuition tridimensionnelle normale est que l'intersection des deux plans soit une ligne alors que dans un espace 4D il s'agit d'un point.

Un ensemble de données multidimensionnel est souvent représenté comme une collection de « plots » aussi appelé « Scatterplot Matrix ».

1.3.2.2.2 Matrice

Le Scatterplot Matrix est une collection de Scatterplots Simples (x-y) ordonnés par paires. Cette représentation fournit une vue d'ensemble des données. La taille de la matrice dépend du nombre de critères (ou variables) visualisés. Les Scatterplots sont dupliqués dans la matrice par rapport à la diagonale (voir la **Figure 24**).

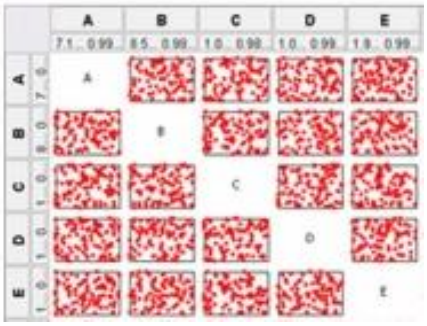


Figure 24 : Exemple de scatterplot matrix issue de (Stump *et al.*, 2009)

Il est possible d'utiliser des marqueurs de couleur ou de forme avec les Scatterplot afin d'ajouter un critère (appelé Symbolic-Scatterplot par Chamber *et al*. (1983)) ou d'ajouter un marqueur de taille (appelé Triple-Scatterplot par Anscombe (1973)).

Ces marqueurs sont particulièrement utiles pour palier à un des manques des Scatterplots (la visualisation en deux voire trois dimensions) : ils permettent d'ajouter des dimensions aux graphiques. L'un des inconvénients est que les patterns de données portant sur plus de deux critères ne sont pas facilement visibles. En d'autres termes, il peut y avoir des patterns de données qui sont accessibles uniquement lorsque trois ou plusieurs critères peuvent être vus

[3] En accord avec Simon qui définit l'intuition comme la reconnaissance de patterns stockés dans la mémoire (Simon, 1992)

simultanément. Ware (2004) met en avant que bien que l'emploi de couleur dans les Scatterplots permette l'identification de clusters et de patterns, l'interprétation peut être difficile.

Enfin, nous pouvons présenter ici une amélioration du Scatterplot Matrix de Cui, Ward, et Rundensteiner (2006) qui proposent d'intégrer dans la diagonale des graphiques en barres représentant la distribution de chaque dimension (voir **Figure 25**).

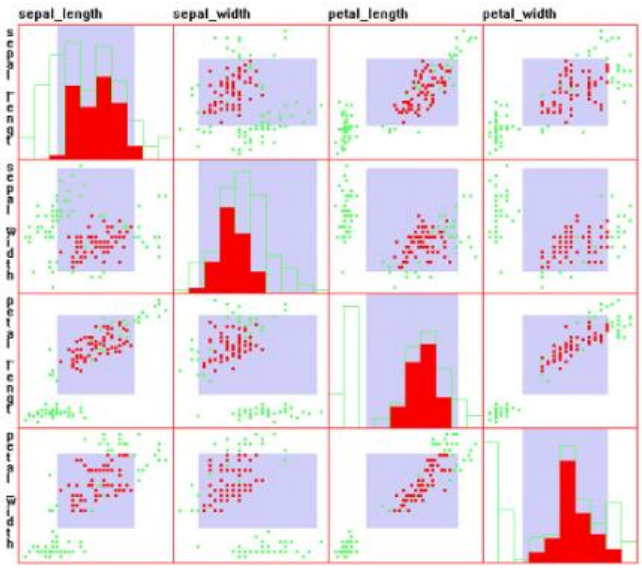


Figure 25 : Impression d'écran de l'amélioration du scatterplot matrix issu (Cui *et al.*, 2006)

1.3.2.3 Chernoff's faces

Le visage de Chernoff est l'une des plus vieilles idées pour visualiser un ensemble de données multidimensionnel (Chernoff, 1973). Le concept est de représenter les valeurs de dix-huit critères comme des caractéristiques d'un visage (voir **Figure 26**). Il s'agit d'une visualisation de l'information de type « iconic display » ainsi chaque critère est paramétré pour représenter une fonction de l'icône, ici le dessin d'un visage. Enfin, il est nécessaire pour chaque alternative qu'un visage soit dessiné.

La justification de Chernoff est que nous sommes habitués à interpréter et différencier les visages. Ainsi, pour lui, il est facile pour les décisionnaires de distinguer les différentes expressions faciales.

Des auteurs se sont intéressés à améliorer les Chernoff's faces comme par exemple Flury et Riedwyl (1981). En effet, ils ont proposé d'ajouter des oreilles aux visages pour obtenir deux dimensions de plus dans ce type de représentation. Mais ils ont également montré quelques inconvénients comme l'affichage de valeurs extrêmes (voir **Figure 27**). Enfin, ils ont défini une modification qui est le visage asymétrique permettant de passer à la visualisation de 36 paramètres ou la visualisation d'alternatives par paires (un seul visage « coupé » en deux).

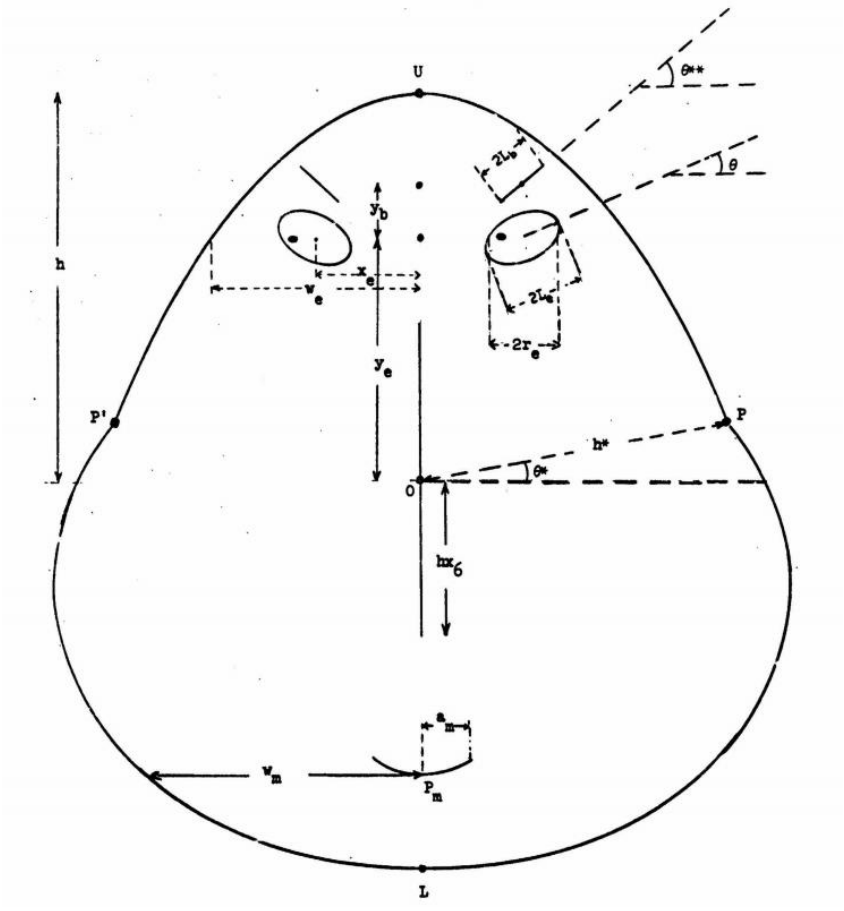


Figure 26 : Paramètres du visage de Chernoff (Chernoff, 1973)

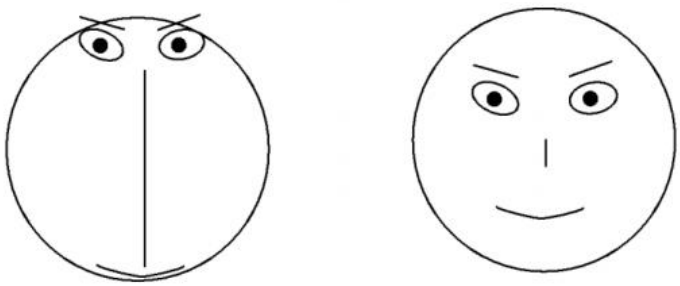


Figure 27 : Impression d'écran de l'inconvénient des valeurs extrêmes (Flury & Riedwyl , 1981)

Miettinen (2014) indique quelques faiblesses de ce type de représentation. Par exemple, il n'y a pas une seule façon d'associer les critères et les paramètres faciaux alors que le « rendu » dépend de la façon dont les paramètres sont sélectionnés. De plus, en raison de la subjectivité de la méthode, comme la préférence de certains paramètres faciaux, cette représentation graphique peut perturber l'interprétation. Et au-delà de ces faiblesses, il reste la mauvaise interprétation engendrée par un visage qui semble heureux par exemple et cela implique que le décisionnaire se focalise sur le critère représenté par la bouche plus que celui représenté par le nez.

1.3.2.4 Diagramme étoile

Dans les années 80, des auteurs ont proposé que les vecteurs de critères soient représentés dans un système de coordonnées en étoile (Chambers et al., 1983). Dans cette représentation, il y a des cercles avec des rayons issus de l'origine qui correspondent aux critères. La longueur des raies représente la valeur du critère.

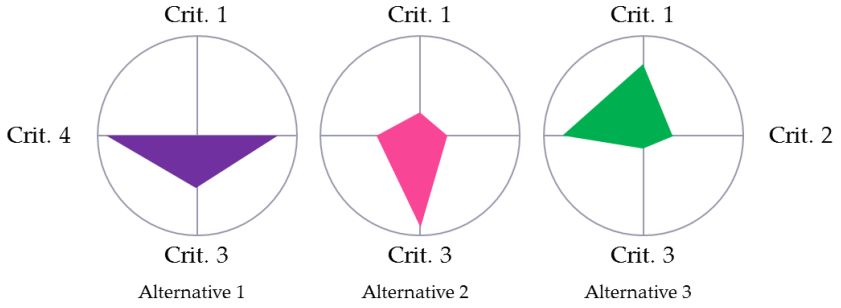


Figure 28 : Diagrammes étoile avec des aires

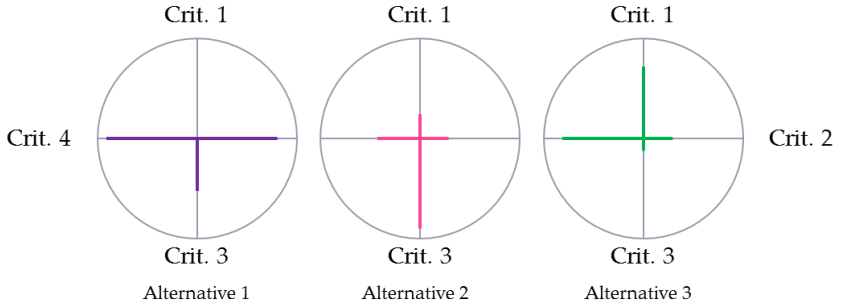


Figure 29 : Diagrammes étoile avec des segments

Dans ce graphique, soit l'idéal (i.e. une solution fictive maximisant les objectifs) se trouve au centre et l'anti-idéal (i.e. l'opposé de l'idéal) sur le cercle soit l'inverse. Force est de constater que ce type de graphique impose que les critères soient tous à maximiser ou tous à minimiser pour une interprétation aisée des objectifs et des performances. Nous trouvons deux façons de « dessiner » les diagrammes en étoile : avec les aires (voir **Figure 28**) et avec les segments (voir **Figure 29**).

Miettinen (2014) avance que le nombre maximum de critères lorsque nous utilisons ce type de graphiques est autour de dix. Elle s'appuie sur les travaux de Ware (2004) qui indique un minimum de 30° entre les raies afin que la forme ne soit pas difficile à interpréter.

Enfin nous pouvons noter que ce type de représentation fait partie des « iconic display ».

1.3.2.5 Spider Graph

Le concept du Spider Graph est similaire à celui du diagramme en étoile. Il est appelé Spider Graph pour sa forme en toile d'araignée. Il est parfois également appelé Radar Chart. Avec ce type de représentation, nous pouvons construire un Spider Graph pour chaque alternative

ou bien en construire un seul et y afficher toutes les alternatives (voir **Figure 30**). Avec ce graphique, les critères sont représentés par des axes positionnés de manière radiale et concourants en un point. Les alternatives sont représentées par des polygones. Chaque sommet d'un polygone se trouve sur un axe et représente la valeur de l'alternative pour le critère correspondant.

De plus, il est possible d'ajouter un profil « objectif » de la même manière que les alternatives mais qui représente les préférences du concepteur.

Few (2005) indique qu'un des problèmes avec ce type de représentation est que le décisionnaire a tendance à préférer les polygones ayant une forme symétrique.

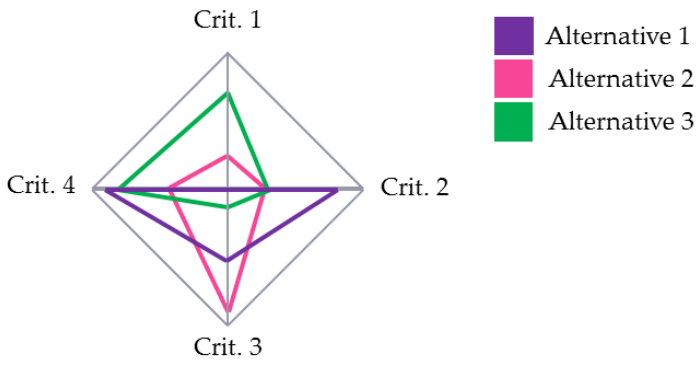


Figure 30 : Illustration du spider graph

1.3.2.6 Value Paths

Le Value Paths aussi appelé Linear Profiles (Mezzich & Worthington, 1978) ou Line Graphs (Wickens, 1992) est une représentation composée de barres qui correspondent aux critères et de lignes qui correspondent aux alternatives (voir **Figure 31**). L'intersection de la ligne sur une barre représente la valeur de l'alternative pour le critère en question. Cette méthode de visualisation est recommandée et est efficace car elle permet aux décisionnaires une interprétation sans surcharge (Miettinen, 2014). Notons que ce type de représentation, au même titre que les graphiques en barres, impose que les critères puissent être sur une même échelle. De plus, il convient de souligner que l'apparence de la visualisation dépend de l'ordre d'affichage des critères.

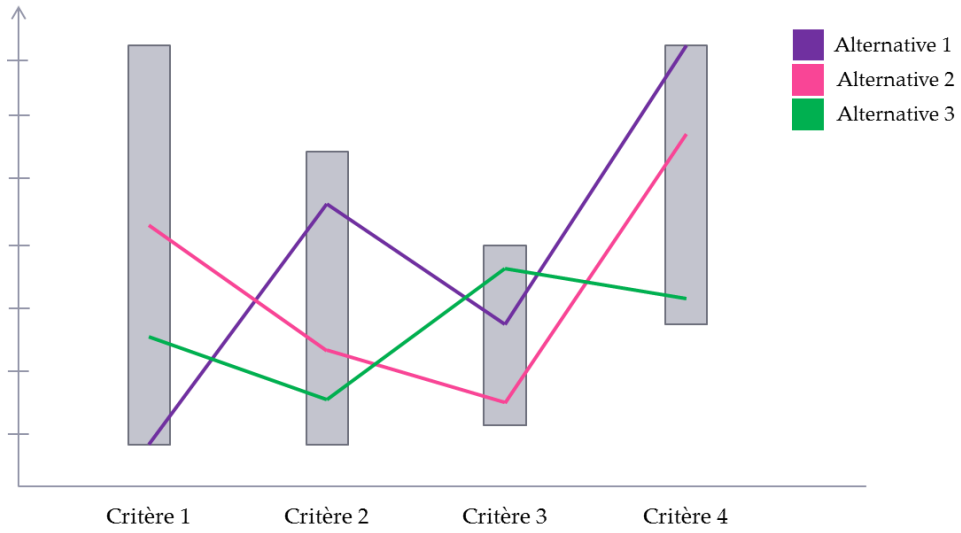


Figure 31 : Illustration du Value Paths

1.3.2.7 Parallel Coordinates Plot

Inselberg (1985) est à l'origine du Parallel Coordinates Plot (PCP). Cette représentation est de type « Geometrically transformed displays ». Ce graphique a été conçu pour travailler sur des problèmes de haute dimension. Il échappe aux limites des systèmes de coordonnées orthogonales en replaçant chaque axe de coordonnées en parallèle. La **Figure 32** illustre le passage d'une représentation en 2D d'un Scatterplot à un PCP.

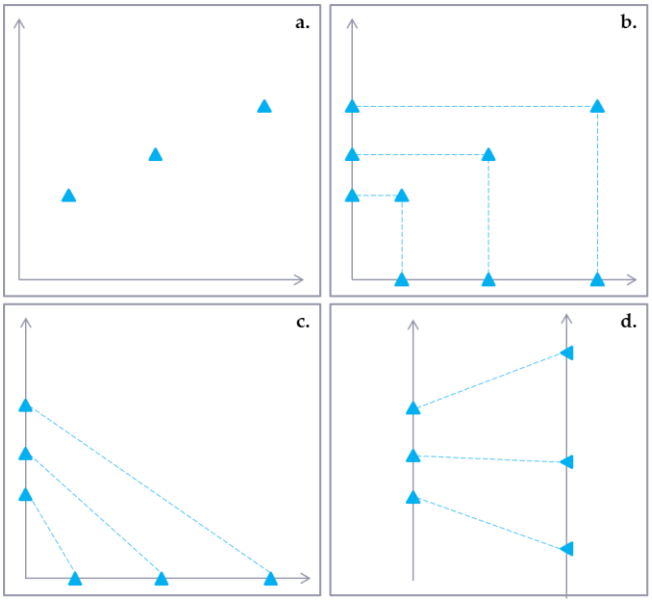


Figure 32 : Construction du Parallel Coordinates Plot

Ainsi chaque dimension ou critère est représenté par un axe (tous mis en parallèle) et les lignes venant couper ces axes sont les alternatives (voir **Figure 33**). L'intersection d'une ligne avec un axe correspond à la valeur de l'alternative pour le critère (représenté par l'axe).

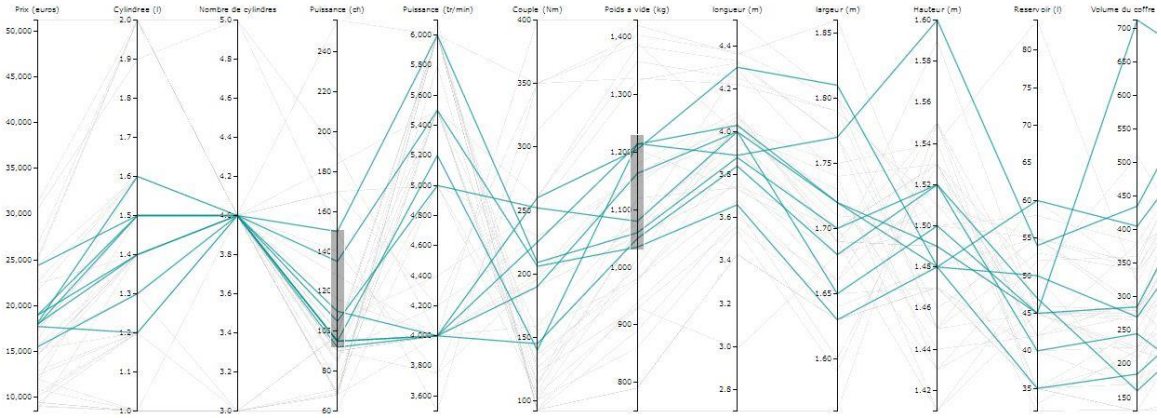


Figure 33 : Illustration d'un Parallel Coordinate

L'utilisation de Parallel Coordinate Plot comme outil d'analyse de données peut paraître à première vue faiblement efficace car il nécessite une interprétation hautement structurée et le développement d'intuitions. En effet, comme mentionné précédemment, les représentations Cartésiennes ont une longue histoire et il en résulte un fort développement d'intuitions sur l'apparence de ce type de représentations. Des intuitions similaires pour le Parallel Coordinate Plot doivent nécessairement être développées (Wegman, 1990). Les **Figure 34**, **Figure 35** et **Figure 36** illustrent trois cas de clusters projetés sur un Scatterplot d'un côté et sur un Parallel Coordinate de l'autre. Ces Figures soulignent bien le besoin de développer des intuitions similaires à celles que nous avons pour des représentations Cartésiennes (Scatterplot).

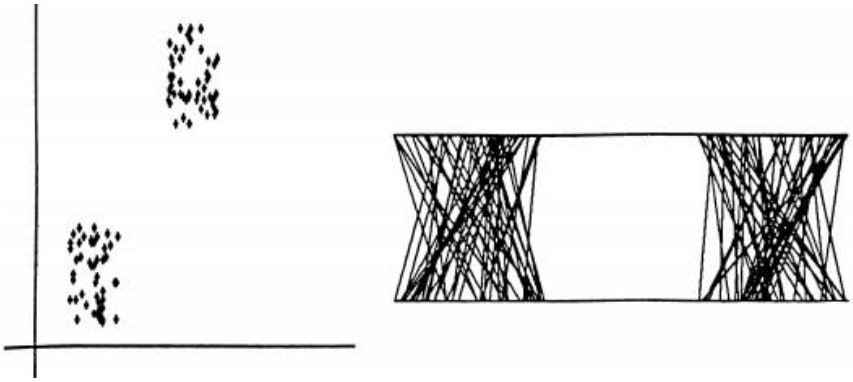


Figure 34 : Impression d'écran illustrant des clusters séparés en x et en y issue de (Wegman, 1990)

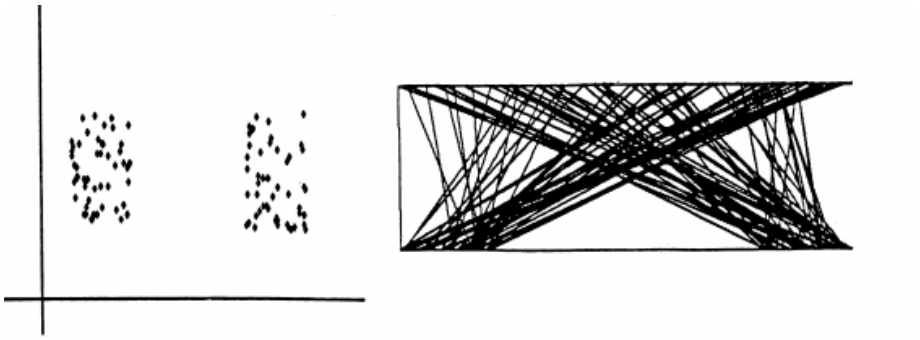


Figure 35 : Impression d'écran illustrant des clusters séparés en x issue de (Wegman, 1990)

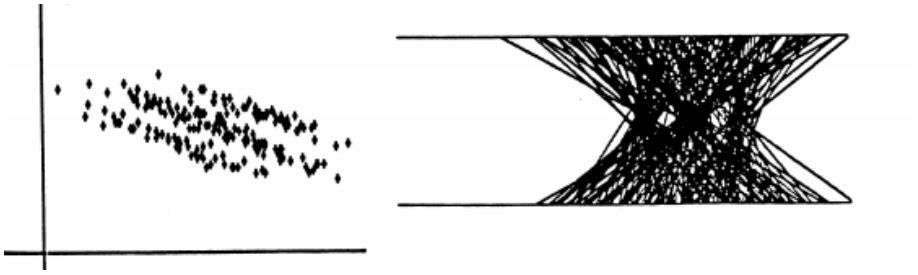


Figure 36 : Impression d'écran illustrant 3 clusters suivant une corrélation négative issue de (Wegman, 1990)

Wegman (1990) ajoute que pour l'analyse exploratoire de données, les techniques de projections incluant deux ou trois dimensions peuvent être trompeuses. Pour lui, il serait hautement désirable d'utiliser une représentation de l'ensemble des coordonnées d'un vecteur de données simultanément surtout si la représentation traite tous les composants d'une manière similaire.

Enfin, pour ce graphique, il est nécessaire de réaliser (n+1)/2 permutations pour la comparaison par paires de toutes les dimensions.

1.3.2.8 Circle Segment

Le Circle Segment est une représentation d'ensembles de données multidimensionnels de type « pixel-oriented visualization ». La technique consiste en l'utilisation d'un pixel de couleur par valeur. L'idée de base de ce graphique est d'afficher chaque dimension de l'ensemble de données comme un segment d'un cercle. Si l'ensemble de données est composé de k-dimensions, le cercle est partitionné en k segments qui représentent chacun une des dimensions de l'ensemble de données. A l'intérieur des segments, les valeurs appartenant à une dimension sont disposées à partir du centre du cercle vers l'extérieur dans un va-et-vient de manière orthogonale par rapport à la ligne qui divise par deux le segment comme illustré sur la **Figure 37** (Ankerst *et al.*, 2011).

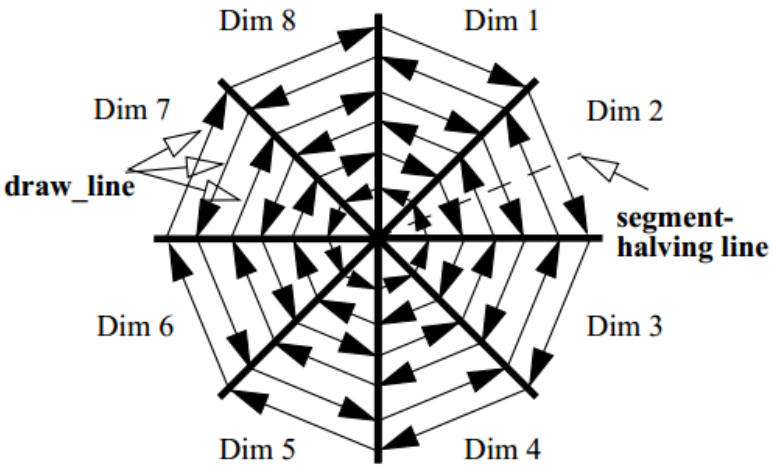


Figure 37 : Impression d'écran issu de (Ankerst et al., 2011) illustrant la construction du circle segment

Ce graphique est particulièrement utile pour détecter des clusters et leurs propriétés (Keim, 2000). La **Figure 38** illustre un exemple de 50 prix de stocks différents. Notons également que ce type de visualisation est surtout utile pour des ensembles comprenant un nombre colossal de données (deux cent mille à deux millions) car une donnée correspond à un pixel. En effet si l'ensemble à visualiser est composé de mille données, le rendu visuel se retrouverait amoindri, voire pauvre.

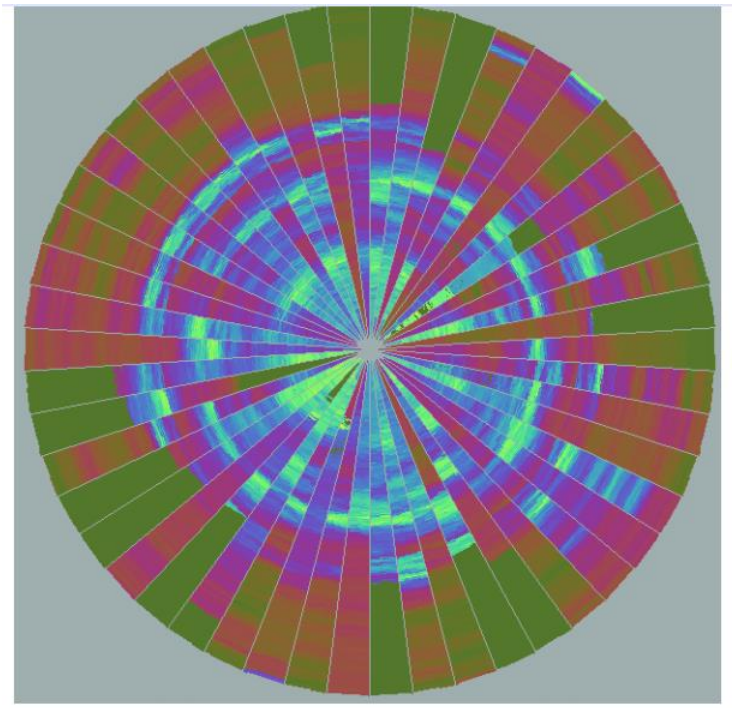


Figure 38 : Exemple d'un circle segment issu de (Ankerst *et al.*, 2001)

1.3.2.9 Le tableau combiné d'une carte de chaleur

Un tableau de valeurs n'est pas une présentation graphique d'un ensemble de données multidimensionnel. Une variante innovante du tableau traditionnel est le Heatmap (carte de chaleur). La structure est similaire à celle d'un tableau de valeurs mais le Heatmap permet une perspective plus holistique, c'est-à-dire une vision globale de l'ensemble de données ce qui est efficace pour des problèmes comptant de nombreuses alternatives. Par essence, les Heatmaps sont des matrices dans lesquelles les cellules sont colorées par rapport à leurs valeurs (Cook *et al.*, 2007). Récemment, Gettinger et al. (2013) proposent de combiner le tableau de valeurs avec le Heatmap (Combined Table). Pour eux, la forte densité d'information de cette représentation facilite l'identification de patterns comme des corrélations et les compromis entre les critères. Nous pouvons noter que, dans la dernière décade, le Heatmap a été proposé pour visualiser la frontière de Pareto (Pryke *et al.*, 2007 ; Lotov et Miettinen, 2008).

1.3.2.10 Le Treemap

Pour finir nous présentons ici le Treemap qui est un graphique utilisant une approche de remplissage d'espace 2D inventé par Shneiderman (1992). Ce graphique est de type « Stacked displays », il est donc utilisé pour des données hiérarchisées. En effet, le Treemap permet de représenter une structure de données en « arbre » qui comprend des poids ou des tailles sur ses nœuds les plus bas (leaf nodes). La **Figure 39** montre un exemple d'un Treemap où chaque carré représente une feuille de l'arbre. Jusqu'à deux dimensions peuvent être mappées sur cette représentation de données hiérarchisées représentées par la taille des carrés et une échelle de couleur (nuances de gris sur la **Figure 39**).

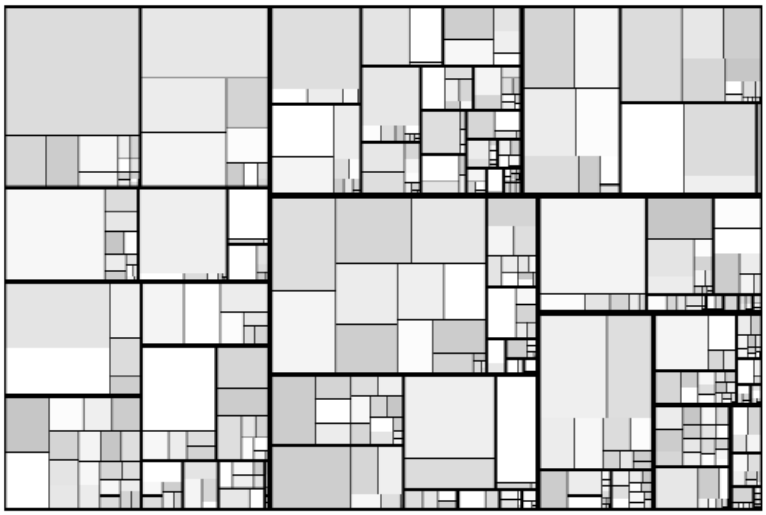


Figure 39 : Exemple d'un Treemap issu (Shneiderman & Wattenberg, 2001)

Sur la **Figure 40** on trouve un parallèle entre une représentation Treemap et une représentation en arbre afin de mieux comprendre la méthode de remplissage de l'espace 2D utilisée par le Treemap.

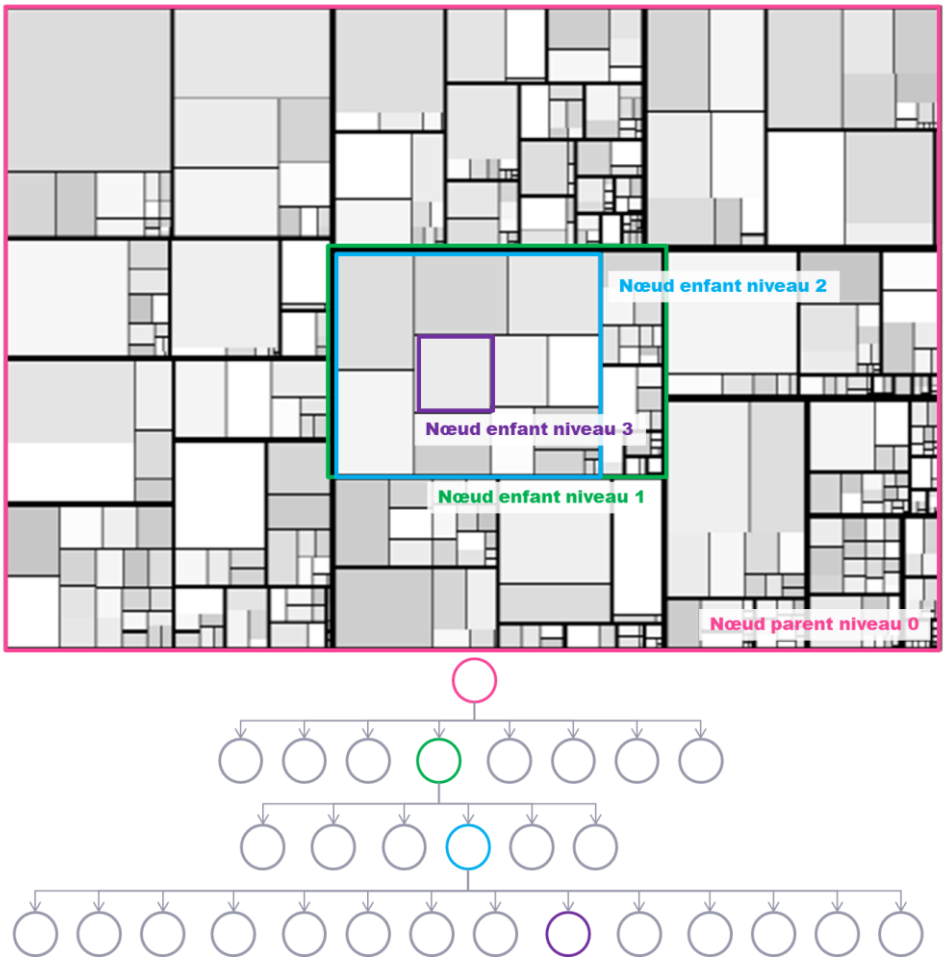


Figure 40 : Illustration Treemap Vs. Structure arbre

Pour clore cette section, nous proposons de positionner les graphiques décrits ci-dessus et de souligner notre contexte sur le modèle de Keim (2001) déjà présenté dans la première section.

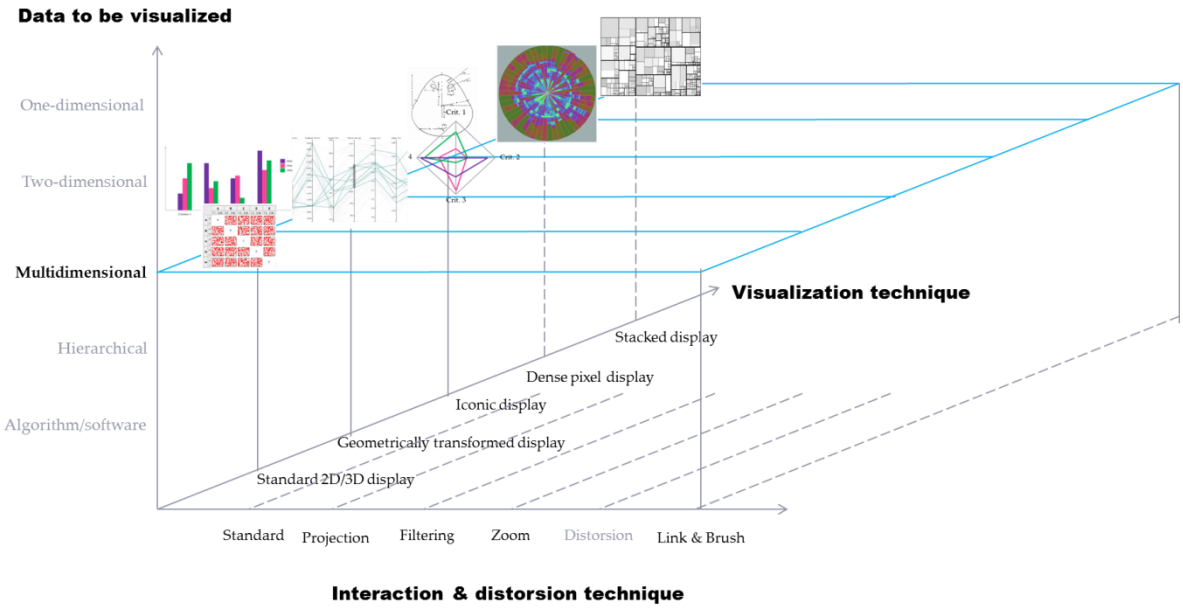


Figure 41 : Positionnement sur le modèle de Keim (2001)

Ainsi, dans notre cas nous nous intéressons à la représentation d'ensembles de données multidimensionnels ce qui correspond au plan souligné en bleu sur la **Figure 41**. De plus, nous retrouvons sur cette illustration certains graphiques que nous avons présentés dans cette section et qui permettent l'exploration d'un ensemble de données multidimensionnel. On remarque que les graphiques sont positionnés en fonction des techniques de visualisation. En effet, chaque graphique présenté utilise une technique de visualisation répertoriée par Keim (2001). Notons que nous considérons l'ensemble des techniques d'interactions (d'où le plan en bleu sur la **Figure 41**).

1.3.3 Comparaison

Nous venons de voir que de nombreux graphiques existent pour représenter un ensemble de données multidimensionnel. Ces différentes représentations peuvent être classées par les techniques qu'elles utilisent mais également par les activités qu'elles « autorisent » ou encore par le type de données qu'elles peuvent afficher.

Miettinen (2014) propose une synthèse des différentes techniques de visualisation. Elle propose de comparer onze types de graphiques permettant la visualisation d'ensembles de données multidimensionnels et six critères de comparaison (voir **Figure 42**). Les critères qu'elle a choisis sont :

- Le nombre maximum de critères
- La sensibilité à l'ordonnancement
- Le nombre maximum d'alternatives
- Le besoin d'échelle
- Le type de comparaison par paires (par paires de critères ou d'alternatives)
- La possibilité de voir des niveaux

	Max number of criteria	Sensitive to order	Max number of alternatives	Ranges needed	Pair-wise	Show asp. levels
Bar charts	Not many	Some versions	Not many	No		As dots
Scatter plots	~10	No		Yes	Cr	No
Value paths		Maybe	Not many	No		Yes
Star coord. syst.	~10	Yes	∞	Yes		Yes
Spider web chart	~10	Yes	∞	Yes		Yes
Petal diagram	~10	No	∞	Yes		Yes
Chernoff's faces	18/36	Yes	∞	Yes		No
Harmon. houses	16	Yes	∞	Yes	Alt	No
Metroglyphs	7	Yes	∞	Yes		As dots
Stick figures	5	Yes	∞	Yes		No
Boxes	~ 10	Yes	∞	Yes		No

Figure 42 : Tableau de synthèse des différentes techniques de visualisation d'après Miettinen (2014)

La synthèse proposée par Miettinen est intéressante pour nous, elle permet de faire un premier filtre sur les graphiques qui ne répondent pas à nos attentes. En effet, dans notre cas, il s'agit de l'exploration d'un nombre de points de conception proche de mille (alternatives). Nous pouvons ainsi éloigner les graphiques de type « Bar Chart » et « Value Paths » qui sont limité en nombre d'alternatives.

Notons que le tableau (**Figure 42**) ne propose pas l'ensemble des graphiques détaillés dans la section précédente. De plus, il n'y a pas d'indications sur l'emploi ou non de techniques d'interactions.

Keim dans ses travaux (Keim, 2000) propose également un tableau comparatif de différents graphiques rangés par techniques de visualisation (voir **Figure 43**). Il indique cependant qu'il s'agit d'une comparaison préliminaire et subjective. Il utilise dans son tableau sept critères :

- Caractéristiques de tâche :
  - Le clustering
  - Zone d'intérêt multivariée
- Caractéristiques de données :
  - Nombre de variables
  - Nombre d'alternatives
  - Données catégorisées
- Caractéristiques de visualisation :
  - Chevauchement visuel
  - Courbe d'apprentissage

		Clust-ering	multi-variate hot spot	no. of vari-ates	no. of data items	cate-gorical data	visual over-lap	learn-ing curve
Geometric Techniques	Scatterplot Matrices	++	++	+	+	-	o	++
	Landscapes	+	+	-	o	o	+	+
	Prosection Views	++	++	+	+	-	o	+
	Hyperslice	+	+	+	+	-	o	o
	Parallel Coordinates	o	++	++	-	o	--	o
Icon-based Techniques	Stick Figures	o	o	+	-	-	-	o
	Shape Coding	o	-	++	+	-	+	-
	Color Icon	o	-	++	+	-	+	-
Pixel-oriented Techniques	Query-Independent	+	+	++	++	-	++	+
	Query-Dependent	+	+	++	++	-	++	-
Hierarchical Techniques	Dimensional Stacking	+	+	o	o	++	o	o
	Treemap	+	o		o	++		o
	Cone Trees	+	+	o	+	o	+	+
Graph-based Techniques	Basic Graphs	o	o	-	+	o	o	+
	Specific Graphs	++	+	-	+	o	+	+

Fig. 2. An attempt at comparing multidimensional visualization techniques (++: very good, +: good, o: neutral, -: bad, –: very bad).

Figure 43 : Tableau de comparaison préliminaire et subjective de Keim (2000)

Notons que les graphiques utilisant les techniques « hierarchical techniques » et « Graph-based techniques » (grisés dans la **Figure 43**) correspondant à l'affichage d'ensembles de données hiérarchisées et en réseau ne rentrent pas dans notre analyse. En effet, notre cas d'étude ne met pas en jeu ce type d'ensemble de données.

Dans ce tableau nous devons également tenir compte du fait qu'aucun graphique n'est « hybridé ». Il n'y a pas non plus d'ajouts d'une dimension telle que la couleur. Ainsi, certaines notes attribuées peuvent être relaxées. Par exemple, si nous ajoutons un marqueur de couleur au Scatterplot Matrix, les données catégorisées sont visibles. Aussi, les techniques d'interactions diminuent les problèmes de chevauchements visuels du Parallel Coordinates Plot.

Enfin, il manque un critère important dans la section « caractéristiques de tâche » : la comparaison. En effet, la comparaison est la tâche permettant l'identification de corrélation (seul le « clustering » est indiqué dans le tableau). Il convient de nous intéresser aux tâches réalisables avec les différents graphiques afin d'identifier ceux qui sont potentiellement utiles à l'exploration de l'espace de conception.

1.3.4 Les tâches

Comme cela a pu transparaître dans les sections précédentes, certaines tâches sont réalisables avec les graphiques soit grâce aux techniques de visualisation soit par des techniques d'interactions.

Ben Shneiderman est l'un des premiers à avoir réalisé ce travail d'analyse et de synthèse lui permettant de créer une taxonomie « task by data type » listant sept actions d'information :

- « Overview: a view of the total collection
- Zoom: a view of a single item. This may be either at the object or the attribute level.
- Filter: removing unwanted items from the displayed set.
- Detail-on-demand: getting the details of a selected group, sub-group or item.
- Relate: viewing the relationships between a selected group, sub-group or item.
- History: the action of undoing, replaying, and refining using a store of historic information.
- Extract: the extraction or focusing in on sub-collection and other parameters from a given set." (Shneiderman, 1996)

L'auteur indique que la plupart des bases de données relationnelles et statistiques sont facilement manipulées comme des données multidimensionnelles dans lesquelles les éléments à n attributs deviennent des points dans un espace à n dimensions. Mais par exemple, il est possible de commander des dimensions supplémentaires à un Scatterplot à l'aide d'un curseur ou de boutons. Il souligne que les tâches comprennent la découverte/recherche de patterns, de clusters, de corrélations entre paires de variables, des écarts et des valeurs aberrantes.

Il convient de noter que l'action « history » dans la taxonomie de Shneiderman peut être écartée car il n'y a pas de données historiques dans notre contexte. En conception, il est possible d'utiliser des informations et/ou données d'un projet passé. Cependant, dans nos travaux, nous nous intéressons au cas de générations de données (et donc la génération de solutions possibles à partir d'un modèle et de simulations).

En 2001, Andrienko & Andrienko (2001) réalisent, dans leurs travaux, une étude du Parallel Coordinates Plot afin de proposer les actions et tâches qui sont et qui doivent être possibles avec ce graphique pour la résolution de problème.

Le résultat de leur étude fait émerger dix tâches :

1. Etude de la distribution des caractéristiques sur un ensemble d'objets;
2. Comparaison des caractéristiques d'un objet sur l'ensemble des caractéristiques;
3. Comparaison d'objets par paires;
4. Comparaison des valeurs des attributs associés à un objet sélectionné;
5. Comparaison des plages de valeurs des attributs;

6. Comparaison des variations de valeurs des différents attributs;
7. La recherche de corrélation entre les attributs;
8. Estimation du degré de similarité entre les objets ;
9. Affectation d'objets à une ou deux classes spécifiées par la sélection d'échantillons représentatifs;
10. Évaluation multicritère des objets.

Ils proposent une synthèse de leur résultat. Les tâches sont :

- Observation des caractéristiques des objets,
- Comparaison des objets,
- Exploration des relations entre les attributs,
- Recherche d'objets similaires à un échantillon donné,
- Classification d'objets selon leur similarité avec deux classes représentatives,
- L'évaluation multicritères des objets

Wilkins propose dans ses travaux de définir les "Concrete user-level tasks" (wilkins, 2003) **:**

- Trouver des objets,
- Regarder les détails d'un objet,
- Comparer les objets,
- Découvrir des relations entre les objets,
- Grouper ou agréger des objets,
- Effectuer des calculs,
- Identifier les tendances

Le modèle de Wilkins montre une avancée par rapport aux deux autres puisque l'auteur ajoute une tâche sensiblement différente : « effectuer des calculs ». Cette tâche est importante mais pour nous elle n'a pas sa place dans l'ensemble des tâches qui doivent être possible avec un graphique (i.e. la visualisation d'information). En effet, il s'agit d'une tâche autorisant la manipulation et des changements directement sur les données. Par exemple, de simples additions ou soustraction entre 2 attributs entrainent soit la modification d'un des attributs soit la création d'un nouvel attribut. Les tâches réalisables sur les graphiques permettent la modification de « point de vue » (à l'aide de techniques d'interactions) alors qu'effectuer des calculs permet la modification du modèle de données (même si elle s'opère éventuellement à l'aide de techniques d'interactions).

La différence entre ces trois modèles s'explique principalement par le contexte d'utilisation des graphiques.

Nous retiendrons de ces modèles :

- La vue d'ensemble
- La vue d'une alternative (ou sous-groupe) sur l'ensemble

- La vue de l'ensemble filtré (retirer les non-désirés)
- La vue du détail d'une alternative
- La vue des relations entre attributs (la comparaison) : clusters, corrélation, écarts et aberrants.
- La vue des similarités entre sous-ensembles (regroupement/classification)
- La vue des tendances sur l'ensemble (patterns et sensibilités)

Enfin rappelons que les tâches sont réalisables :

- Soit grâce aux techniques de visualisation : par exemple la corrélation entre deux attributs sur un scatterplot
- Soit grâce aux techniques d'interactions : par exemple la vue d'un sous-groupe en filtrant l'ensemble de données à l'aide de slider ou méthode de brush.

Nous avons décrit des représentations graphiques à travers les fonctions qu'elles proposent et les tâches réalisables. Cet état de l'art nous permet, à présent, d'identifier les graphiques potentiellement utiles pour l'exploration de l'espace de conception.

1.3.5 Bilan

Nous présentons ici un bilan sur les graphiques permettant l'exploration d'un espace de conception multidimensionnel.

Tableau 2 : Comparaison des graphiques selon le nombre de critères et d'alternatives

Graphique	Critères	Alternatives
Graphique en barre	Un-à-un ou plusieurs avec la même échelle	Plusieurs
Scatterplot simple	2 ou 3	Toutes
Scatterplot matrice	Tous (par pairs)	Toutes
Chernoff's faces	Plusieurs (~18)	Une-à-une
Diagramme en étoile	Plusieurs (~18)	Une-à-une
Radar Chart	Plusieurs (~18)	Un grand nombre
Value path	plusieurs avec la même échelle	Plusieurs
Parallel coordinate	Tous	Toutes
Circle segment	Tous	Toutes
Treemap	2	Toutes
Combined Table	Tous	Toutes

Nous présentons un premier tableau comparatif des graphiques présentés en 2ième section du chapitre. Dans ce tableau (**Tableau 2**) nous proposons deux critères de comparaison : le nombre de critères (ou attributs) et le nombre d'alternatives qui peuvent être affichés. Nous utilisons ces deux critères car ils sont suffisants pour réaliser un premier filtre et « éliminer » les graphiques qui ne nous sont pas utiles.

Premièrement, nous retirons le graphique Treemap de notre étude car il permet la présentation de données hiérarchisées et que ce type de structure ne correspond pas à notre contexte.

Aussi, nous proposons d'exclure du tableau les graphiques ne permettant pas de visualiser toutes les alternatives à l'exception du Radar Chart permettant d'en visualiser un grand nombre (au même titre que nous conservons le Circle Segment qui lui ne peut satisfaire que la visualisation d'un très grand nombre de données).

De plus, nous pouvons relever que les graphiques permettant la vue d'ensemble sont : Le Scatterplot Matrix, le Parallel Coordinate, le Circle Segment et le Combined Table.

Suite à ce premier filtre, nous proposons un deuxième tableau de comparaison des graphiques restant (**Tableau 3**). Nos critères de comparaisons sont « clusters », « corrélations », « écarts », « aberrant », « tendance » et « apprentissage ».

Nous avons déjà défini le cluster comme un groupement de solutions (ou d'alternatives) homogène, des solutions qui partagent des caractéristiques communes qui, par exemple, correspond à un critère.

Une corrélation est une relation de « dépendance » entre des variables. Par exemple, deux variables A et B, si les valeurs de la variable A augmentent et que les valeurs de la variable B augmentent de manière proportionnelle alors il y a corrélation positive entre les variables A et B.

Un écart correspond à des valeurs d'une variable qui ont un « décalage » avec les valeurs attendues.

Une valeur aberrante est une valeur sur une variable qui montre un comportement totalement différent du comportement des autres valeurs. Par exemple, soit deux variables A et B en corrélation positive (tel que pour A=a, B=b) et que pour A = a, la valeur de la variable B ≠ b.

Nous décrivons une tendance comme un motif reconnaissable entre des variables. Par exemple, pour deux variables A et B, nous observons une tendance si en augmentant les valeurs sur la variable A et que passé une certaine valeur (a) les valeurs de la variable B tendent vers une valeur fixe (b).

Tableau 3 : Comparaison des graphiques selon les différentes « découvertes »

Graphique	Clusters	Corrélations	Ecarts	Aberrant	Tendance	Apprentis sage	Note
Scatterplot Simple	++	+	+	++	0	++	8
Scatterplot Matrix	++	++	+	+	+	+	8
Radar Chart	0	0	++	++	+	+	6
Parallel Coordinate	+	++	++	++	++	-	8
Combined Table	0	++	+	++	0	++	7
Circle Segment	++	- -	0	- -	+	0	-1
(++ très bon = 2 ; + bon = 1 ; 0 neutre = 0 ; - mauvais = -1 ; - - très mauvais = -2)							

Notons que les Scatterplots bénéficient d'un bon indicateur d'apprentissage car ils héritent des intuitions développées sur les systèmes de coordonnées cartésiens. La matrice obtient un moins bon score car elle nécessite un peu de pratique pour relier mentalement les différentes découvertes entre elles. Vous pouvez constater cette « gymnastique » mentale à travers l'exemple en **Figure 44** avec l'observation 1.

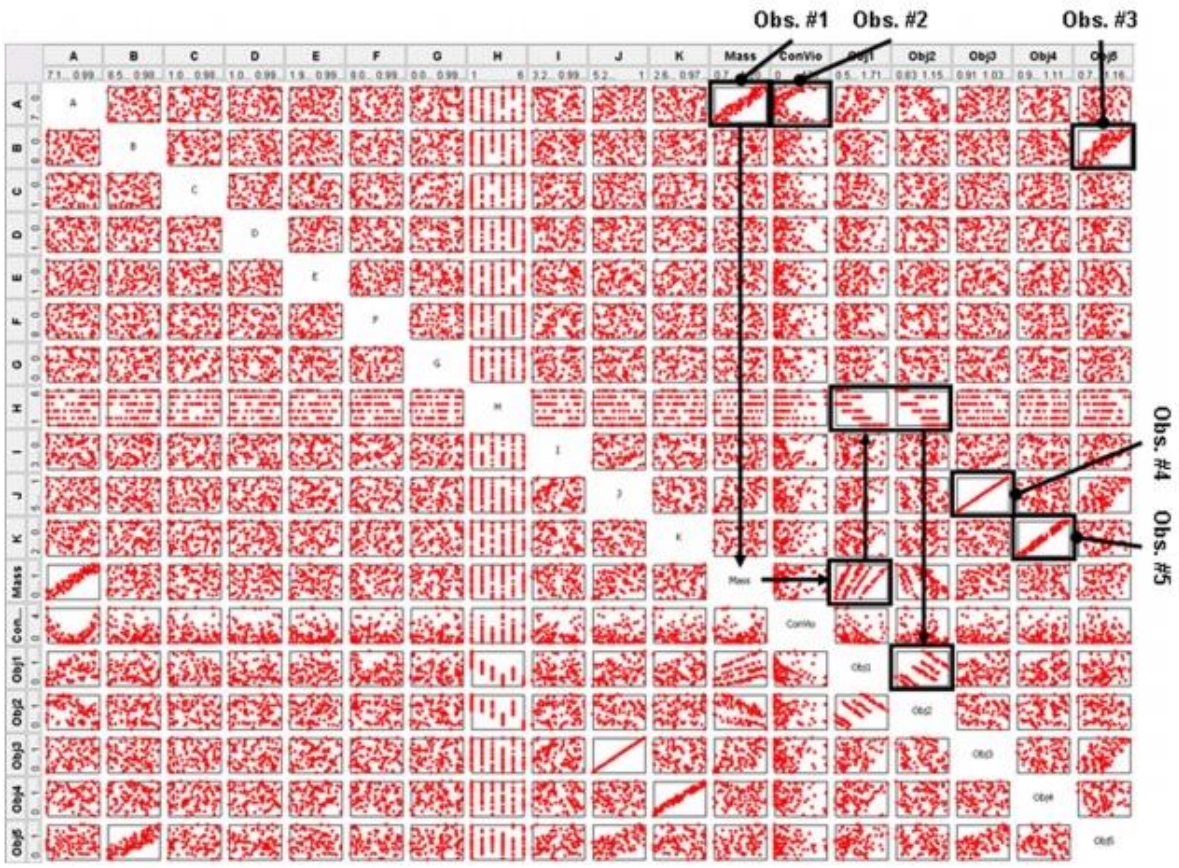


Figure 44 : Exemple de visualisation d'un scatterplot matrix issu de (Stump *et al.*, 2009)

Le Radar Chart obtient également un bon indicateur d'apprentissage. Ce graphique est de type « Icon-based displays » ainsi sa lecture est basée sur la « forme ». Il ne demande donc pas beaucoup de pratique. De plus, il ne permet pas l'identification de clusters ou de corrélations. Il n'est donc pas nécessaire d'apprendre à reconnaître ce type d'interactions entre les variables. Enfin, il est utile pour observer des « profils » pouvant dévoiler des écarts, des valeurs aberrantes et des tendances.

Il est assez simple d'identifier des clusters avec le Circle Segment mais l'interprétation avec ce type de graphique n'est pas nécessairement simple c'est pourquoi son indicateur d'apprentissage est neutre.

Enfin le Parallel Coordinates Plot ne révèle pas un bon indicateur pour l'apprentissage. En effet, bien que ce type de graphique soit assez complet pour les « découvertes » il demande tout de même un peu de pratique afin de développer des intuitions et identifier les clusters et les corrélations.

Le Combined Table montre un très bon indicateur pour l'apprentissage car sa structure traditionnelle de tableau de valeurs lui permet d'être rapidement interprétable.

Au vu des indicateurs affichés dans le **Tableau 3** pour le Circle Segment, nous proposons de l'exclure de notre étude. En effet, nous nous restreindrons à ne garder que les graphiques

obtenant une note supérieure à 0. Le Circle Segment est tout de même un graphique intéressant pour de très grands ensembles de données et particulièrement pour des données de type « time series » ce qui est légèrement en marge de notre situation. Cependant nous conservons la technique de marqueur de couleur pour ajouter une dimension aux autres graphiques (les Scatterplots, le Radar Chart et le Parallel Coordinates Plot).

Pour finir nous proposons d'illustrer les relations pour chaque graphique : les Scatterplot 2D, 3D (SSP) et Matrix (SPM) ; le Radar Chart (RC), le Parallel Coordinates Plot (PCP) et le Combined Table (CT) dans un espace à six dimensions (**Figure 45**). Nous présentons dans la **Figure 45** la modélisation des relations (traits bleus) possibles pour les graphiques dans un cas statique et dans un cas interactif et nous schématisons dans un 3ième cas la visualisation d'une alternative.

Nous observons que le Scatterplot Matrix (SPM) est le graphique le plus utile dans un cas statique. Il permet en effet de visualiser l'ensemble des relations entre les attributs (dimensions). Dans un cas interactif, le Radar Chart (RC), le Combined Table (CT) et le Parallel Coordinates Plot (PCP) dévoilent le même motif et se retrouvent donc au même niveau que le Scatterplot Matrix (SPM). Dans les deux cas les Scatterplots dits simples (SSP) n'autorisent qu'une seule ou deux relations mais permettent un niveau de détail plus important que le Scatterplot Matrix (pas visible dans la **Figure 45** mais un aperçu est disponible dans le **Tableau 3**).

Enfin, le 3ième cas reflète bien la « gymnastique » mentale que les concepteurs doivent effectuer avec le Scatterplot Matrix (SPM). En effet, on remarque qu'une seule alternative n'est pas affichée à l'aide d'un seul item comme pour le Radar Chart (l'alternative est représentée par un polygone), le Combined Table (l'alternative est représentée par une ligne du tableau) et le Parallel Coordinates Plot (l'alternative est représentée par une seule ligne). Le Scatterplot Matrix duplique la « matérialisation » de l'alternative autant de fois qu'il y a de comparaisons par paires d'attributs (ou dimensions).

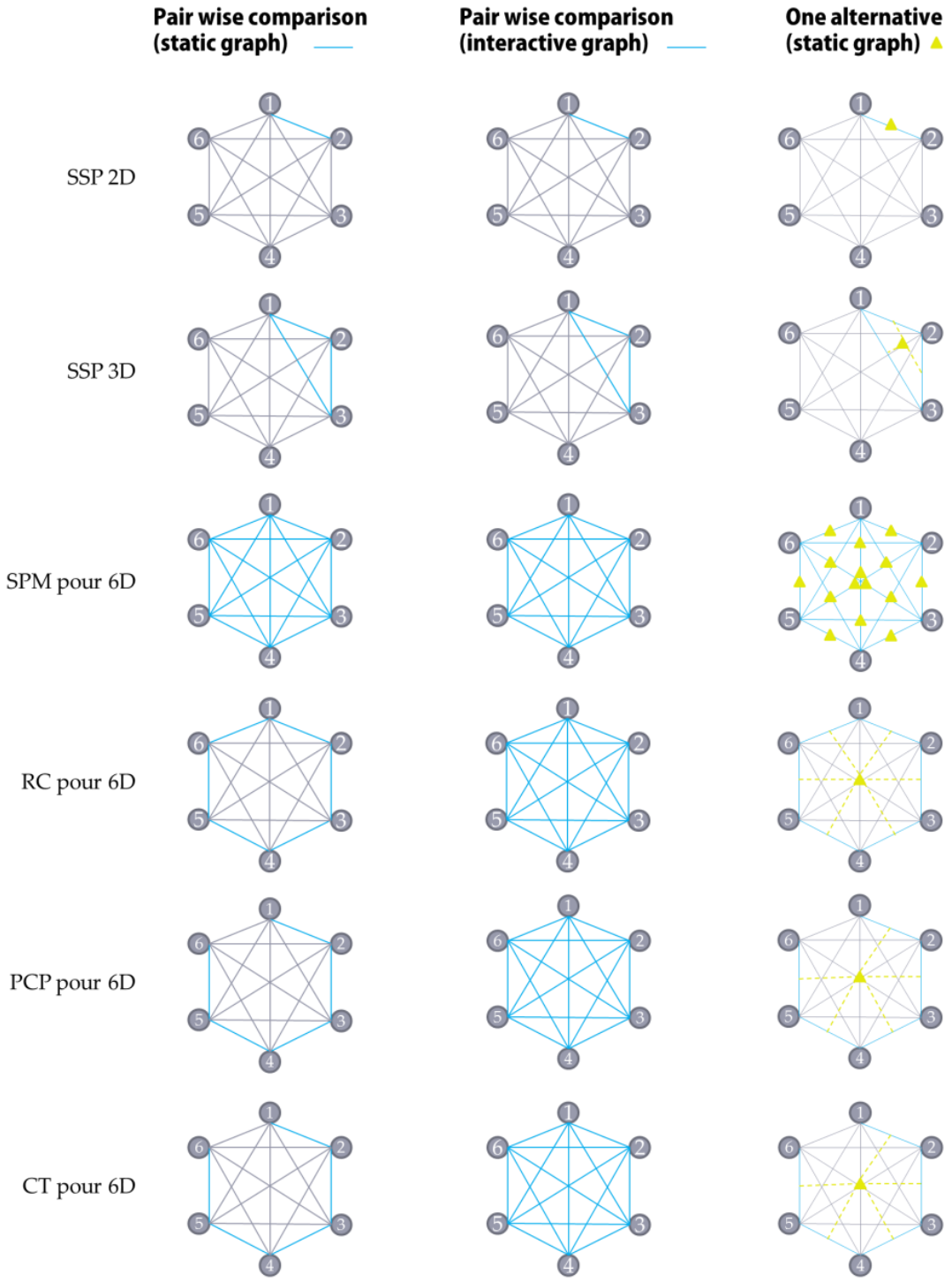


Figure 45 : Illustration des relations pour chaque graphique dans espace à 6 dimensions

Compte tenu des Tableaux comparatifs (**Tableau 2** et **Tableau 3**) et de la **Figure 45** nous pouvons postuler que pour l'exploration d'un espace de conception multidimensionnel :

- Le Scatterplot Matrix est plus efficace pour identifier les clusters et les corrélations entre les attributs (**Figure 46**).

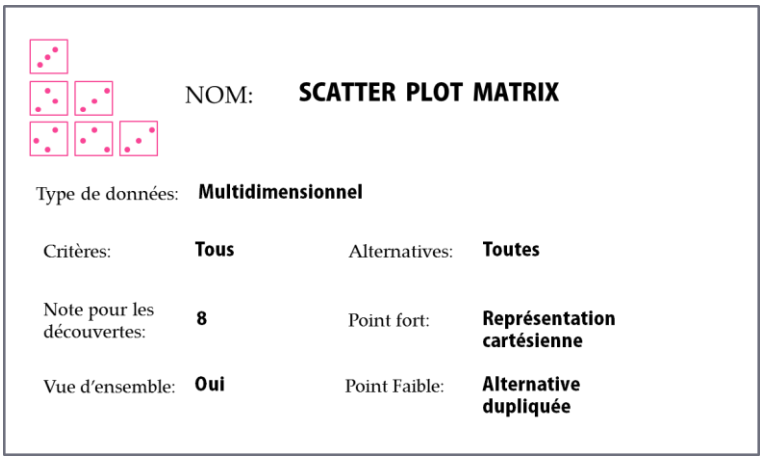


Figure 46 : Fiche identité du Scatter Plot Matrix

- Le Simple Scatter Plot (2D ou 3D) est efficace pour obtenir du détail entre deux ou trois attributs et identifier les valeurs aberrantes par exemple (voir **Figure 47**).

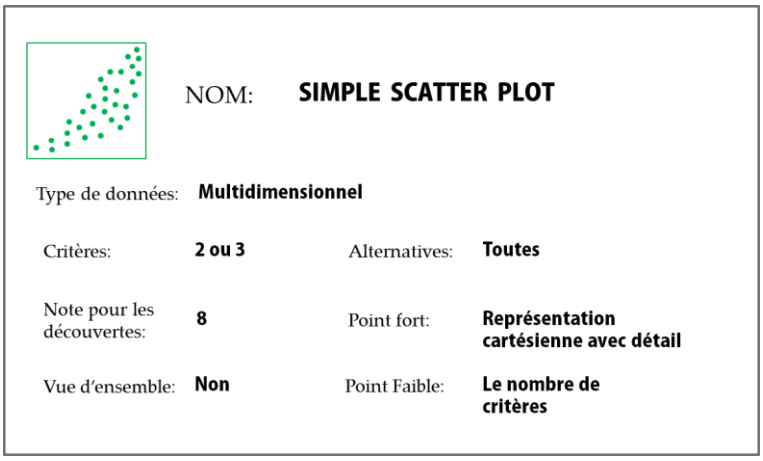


Figure 47 : Fiche identité du Simple Scatter Plot

- Le Parallel Coordinates Plot est plus efficace pour identifier les tendances (les patterns) et particulièrement les sensibilités. Il est également efficace pour détecter les écarts sur l'ensemble de données (**Figure 48**).

NOM: **PARALLEL COORDINATES PLOT**

Type de données: **Multidimensionnel**

Critères:	**Tous**	Alternatives:	**Toutes**
Note pour les découvertes:	**8**	Point fort:	**Unicité de l'alternative**
Vue d'ensemble:	**Oui**	Point Faible:	**Apprentissage**

Figure 48 : Fiche identité du Parallel Coordinates Plot

- Le Radar Chart a plus sa place dans une situation MADM car il ne peut pas afficher un très grand nombre d'alternatives, il demande donc un filtrage en amont et ne montre pas de très bons indicateurs pour la reconnaissance de clusters et de corrélations. Mais ce graphique permet de distinguer aisément des profils (voir **Figure 49**).

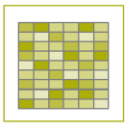
NOM: **RADAR CHART**

Type de données: **Multidimensionnel**

Critères:	**Beaucoup**	Alternatives:	**Beaucoup**
Note pour les découvertes:	**6**	Point fort:	**Taille**
Vue d'ensemble:	**Oui**	Point Faible:	**Asymétrie**

Figure 49 : Fiche identité du Radar Chart

- Le Combined Table a également plus de pertinence en situation MADM car avec le nombre d'alternatives à afficher sa taille peut vite devenir un handicap. Par ailleurs, son interprétation est aisée et son niveau de détail (précision) est important (**voir Figure 50**)



NOM: **COMBINED TABLE**

Type de données: **Multidimensionnel**

Critères:	**Beaucoup**	Alternatives:	**Beaucoup**
Note pour les découvertes:	**7**	Point fort:	**Niveau de détail**
Vue d'ensemble:	**Oui**	Point Faible:	**Taille**

Figure 50 : Fiche identité du Combined Table

1.3.6 L'essentiel du chapitre 3

Dans ce chapitre nous dressons un état de l'art, non exhaustif, des représentations graphiques d'ensembles de données multidimensionnels. Nous réalisons un comparatif de ces graphiques afin d'identifier ceux qui semblent pertinents dans notre contexte. Nous nous appuyons notamment sur :

- Des caractéristiques « physiques » : nombre d'alternatives et nombre de critères (et donc la taille du graphique)
- Des caractéristiques « fonctionnelles » : comparaison inter et intra critères. Interaction entre les variables, la vue d'ensemble, etc.
- Des caractéristiques « d'apprentissage » : les découvertes réalisables (et le degré d'aisance pour les faire)

Enfin, ce chapitre établit les graphiques potentiellement pertinents pour l'optimisation multicritère en Design by Shopping. Cinq graphiques sont retenus et une fiche identité est proposée pour chacun d'entre eux.

CONCLUSION DE LA PREMIERE PARTIE

Dans la première partie nous présentons trois chapitres décrivant notre positionnement scientifique (chapitre 1) et notre état de l'art (chapitre 2 et chapitre 3).

Le premier chapitre positionne nos travaux de recherche en exploration d'espace de conception. L'exploration d'espace de conception est une méthode d'optimisation multicritère en conception (multi-attribut ou multi-objectif) en Design by Shopping. Le processus de conception, avec cette méthode, est décomposé en trois phases principales : la découverte, l'optimisation et la sélection. Nous faisons le constat que la phase de sélection peut être dans une situation multi-objectif (intégré dans le processus complet) ou dans une situation multi-attribut.

Le deuxième chapitre confirme la formulation d'une décision informée pour la conception en Design by Shopping notamment en soulignant la nécessité d'inclure l'humain dans le processus, l'activité d'analyse et de traitement de données et le besoin de transformation de données en information et par la suite en connaissance. Nous faisons le constat que les moyens utiles à la décision informée sont la visualisation d'information.

Le troisième chapitre propose un état de l'art des représentations graphiques d'ensembles de données multidimensionnels. Nous réalisons un comparatif de ces graphiques afin d'identifier ceux qui semblent pertinents dans notre contexte. Nous nous appuyons notamment sur : des caractéristiques physiques, fonctionnelles et d'apprentissage. Ce chapitre établit les graphiques potentiellement pertinents pour l'optimisation multicritère en Design by Shopping. Cinq graphiques sont retenus.

Deuxième partie

Problématique et démarche de recherche

2 DEUXIEME PARTIE : Problématique et démarche de recherche

Dans cette deuxième partie, nous présentons notre problématique et démarche de recherche. Nous proposons de la diviser en trois sections distinctes.

La première section positionne nos travaux de recherche. Nous y décrivons la logique de construction de notre problématique. Notre problématique agit à deux niveaux et est donc composée de deux questions de recherche : l'une est liée à la performance en Exploration d'Espace de Conception et l'autre est liée au concepteur engagé dans une approche de Design by Shopping. Ces questions de recherche sont constituées à partir des constats exposés dans notre état de l'art.

Dans la deuxième section, nous proposons la définition de nos hypothèses de résolution de notre problématique de recherche.

La dernière section présente le plan d'actions de notre recherche qui est composé de deux grandes phases.

2.1 Positionnement de nos travaux de recherche

Nos travaux de recherche ont pour ambition d'assister l'activité d'optimisation multicritère en conception lorsque le concepteur est engagé dans une approche de Design by Shopping. Notre recherche porte plus spécifiquement sur l'identification de graphiques pertinents adaptés aux phases de découverte, d'optimisation et de sélection du processus d'exploration d'espace de conception. Nos travaux de recherche portent sur l'exploration d'espace de conception en situation de décision multi-objectif (MODM) mais aussi en situation de décision multi-attribut (MADM) notamment pour la phase de sélection (**premier chapitre**).

La deuxième section de notre première partie (**p. 37**) met en évidence l'implication d'une décision informée dans le processus de conception. Dans cette partie, après un état de l'art, nous définissons la décision informée comme une réponse à une question posée, l'élection d'une alternative parmi plusieurs qui permettra d'obtenir les avantages optimaux et les inconvénients minimaux suite à un processus de traitement et d'analyse itératif et interactif durant lequel l'acteur gagne en compréhension, connaissance et « insights » grâce à la visualisation et la manipulation de grands ensembles et/ou modèle de données. (**deuxième chapitre**).

Dans la dernière section de notre état de l'art, nous présentons de nombreux graphiques utiles pour représenter des ensembles de données multidimensionnels. Notre état de l'art montre que les graphiques ne dévoilent pas la même quantité de données, qu'ils ne se caractérisent pas par les mêmes fonctions et qu'ils ne sont pas, a priori, à utiliser dans les mêmes situations ou micro-situations de l'exploration d'espace de conception (**troisième chapitre**).

L'utilisation des graphiques est inhérente à la conception et l'optimisation multicritère en Design by Shopping. Elle est visible dans les différents outils dédiés à l'exploration d'espace de conception et dans les moyens nécessaires à la décision informée. Seulement, l'utilisation des graphiques n'est pas définie ni guidée pour l'activité de conception multicritère en Design by Shopping. Nous résumons les constats que nous avons identifiés en quatre points :

- La méthode est composée de trois phases différentes elles-mêmes composées de micro-situations nécessitant, a priori, différentes ressources pour le concepteur.

- La phase de sélection peut être exécutée en situation multi-attribut (i.e. espace réduit et nombre limité de solutions possibles) ou en situation multi-objectif (i.e. espace continu et nombre infini de solutions possibles).

- Une décision informée nécessite un gain de compréhension, de connaissance et d'insights à l'aide de visualisation de données.

- Plusieurs graphiques sont identifiés pour être pertinents pour représenter un ensemble de données multidimensionnel (i.e. un même jeu de données) mais ne revêtent pas les mêmes caractéristiques (i.e. des fonctions différentes).

Ainsi, nous formulons la question de recherche qui nous motive comme suit :

« Quel(s) graphique(s) permet(tent) aux concepteurs d'aboutir à une décision informée en exploration d'espace de conception ? »

2.1.1 Définition du périmètre

Nos travaux de recherche concernent une des méthodes utilisée pour la conception multicritère en Design by Shopping : l'Exploration d'espace de conception. Plus précisément, cette méthode intègre trois aspects qui sont les suivants :

- La visualisation d'information : des graphiques permettant la représentation d'un ensemble de données multidimensionnel.
- La décision informée : l'élection d'une alternative parmi plusieurs afin d'obtenir les avantages optimaux et les inconvénients minimaux suite à un processus de traitement et d'analyse itératif et interactif durant lequel l'acteur gagne en compréhension, connaissance et « insights ».
- Le processus de conception en Design by Shopping, i.e. trois phases: la découverte, l'optimisation et la sélection, cette dernière pouvant être exécutée en situation multi-attribut ou en situation multi-objectif.

La **Figure 51** illustre le périmètre de notre recherche. Ce périmètre est décrit comme l'interaction entre plusieurs champs de recherche. En effet, nous le retrouvons au carrefour de l'optimisation multicritère en conception, la décision informée et la visualisation d'information. Cette illustration est utile pour représenter l'interconnexion des trois champs indissociables pour nos travaux de recherche et l'utilisation de la visualisation d'information comme élément clé aux deux composantes de notre problématique de recherche.

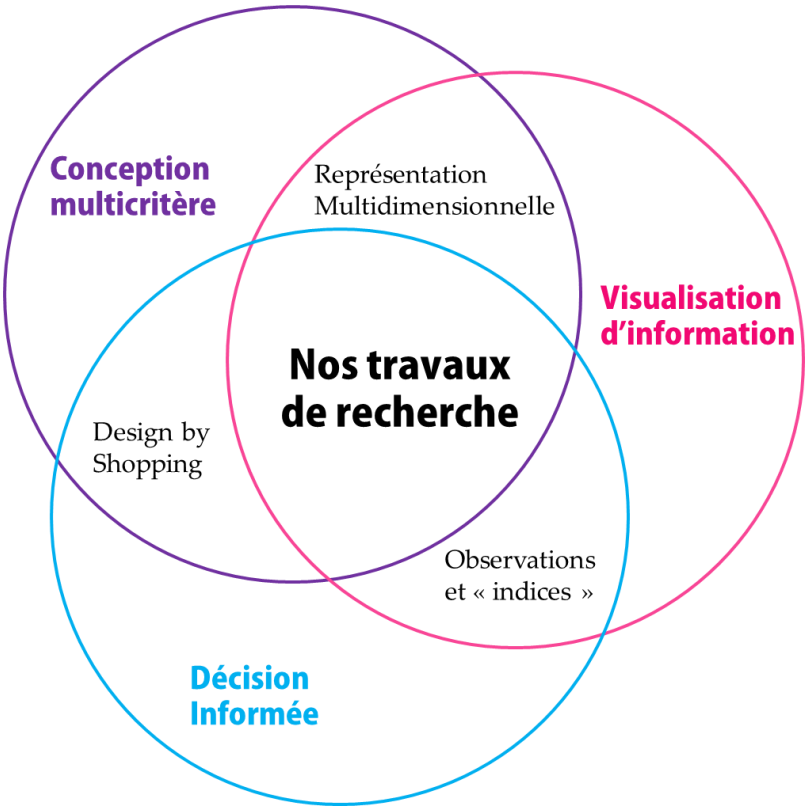


Figure 51 : Proposition du périmètre de nos travaux

2.1.2 Définition des deux composantes de notre problématique

Notre problématique agit à deux niveaux et est donc composée de deux questions de recherche : l'une est liée à la performance en exploration d'espace de conception et l'autre est liée au concepteur engagé dans une approche de Design by Shopping.

Le premier niveau est lié à la performance de la conception multicritère lors de l'exploration d'espace de conception. Ainsi, notre premier objectif est la définition et l'identification de graphiques pertinents pour la résolution de problèmes de conception multicritère dans une approche de Design by Shopping.

Le deuxième niveau est lié au concepteur ; la confiance qu'il porte à sa décision et sa capabilité[4] à aboutir à une décision informée. Cette dernière notion est particulièrement importante car dans cette méthode le concepteur est engagé dans une approche de Design by Shopping.

L'identification de la première composante de notre problématique (**Figure 52**) qui a pour objectif une finalité de performance, nous a conduite à définir des questions de recherche sur lesquelles nous nous positionnons :

- Quel(s) graphique(s) est(sont) petinent(s) pour l'exploration d'espace de conception ?

[4] La capabilité est un pouvoir d'être ou de faire. Elle se différencie de la "capacité" en ce qu'elle est réelle et actuelle (Fernagu Oudet, 2012).

- Y-a-t-il plusieurs graphiques pertinents ou bien un seul est vraiment suffisant ?
- S'il y a plusieurs graphiques pertinents, à quel moments du processus le sont-ils ?
- Existe-t-il une différence de graphiques pour la sélection d'une solution optimale entre le MADM et le MODM ?

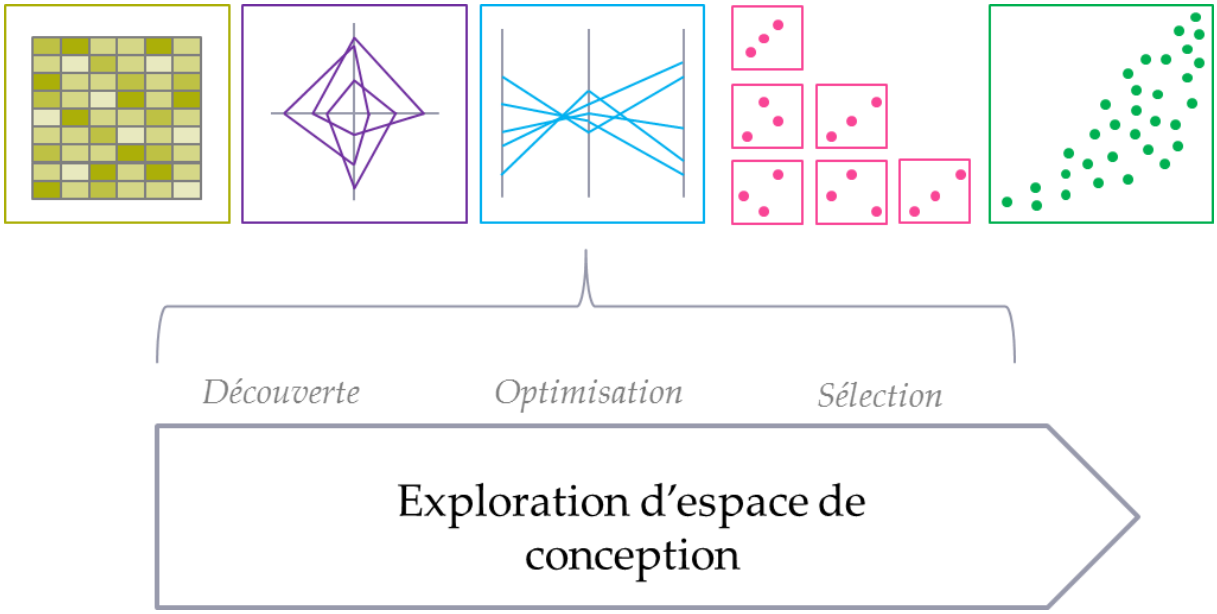


Figure 52 : Illustration de la première composante de notre problématique

Le deuxième objectif, représenté par la deuxième composante de notre problématique de recherche (**Figure 53**), concerne la décision informée, i.e. le concepteur dans un rôle de décisionnaire en exploration d'espace de conception. De même que pour le premier objectif, nous nous sommes positionnés par rapport à certaines questions de recherche :

- Quel(s) graphique(s) permet(tent) d'aboutir à une décision informée en exploration d'espace de conception ?
- Y-a-t-il plusieurs graphiques qui favorisent la décision informée ou bien un seul est vraiment efficace ?
- S'il existe un graphique favorisant la décision informée, permet-il une décision avec relativement de confiance ?
- La phase de découverte est celle où sont faites les « découvertes », ainsi existe-t-il un lien entre cette phase et le graphique favorisant la décision informée ?

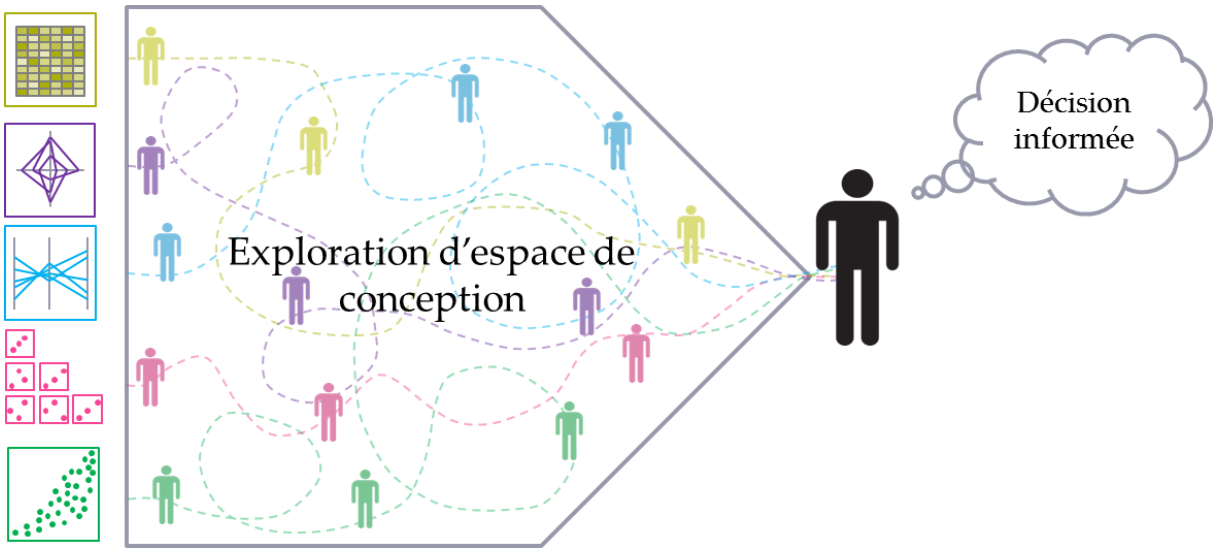


Figure 53 : Illustration de la deuxième composante de notre problématique

Comme nous l'avons montré sur la **Figure 51**, nos travaux de recherche sont à l'interconnexion de trois champs indissociables. Ainsi, les deux composantes de notre problématique de recherche ne peuvent pas être traitées indépendamment.

2.2 Les hypothèses de résolution

Comme nous l'avons présenté précédemment, la première composante de notre problématique concerne la performance en conception multicritère lors de l'exploration d'espace de conception. L'objectif est la définition et l'identification de graphiques pertinents pour la résolution de problèmes multicritères en conception avec une approche de Design by Shopping. Le processus est composé de trois phases qui sont, elles-mêmes, composées de micro-situations. Notre travail, pour cette composante, vise à vérifier trois hypothèses.

Les micro-situations de la « sélection » sont la comparaison de solutions performantes afin d'identifier celles qui sont un bon compromis et la sélection d'une solution qui réponde à certains critères et/ou préférences. Dans la première micro-situation, il s'agit de comparer les solutions entre elles alors que dans la deuxième il s'agit de comparer une solution avec les autres sur l'ensemble des critères.

Nous le traduisons par un besoin en visualisation d'information :

- Une vue d'ensemble de l'espace de conception et de l'ensemble des solutions possibles.
- Une facilité d'identification des solutions les unes par rapport aux autres.

Hypothèse 1 : il est préférable d'utiliser le Parallel Coordinates Plot pour être performant dans la phase de sélection (H1.1 dans une situation multi-attribut et H1.2 dans une situation multi-objectif).

« L'optimisation » se compose également de « micro-situations » telles que la réduction de l'espace de conception vers une zone de performance de l'espace de conception en s'appuyant sur un vecteur de préférence et/ou sur le sous-ensemble de points appelés solutions Pareto-optimales et l'analyse de sensibilité des variables en vue de limiter l'espace de conception à un ensemble de solutions performantes.

Nous le traduisons par un besoin en visualisation d'information :

- Une vue détaillée de l'espace de conception (principalement sur les zones de performance identifiées)
- Une facilité d'interprétation des interactions entre les variables (principalement pour les solutions Pareto-Optimales)

Hypothèse 2 : il est préférable d'utiliser le Simple Scatter Plot pour être performant dans la phase d'optimisation

« La découverte », rassemble plusieurs « micro-situations », à savoir la compréhension du problème en observant les interactions entre les variables d'un point de vue global, l'identification d'une zone de performance, et l'interaction entre les variables d'un point de vue local, c'est-à-dire centré sur les solutions dites optimales.

Nous le traduisons par un besoin en visualisation d'information :

- Une vue d'ensemble de l'espace de conception et de l'ensemble des solutions possibles.
- Une facilité d'identification et d'interprétation des interactions entre les variables (ex. corrélation)

Hypothèse 3 : Il est préférable d'utiliser le Scatter Plot Matrix pour être performant dans la phase de découverte

La deuxième composante de notre problématique de recherche concerne la décision informée i.e. le concepteur comme décisionnaire en exploration d'espace de conception. Les décisions de ce type sont prises suite à un processus de traitement et d'analyse itérative et interactive durant lequel le concepteur gagne en compréhension, connaissance et « insights ».

Nous avons vu précédemment que le gain de compréhension du problème de conception, la connaissance et les insights sont des observations réalisées durant la phase de découverte de l'exploration d'espace de conception. Ainsi, notre travail, pour la deuxième composante, vise à vérifier une quatrième hypothèse :

Hypothèse 4 : L'utilisation du Scatter Plot Matrix permet d'aboutir à une décision informée. Du moins, il est nécessaire d'être performant durant la phase de découverte pour aboutir à une décision informée.

2.3 Notre plan d'actions

Nous présentons dans cette dernière section le plan d'actions de nos travaux de recherche nous ayant permis de répondre à notre problématique de recherche. Nous décidons de traiter les situations multi-attributs et multi-objectifs indépendamment. A travers nos expérimentations, nous découpons nos travaux en deux parties.

La première partie traite la spécificité de la phase de sélection dans une situation multi-attribut. Trois graphiques ont été identifiés comme potentiellement pertinents pour la sélection dans une situation de prise de décision multi-attribut (voir **Figure 54**). Ainsi, la première partie tente de vérifier la sous-hypothèse H1.1.

La deuxième partie est consacrée aux trois phases de l'exploration d'espace de conception dans une situation multi-objectif. Trois graphiques ont été identifiés comme potentiellement pertinents dans une situation de prise de décision multi-objectif (voir **Figure 54**). Cette partie est la plus importante puisqu'elle tente de vérifier la sous hypothèse H1.2 et les hypothèses H2, H3 et H4.

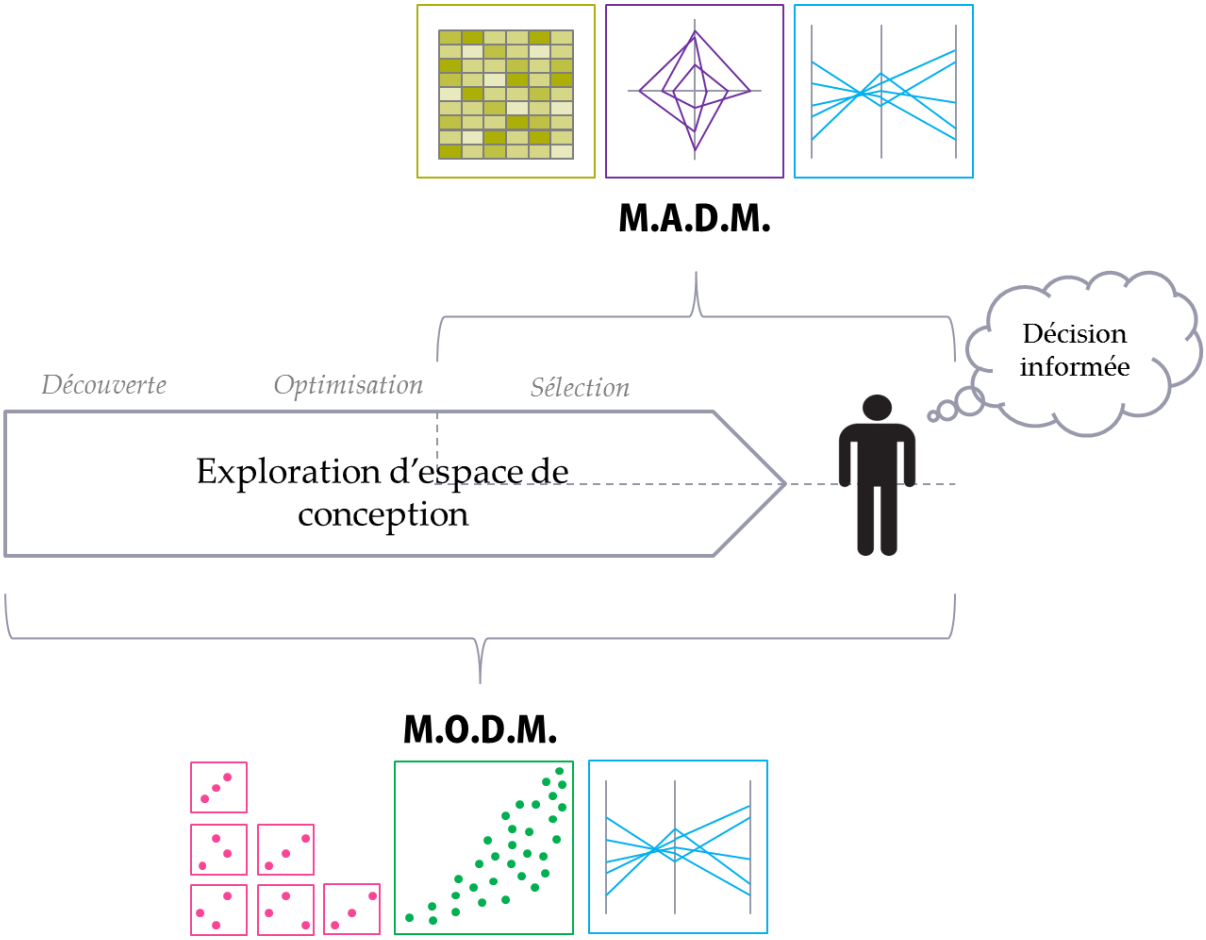


Figure 54 : Découpage de notre plan d'actions

CONCLUSION DE LA DEUXIEME PARTIE

Dans la deuxième partie nous avons décrit notre problématique de recherche qui est :

« Quel(s) graphique(s) permet(tent) aux concepteurs d'aboutir à une décision informée lors de l'exploration de l'espace de conception ? »

Cette problématique est constituée de deux composantes :

- La première composante est liée à la performance en conception multicritère lors de l'exploration d'espace de conception. Le premier objectif est la définition et l'identification de graphiques pertinents pour la résolution de problèmes de conception multicritères dans une approche de Design by Shopping.
- La deuxième composante est liée au concepteur. Elle concerne la décision informée i.e. le concepteur comme décisionnaire en exploration d'espace de conception. Le deuxième objectif est l'identification de graphiques permettant d'aboutir à une décision informée i.e. une décision prise principalement à partir de l'information acquise durant l'exploration d'espace de conception.

A partir de ces deux composantes de notre problématique de recherche découlent nos quatre hypothèses de recherche.

Enfin, nous terminons ce chapitre avec la description d'un plan d'actions pour lequel nous proposons un découpage de nos travaux en deux parties pour vérifier nos 4 hypothèses de recherche.

Troisième partie

EXPERIMENTATIONS

3 TROISIEME PARTIE : Expérimentations

Cette partie présente les expérimentations réalisées dans nos travaux pour répondre à nos hypothèses de recherche.

L'expérimentation n°1 traite des graphiques pertinents pour la phase de sélection dans une situation de prise de décision multi-attribut.

L'expérimentation n°2 traite deux aspects :

- Les graphiques performants pour la résolution de problèmes multicritères en exploration d'espace de conception dans une situation de prise de décision multi-objectif (Expé 2.1 à 2.3 sur la **Figure 55**)
- Les graphiques pour la décision informée en exploration d'espace de conception (Expé 2.4 sur la **Figure 55**)

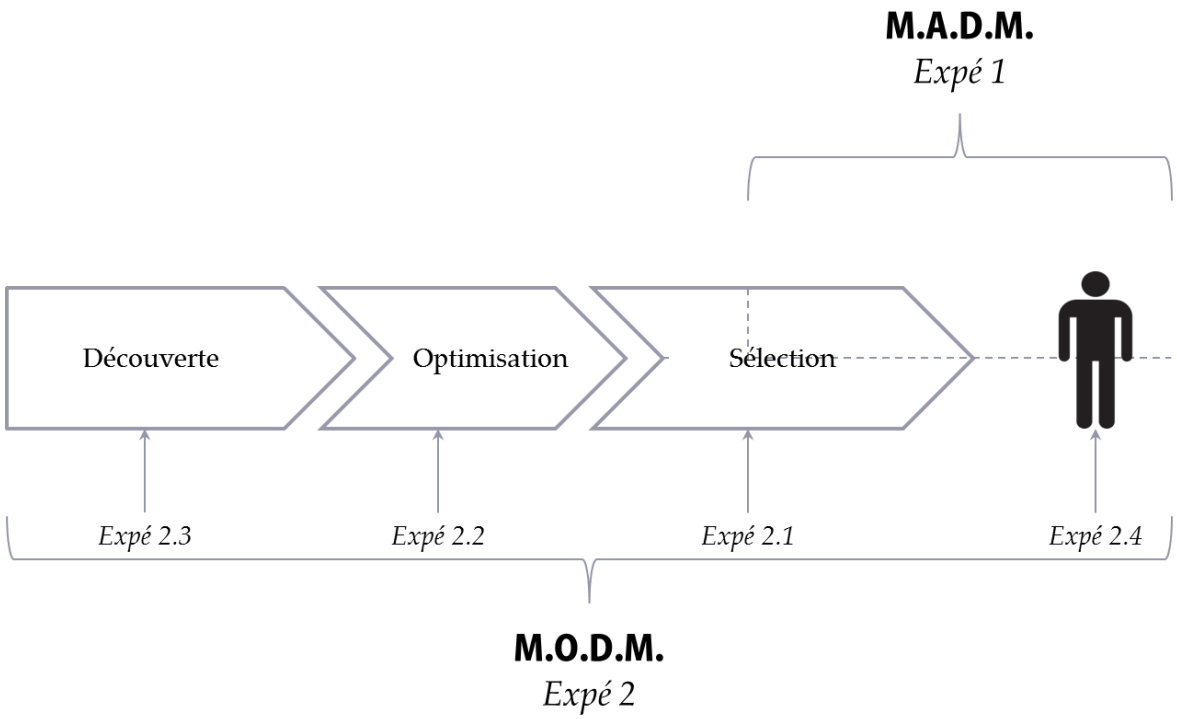


Figure 55 : Illustration du découpage et positionnement des expérimentations

Pour l'expérimentation n°2 nous décidons de découper la présentation en 4 sous-parties :

- Expérimentation 2.1 qui traite la phase de sélection (section 0 p. 159)
- Expérimentation 2.2 qui traite la phase d'optimisation (section 3.7 p. 163)
- Expérimentation 2.3 qui traite la phase de découverte (section 3.8 p. 169)
- Expérimentation 2.4 qui traite la décision informée (section 3.10 p. 179)

Pour chacune d'entre elles, nous décrivons les mesures utilisées et le traitement des résultats. Nous proposons une conclusion intermédiaire à la fin de chaque sous-partie.

Un guide de lecture de la troisième Partie du manuscrit est illustré dans la **Figure 56**.

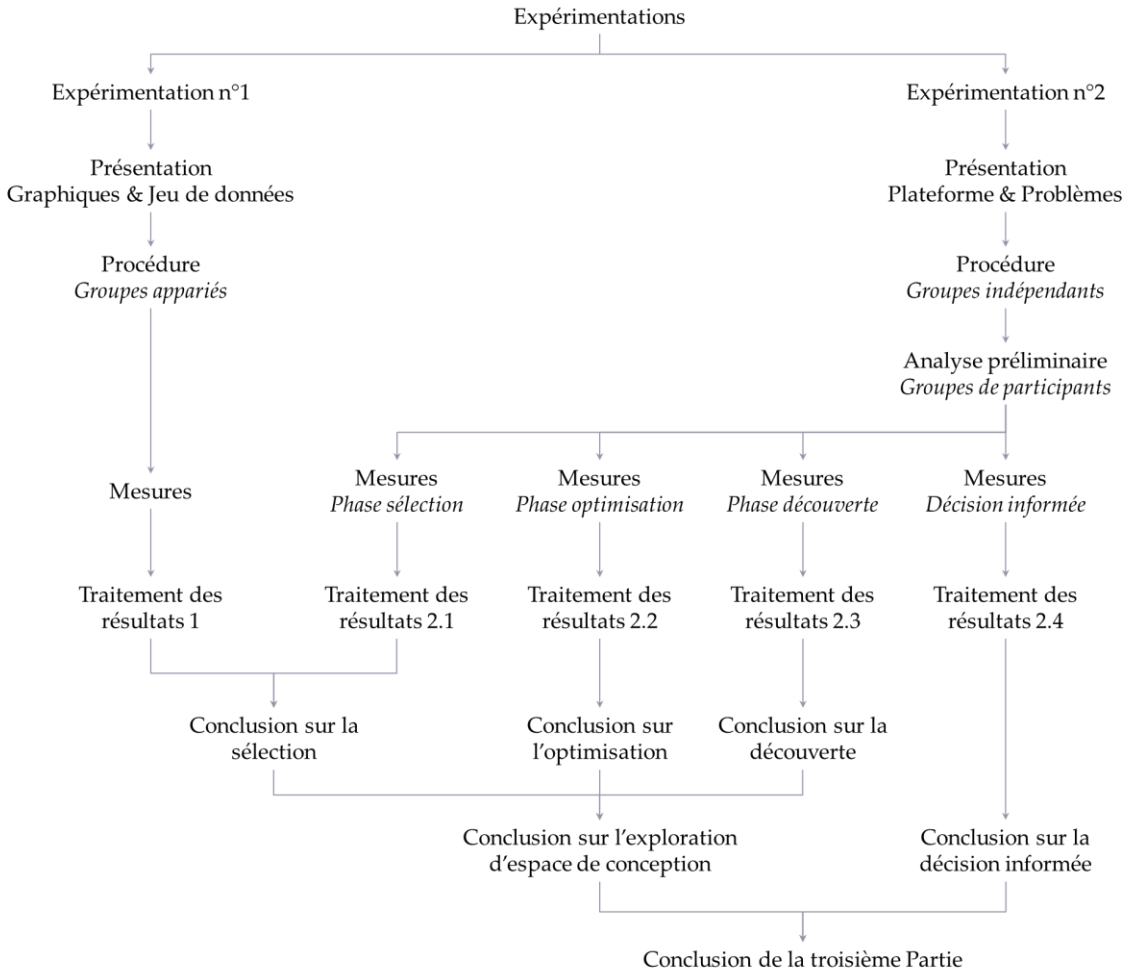


Figure 56 : Guide de lecture de la troisième Partie

Les analyses statistiques réalisées pour les 2 expérimentations ont été effectuées à l'aide du site de computation statistique : http://vassarstats.net/ et le choix des tests statistiques se sont fait à partir du cours en ligne de Hunault Gilles de l'Université d'Angers (Hunault, consulté en 2015).

Enfin, l'ensemble des supports utilisés pour les expérimentations (graphiques et plateforme d'expérimentation) sont disponibles en ligne :

- Pour l'expérimentation n°1 : http://these.aaa.alwaysdata.net/expe1/
- Pour l'expérimentation n°2 : http://these.aaa.alwaysdata.net/expe2/

3.1 INTRODUCTION DE L'EXPERIMENTATION N° 1

L'objectif de l'expérimentation n°1 est l'identification d'un graphique qui rend la sélection performante en conception dans une situation de prise de décision multi-attribut (notamment en *Design by Shopping*).

Trois graphiques ont été identifiés pour permettre d'être performant durant la phase de sélection lors de l'exploration d'espace de conception : Le Radar Chart (RC), le Combined Table (CT) et le Parallel Coordinates Plot (PCP).

Nous proposons de mesurer durant les tests trois aspects que nous détaillons dans la section 3.1.3 :

- La recherche de compromis
- La solution sélectionnée (la réalisation du compromis)
- Le ressenti du participant

3.1.1 Description des graphiques et du jeu de données

3.1.1.1 Le Radar Chart

Dans un Radar Chart, les valeurs des variables sont affichées sur des axes distincts disposés de manière radiale concourants en un point (voir **Figure 57**). Les points de conception sont représentés par des lignes qui relient les points sur les axes. Les points sur les axes représentent les valeurs des paramètres de conception et des variables de performance des solutions possibles.

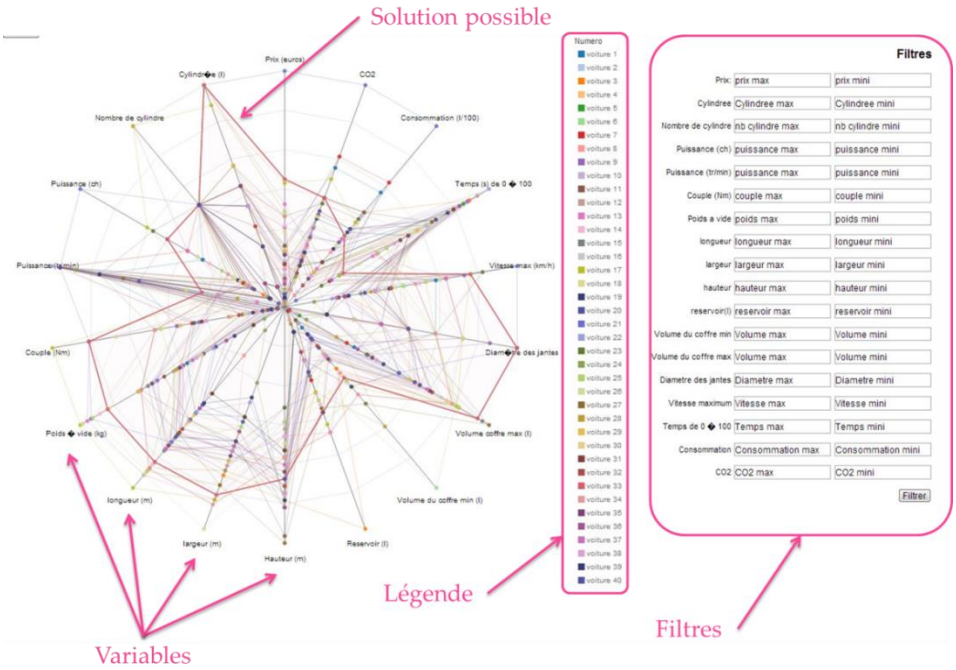


Figure 57 : Description du Radar Chart

Dans notre implémentation, l'interaction est obtenue par la souris premièrement sur les points (donc sur les axes) affichant la valeur (de la solution pour la variable en question) et deuxièmement sur les lignes affichant le numéro du point de conception (solution possible). L'action de filtrage est effectuée en utilisant un formulaire à droite du graphique. Nous avons choisie d'utiliser un formulaire pour ne pas surcharger le graphique.

3.1.1.2 Le Combined Table

Le Combined Table est la combinaison d'un tableau de valeurs et d'une carte de chaleur (voir **Figure 58**). Dans notre implémentation, chaque ligne représente un point de conception et chaque colonne représente une variable (paramètre de conception ou variable de performance). L'interaction avec le graphique s'opère sur les labels (noms des variables) en haut de chaque colonne. En cliquant sur ces labels, les valeurs sont triées de la plus grande à la plus petite, ainsi les solutions sont réordonnées dans le tableau.

Outre les valeurs de la table, nous avons ajouté une carte de chaleur. En principe, les cartes carte de chaleur sont dans des matrices dans lesquelles les cellules sont colorées en fonction des valeurs. Dans chaque colonne, la valeur la plus élevée a une cellule rouge vif. Le rouge est dégradé jusqu'à la couleur blanche pour la plus faible valeur de la colonne. De plus, au survol de la souris, la ligne entière (correspondant à une solution) est soulignée (emphase) et le numéro de la solution apparait dans un encadré bleu en haut à gauche de l'écran.

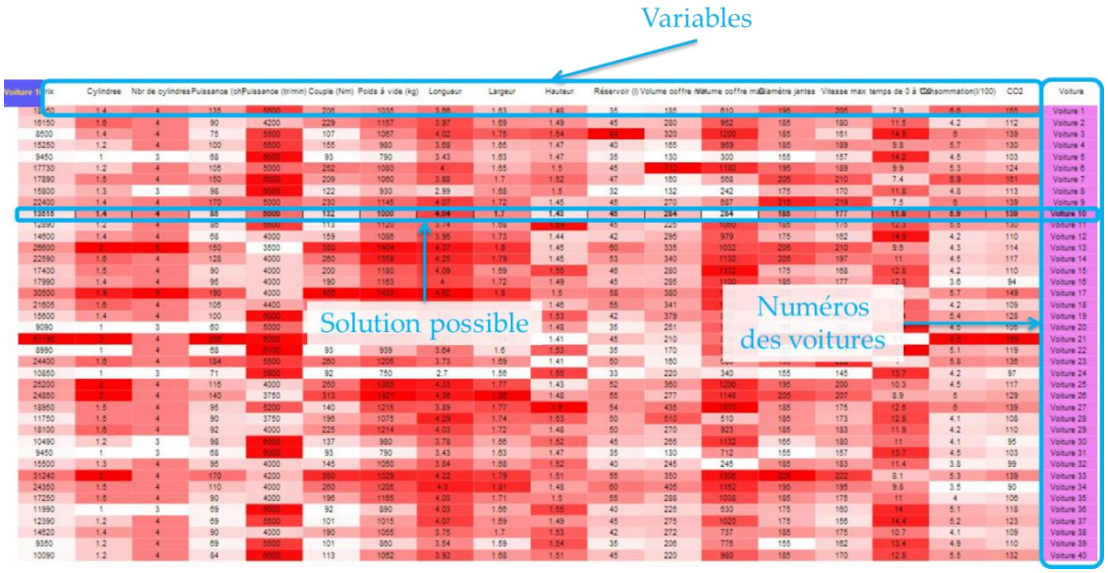


Figure 58 : Description du Combined Table

3.1.1.3 Le Parallel Coordinates Plot

Dans le Parallel Coordinates Plot, les valeurs des variables sont affichées sur des axes séparés disposés en parallèle. Les points de conception (ou solutions) sont représentés par des lignes qui relient les points sur les axes. Les points sur les axes représentent les valeurs de la variable (représentée par l'axe). Dans notre implémentation, l'interaction est directe sur le

graphique. Pour filtrer les valeurs, il suffit de placer la souris sur l'un des axes et, en un clic, faire glisser le curseur pour obtenir une plage désirée du critère comme une fenêtre où les vecteurs de valeurs à l'extérieur sont exclus. Enfin, lorsque la souris est placée au-dessus d'une ligne (c'est-à-dire un point de conception), le numéro de la solution est affiché en haut à gauche de l'interface (voir **Figure 59**).

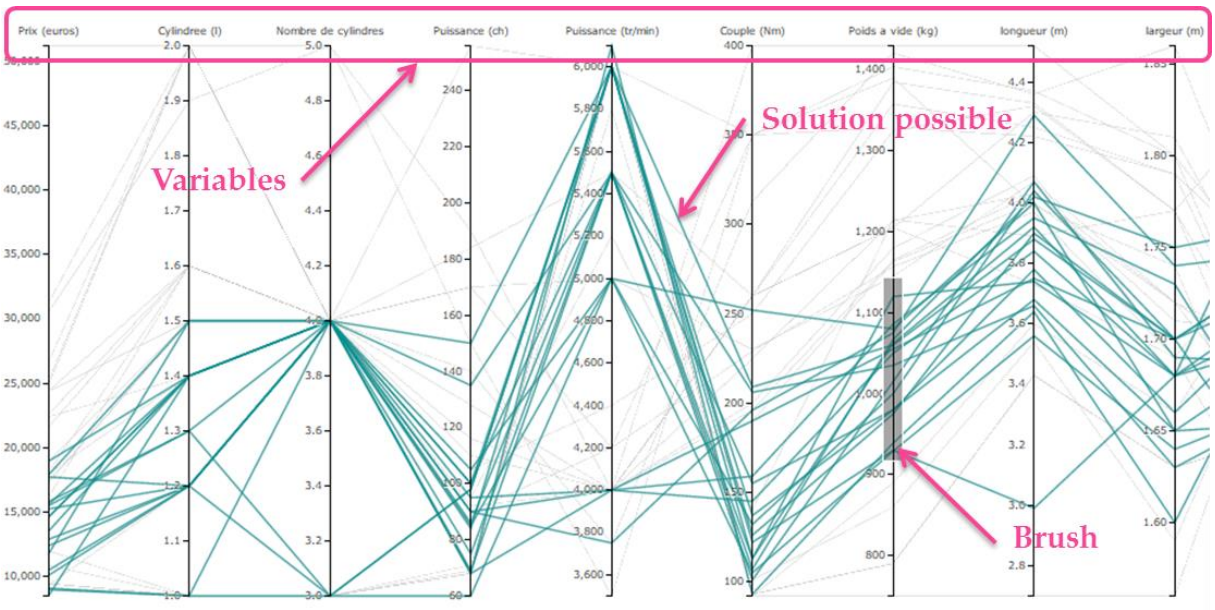


Figure 59 : Description du Parallel Coordinates Plot

3.1.1.4 Jeu de donnée

Le jeu de données a été créé à partir de caractéristiques de voitures (tiré du site web : http://comparatif.larevueautomobile.com/). Les caractéristiques de quarante voitures ont été relevées et sur chacune des voitures plus de dix-huit attributs sont disponibles. Nous obtenons donc un jeu de 720 données exploitables pour les tests (voir **Annexes p.210**).

Notons que pour l'expérimentation nous retirerons la marque et le modèle des voitures pour éviter l'influence sur le participant liée à l'image qu'il se fait de la voiture (ex : si je dois chercher une voiture avec un bon rapport qualité/prix je vais probablement regarder la Dacia Sandero).

Comme exprimé précédemment les graphiques sont disponibles en ligne à cette adresse : http://these.aaa.alwaysdata.net/expe1/

3.1.2 Déroulement de l'expérimentation n° 1

L'expérimentation est découpée en trois étapes (voir **Figure 60**). Dans la première étape, nous proposons tout d'abord une présentation des objectifs de l'expérimentation au participant puis un questionnaire lui est proposé. Ce premier questionnaire nous permet d'obtenir les informations générales du participant (ex : son âge, sexe, etc.) et sa connaissance des trois graphiques qui sont testés.

La deuxième étape de l'expérimentation débute par la présentation du graphique avec lequel le participant va interagir i.e. l'organisation de l'information et les interactions. Nous poursuivons la séance avec une phase d'entrainement. Dans cette phase, le participant est amené à interagir avec le graphique (présenté précédemment) afin de trouver quatre items (une solution possible, une valeur, etc.). Une fois la phase d'entrainement terminée, le participant réalise la partie test i.e. la sélection d'une solution en explorant l'espace des possibles à l'aide du graphique. L'étape 2 est répétée trois fois. En effet, pour l'expérimentation nous suivons une approche par groupes appariés i.e. chaque participant test les trois graphiques.

Enfin, dans la troisième étape, les participants répondent à un dernier questionnaire. Ce questionnaire post-test nous permet de relever la charge cognitive et la confiance du participant.

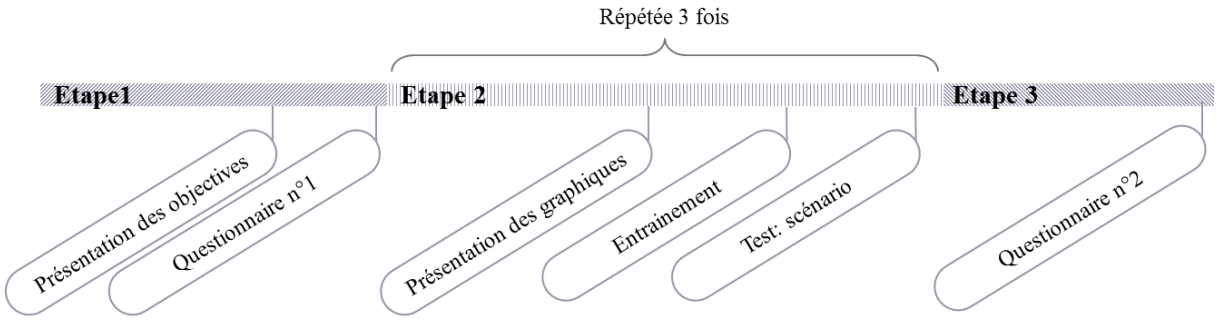


Figure 60 : Illustration du déroulement de l'expérimentation n°1

Pour favoriser les résultats de l'expérimentation, comme mentionné précédemment, nous optons pour une approche appariée c'est-à-dire que chaque participant teste les trois graphiques à la suite. Notons que cela implique la construction de trois scénarios de conception différents et l'utilisation d'une méthode pour éviter les effets d'ordre (et de rang) et l'apprentissage.

3.1.2.1 Effet d'ordre et d'apprentissage

Afin de ne pas introduire de biais dans notre expérimentation induit par l'ordre d'apparition des graphiques qui peut entrainer un phénomène d'apprentissage chez le participant (i.e. peu importe l'ordre des graphiques, le participant est de plus en plus performant). Nous utilisons la méthode de contre-balancement (complet). Dans notre cas nous avons trois graphiques à tester donc 3!=6 séquences à mettre en place (**Tableau 4**).

Tableau 4 : Contrebalancement des graphiques

N° de la séquence	Nom du graphique		
	Graph A	Graph B	Graph C
1	CT	RC	PCP
2	CT	PCP	RC
3	RC	CT	PCP
4	RC	PCP	CT
5	PCP	RC	CT
6	PCP	CT	RC

3.1.2.2 La phase d'entrainement

Comme nous l'avons précisé dans la section « déroulement » (section 3.1.2), pour chaque graphique, le participant débute par une phase d'entrainement. Cet entrainement consiste principalement en la prise en main de l'interface graphique. Ainsi, nous proposons un exercice dans lequel il faut trouver quatre items. Cet exercice est inspiré des travaux de Ghoniem et al. (2005) que nous avons adapté à notre cas car nous ne traitons pas le même type de données. Dans leurs travaux, ils manipulent un ensemble de données en réseaux et sept tâches sont demandées aux participants:

1. Estimation du nombre de nœuds
2. Estimation du nombre de liens
3. Trouver le nœud le plus connecté
4. Trouver le nœud (à partir de son label)
5. Trouver un lien entre deux nœuds spécifiés
6. Trouver un voisin commun entre deux nœuds spécifiés
7. Trouver un chemin entre deux nœuds spécifiés

Il s'agit donc dans notre cas de trouver une valeur spécifique, une alternative, un maximum et des équivalences.

De même que nous concevons trois scénarios de conception pour la phase de test, notre approche par groupes appariés nous impose de concevoir trois entrainements différents :

- Premier *Training* :
  - Trouvez la voiture qui a le plus gros réservoir
  - Trouvez une voiture qui a un poids égale à 1120kg
  - Trouvez le nombre de chevaux (« puissance(ch) ») de la voiture 5

    - Trouvez deux voitures qui ont la même cylindrée, le même nombre de cylindres et un diamètre de jantes identiques
- Deuxième *Training* :
    - Trouvez la voiture qui a le plus grand coffre min
    - Trouvez une voiture qui a un poids égale à 1062kg
    - Trouvez le nombre de chevaux (« puissance(ch) ») de la voiture 38
    - Trouvez deux voitures qui ont la même cylindrée, le même nombre de cylindres et la même puissance (tr/min)
- Troisième *Training* :
    - Trouvez la voiture qui a la plus grande taille en hauteur
    - Trouvez une voiture qui a un poids égale à 1325kg
    - Trouvez le nombre de chevaux (« puissance(ch) ») de la voiture 9
    - Trouvez deux voitures qui ont la même cylindrée, le même couple (Nm) et la même hauteur

3.1.2.3 Test : Scénarios

Nous proposons une expérimentation composée de trois scenarios conduisant à la sélection d'une alternative dans un problème de choix d'une voiture suivant des critères fixés. Les trois scénarios ont été imaginés pour imiter l'activité de conception dans un cadre simplifié afin de se concentrer sur la phase de sélection du processus de conception. Les scénarios sont conçus comme des personnas[5] ou encore des *brief marketing* dans lesquels est décrit un profil-cible. Pour les trois cas, cinq objectifs en conflit doivent être satisfaits afin de sélectionner une alternative. Autrement dit, il est nécessaire de réaliser un compromis afin de choisir un idéal dans une situation de conception/sélection de voiture pour un profil-cible (défini dans les trois scénarios).

Ainsi, nous avons conçu trois scénarios : Le scénario « Taille », le scénario « Puissance » et le scénario « Trajet ». Les scénarios sont définis tels que :

Scénario « **Taille** » : Le profil-cible a une famille composée de trois enfants et un(e) conjoint(e). Il travaille tous les jours dans une grande ville où il est difficile de se garer. C'est pourquoi vous cherchez une voiture qui soit la plus petite possible pour qu'il puisse se garer facilement (largeur et longueur) mais avec un volume intérieur suffisant pour lui et sa famille (hauteur, capacité du coffre et place à l'arrière).

Scénario « **Puissance** » : Le profil-cible aime les voitures puissantes. La vitesse maximum du véhicule doit être pour lui la plus grande avec la meilleure accélération (temps pour passer de 0 à 100km/h). La voiture doit avoir aussi un nombre important de chevaux et un couple

[5] Le persona n'est pas la représentation d'une personne réelle mais une abstraction simplifiée d'un utilisateur archétypique, fictif, imaginaire, du futur système. L'expression d'un persona peut être formulée sous une forme narrative, propice à être intégrée dans un scénario. (Ruault et al., 2012)

moteur important. Enfin, pour assurer que la voiture soit la plus puissante, vous choisirez une voiture qui soit légère (poids de la voiture).

Scénario « **Trajet** » : Le profil-cible est une personne qui part quasiment tous les week-ends et qui utilise sa voiture pour ses trajets. Ainsi, vous lui cherchez une voiture avec un grand réservoir et le minimum de consommation car le profil-cible ne veut pas s' arrêter plusieurs fois sur le trajet pour faire le plein d'essence et ne pas payer trop cher ses déplacements. Comme il part en voyage souvent avec sa voiture, celle-ci doit être grande (longueur, largeur, hauteur, taille coffre = max) et enfin, vous choisirez une voiture avec une bonne tenue de route (diamètre des jantes=max).

Une synthèse des cinq objectifs par scénario est proposée dans la **Figure 61**.

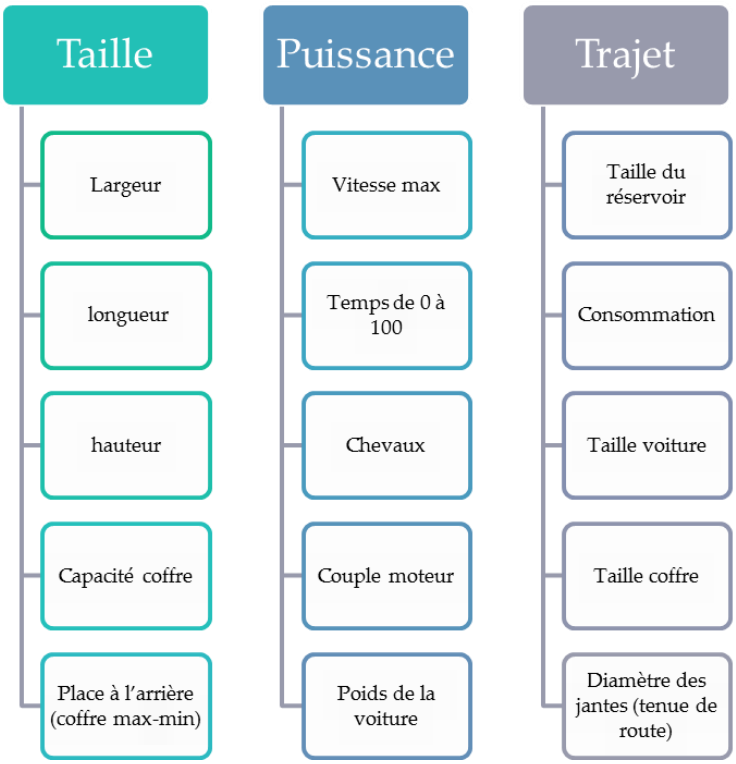


Figure 61 : Synthèse des cinq objectifs des trois scénarios

3.1.3 Les mesures

Dans cette section nous décrivons les différentes mesures utilisées dans l'expérimentation.

3.1.3.1 Etape 1 : Indicateurs sur les participants

Dans la première étape, outre la présentation des objectifs de l'expérimentation au participant, un questionnaire lui est proposé. Ce premier questionnaire nous permet d'obtenir les informations générales du participant (ex : son âge, sexe, etc.) sa représentation des voitures et sa connaissance des 3troisgraphiques qui sont testés :

Information générale : l'âge, sexe et activité professionnelle

Image de la voiture :

- Quels sont les critères les plus importants pour vous lors du choix d'une voiture neuve ? (Donnez les trois critères les plus importants par ordre d'importance)
- Même question, mais en considérant que le budget dont vous disposez pour l'achat de la voiture est illimité.
- La puissance de la voiture pour vous se traduit par (sélectionner les critères).

Connaissance des graphiques (questions posées pour les trois graphiques) :

- Avez-vous déjà vu ce type de graphique ?
- Avez-vous déjà utilisé/travaillé avec ce type de graphique ?
- Avez-vous déjà interagie avec ce type de graphique ?
- Si oui, depuis combien de temps utilisez-vous ce type de graphique ?

Notons que l'intérêt pour nous des questions portant sur la voiture est de contrôler une éventuelle décision aberrante du participant. C'est-à-dire la sélection d'une voiture en accord avec les préférences du participant et non du scénario proposé.

3.1.3.2 Etape 2 : Test

Durant le test sur les trois graphiques, nos mesures portent sur deux aspects :

- La recherche de compromis
- La solution sélectionnée (i.e. réalisation du compromis)

3.1.3.2.1 La recherche de compromis

Afin de mesurer l'influence d'un graphique sur la recherche de compromis nous utilisons deux indicateurs : le nombre d'actions réalisées durant l'exploration et le temps écoulé entre les actions.

Nous appelons « actions » toutes interactions avec le graphique. Il peut s'agir de l'utilisation des fonctions de « brush » avec le PCP, de filtre avec le RC ou de tri avec le CT. Chaque action est une indication de l'activité du concepteur et donc de sa recherche d'un compromis.

3.1.3.2.2 La solution sélectionnée

Le deuxième aspect mesuré durant le test sur les graphiques est composé de deux indicateurs également : le temps total et la qualité de la solution sélectionnée.

Ainsi, nous mesurons l'influence d'un graphique sur le compromis. Pour nous, le compromis est représenté par la solution sélectionnée par le participant. Nous mesurons donc :

- Le temps total écoulé pour prendre une décision i.e. pour sélectionner une solution optimale.

- La qualité de la solution sélectionnée mesuré à l'aide du calcul automatisé d'un indice (la construction de cet indice est décrite dans la section 3.2).

3.1.3.3 Etape 3 : Ressenti du participant

Enfin, dans la dernière partie de l'expérimentation, nous nous intéressons à l'influence du graphique sur le ressenti du concepteur notamment en mesurant la charge cognitive perçue et la confiance du participant en sa décision. Ces mesures sont relevées à l'aide d'un questionnaire.

Pour mesurer la charge cognitive nous choisissons la méthode DALI (Pauzié, 2008). Nous optons pour cette méthode car elle offre un plus large éventail de questions en comparaison aux méthodes NASA TLX ou SWAT. Il est entendu que les questions ont été adaptées à notre contexte de recherche de compromis et de sélection d'une solution optimale (voir **Tableau 5**).

Pour mesurer la confiance du participant, nous utilisons la méthode de Aloysius et al. (Aloysius et al., 2006) scindée en deux variables : le niveau de certitude que le participant porte à la solution sélectionnée et sa préférence d'utilisation entre les trois graphiques (préférence pour la recherche de compromis et préférence pour réaliser le compromis).

Dans ce questionnaire les réponses aux questions sont données sur une échelle comprise entre 0 et 100.

Tableau 5 : Mesures et objectifs du questionnaire post-test

Objectifs	Méthodes	Variables
Influence du graphique sur la charge cognitive	DALI (Pauzié 2008) adapté à notre cas	Effort d'attention pour la recherche de compromis
		Effort d'attention pour la sélection d'une solution
		Effort visuel pour la recherche de compromis
		Effort visuel pour la sélection d'une solution
		Stress perçu
Influence du graphique sur la confiance	Aloysius et al. (2006) adapté à notre cas	Niveau de certitude dans la solution sélectionnée
		Préférence du graphique

3.1.4 Synthèse

Cette expérimentation nous permet de vérifier l'influence de trois graphiques sur la sélection d'une solution optimale en *Design by Shopping* à travers trois aspects :

- La recherche de compromis
- La solution sélectionnée
- Le ressenti du participant

Une synthèse des différentes mesures est illustrée dans la **Figure 62**.

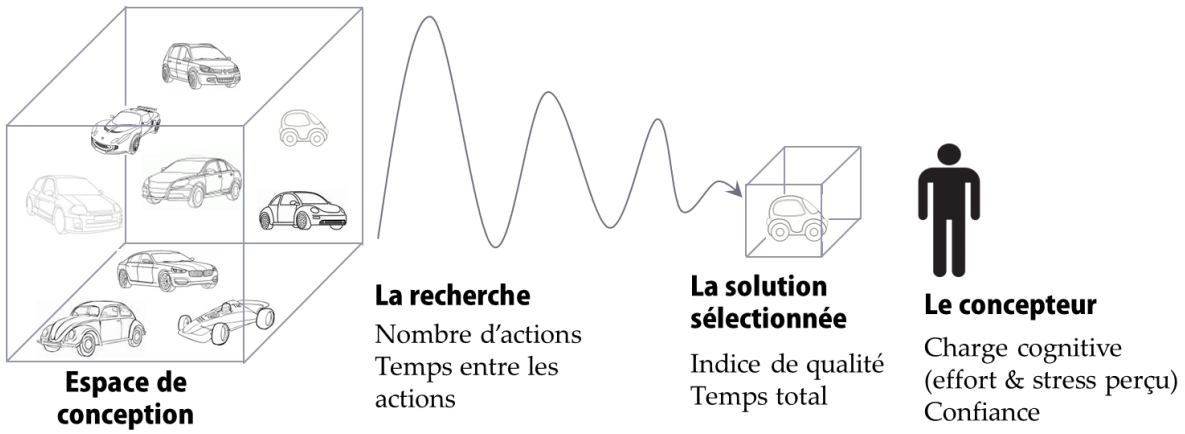


Figure 62 : Synthèse des mesures de l'expérimentation n°1

3.2 INDICE DE QUALITE

Dans cette section nous décrivons la construction d'un indice de qualité relié à un scénario de conception. Nous cherchons ici à calculer de manière automatique un indice pour déterminer quel(s) graphique(s) permet(tent) aux concepteurs de prendre les « meilleures » décisions.

Dans l'expérimentation n°1, nous avons trois graphiques (PCP, RC, et CT) pour l'exploration et la sélection d'une voiture parmi quarante. Trente participants ont chacun joué trois rôles décrits dans trois scénarii (TAILLE, PUISSANCE et TRAJET) en utilisant un des graphiques pour chaque scénario.

Dans chacun des scenarii, il y a cinq variables de performance, ayant des poids de préférence équivalent, à maximiser ou minimiser.

Pour le scénario TAILLE, les variables à maximiser sont : la hauteur, la taille du coffre « min » et la différence entre la taille du coffre « max » et la taille du coffre « min ». Les variables à minimiser sont : la longueur et la largeur.

Pour le scénario PUISSANCE, les variables à maximiser sont : le nombre de chevaux, le couple moteur et la vitesse « max ». Les variables à minimiser sont : le poids du véhicule et le temps nécessaire pour passer de 0 à 100 km/h.

Pour le scénario TRAJET, les variables à maximiser sont : la taille du réservoir, la largeur la taille du coffre « min » et le diamètre des jantes. La variable à minimiser est la consommation.

Afin d'évaluer la performance de la solution qui est sélectionnée par le participant, nous proposons l'élaboration d'un indice de qualité à partir de méthodes d'optimisation multicritère (ou multi-objectif). Ainsi, nous présentons premièrement le principe de dominance au sens de Pareto, puis la construction d'un Indice Basique utilisé comme critère de synthèse pour finir par la présentation de trois indices de qualité applicables à notre contexte et les résultats.

3.2.1 Recherche des solutions non-dominées

Nous utilisons le principe de Pareto pour identifier dans chacun des trois scénarii les solutions dites non-dominées (ou Pareto-optimales).

3.2.1.1 Dominance au sens de Pareto

Considérons un problème multi-objectif tel que :

$min(F(x) = (f_1(x), f_2(x), \ldots, f_n(x)), n \geq 2$

Où :

- $x \in X$
- X est l'espace de décision
- Y est l'espace objectif (ou de performance)
- Et $Y = F(X)$

On définit la dominance tel que :

y domine z ssi $\forall\, i \in [1 \ldots n]$, $f_i(y) \leq f_i(z)$ et $\exists\, k \in [1 \ldots n] \mid f_k(y) < f_k(z)$.

Et la dominance au sens de Pareto se définit tel que :

Une solution $x^p \in X$ est Pareto-Optimale ssi $\nexists\, x \in X$ tel que x domine x^p

L'image de l'ensemble des solutions Pareto-Optimales dans l'espace objectif est appelé front ou frontière de Pareto (ou surface de compromis).

Enfin, l'évaluation des performances des solutions dominées s'effectue en calculant la distance euclidienne entre la solution dominée et la frontière de Pareto. Dans notre cas, nous calculons la distance minimum entre la solution dominée et une solution non-dominée (plus proche voisin non-dominé).

$$d = \min_x \left(\sqrt{\sum_{i=1}^{n} \left(x_i - x_i^p\right)^2} \right)$$

- $\forall\, i \in [1 \ldots n]$, n étant le nombre de variables de performance
- $\forall\, x \in X$ solution dominée et $x^p \in X$ solution non-dominée

3.2.1.2 Pareto appliqué au Scénario TAILLE

Dans ce scenario, il y a vingt-et-une solutions non-dominées au sens de Pareto. Nous remarquons le nombre important de solutions non-dominées sur la **Figure 63** où en rose sont représentées les solutions non-dominées et en gris les solutions dominées. La liste des solutions avec le nombre de dominations et la distance au front de Pareto pour le scénario TAILLE est disponible en annexes p.210.

Nous pouvons noter que la solution n°27 est celle qui domine le plus de solutions (voir **Figure 64**).

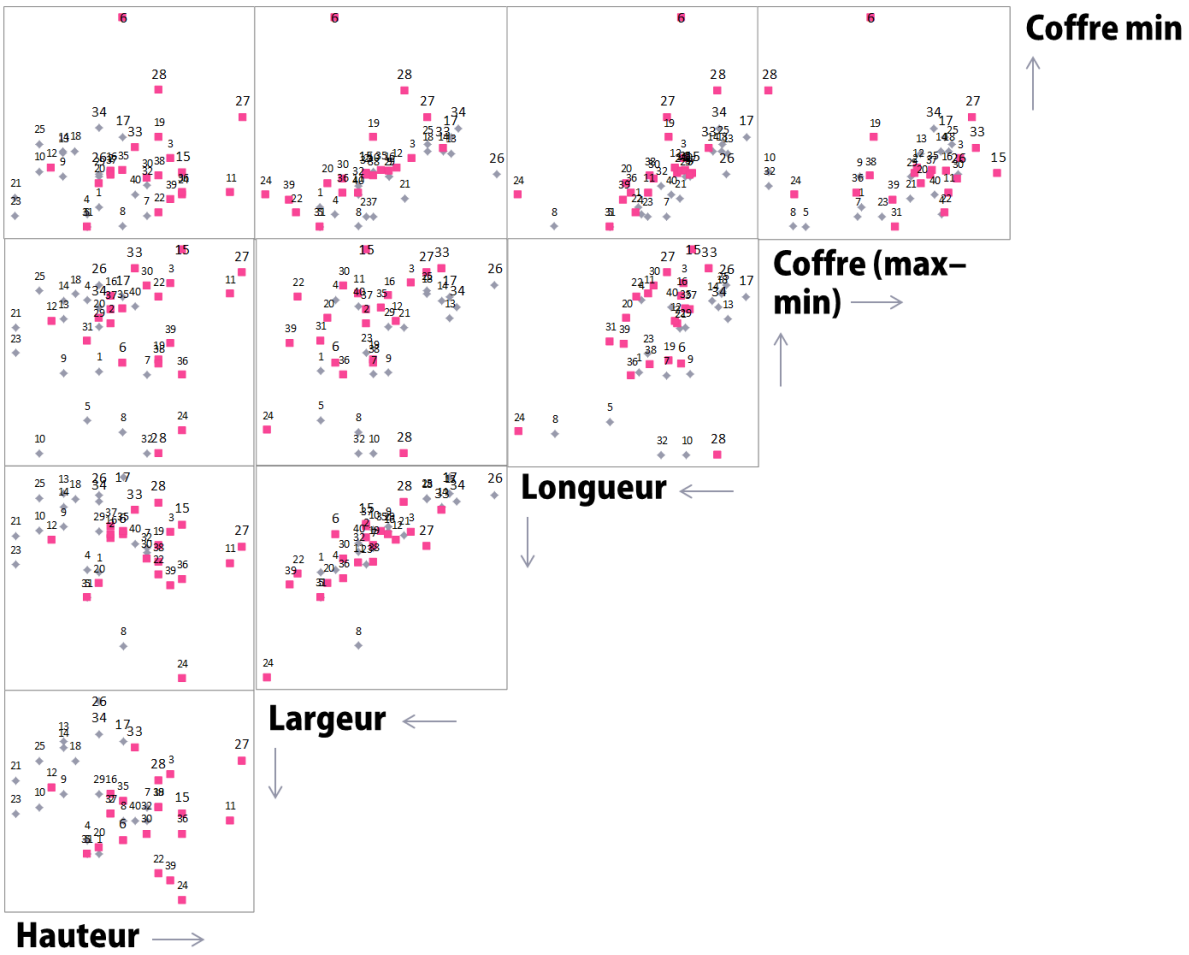


Figure 63 : Exemple scénario TAILLE (en rose les solutions dites non-dominées au sens de Pareto)

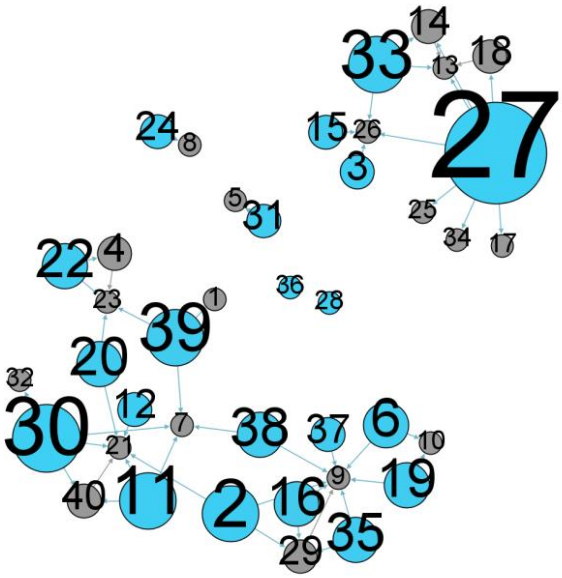


Figure 64 : Carte des dominations pour le scénario Taille. Les ronds gris représentent les solutions dominées et la taille des ronds est proportionnelle au nombre de dominations qu'effectue la solution.

3.2.1.3 Pareto appliqué au Scénario PUISSANCE

Dans ce scenario, il y a quinze solutions non-dominées au sens de Pareto. La solution n°1 est celle qui domine le plus de solutions suivie de près par les solutions n°6, 7 et 4 (voir **Figure 65).** La liste des solutions avec le nombre de dominations et la distance au front de Pareto pour le scénario PUISSANCE est disponible en annexes p.211.

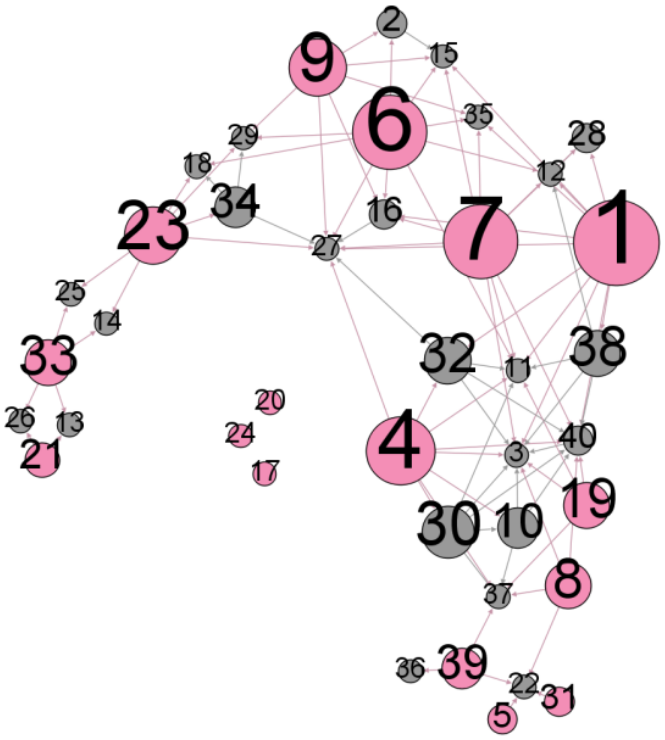


Figure 65 : Carte des dominations pour le scénario Puissance. Les ronds gris représentent les solutions dominées et la taille des ronds est proportionnelle au nombre de dominations qu'effectue la solution.

3.2.1.4 Pareto appliqué au Scénario TRAJET

Dans ce scenario, il y a dix solutions non-dominées au sens de Pareto. La solution n°34 est celle qui domine le plus de solutions (avec vingt-sept dominations). Nous pouvons observer sur la **Figure 67** que la solution n°18 domine beaucoup de solutions (taille du nœud) alors qu'elle est une solution dominée. La voiture 34 (i.e. la Renault Megane 3 Energy dCi 110) obtient ce résultat car elle est celle qui correspond au mieux au scénario. La **Figure 66** nous permet de remarquer que la voiture 34 est en 2ième position pour la taille du réservoir, en 2ième position pour la largeur, en 4ième position pour la taille du coffre min et bien qu'elle ne soit pas très performante pour la variable « diamètre de jantes », elle est celle qui à la consommation d'essence la plus basse.

La liste des solutions avec le nombre de dominations et la distance au front de Pareto pour le scénario TRAJET est disponible en annexes p.211.

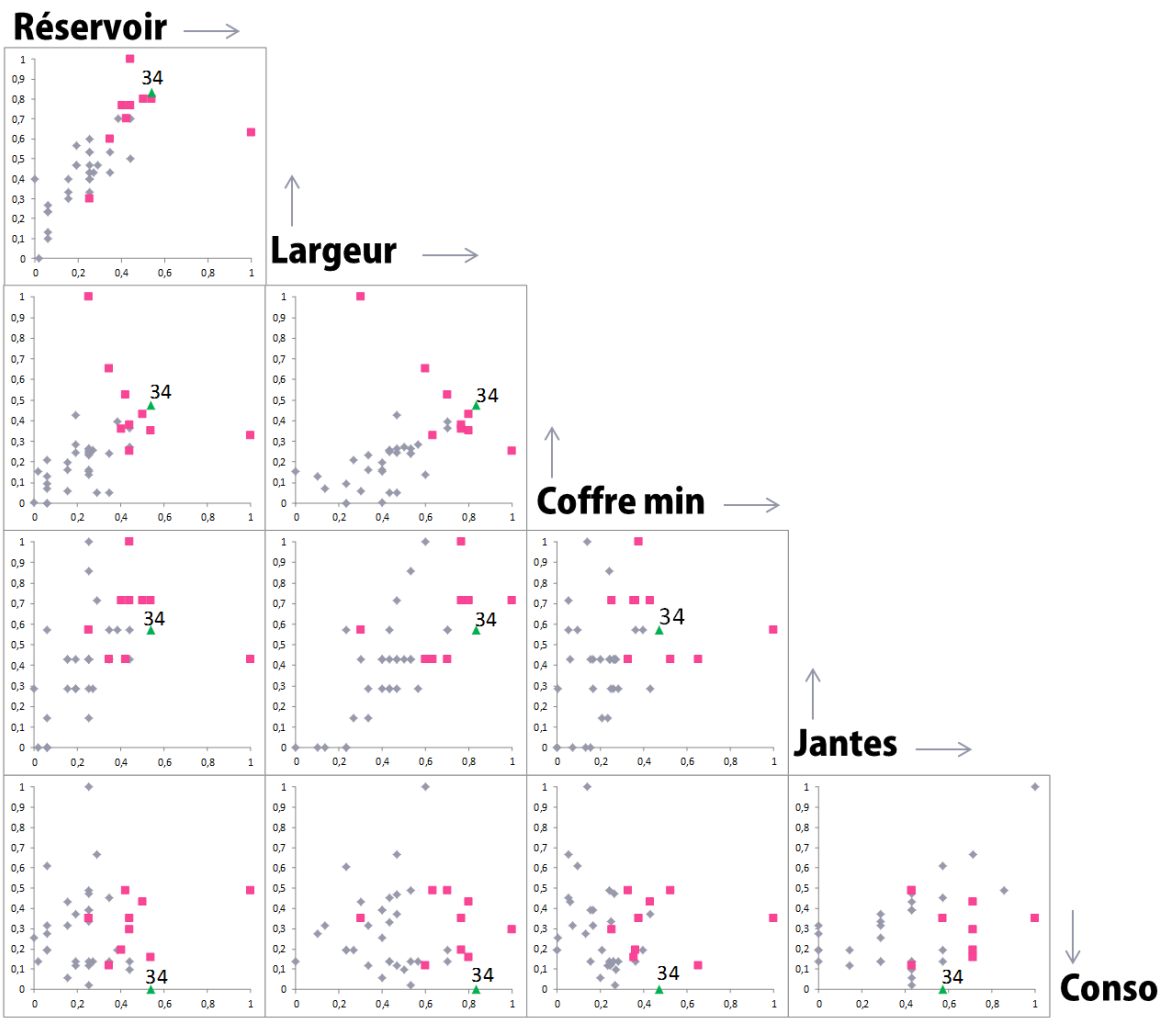


Figure 66 : Matrice de nuages de points des variables de performance du scénario TRAJET (en rose les solutions dites non-dominées au sens de Pareto et en vert la solution 34)

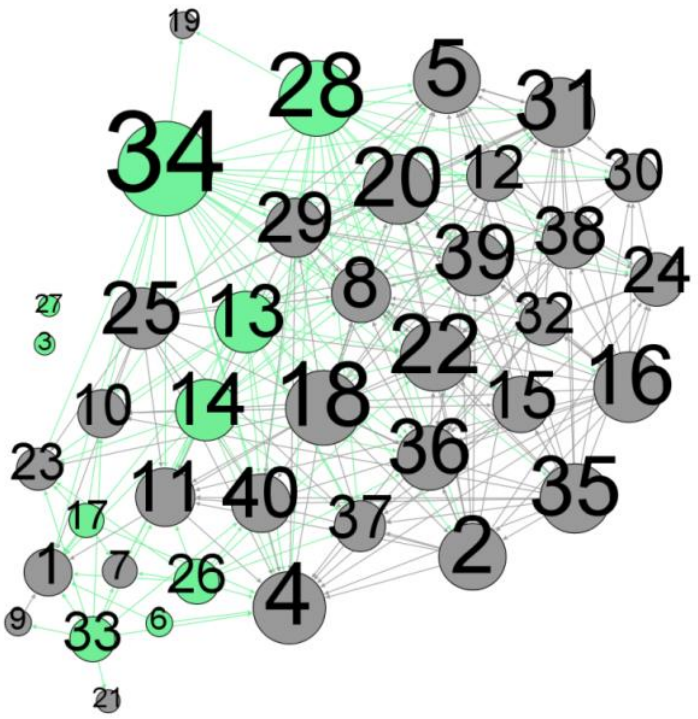


Figure 67 : Carte des dominations pour le scénario Trajet. Les ronds gris représentent les solutions dominées et la taille des ronds est proportionnelle au nombre de dominations qu'effectue la solution.

Afin d'expliquer la cascade de dominations qui se manifeste dans ce scénario nous proposons d'illustrer les quarante voitures dans un cadre simplifié où seules la largeur et la taille du réservoir sont à maximiser (voir **Figure 68**). Dans ce cadre, nous relevons quatre solutions Pareto-Optimales en carrés roses et triangles verts sur la **Figure 68**. La cascade de dominations provient de la « position » des solutions sur le nuage de points Réservoir-Largeur. En effet, nous observons une corrélation positive « grossière » entre ces deux variables qui sont à maximiser. Ainsi les solutions ont tendance à se dominer les unes derrière les autres : La voiture 34 domine trente-six solutions et la solution dominée n°17 proche de la 34 domine quant à elle trente-cinq solutions.

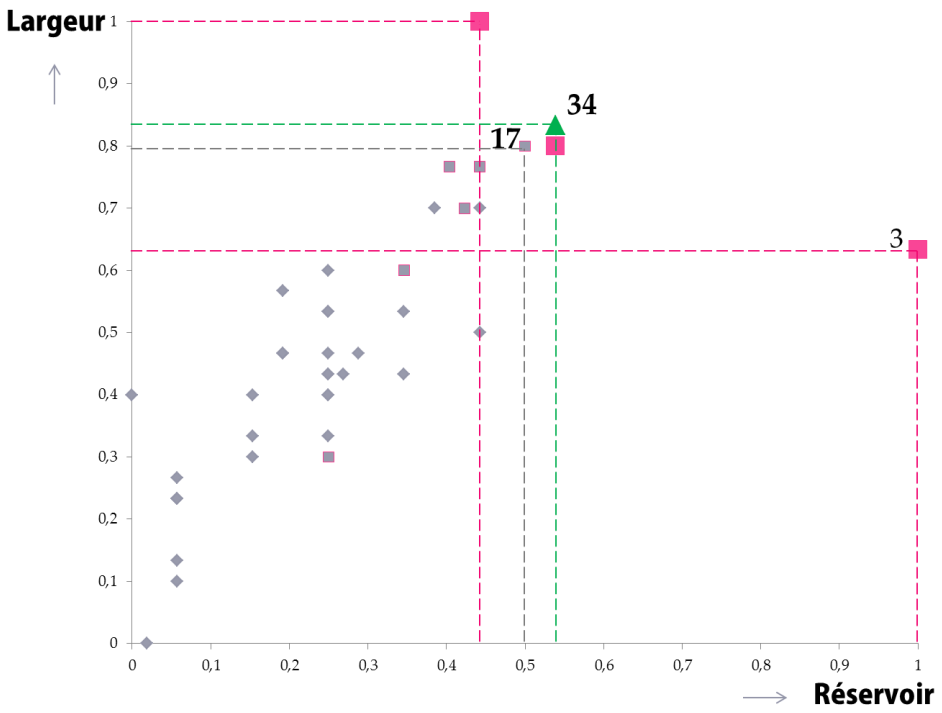


Figure 68 : Illustration de la « cascade » de dominations avec un cas bi-objectif (en rose et vert les solutions Pareto-Optimales)

3.2.1.5 Bilan

Il est important de noter que les solutions Pareto-Optimales ne dominent pas toutes le même nombre de solutions, voire certaines n'en dominent pas. Par exemple sur la **Figure 69** où les variables g1 et g2 sont à minimiser, les solutions A, B et C sont Pareto-Optimales, la solution C domine la solution D mais la solution B ne domine pas la solution D et la solution A ne domine aucune solution.

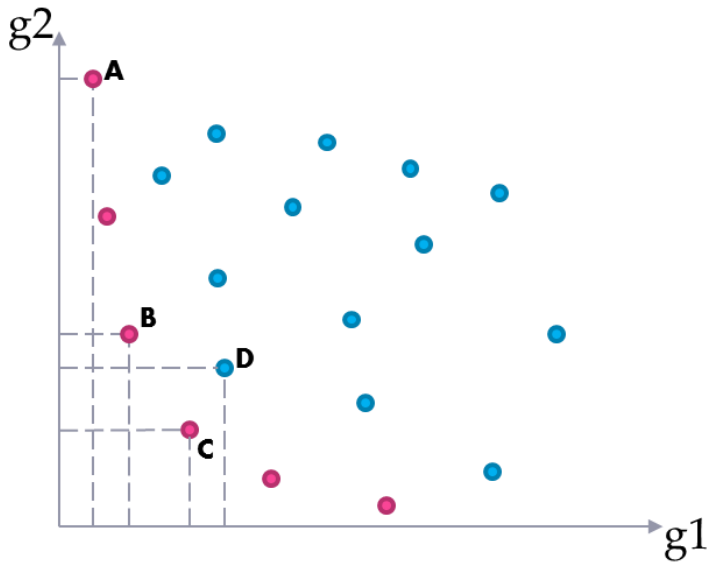


Figure 69 : Exemple de solutions pour un problème bi-objectif où il faut minimiser g1 et g2

Nous observons qu'un très grand nombre de solutions sont sur le front de Pareto. En effet plus de la moitié des solutions sont non-dominées pour le scénario TAILLE et jusqu'à un quart pour le scénario TRAJET (**Tableau 6**).

Tableau 6 : Nombre de solutions non-dominées pour les trois scénarios

Scénario	Nombre de solutions non-dominées
TAILLE	21
PUISSANCE	15
TRAJET	10

Habituellement, une bonne solution est évaluée, sans aucune information particulière sur les préférences entre objectifs, par la distance au front de Pareto. Mais dans nos trois scénarios nous avons laissé entendre que nos objectifs de performance choisis pour chaque scénario étaient d'égales importances parmi les cinq performances de voitures.

Ainsi la meilleure solution trouvée par un sujet (d'après la consigne) est celle qui est à l'intersection (*point optimal*) de la frontière de Pareto et de l'hyperplan « objectif ». Nous définissons l'hyperplan « objectif » comme l'orthogonale à la droite diagonale de l'hypercube dans l'espace des objectifs contenant tous les points. La diagonale est la droite qui relie la solution Idéale et la solution Anti-idéale (ces solutions fictives sont définies dans la section précédente).

La **Figure 70** illustre en quatre étapes ces différentes notions. La **Figure 70 (a.)** représente l'espace objectif (où g1 et g2 sont à minimiser) et les points violets représentent le front de Pareto. La **Figure 70 (b.)** illustre le « cadrage » de l'espace à l'aide des valeurs maximums et minimums des fonctions g1 et g2. La vignette c. (**Figure 70**) symbolise un hyperplan « objectif » (en vert) de l'espace respectant des préférences de mêmes importances. Enfin, la **Figure 70 (d.)** décrit l'identification du point Optimal à l'intersection de l'hyperplan « objectif » (en vert) orthogonal à la diagonale rose (reliant les solutions Idéale et Anti-idéale) et le front de Pareto (représenté par les points violets).

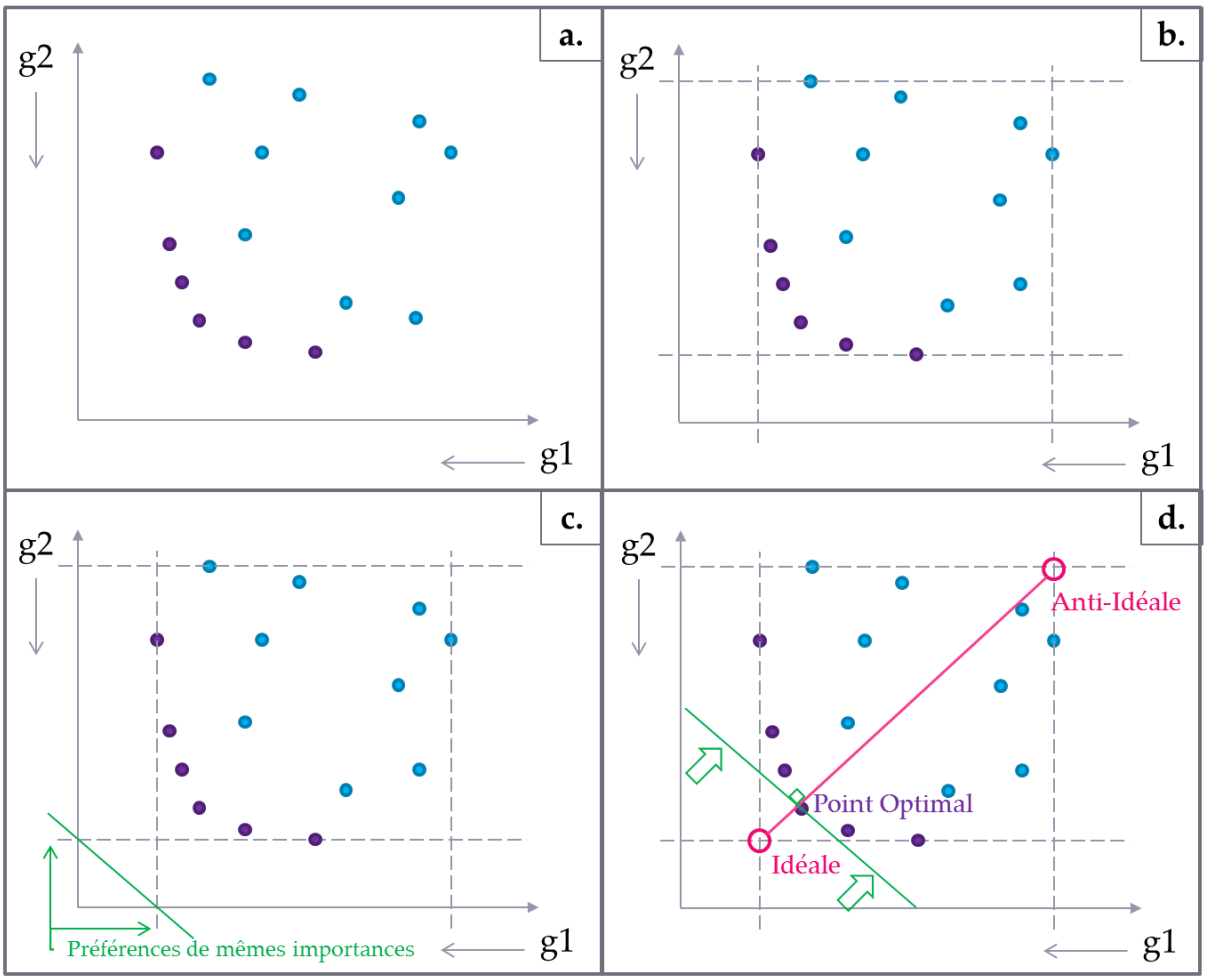


Figure 70 : Illustration des étapes pour l'identification du point Optimal

L'indice de qualité des réponses des sujets, que nous proposons par la suite a pour but de trouver l'hyperplan « objectif » de valeur minimale et de pentes égales dans toutes les dimensions (le point optimal).

3.2.2 Indice Basique

Nous proposons le calcul d'un indice proche d'une méthode d'optimisation mathématique multicritère : le « compromise programming ». De cette manière, nous restons en accord avec la consigne donnée aux participants car c'est une méthode qui s'articule autour d'une optique de minimisation des écarts par rapport à la solution Idéale (fictive). Sa formulation algébrique est la suivante :

$$\min_{x}\left[\sum_{j=1}^{n}\alpha_j\left[\frac{g_j(x^*)-g_j(x)}{g_j(x^*)-g_j(x_*)}\right]\right]$$

Equation 1

Où :

- j=1 … n
- α_j est le poids de préférence du j^ième^ critère g
- $g_j(x)$ est la valeur que prend la solution x pour le j^ième^ critère g
- $g_j(x^*)$ est la valeur que prend la solution Idéale pour le j^ième^ critère g
- $g_j(x_*)$ est la valeur que prend la solution Anti-idéale pour le j^ième^ critère g

Dans notre cas, la solution Idéale et la solution Anti-idéale sont des solutions fictives. En effet, les coordonnées de la solution Idéale sont obtenues en optimisant chaque fonction objectif (g_j) séparément (voir une illustration en **Figure 71**). Les solutions Idéale et Anti-idéale définissent une hyper-boite qui cadre ou contient tous les points solutions auxquels font face les participants. On peut donc prendre l'origine des graphiques aux coordonnées du point Anti-idéal (ou Idéal si la variable est à minimiser) et non en zéro.

Nous pouvons souligner ici que la solution Anti-idéale pour l'ensemble des solutions Pareto-Optimales s'appelle le Nadir (voir une illustration en **Figure 71**).

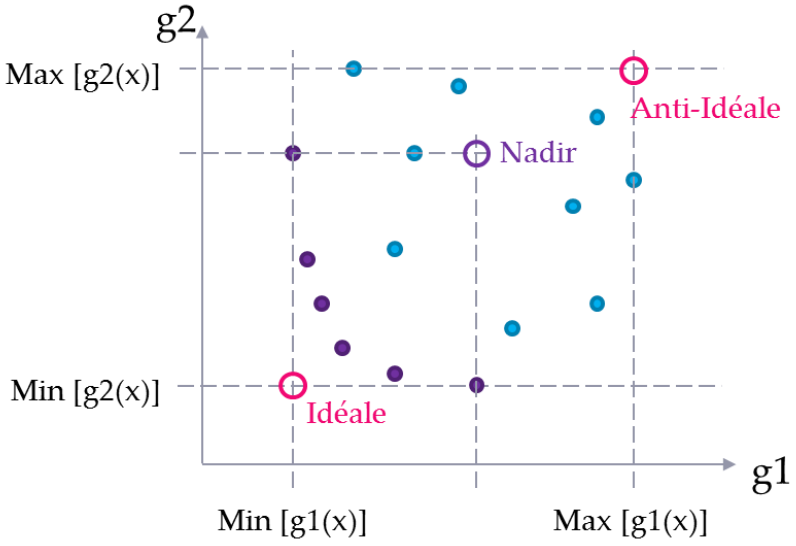


Figure 71 : Exemple de solution Idéale et Anti-idéale dans un problème où il faut minimiser g1 et g2

Notre Indice Basique revient à privilégier des solutions à la fois les plus proches de la solution Idéale et les plus éloignées de la solution Anti-idéale. C'est à dire à trouver l'hyperplan « objectif » de valeur minimale et de pentes égales dans toutes les dimensions. Dans notre cas les pentes sont égales dans toutes les dimensions car les poids de préférence sont tous à 1. Ainsi, l'hyperplan « objectif » est orthogonal à la diagonale de boite, reliant l'Idéale et l'Anti-idéale, de l'hypercube dans l'espace des objectifs.

Dans le cas où la frontière de Pareto est relativement orthogonale à la diagonale, le point optimal est bien défini par la minimisation de la distance à l'Idéal sur la distance à l'Anti-idéal (voir une illustration en **Figure 72** avec un front de Pareto convexe). Mais dans un autre

cas, la minimisation de la distance à l'Idéal sur la distance à l'Anti-idéal peut se trouver assez éloignée du point à l'intersection du front de Pareto et de la diagonale de boite (voir une illustration en **Figure 73** avec un front de Pareto concave).

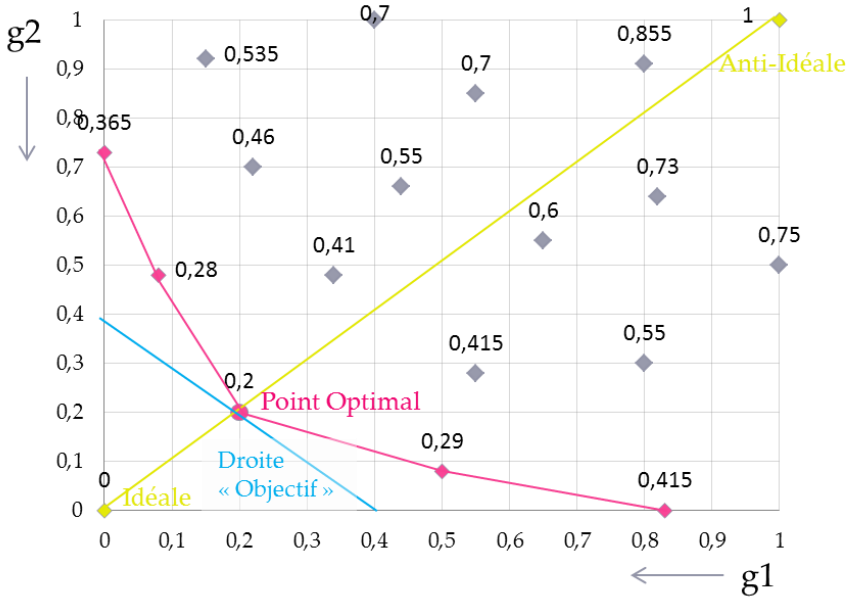


Figure 72 : Illustration du Point Optimal dans le cas d'un front de Pareto convexe. En rose le front de Pareto, en jaune la diagonale de boite et les indices dits basiques à coté de chaque point.

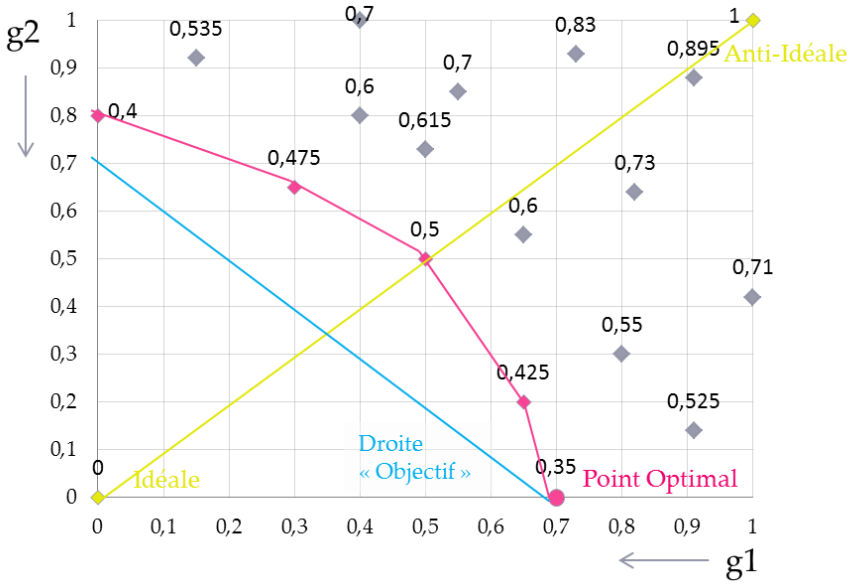


Figure 73 : Illustration du Point Optimal dans le cas d'un front de Pareto concave. En rose le front de Pareto, en jaune la diagonale de boite et les indices dits basiques à coté de chaque point.

La différence entre le « compromise programming » (**Equation 1**) et notre Indice Basique (**Equation 2**) se caractérise par :

- Les variables à maximiser/minimiser : elles ont des poids de préférence égales à 1. Nous pouvons donc nous affranchir de ce coefficient (α_j) dans l'équation.

- Un indice à maximiser : ce qui inverse le numérateur de l'équation.
- La division de l'équation par le nombre de variables de performance n=5 afin que l'indice soit compris entre 0 et 1.

Nous proposons donc le calcul d'un indice i_j (**Equation 2**) pour chaque scénario :

$$i_j = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{f_k(P)-f_k(A)}{f_k(I)-f_k(A)}\right|, k=1 \dots n$$

Equation 2

$$i = \left|\frac{i_j - i_{j_{min}}}{i_{j_{max}} - i_{j_{min}}}\right|$$

Equation 3

Où f_k est la k^ième^ variable de performance (n=5 pour chaque scénario). $f_k(P)$ est la valeur de la solution *P* pour la variable f_k. et:

- Si f_k est à maximiser: $f_k(A)$= la valeur minimum de f_k et $f_k(I)$= la valeur maximum.
- Si f_k est à minimiser: $f_k(A)$= la valeur maximum de f_k et $f_k(I)$= la valeur minimum.

Nous notons *I* pour la solution Idéale et *A* pour la solution Anti-idéale.

Puis, nous calculons l'indice dit « basique » *i* avec **l'Equation 3** afin que la solution la plus proche de l'Idéale ait un indice égal à 1 et la solution la plus proche de l'Anti-idéale ait un indice égal à 0 pour chacun des scénarios. De cette manière, nous pouvons comparer les solutions obtenues avec les trois scénarios. En **Figure 74 (a.)**, nous observons qu'avec l'indice i_j les solutions sont difficilement comparables entre les trois scénarios puisque les solutions aux extremums ne prennent pas les mêmes valeurs d'indice. Par contre, avec l'Indice Basique *i* les solutions montrent des indices étendus entre 0 et 1 (**Figure 74 b**.).

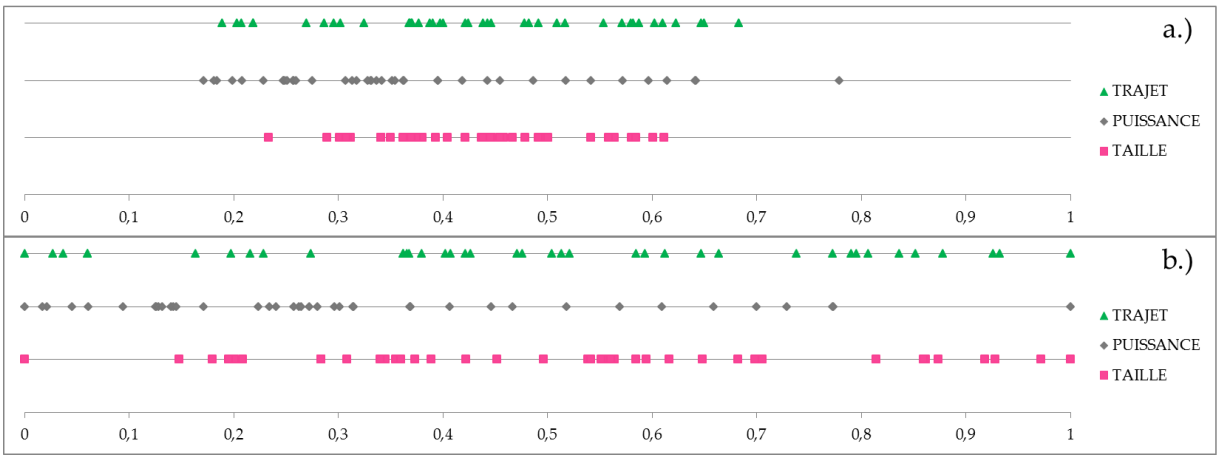


Figure 74 : Illustration des indices que prennent les solutions dans les trois scénarios : a.) Indice i_j b.) Indice *i*

Comme exprimé précédemment, l'utilisation unique de l'Indice Basique n'est pas envisageable. Cet indice est une composante utilisée conjointement avec le principe de dominance au sens de Pareto (et distance au front de Pareto) pour la construction de notre indice de qualité.

En plus du besoin de traduire en automatique la consigne donnée aux participants (à travers les trois scénarios), l'une des raisons importantes de l'utilisation conjointe de ces deux « composantes » pour la construction de notre indice de qualité est que dans certains cas nous pouvons trouver une solution dite dominée qui obtient un meilleur Indice Basique que certaines solutions dites non-dominées.

Nous observons ce phénomène sur la **Figure 75** où la solution encerclée est dominée mais obtient un Indice Basique de 0.446 qui est un indice supérieur à ceux obtenus par plusieurs solutions non-dominées indiquées par des flèches bleues. Nous remarquons également sur la **Figure 75** que les solutions les plus proches de l'Idéale ne sont pas nécessairement celles qui obtiennent le meilleur indice. Il s'agit évidemment d'une coupe 2D d'un problème en 5D et l'Indice Basique affiché (**Figure 75**) est calculé pour les cinq variables de performance.

Figure 75 : Carte des indices pour le scénario PUISSANCE avec « vitesse » à maximiser et « poids » à minimiser

Ainsi, nous proposons de construire un indice de qualité qui combine la distance au front de Pareto et l'Indice Basique pour mieux traduire la qualité des solutions possibles liées à un

scénario (et sa consigne) et ainsi déterminer quel graphique permet à la majorité des concepteurs de prendre les meilleures décisions.

3.2.3 Notre indice de qualité

Dans cette section, nous décrivons la construction de trois indices de qualité applicables à nos trois scénarii. Nous avons vu dans les sections précédentes qu'il y avait un nombre important de solutions Pareto-Optimales dans chacun des scénarii et que dans certains cas nous observons que l'Indice Basique d'une solution dominée pouvait être supérieur à ceux de solutions Pareto-Optimales. Nous proposons donc la construction de trois indices de qualité applicables à notre contexte.

3.2.3.1 Indice 1

Nous proposons ici de soustraire la distance au front de Pareto à l'Indice Basique pour construire l'indice de qualité. De cette manière, nous différencions les solutions non-dominées grâce aux indices basiques et nous pouvons « sanctionner » les solutions dominées puisque l'Indice Basique se retrouve diminué par la distance au front de Pareto (**Tableau 7**).

Tableau 7 : Construction de l'indice n°1 pour les solutions dominées et non-dominées

Type de solutions	Distance au front de Pareto	Indice Basique	Indice de qualité	Commentaires
Non-dominées	0	I	I - 0	Permet de différencier les solutions dites optimales
Dominées	d		I - d	Permet de « sanctionner » les solutions dominées

Une illustration de la construction de cet indice est présentée en **Figure 76**. Dans cette illustration nous personnifions les solutions possibles en leur donnant la parole. Les solutions sont représentées par des « smileys » contents ou mécontents. La **Figure 76** se lit de haut en bas. Dans la première étape, les solutions sont ordonnées par rapport à leur Indice Basique. Mais ce classement ne convient pas à toutes les solutions. Par exemple la solution violette considère ne pas être bien classée puisqu'elle est une solution non-dominée. Dans la deuxième étape nous proposons donc de séparer les solutions non-dominées de celles dominées. Une nouvelle fois toutes les solutions ne sont pas satisfaites soit parce qu'elles sont proches de la solution Idéale soit parce qu'elles sont proches de la surface de compromis (Pareto Front). Dans la troisième étape nous tentons de mettre les solutions d'accord en les ordonnant à l'aide de l'indice n°1 qui prend en compte l'Indice Basique (distance avec l'Idéale) et la dominance au sens de Pareto (distance avec la surface de compromis).

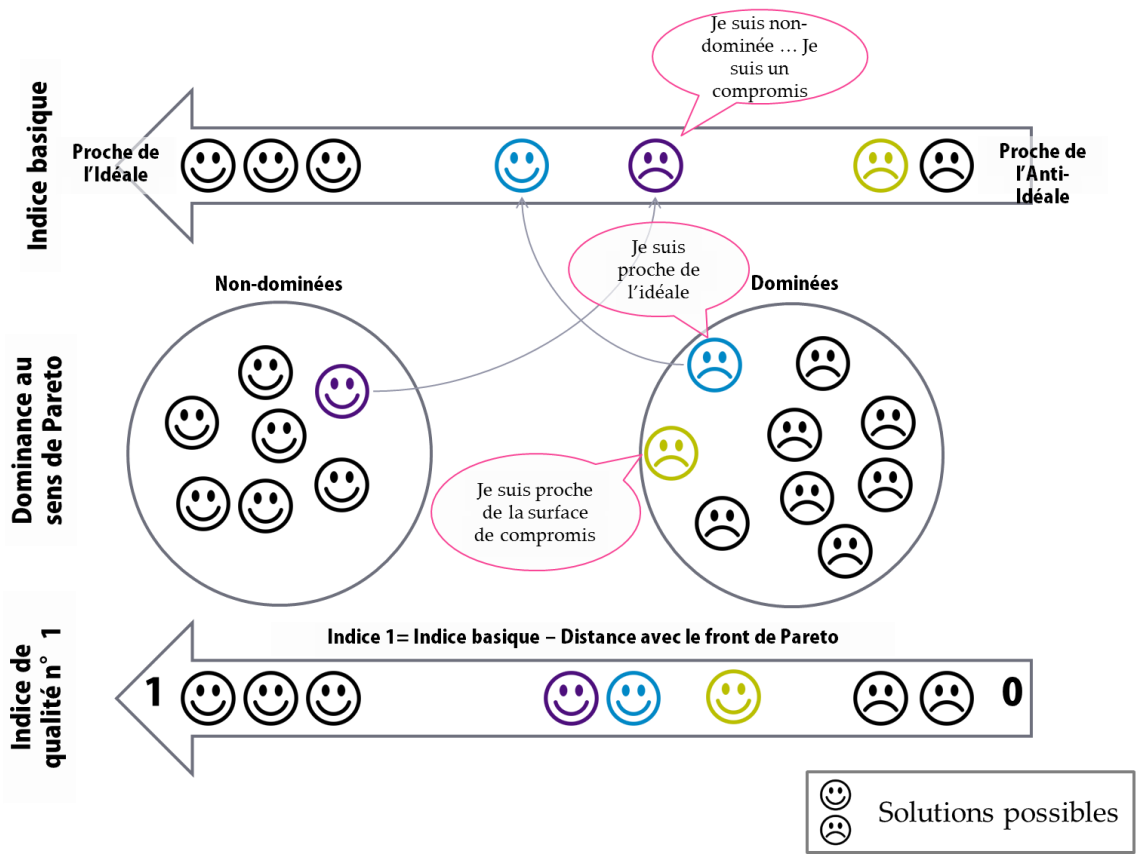


Figure 76 : Illustration de la construction de l'indice n°1

3.2.3.2 Indice 2

Nous proposons ici de distinguer plus largement les solutions non-dominées de celles dominées.

L'indice de qualité pour les solutions non-dominées correspond à l'Indice Basique replacé sur une échelle comprise en 0 et 1. 1 pour la meilleure solution et 0 pour la moins bonne solution dans l'ensemble des non-dominées.

Et de l'autre côté, l'indice de qualité pour les solutions dominées correspond à la distance au front de Pareto multiplié par (-1). De cette manière, les solutions dominées obtiennent un indice de qualité négatif. Plus la solution est proche du front de Pareto et plus son indice de qualité tend vers 0 et plus la solution est « loin » du front et plus son indice de qualité est « négatif » (**Tableau 8**).

Tableau 8 : Construction de l'indice n°2 pour les solutions dominées et non-dominées

Type de solutions	Indice de qualité	Commentaires
Non-dominées	Indice Basique recadré sur une échelle [0 ;1]	Permet de différencier les solutions dites optimales
Dominées	Distance Pareto *(-1)	Permet de « sanctionner » les moins bonnes solutions

Une illustration de la construction de cet indice est présentée en **Figure 77**. De même que pour l'indice n°1, dans cette illustration nous personnifions les solutions possibles en leur donnant la parole et La **Figure 77** se lit de haut en bas. Dans la première étape les solutions sont séparées en deux groupes : les solutions non-dominées et les solutions dominées. Au sein des deux groupes les solutions souhaitent être différenciées. Ainsi, dans la deuxième étape nous proposons de différencier les solutions non-dominées en les ordonnant à l'aide de l'Indice Basique et de différencier les solutions dominées en les classant par rapport à leur distance au front de Pareto. Dans la troisième étape nous ordonnons toutes les solutions sur une échelle comprise entre [1 ;-1]. Les solutions non-dominées ont un indice compris entre 1 et 0 (ce qui correspond à la distance des solutions avec l'Idéale) et les solutions dominées ont un indice compris entre 0 et -1 (ce qui correspond à la distance des solutions avec le front de Pareto).

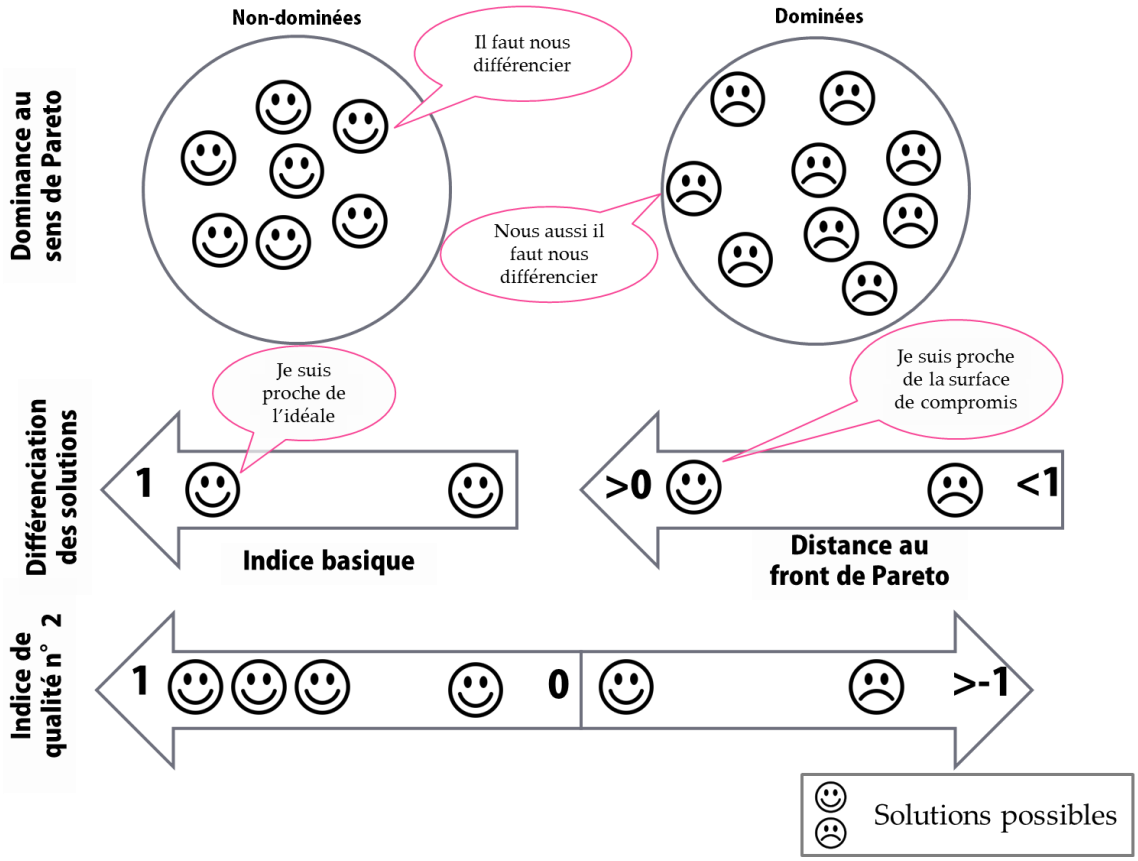


Figure 77 : Illustration de la construction de l'indice n°2

3.2.3.3 Indice 3

Enfin, nous proposons ici d'ajouter un critère à la construction de l'indice de qualité : Le nombre de solutions que domine la solution non-dominées.

En effet, comme présenté dans section 2.2, 2.3 et 2.4, les solutions dites non-dominées ne dominent pas le même nombre de solutions. Nous observons par exemple dans le scenario TRAJET que la solution n°34 domine 27 solutions alors que la solution n°6 en domine 2.

Ainsi, l'indice de qualité est le résultat de l'indice 2 (présenté en section 4.2) multiplié par le (nombre de solutions que la solution domine +2). Nous ajoutons 2 à ce critère afin de ne pas obtenir un indice égale à 0 pour les solutions dites dominées (**Tableau 9**).

Tableau 9 : Construction de l'indice n°3 pour les solutions dominées et non-dominées

Type de solutions	Nbr de dominations	Indice 2	Indice de qualité	Commentaires
Non-dominées	k+2	I2	I2*(k+2)	Permet de différencier les solutions dites optimales en considérant le nombre de dominations
Dominées	1		I2	Permet de « sanctionner » les moins bonnes solutions

Une illustration de la construction de cet indice est présentée en **Figure 78**. De même que pour les indices n°1 et n°2, dans cette illustration nous personnifions les solutions possibles en leur donnant la parole et La **Figure 78** se lit de haut en bas. La première et la deuxième étape sont similaires à celles de la **Figure 77**. Dans cette configuration certaines solutions non-dominées ne sont pas satisfaites du classement. Par exemple la solution bleue souhaite que le nombre de dominations soient pris en compte. Ainsi, dans la troisième étape nous proposons l'indice n°3 qui combine, pour les solutions non-dominées, la distance avec l'Idéale (Indice Basique) et le nombre de dominations de la solution. Le classement des solutions dominées est inchangé par rapport à l'indice n°2.

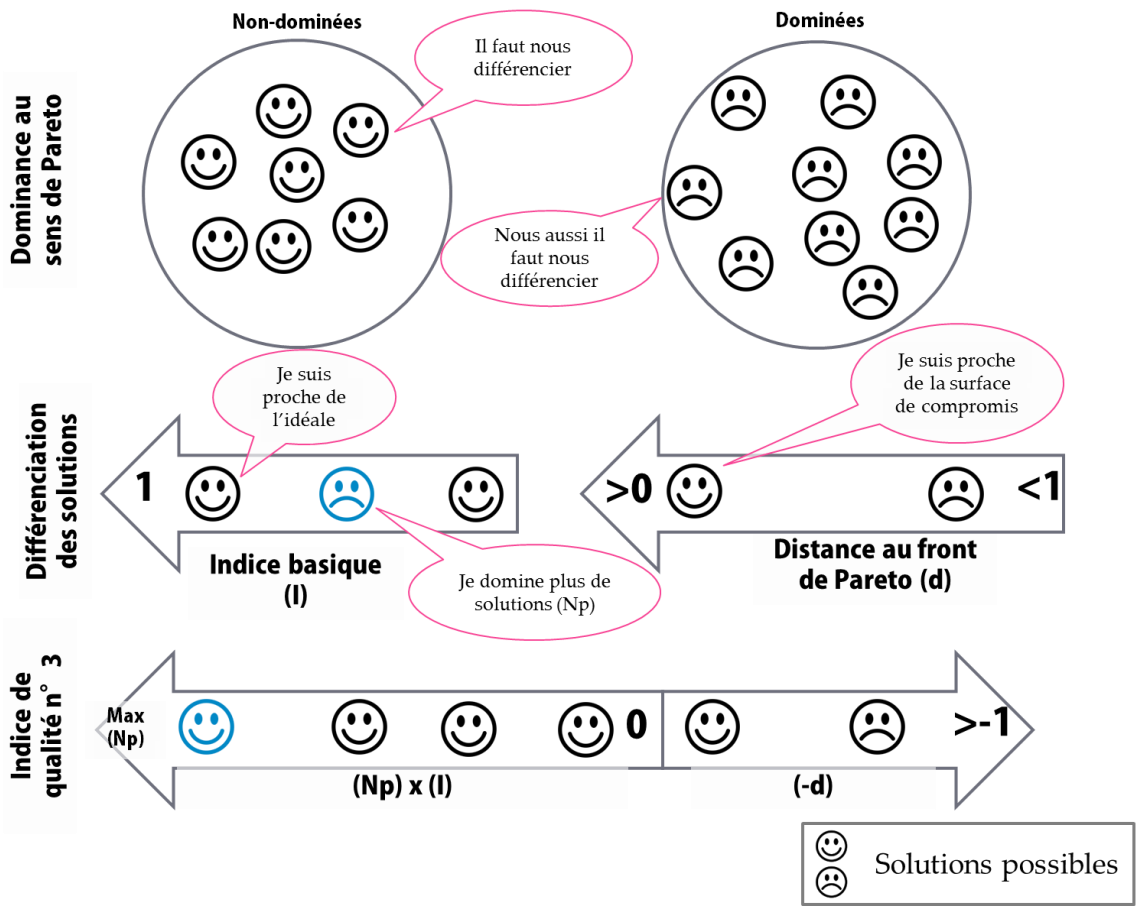


Figure 78 : Illustration de la construction de l'indice n°3

3.2.3.4 Bilan

Nous proposons donc trois indices de qualité pour évaluer la performance des solutions sélectionnées par les participants pour les trois différents scénarii. Nous notons que le classement des solutions entre les trois indices est différent (voir **Figure 79**).

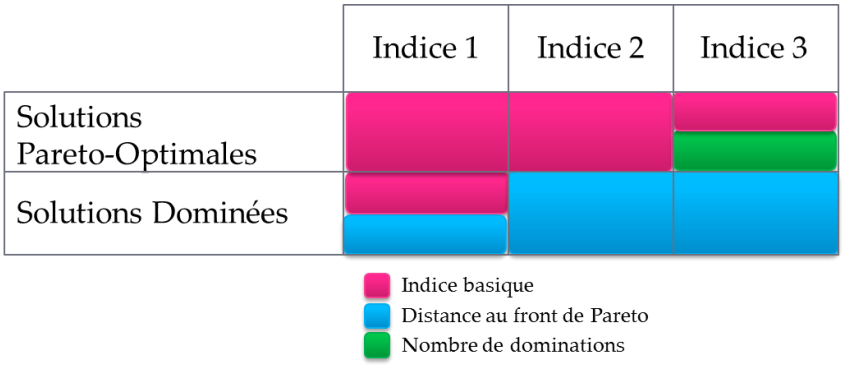


Figure 79 : Illustration des différences de classement des solutions entre les trois indices

Si nous regardons les solutions Pareto-Optimales : il n'y a pas de différences entre l'indice 1 et l'indice 2 puisque les indices que prennent les solutions Pareto-Optimales ne dépendent que de l'Indice Basique. Nous trouvons par contre une différence avec l'indice 3 car l'indice que prennent les solutions Pareto-Optimales est fortement lié au nombre de dominations.

Une illustration des classements des solutions Pareto-Optimales du scénario TRAJET, pour les trois indices, est proposée dans la **Figure 80**.

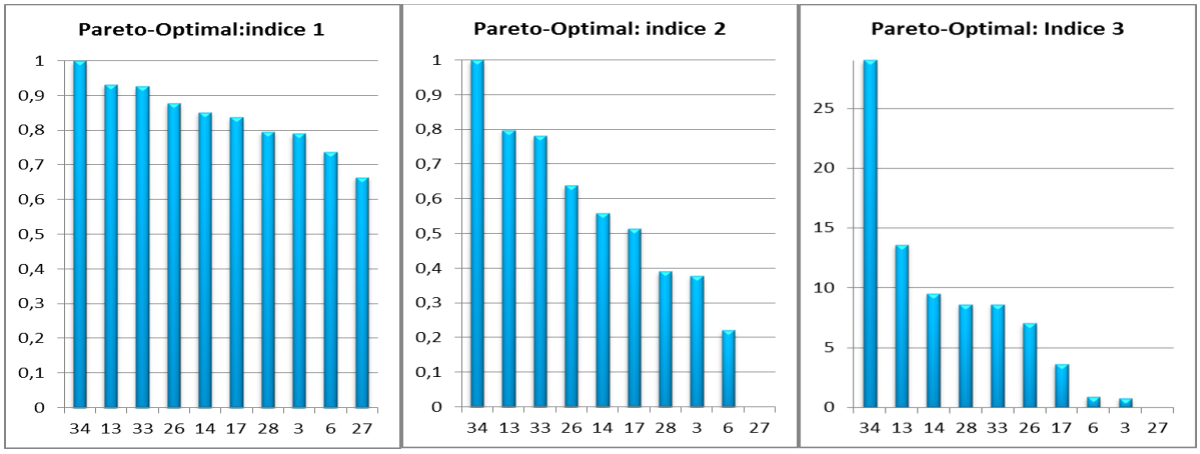


Figure 80 : Illustration des classements des solutions Pareto-Optimales avec les trois indices pour le scénario TRAJET.

Ensuite, si nous observons les solutions dominées, il n'y a pas de différences entre l'indice 2 et 3 car il s'agit exactement du même calcul : la distance au front de Pareto. Par contre, nous trouvons une différence avec l'indice 1 car en plus de la distance au front de Pareto, l'indice que prennent les solutions dépend de l'Indice Basique.

Afin de mieux illustrer la différence entre ces trois indices de qualité, nous présentons trois graphiques en barres qui affichent les indices de qualité n°1 des quarante voitures pour le scénario TRAJET (**Figure 81**), les indices de qualité n°2 des quarante voitures pour le scénario TRAJET (**Figure 82**) et les indices de qualité n°3 des quarante voitures pour le scénario TRAJET (**Figure 83**).

Les **Figure 81, Figure 82** et **Figure 83** mettent en lumière les caractéristiques de ces trois indices de qualité. En effet, nous observons une forme de symétrie pour l'indice n°1 (**Figure 81**), la sanction totale appliquée par l'indice n°2 sur les solutions dominées (**Figure 82**) et la forte valorisation des solutions dites Pareto-Optimales par l'indice n°3 (**Figure 83**).

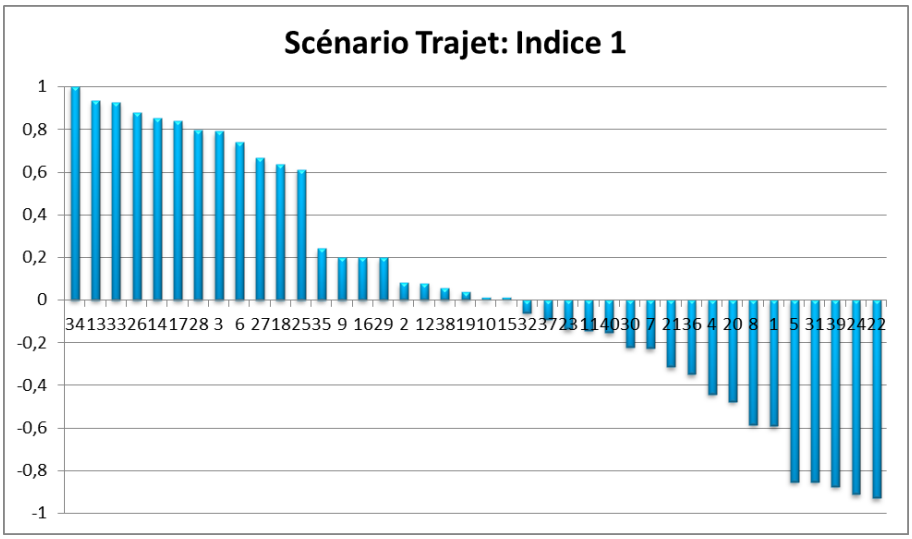


Figure 81 : Illustration des indices de qualité n°1 des quarante voitures pour le scénario TRAJET

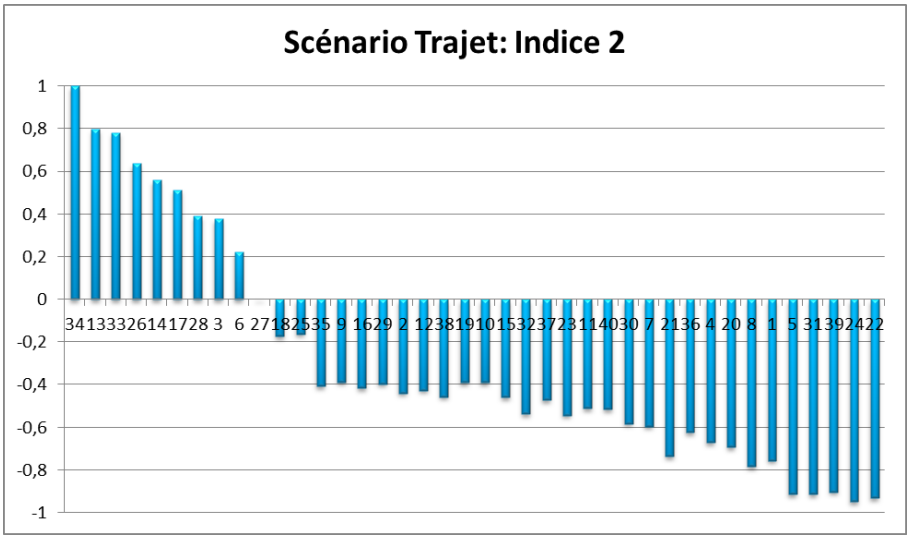


Figure 82 : Illustration des indices de qualité n°2 des quarante voitures pour le scénario TRAJET

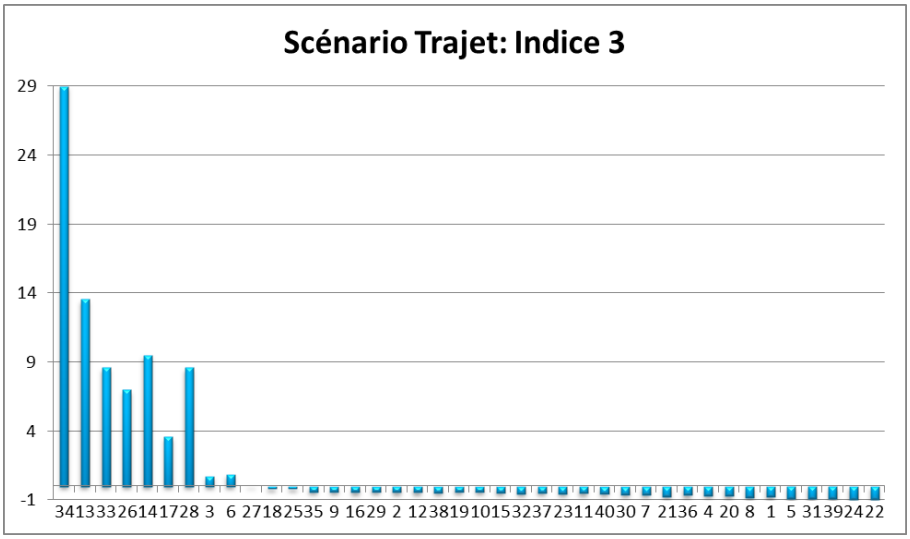


Figure 83 : Illustration des indices de qualité n°3 des quarante voitures pour le scénario TRAJET

Pour conclure cette section nous décidons d'utiliser l'indice n°1 pour mesurer la qualité des solutions sélectionnées. Les indices n°2 et n°3 différencient les solutions non-dominées des solutions dominées. Cela présuppose que les deux types de solutions (dominance au sens de Pareto) soient visuellement distinguables avec les graphiques (PCP, RC et CT) hors ce n'est pas le cas. Il n'est donc pas envisageable, dans cette configuration, d'utiliser l'indice n°2 et/ ou l'indice n°3. Nous devons donc utiliser l'indice n°1 comme indice de qualité car le même calcul est réalisé pour toutes les solutions possibles. Nous proposons dans la **Figure 84** de représenter les meilleures et les pires voitures définies par l'indice de qualité (n°1) pour les trois scénarios (le sens des flèches indique les meilleures voitures).

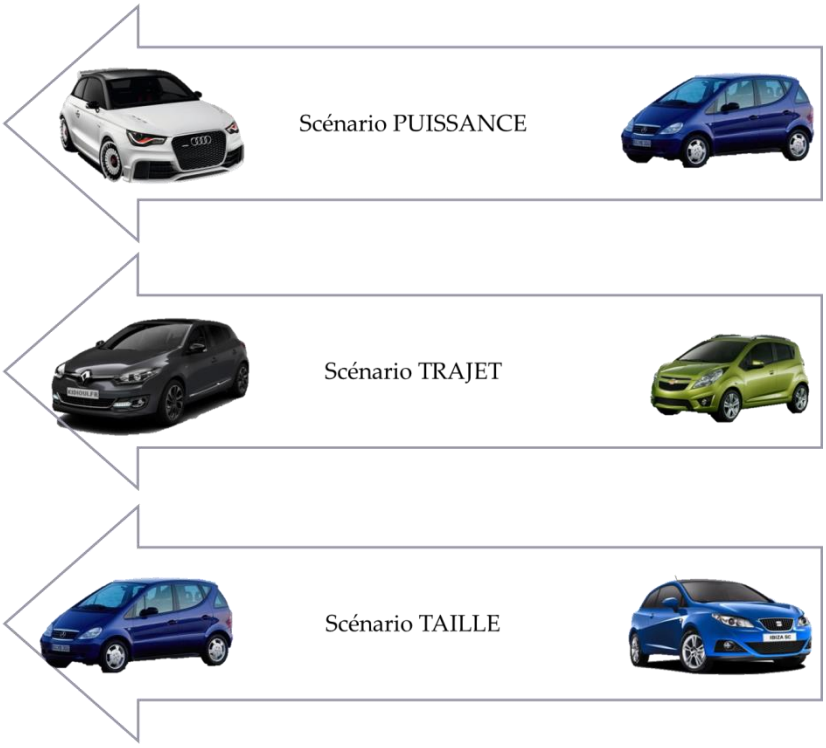


Figure 84 : Meilleures et pires voitures définies par l'indice de qualité pour les trois scénarios

3.3 TRAITEMENT DES RESULTATS 1

L'objectif de l'expérimentation est l'identification d'un graphique performant pour la sélection d'une solution optimale en Design by Shopping. Dans une situation de choix de voiture, trois graphiques sont testés par trente sujets avec une approche par groupes appariés.

Pour la phase de sélection, trois aspects sont mesurés :

- Le compromis composé de deux indicateurs : le temps d'accomplissement et la qualité de la solution sélectionnée.
- La recherche de compromis composé de deux indicateurs : le nombre d'actions réalisées et le temps entre les actions.
- Le ressenti du concepteur composé de deux indicateurs : la charge cognitive et la confiance du participant.

Pour l'analyse de données de cette partie, nous posons les hypothèses statistiques globales suivantes :

- H0 : il n'existe pas de différence en les trois groupes
- H1 : il existe une différence entre les trois groupes (le PCP est le graphique le plus performant pour la sélection)

Pour l'analyse statistique, quatre tests sont utilisés. Pour les variables quantitatives nous opérons une analyse de la variance (ANOVA-*within*) et si nécessaire un test-T par paires. Pour nos variables qualitatives ordinales nous opérons un test de Friedman et si nécessaire un test de Wilcoxon (rang signé) par paires (**Tableau 10**).

Tableau 10 : Tests statistiques utilisés par indicateurs de l'expérimentation n°1

Aspect mesuré	**Indicateur**	**1er test statistique**	**2ième test statistique**
Le compromis	Le temps	ANOVA-*within*	Test-T
	La qualité		
La recherche de compromis	Le nombre d'actions		
	Le temps entre les actions		
Le ressenti	La charge cognitive	Friedman	Wilcoxon
	La confiance		

Nous avons pour les analyses trois groupes appariés : PCP, RC et CT tous les trois composés de n=30 individus.

3.3.1 Le compromis

Pour mesurer l'influence du graphique sur le compromis réalisé par le concepteur nous utilisons deux indicateurs : le temps d'accomplissement i.e. le temps total écoulé pour sélectionner une solution optimale et la qualité de la solution sélectionnée.

3.3.1.1 Le temps d'accomplissement

Pour cet indicateur, nous obtenons une moyenne de 192.80 secondes pour le PCP, 241.87 secondes pour le RC et 210.97 secondes pour le CT. L'ANOVA nous donne comme résultats : $F(2,58)=1.12$ et $p=0.334$. Nous devons donc rejeter l'hypothèse alternative et conclure qu'il n'existe pas de différence significative entre les trois groupes**. Il n'y a pas d'influence du graphique sur le temps nécessaire à la sélection d'une solution optimale.**

3.3.1.2 La qualité

Pour cet indicateur, nous obtenons une moyenne de 0.872 pour le PCP, 0.860 pour le RC et 0.863 pour CT. L'ANOVA nous donne comme résultats : $F(2,58)=0.026$ et $p=0.974$. Nous devons donc rejeter l'hypothèse alternative et conclure qu'il n'existe pas de différence significative entre les trois groupes. **Il n'y a pas d'influence du graphique sur la qualité de la solution sélectionnée.**

Nous devons conclure qu'il n'existe pas de différence entre les trois graphiques pour le compromis.

3.3.2 La recherche de compromis

Pour mesurer, l'influence du graphique sur la recherche de compromis nous utilisons deux indicateurs : le nombre d'actions réalisées pour réaliser le compromis et le temps écoulé entre les actions.

3.3.2.1 Le nombre d'actions

Pour cet indicateur, nous obtenons une moyenne de 9.67 actions pour le PCP, 3.97 actions pour le RC et 3.60 actions pour le CT. L'ANOVA nous donne comme résultats : $F(2,58)=14.1$ et $p=0,000011$. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test-T par paires :

- PCP vs. RC : $t(29)=3.81$ $p< 0,0007$
- PCP vs. CT : $t(29)=4.22$ $p< 0,0002$
- RC vs. CT : $t(29)=0.46$ $p< 0,6496$

Il existe donc une différence significative entre les graphiques PCP et RC ; et PCP et CT. Nous pouvons conclure que **le PCP est le graphique montrant un plus grand nombre d'actions pour réaliser le compromis.**

3.3.2.2 Le temps entre les actions

Pour cet indicateur, nous obtenons une moyenne de 26.19 secondes pour le PCP, 72.96 secondes pour le RC et 50.29 secondes pour le CT. L'ANOVA nous donne comme résultats : F(2,58)=10.7 et p=0,000111. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test-T par paires :

- PCP vs. RC : t(29)=3.90 p< 0.0005
- PCP vs. CT : t(29)=5.02 p< 0.0001
- RC vs. CT : t(29)=1.92 p< 0.0652

Il existe donc une différence significative entre les graphiques PCP et RC ; et PCP et CT. Nous pouvons conclure que **le PCP est le graphique montrant un temps plus court entre les actions.**

Nous devons conclure qu'il existe une différence entre les trois groupes. Il y a donc une influence du graphique sur la recherche de compromis : le graphique PCP est le plus pertinent car il est le graphique montrant un plus grand nombre d'actions pour réaliser le compromis et celui montrant un temps plus court entre les actions.

3.3.3 Le ressenti

Pour mesurer le ressenti du concepteur nous utilisons deux indicateurs : la charge cognitive perçue par le sujet et la confiance.

Notons que pour l'analyse de ces variables, nous opérons un test de Friedman et si nécessaire un test de Wilcoxon.

3.3.3.1 La charge cognitive

Ce premier indicateur est composé de l'effort d'attention ressenti, l'effort visuel ressenti et le stress (fatigue).

3.3.3.1.1 Effort d'attention

Nous réalisons un test de Friedman pour répondre à nos hypothèses statistiques avec un risque de première espèce égale à 5%. Nous obtenons une moyenne de rang de 1.9 pour le PCP, 2.4 pour le RC et 1.7 pour le CT. Les résultats du test sont csqr=8.87, df=2 et p=0.0119. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test de Wilcoxon par paires :

- PCP vs. RC : W=-193 et p=0.024
- PCP vs. CT : W=57 et p=0.281
- RC vs. CT : W=201 et p=0.039

Il existe donc une différence significative entre les graphiques PCP et RC ; et RC et CT. Nous pouvons conclure que **le RC est le graphique avec lequel les participants ont ressenti un plus grand effort d'attention**.

3.3.3.1.2 Effort Visuel

Nous réalisons un test de Friedman pour répondre à nos hypothèses statistiques avec un risque de première espèce égale à 5%. Nous obtenons une moyenne de rang de 1.7 pour le PCP, 2.4 pour le RC et 1.9 pour le CT. Les résultats du test sont csqr=8.15, df=2 et p=0.017. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test de Wilcoxon par paires :

- PCP vs. RC : W=-270 et p=0.002
- PCP vs. CT : W=6 et p=0.476
- RC vs. CT : W=205 et p=0.018

Il existe donc une différence significative entre les graphiques PCP et RC ; et RC et CT. Nous pouvons conclure que **le RC est le graphique avec lequel les participants ont ressenti un plus grand effort visuel.**

3.3.3.1.3 Stress

Nous réalisons un test de Friedman pour répondre à nos hypothèses statistiques avec un risque de première espèce égale à 5%. Nous obtenons une moyenne de rang de 1.9 pour le PCP, 2.5 pour le RC et 1.5 pour le CT. Les résultats du test sont csqr=19.27, df=2 et p=0.0013. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test de Wilcoxon par paires :

- PCP vs. RC : W=-254 et p=0.005
- PCP vs. CT : W=144 et p=0.052
- RC vs. CT : W=296 et p=0.001

Il existe donc une différence significative entre les graphiques PCP et RC ; et RC et CT. Nous pouvons conclure que **le RC est le graphique avec lequel les participants ont ressenti un plus grand stress.**

Nous devons conclure qu'il existe une différence entre les trois groupes. Il y a donc une influence du graphique sur la charge cognitive du concepteur : le graphique RC provoque une charge cognitive plus élevée.

3.3.3.2 La confiance

Ce deuxième indicateur est composé de la certitude des participants et leur préférence pour un graphique.

3.3.3.2.1 La certitude

Nous réalisons un test de Friedman pour répondre à nos hypothèses statistiques avec un risque de première espèce égale à 5%. Nous obtenons une moyenne de rang de 2 pour le PCP, 1.9 pour le RC et 2.1 pour le CT. Les résultats du test sont csqr=1.22, df=2 et p=0.543. Nous devons donc rejeter l'hypothèse alternative et inférer qu'**il n'existe pas de différences significatives entre les trois groupes pour l'indicateur « certitude ».**

3.3.3.2.2 La préférence

Nous séparons ici la préférence des participants pour un graphique : pour la recherche de compromis et pour le compromis.

Nous réalisons un test de Friedman pour répondre à nos hypothèses statistiques avec un risque de première espèce égale à 5%.

3.3.3.2.2.1 Rechercher un compromis

Nous obtenons une moyenne de rang de 1.6 pour le PCP, 2.6 pour le RC et 1.8 pour le CT. Les résultats du test sont csqr=16.47, df=2 et p=0.0003. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test de Wilcoxon par paires :

- PCP vs. RC : W=-335 et p=0.0003
- PCP vs. CT : W=-53 et p=0.2946
- RC vs. CT : W=295 et p=0.0012

Il existe donc une différence significative entre les graphiques PCP et RC ; et RC et CT. Nous pouvons conclure que **le RC est le graphique avec lequel les participants ont moins apprécié la recherche de compromis.**

3.3.3.2.2.2 Réaliser un compromis

Nous obtenons une moyenne de rang de 1.6 pour le PCP, 2.4 pour le RC et 1.9 pour le CT. Les résultats du test sont csqr=9.8, df=2 et p=0.0074. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les trois groupes. Nous opérons ensuite un test de Wilcoxon par paires :

- PCP vs. RC : W=-283 et p=0.0018
- PCP vs. CT : W=-87 et p=0.1867
- RC vs. CT : W=201 et p=0.0197

Il existe donc une différence significative entre les graphiques PCP et RC ; et RC et CT. Nous pouvons conclure que **le RC est le graphique avec lequel les participants ont moins apprécié réaliser un compromis.**

Nous devons conclure qu'il existe une différence entre les trois groupes. Il y a donc une influence du graphique sur la préférence du concepteur : le graphique RC est le moins apprécié par les participants car il est le graphique avec lequel les participants ont moins apprécié la recherche de compromis et celui avec lequel les participants ont moins apprécié réaliser un compromis.

De manière plus globale, nous pouvons conclure que le graphique influence la confiance du participant et que le graphique RC est le moins performant pour cet indicateur.

3.3.4 Conclusion pour la sélection en M.A.D.M.

Nos résultats montrent, en considérant les conditions expérimentales, la configuration des graphiques et les consignes données aux participants, que le graphique Parallel Coordinates Plot (PCP) est le graphique qui permet d'être performant pour la sélection dans une situation de décision multi-attribut (**Tableau 11**).

Le PCP est le plus performant pour la recherche d'un compromis. Il montre également un très bon indicateur pour le compromis (le PCP obtient de meilleures moyennes pour le temps d'accomplissement et l'indice de qualité) bien qu'il ne soit pas significativement différent du RC et du CT (selon les analyses statistiques).

Enfin nous déconseillons l'utilisation du Radar Chart (RC) qui est le graphique le moins « apprécié » par les concepteurs notamment pour les indicateurs de charge cognitive et de confiance.

Tableau 11 : Synthèse des résultats pour la sélection en M.A.D.M.

	La recherche	Le compromis	Le ressenti
Meilleur graphique	PCP	PCP ≡ CT ≡ RC	PCP ≡ CT
Pire graphique	CT ≡ RC		RC

3.4 INTRODUCTION DE L'EXPERIMENTATION N° 2

L'objectif de l'expérimentation n°2 est l'identification d'une combinaison de graphiques permettant aux concepteurs d'aboutir à une décision informée lors de l'exploration d'un espace de conception.

Plus précisément, nous cherchons, premièrement, à tester l'influence des graphiques sur les trois phases de l'exploration d'un espace de conception i.e. identifier un graphique permettant d'être performant pour chacune des phases. Et deuxièmement, nous cherchons à identifier un ou plusieurs graphiques permettant aux concepteurs d'aboutir à une décision informée.

Trois graphiques ont été identifiés pour être pertinents lors de l'exploration d'un espace de conception : Le Scatter Plot Matrix (SPM), le Simple Scatter Plot (SSP) et le Parallel Coordinates Plot (PCP).

Afin de réaliser l'expérimentation nous avons développé une « plateforme » d'expérimentation web. Cette plateforme permet entre autre la génération de *design point* et la réduction de l'espace de conception.

Comme exprimé précédemment, la plateforme d'expérimentation est disponible en ligne à l'adresse suivante : http://these.aaa.alwaysdata.net/expe2/

3.4.1 Présentation de la plateforme d'expérimentation

Dans cette section nous présentons notre plateforme d'expérimentation qui est un outil d'exploration de points de conception.

L'outil développé permet l'exploration d'espace de conception à travers différents graphiques et fonctionnalités / interactions. Il est composé de trois graphiques différents et propose quatre problèmes de conception.

Ainsi dans cette section nous décrivons l'interface puis les menus et les interactions qui autorisent la manipulation de données et les problèmes de conception.

3.4.1.1 L'interface

L'interface de l'outil est composée de quatre zones. Comme le montre la **Figure 85**, nous avons :

- Une zone appelée « Menu » où se trouve les différents onglets et fonctionnalités de l'outil.
- Une zone appelée « Graphique » où se trouve la projection des points de conception.
- Une zone appelée « Détail » où s'affiche le détail d'un point de conception lorsque la fonction « selection » est utilisée.

- Une zone appelée « Info » où se trouvent des indications comme une échelle de couleur et le nombre de points de conception projetés dans le graphique.

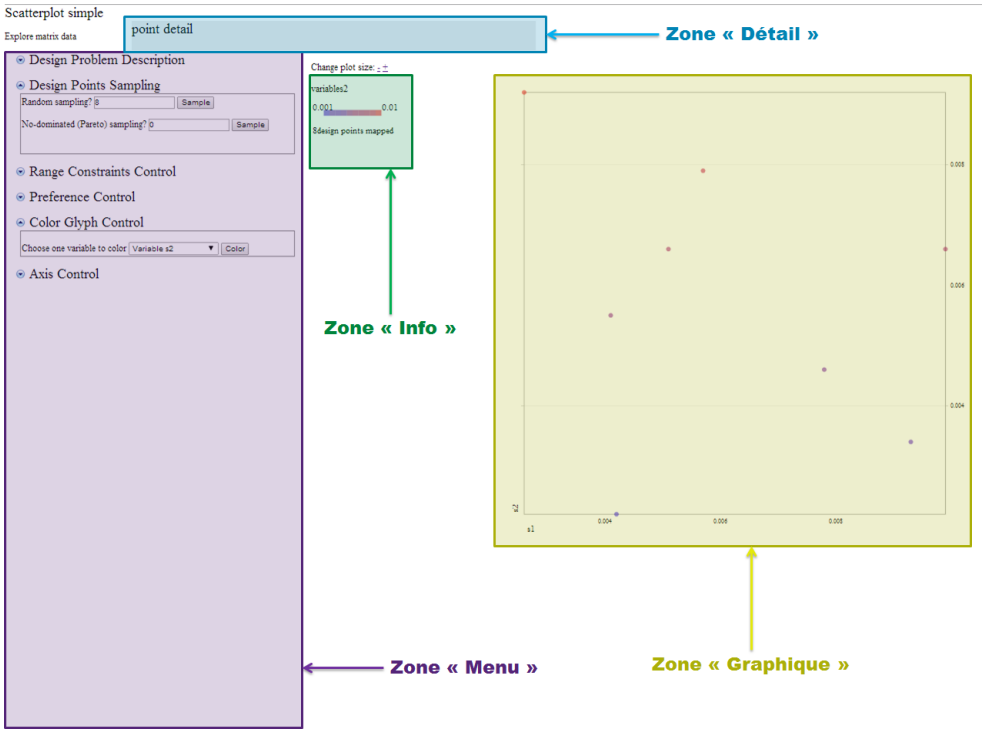


Figure 85 : Composition de l'interface – Exemple du Simple Scatter Plot

L'organisation de l'interface reste inchangée quel que soit le graphique ou le problème de conception (voir **Figure 86**).

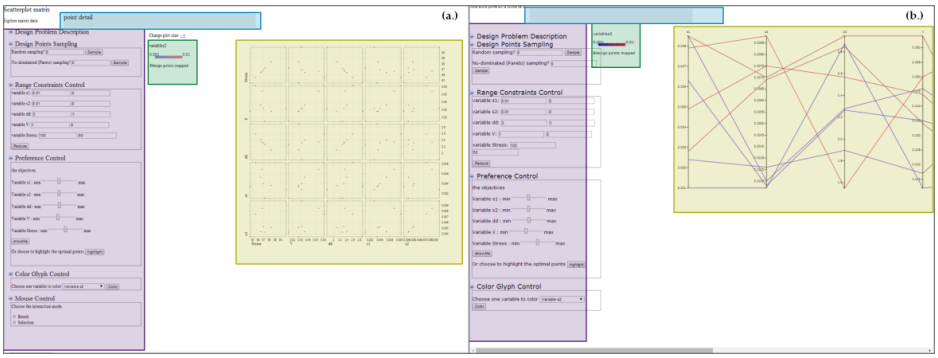


Figure 86 : Interface avec le graphique Scatter Plot Matrix (a.) et avec le Parallel Coordinates Plot (b.)

3.4.1.2 Menus

Les menus sont composés d'une base de cinq fonctionnalités (une de plus pour le SSP et le SPM). Les menus sont positionnés en haut à gauche de l'écran et sont tous étirables (« expand ») et rétractables (« collapse ») simplement en cliquant sur l'intitulé de la fonctionnalité ou sur la petite flèche à sa gauche, dans le petit rond bleu (voir la **Figure 87**).

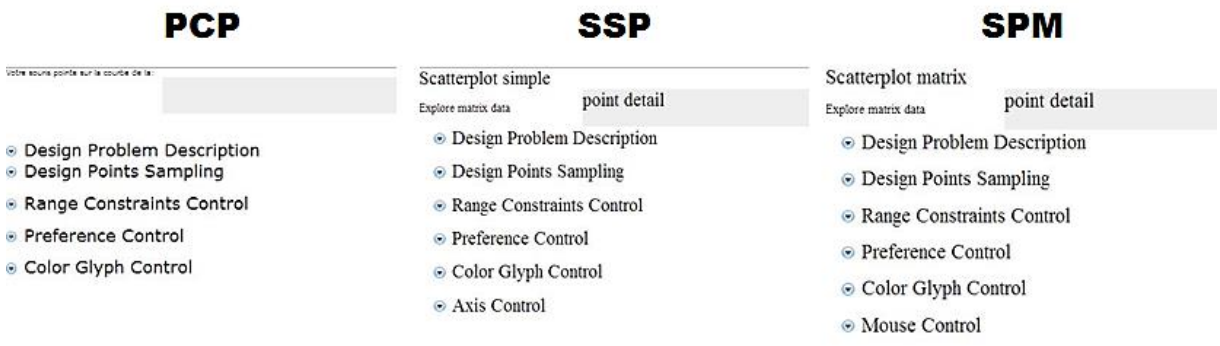


Figure 87 : Les trois menus

3.4.1.2.1 Onglet « Design Problem Description »

Le premier onglet du menu, intitulé « design problem description », propose l'énoncé d'un problème de conception. Dans cette description, nous trouvons un schéma représentant le problème, la présentation des variables de conception, des variables de performance et les objectifs et les contraintes (voir **Figure 88**).

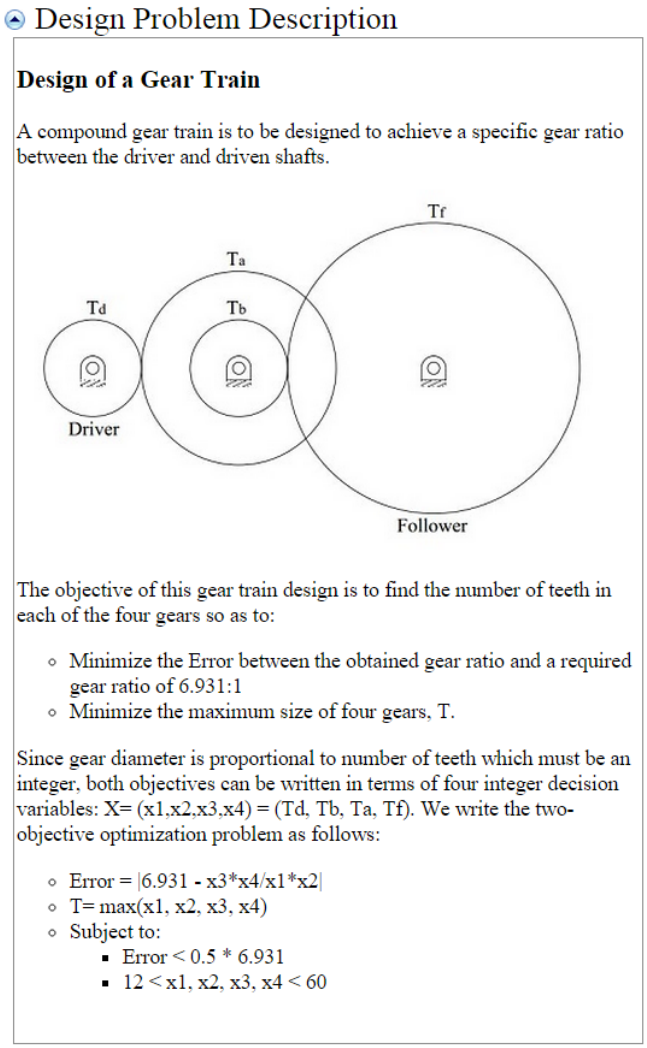


Figure 88 : Onglet « Design Problem Description »

3.4.1.2.2 Onglet « Design Points Sampling »

Le deuxième onglet du menu intitulé « design points sampling » permet la génération de points de conception faisables dans l'espace de conception. Deux générateurs de points sont disponibles comme le montre la **Figure 89**. Le « random sampling » produit le nombre de points désiré dans l'espace de conception de façon aléatoire. Le « no-dominated (Pareto) sampling » produit des points dits non-dominés au sens de Pareto.

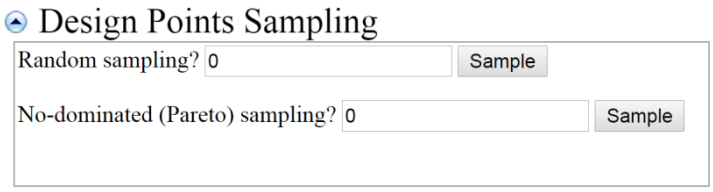


Figure 89 : Onglet « Design Points Sampling »

Attention, ce générateur ne produit pas d'optimums au sens de Pareto mais des « non-dominés », c'est-à-dire que plus il y a de non-dominés générés et plus les points seront proches de la frontière de Pareto (voir **Figure 90**). De cette manière, cette fonction n'est disponible que s'il existe des points générés au préalable.

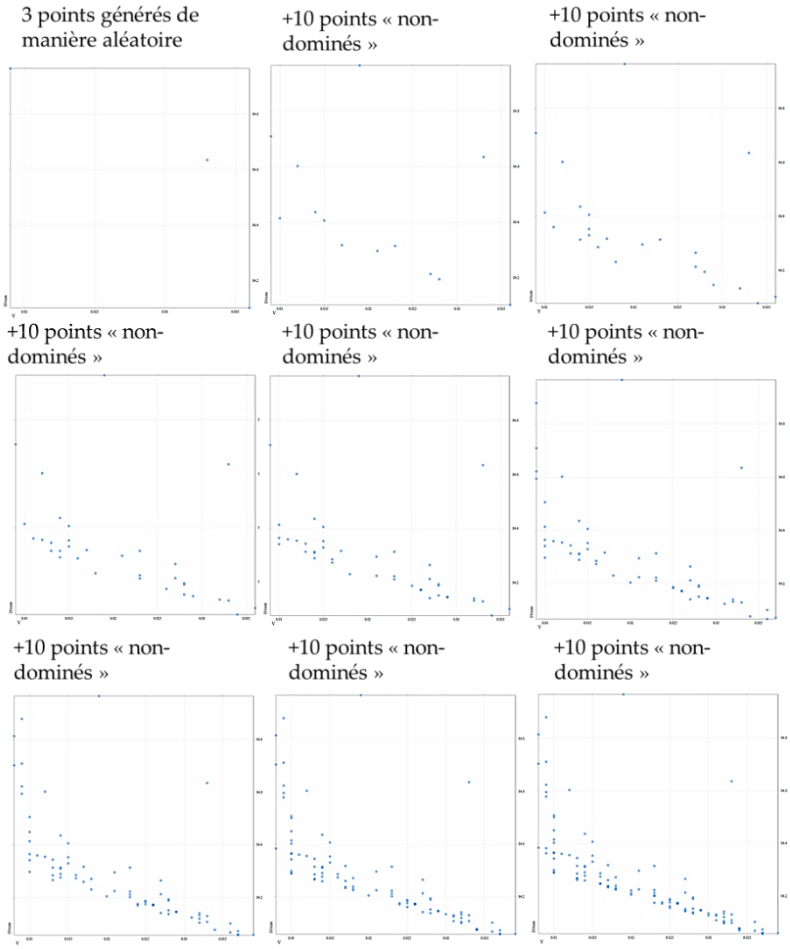


Figure 90 : Illustration du générateur de points non-dominés au sens de Pareto

3.4.1.2.3 Onglet « Range Constraints Control »

Le troisième onglet intitulé « range constraints control » permet de réduire l'espace de conception en contrôlant les valeurs maxi et mini de chaque variable (voir **Figure 91**).

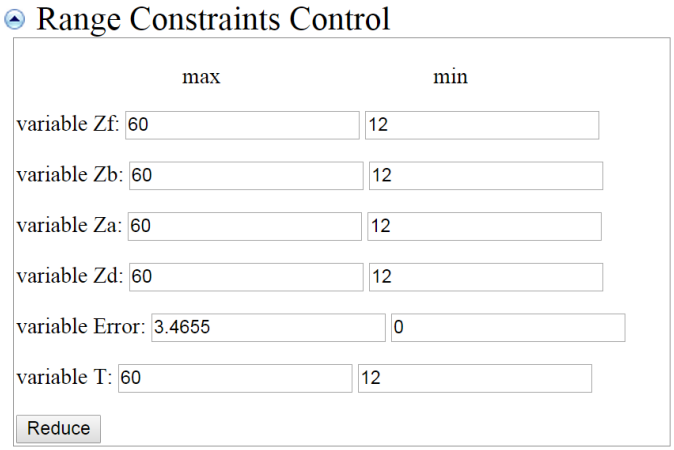


Figure 91 : Onglet « Range Constraints Control »

Les bornes maxi et mini par défaut sont affichées dans les champs modifiables.

3.4.1.2.4 Onglet « Preference Control »

Le quatrième onglet intitulé « preference control » permet de mettre en lumière les points de conception selon deux fonctions dites de préférence.

La première fonction permet de colorer les points de conception sur une échelle de couleur en accord avec les objectifs du problème (il s'agit de l'Indice Basique). Pour chaque variable, il est possible d'indiquer si elle est à maximiser ou minimiser à l'aide de « sliders ». Le slider doit rester au milieu si la variable ne doit être ni maximiser ni minimiser (voir **Figure 92**). Ainsi lorsque le bouton « showMe » est cliqué, les points sont coloriés en tenant compte des objectifs qu'ont été indiqués : du rouge pour les points qui répondent au mieux aux objectifs au bleu pour les points qui sont les plus loin des objectifs.

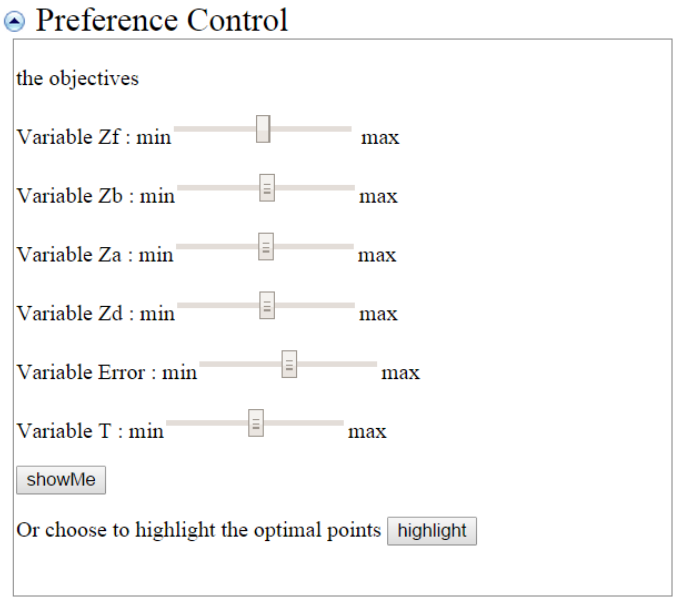


Figure 92 : Onglet « Preference Control »

La deuxième fonction permet de mettre en lumière les solutions (points) dites optimales (au sens de Pareto). Lorsque le bouton « highlight » est cliqué, les points dits optimums sont colorés en rouge et les autres sont grisés (voir **Figure 93**).

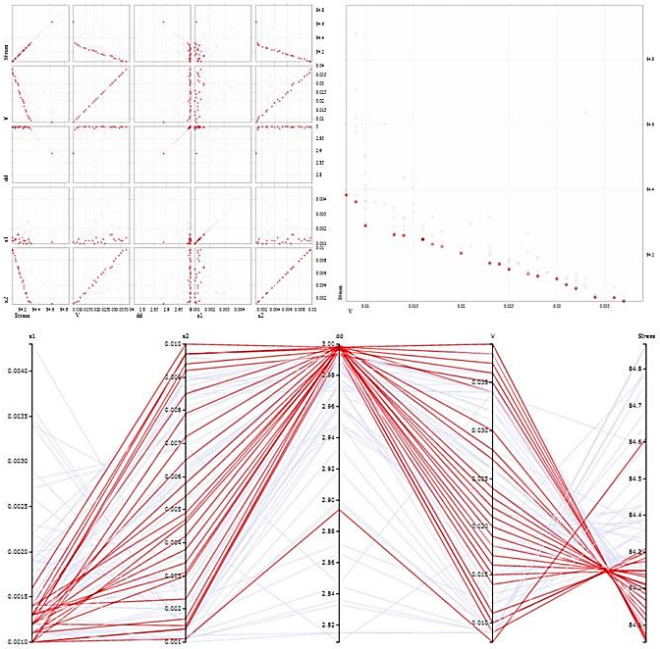


Figure 93 : Illustration de la mise en lumière des optimums

3.4.1.2.5 Onglet « Color Glyph Control »

Le cinquième onglet intitulé « color glyph control » permet d'ajouter un marqueur de couleur sur une variable au choix (voir la liste déroulante sur la **Figure 94**).

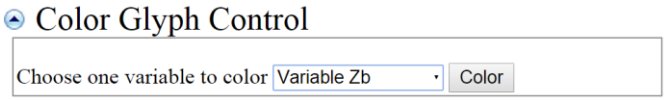


Figure 94 : Onglet « Color Glyph Control »

3.4.1.2.6 Onglet « Axis Control » (spécifique au SSP)

Le sixième onglet pour le SSP intitulé « axis control » permet de changer les variables à afficher sur le nuage de points. Seules deux variables peuvent être affichées car il n'y a que deux axes (X et Y). Il suffit donc de sélectionner la variable en X et la variable en Y (voir les « boutons radios » sur la **Figure 95**).

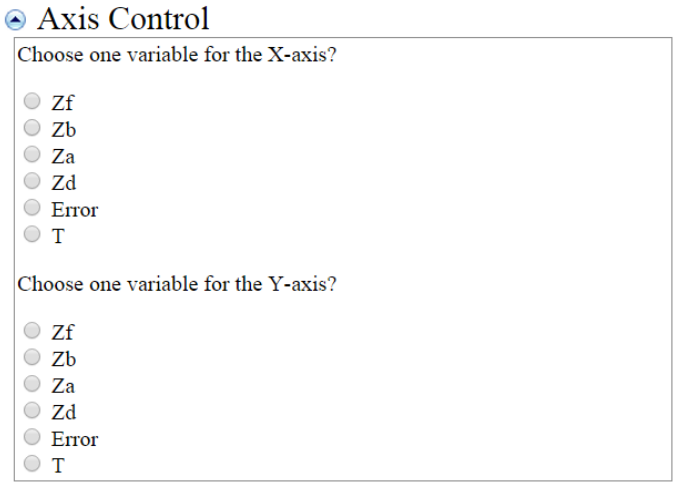


Figure 95 : Onglet « Axis Control »

3.4.1.2.7 Onglet « Mouse Control » (spécifique au SPM)

Le sixième onglet pour le SPM intitulé « mouse control » permet de changer l'interaction autorisé par la souris à l'aide de « radio button » (voir **Figure 96**).

Le mode « brush » (par défaut) permet la fonction de brush (sélection de points dans une vue et mise en lumière de ces mêmes points dans les autres vues).

Le mode « selection » permet d'afficher le détail d'un point lors du survol du point avec la souris.

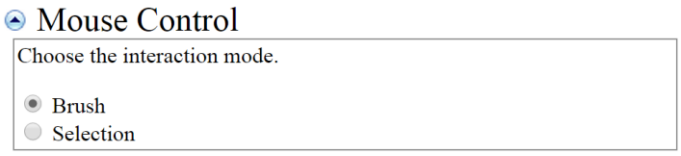


Figure 96 : Onglet « Mouse Control »

3.4.1.3 Interactions

Outre les fonctionnalités disponibles dans le menu, il est possible d'interagir directement sur les graphiques. Les interactions permettent l'utilisation de trois fonctions la *selection* , le *brush* et le *drag&drop*.

3.4.1.3.1 La fonction « selection »

La fonction « selection » est à disposition avec l'utilisation des trois graphiques. Cette fonction permet d'afficher le détail d'un point de conception. Pour ce faire, il suffit de survoler un point (pour le SSP et le SPM) ou une ligne (pour le PCP). L'interaction est donc en direct sur le graphique lorsque la souris est en mode *pointer* (voir **Figure 97**).

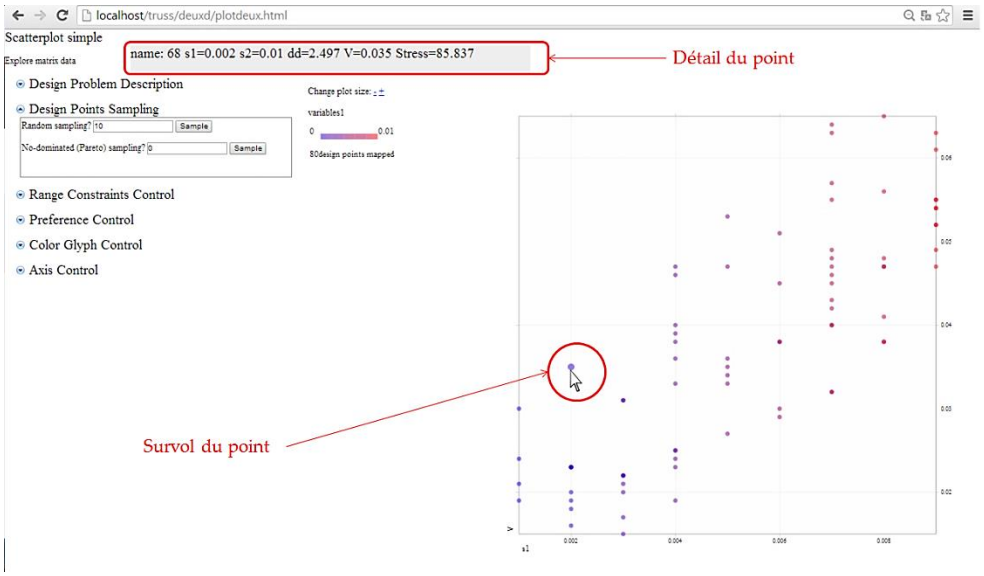


Figure 97 : Illustration de la fonction « selection » avec le SSP

3.4.1.3.2 La fonction « brush »

La fonction *brush* permet la sélection de points dans une vue afin de mettre en lumière ces même points dans les autres vues (les points non-sélectionnés sont grisés). Cette fonction est à disposition avec l'utilisation du SPM et du PCP (le SSP étant composé d'une seule vue, le *brush* n'a pas raison d'être). L'interaction est donc en direct sur le graphique lorsque la souris est en mode *crosshair* (voir **Figure 98**).

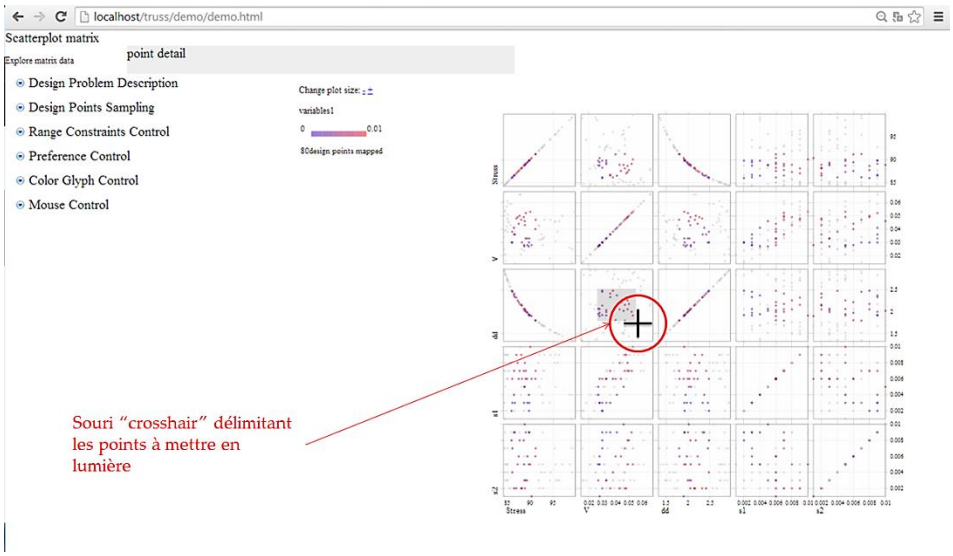


Figure 98 : Illustration de la fonction « brush » avec le SPM

3.4.1.3.3 La fonction « drag & drop »

La fonction *drag & drop* n'est disponible qu'avec l'utilisation du PCP. Cette fonction permet de déplacer les axes du graphique. L'interaction est donc en direct sur le graphique lorsque la souris est en mode *move* (voir **Figure 99**). L'interaction s'opère au niveau des labels (noms des variables).

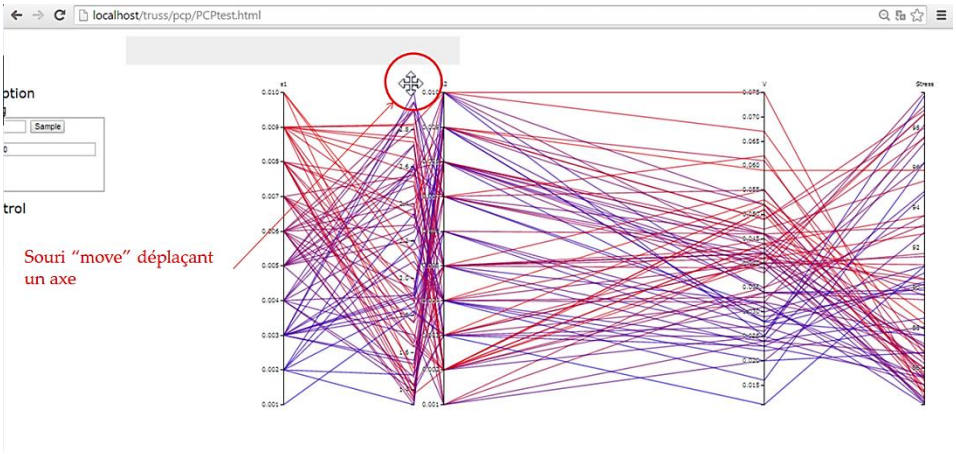


Figure 99 : Illustration de la fonction « drag & drop » avec le PCP

3.4.1.4 Les problèmes de conception

Pour la plateforme d'expérimentation, quatre problèmes de conception bi-objectifs ont été implémentés :

- Le *Two-members Truss design* (Truss)
- Le *Gear train design* (Gear)
- Le *Multiple-disk clutch brake design* (Disk)
- Le *Pressure vessel design* (Vessel)

Les trois premiers problèmes sont issus des travaux de Deb et Srinivasan (Deb et Srinivasan, 2006) et le quatrième problème est issu de (Karandikar et Mistree, 1992 ; Lewis et Mistree, 1998 ; Canbaz, 2013).

3.4.1.4.1 Le problème Truss

Une structure triangulée à deux membres doit être dimensionnée afin de supporter une certaine charge sans rupture élastique (voir **Figure 100**). Ce problème considère deux objectifs :

- Minimiser le volume total de la structure (noté V)
- Minimiser la contrainte développée dans les deux membres (noté *Stress*).

La contrainte développée dans les membres provient de l'application d'une charge de 100kN.

Il y a trois variables de décision (ou paramètres de conception) :

- L'aire de la section du premier membre $x1$ mesurée en m²
- L'aire de la section du deuxième membre $x2$ mesurée en m²
- La longueur verticale y mesurée en m.

Le problème d'optimisation non-linéaire est donné par :

- $V = x1\sqrt{(16+y^2)} + x2\sqrt{(1+y^2)}$
- $Stress = \max(\sigma 1 , \sigma 2)$
- Sujet à:
  - $\max(\sigma 1 , \sigma 2) < S_{max}$
  - $0 < x1, x2 < A_{max}$
  - $1 < y < 3$

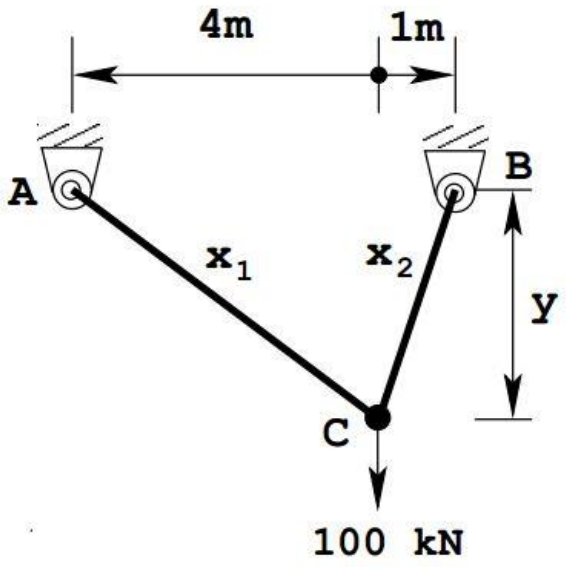


Figure 100 : Illustration du problème Truss

En utilisant les dimensions et la charge spécifiée, nous observons que le membre n°1 est sujet à une charge de $[20\sqrt{(16+y^2)}/y]$ kN et le membre n°2 à une charge de $[80\sqrt{(1+y^2)}/y]$ kN. Les contraintes sont calculées par :

- $\sigma 1 = 20\sqrt{(16+y^2)}/y*x1$
- $\sigma 2 = 80\sqrt{(1+y^2)}/y*x2$

Ici, nous limitons les contraintes à $S_{max} = 10^5$ kPa et les aires des deux sections à A_{max}=0.01 m²

3.4.1.4.2 Le problème Gear

Un train d'engrenages doit être conçu pour obtenir un rapport d'engrenages spécifique entre le conducteur et les arbres entraînés (voir **Figure 101**).

L'objectif de la conception du train d'engrenages est de trouver le nombre de dents de chacune des quatre roues de manière à:

- Minimiser l'erreur entre le rapport obtenu et un rapport requis de 6.931:1 (variable notée Error)
- Minimiser la taille maximum des quatre rapports, T.

Comme le diamètre des roues dentées est proportionnel au nombre de dents (qui est un entier), les deux objectifs peuvent être écrit à l'aide de quatre variables: X= (x1,x2,x3,x4) = (Td, Tb, Ta, Tf). Nous écrivons le problème d'optimisation bi-objectif:

- Error = |6.931 – (x3*x4)/(x1*x2)|
- T= max(x1, x2, x3, x4)
- Sujet à:
    - Error < 0.5 * 6.931

- 12 < x1, x2, x3, x4 < 60

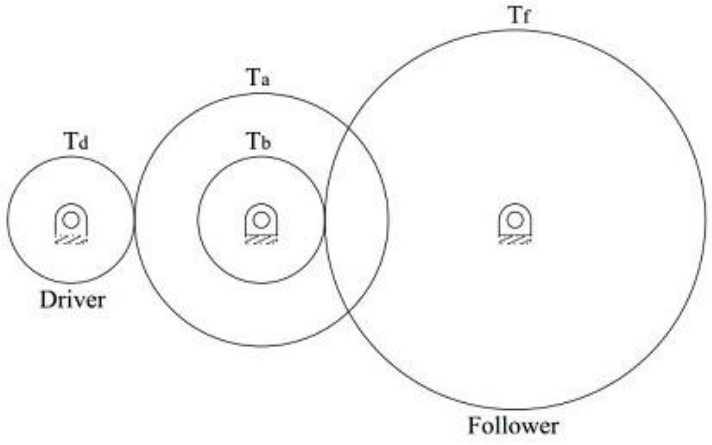


Figure 101 : Illustration du problème Gear

3.4.1.4.3 Le problème Disk

Dans ce problème, un frein à disque multiple doit être conçu. Deux objectifs en conflit sont considérés:

- Minimiser la masse (M) du système en kg.
- Minimiser le temps d'arrêt (S) en secondes

Il existe cinq variables de décision Ri, Ro, t, F, et Z, où Ri est le rayon interne en mm, Ro est le rayon extérieur en mm, t est l'épaisseur des disques en mm, F est la force d'actionnement en N et Z est le nombre de disques (voir **Figure 102**). Les cinq variables sont considérées distinctes et leurs valeurs admissibles sont données ci-dessous:

- 60 < Ri < 80
- 90 < Ro < 110
- 1 < t < 3
- 600 < F < 1000
- 2 < Z < 10

Le problème d'optimisation est formulé ci-dessous:

- $M = \Pi * (Ro^2 - Ri^2)*t*(Z+1)0.0000078$
- $S = Iz*(\Pi*250/30)/(3+Mh)$
- Sujet à:
  - C1: $Ro-Ri-20 \geq 0$
  - C2: $30-(Z+1)*(t+0.5) \geq 0$
  - C3: $1-(F/ \Pi *(Ro^2 - Ri^2))\geq 0$
  - C4: $10000- (F/ \Pi *(Ro^2 - Ri^2))*\Pi*250*Rsr/30 \geq 0$
  - C5: $10000-\Pi*250*Rsr/30 \geq 0$
  - C6: $Mh-1.5*40 \geq 0$

- C7: $S \geq 0$
- C8: $15-S \geq 0$

Et les paramètres sont donnés ci-dessous:

- $Iz = 55$ kg.m²
- $Mh = (2/3) * 0.5 * F * Z* (Ro^3 - Ri^3) / (Ro^2 - Ri^2)$
- $Rsr = (2/3) * (Ro^3 - Ri^3) / (Ro^2 - Ri^2)$

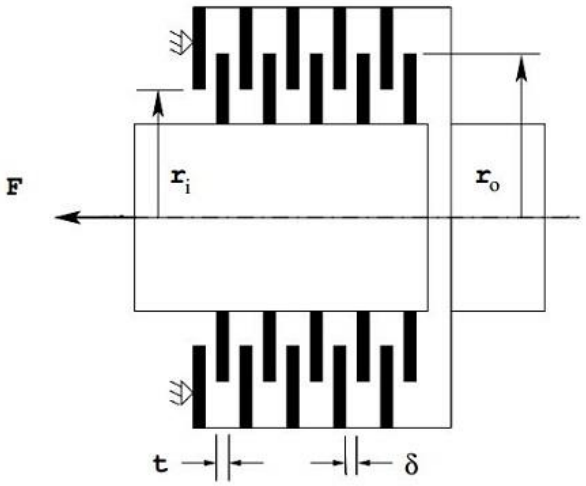


Figure 102 : Illustration du problème Disk

3.4.1.4.4 Le problème Vessel

Il s'agit d'un problème de conception d'une cuve cylindrique sous pression à paroi mince avec des extrémités hémisphériques (voir **Figure 103**). Il y a trois paramètres de conception (R, T, L) et deux variables de performances (W, V). Les objectifs sont de minimiser W et maximiser V en contrôlant R, T et L tout en satisfaisant des contraintes C1 à C7.

Nomenclature et constantes du problème:

- W, Poids de la cuve sous pression, lbs.
- V, Volume de la cuve sous pression.
- R, Rayons.
- T, Epaisseur de la paroi.
- L, Longueur du cylindre.
- Constantes :
  - P, Pression à l'intérieur du cylindre.
  - UTS, Résistance à la traction du matériau, 35 klb.
  - d, Densité du matériau de la cuve, 0.283.
  - Circ, Contrainte circonférentielle.

Formules de performance:

- $W = d[\ 4/3\ \Pi(R + T)^3 + \Pi(R + T)^2\ L - (4/3\ \Pi\ R^3 + \Pi\ R^2\ L)]$
- $V = 4/3\ \Pi\ R^3 + \Pi\ R^2\ L$

Contraintes initiales:

- Contrainte d'effort:
  - C1: $(P\ R\ /T) - UTS < 0$
- Contraintes géométriques:
  - C2: $5T - R < 0$
  - C3: $R + T - 40 < 0$
  - C4: $L + 2R + 2T - 150 < 0$
- Limites:
  - C5: $0.1 < R < 36$
  - C6: $0.1 < L < 140$
  - C7: $0.5 < T < 6$

Limites des performances:

- $13.7288 < W < 56907.6$
- $64.4138 < V < 480404$

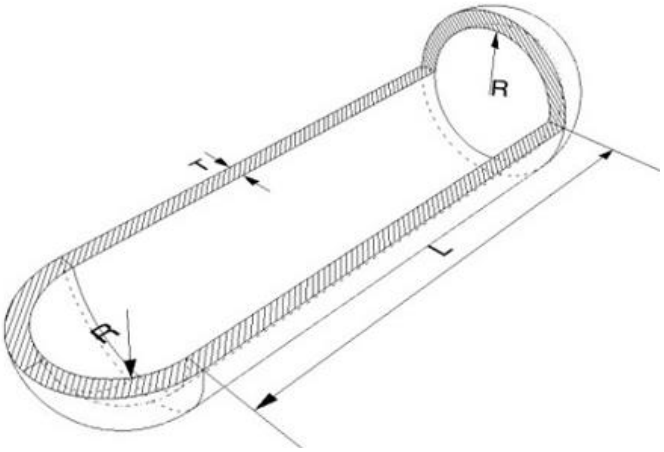


Figure 103 : Illustration du problème Vessel

3.4.2 Déroulement de l'expérimentation n° 2

Dans cette expérimentation, nous décidons d'adopter une approche par groupes indépendants[6]. Cela implique que chaque participant teste un seul graphique.

Nous proposons de segmenter le déroulement de l'expérimentation en deux parties principales : la partie Formation et la partie Test (respectivement en vert et bleu + violet sur la **Figure 104**). De plus, nous intégrons un type de jalon durant la séance : les Q.C.M. (en rose sur la **Figure 104**).

[6] L'approche par groupes appariés a été envisagée mais la séance atteignait un temps nécessaire de 3h. Elle a donc dû être abandonnée.

L'expérimentation est dimensionnée pour être bornée à une séance de deux heures. La partie Formation nécessite environ une heure, la partie Test est tablée sur quarante minutes et les vingt minutes restantes sont absorbées par les différents Questionnaires à Choix Multiples (Q.C.M.) de l'expérimentation.

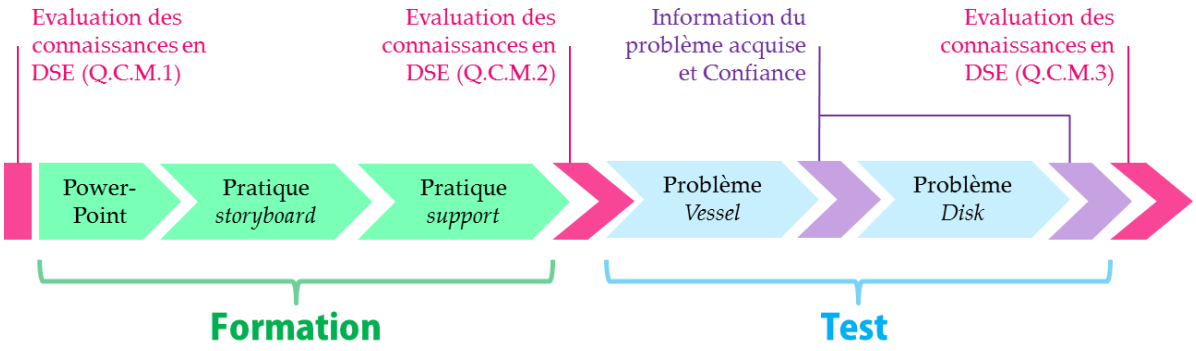


Figure 104 : Illustration du déroulement de l'expérimentation n°2

3.4.2.1 La phase de Formation

Cette partie permet la mise à niveau des participants. Elle est scindée en trois étapes (voir **Figure 105**) :

- La première est sous forme d'un cours accéléré avec un support de présentation
- Les deuxième et troisième étapes servent de prise en main de l'outil

Le mini cours introduit les notions théoriques liées à l'exploration et la visualisation d'espace de conception (cours accéléré en annexes p.213). Ainsi, une partie du support est dédiée à la méthode de conception (définition de l'espace de conception, la dominance au sens de Pareto, etc.). Une deuxième partie est dédiée au graphique, son utilisation et interprétation. L'expérimentation étant avec des groupes indépendants, la partie visualisation du cours accéléré ne décrit l'utilisation que d'un seul graphique.

La formation se poursuit avec deux étapes dites de pratiques permettant premièrement la prise en main de l'outil (i.e. la plateforme d'expérimentation) et deuxièmement la mise en pratique des notions apprises dans le mini cours.

La prise en main de l'outil consiste à résoudre un problème de conception, le problème Truss, à l'aide d'un tutoriel sous forme d'un storyboard. Trois storyboards ont été imaginés pour l'expérimentation, soit un storyboard par graphique. Les storyboards sont conçus de manière à utiliser toutes les fonctionnalités et interactions disponibles sur la plateforme d'expérimentation (storyboard en annexes p.223).

La mise en pratique consiste à résoudre un autre problème de conception, le problème Gear, avec le manuel d'utilisation de la plateforme d'expérimentation comme support. Durant

cette partie le participant est autorisé à poser des questions. Cette étape se termine lorsque le participant propose une solution optimale.

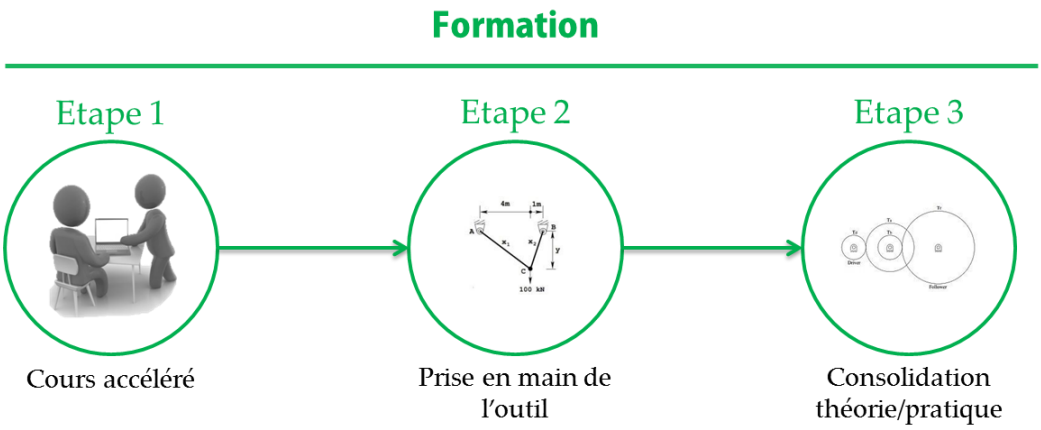


Figure 105 : Illustration des trois étapes de la phase Formation

3.4.2.2 La phase de Test

Cette partie est l'étape de test du graphique lors de l'exploration d'un espace de conception. Dans cette phase de l'expérimentation, il y a deux problèmes de conception à résoudre à l'aide d'un des trois graphiques :

- Le problème Vessel : composé de cinq variables, i.e. trois paramètres de conception et deux variables de performance.
- Le problème Disk : composé de sept variables, i.e. cinq paramètres de conception et deux variables de performance.

Le participant a une limite de temps de dix minutes pour chaque résolution de problème sachant que le chronomètre se déclenche lorsque le participant génère une première fois des points de conception. Durant cette partie de l'expérimentation, les questions et les supports d'aide ne sont plus autorisés. L'interface étant en anglais, la description du problème est donnée en français sur un support papier.

Pour la résolution des deux problèmes de conception, le participant génère des solutions possibles (en utilisant les deux *samplers* au choix) et réduit l'espace de conception vers une « zone » de performance de l'espace où il sélectionne une solution optimale. Pour observer son exploration (et relever certains de nos indicateurs), nous utilisons un eye-tracker (Tobii X2).

La consigne pour les deux problèmes est de résoudre le problème d'optimisation bi-objectif en utilisant la méthode d'exploration d'espace de conception afin de sélectionner une solution optimale comportant si possible des paramètres standardisables.

La notion de paramètres standardisables est issue des travaux de Deb et Srinivasan (2006) et est décrite dans la section 3.6.1.

De plus, la phase de test est jalonnée par des questionnaires post-problème (en violet sur la **Figure 104** et **Figure 106**). Ces questionnaires post-problèmes nous permettent de connaitre la confiance du participant vis-à-vis de sa décision et les informations acquises relatives au problème.

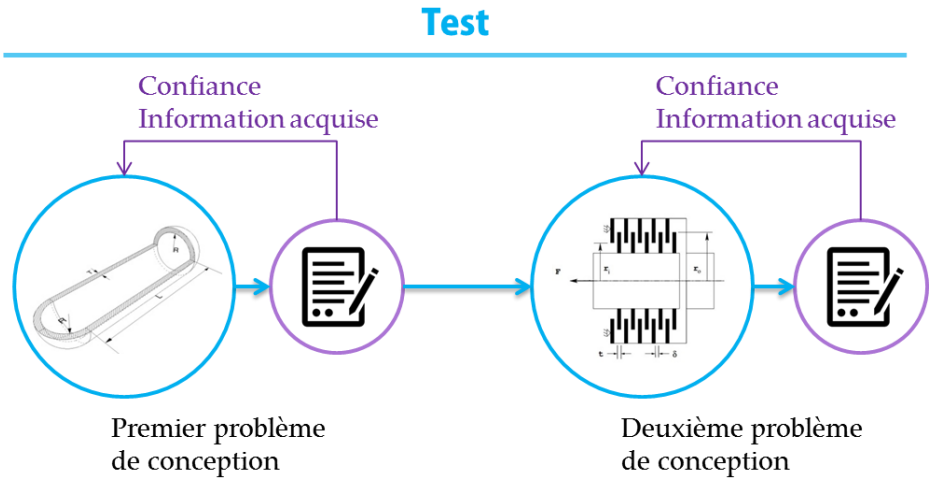


Figure 106 : Illustration des différentes étapes de la phase Test

3.4.2.3 Les Questionnaires à Choix Multiples (Q.C.M.)

Pour concevoir un questionnaire, deux types de questions peuvent être utilisés[7] :

- la question ouverte ou non-structurée : elle laisse l'individu interrogé totalement libre du choix de sa réponse dans sa forme et sa longueur. Elle est généralement utilisée pour les études exploratoire/qualitatives lorsque la recherche est l'approfondissement ou l'exploration.
- la question fermée ou structurée (dichotomique, à choix multiple ou avec échelle): elle précise la série de réponses possibles. Ses avantages sont la précision et la facilité de traitement. Elle est généralement utilisée pour les études descriptives/quantitatives lorsque la recherche est le contrôle ou la vérification.

Dans notre expérimentation, nous utilisons les Q.C.M. qui font offices d'indicateurs de contrôle. Chaque Q.C.M. est composé de dix questions portant sur l'exploration et la visualisation d'espace de conception. Pour chaque graphique, il y a trois Q.C.M. soit neuf au total. Ainsi, les questions portant sur la visualisation sont contextualisées au graphique mais sont du même style pour éviter un biais (un groupe de participants favorisé par des questions plus simples par exemple).

Vous trouverez ci-dessous un exemple de la question n°8 du Q.C.M. 1. La question est la même mais l'illustration diffère (voir **Figure 107**).

[7] http://www.unilim.fr/pages_perso/vincent.jalby/m1aes/cours/m1aes_03.pdf

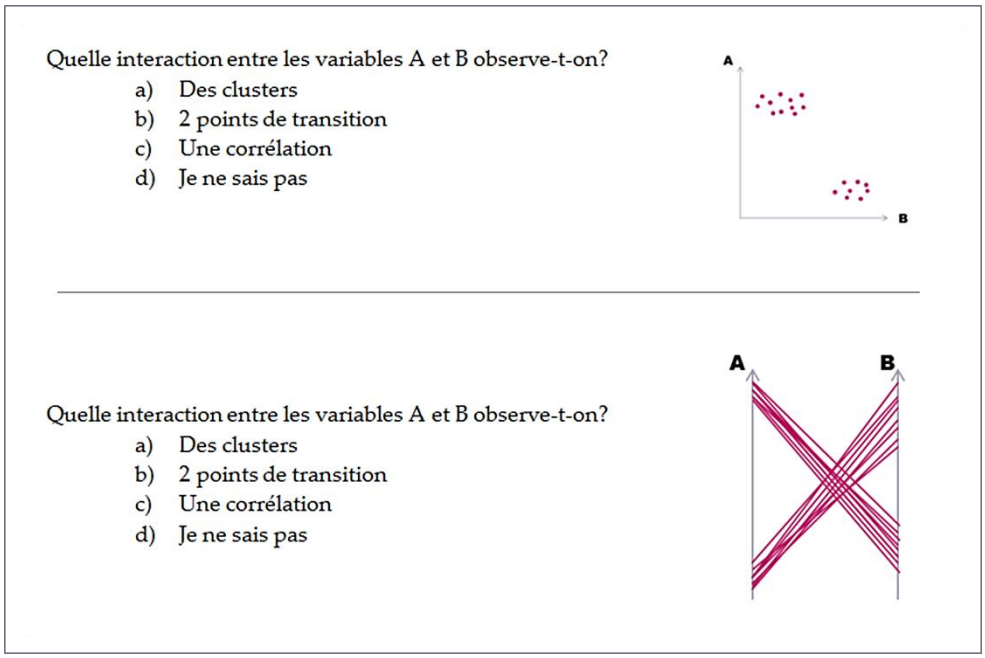


Figure 107 : Exemple de la question n°8 du Q.C.M. 1

Dans l'expérimentation, chacun des trois Q.C.M. ont un rôle bien spécifique. Le premier est réalisé avant de commencer la séance et nous permet de connaitre le profil « concepteur » du participant à partir de son niveau de connaissance. Nous définissons trois profils en comptant le nombre de bonnes réponses données au premier QCM :

- Expert pour un nombre de bonnes réponses compris entre 8 et 10
- Intermédiaire pour un nombre de bonnes réponses compris entre 4 et 7
- Novice pour un nombre de bonnes réponses compris entre 0 et 3

Le deuxième Q.C.M. nous permet principalement de contrôler l'évolution de connaissance du participant et de vérifier qu'il atteint au minimum un niveau Intermédiaire avant de passer à la phase de Test.

Le troisième Q.C.M. nous permet de vérifier que le graphique utilisé par les participants n'ait pas créé de confusion. Plus particulièrement, nous vérifions que le nombre de bonnes réponses entre le deuxième et le troisième n'ait pas diminué.

Enfin, les trois Q.C.M. sont pour nous le moyen de vérifier qu'il n'existe pas de différences entre les trois groupes ou plus largement nous assurent de comparer des groupes comparables. En effet, nous avons trois groupes de participants :

- Les participants réalisant l'expérimentation avec le graphique SPM
- Les participants réalisant l'expérimentation avec le graphique SSP
- Les participants réalisant l'expérimentation avec le graphique PCP

Et potentiellement nous avons des participants avec différents niveaux d'expertise.

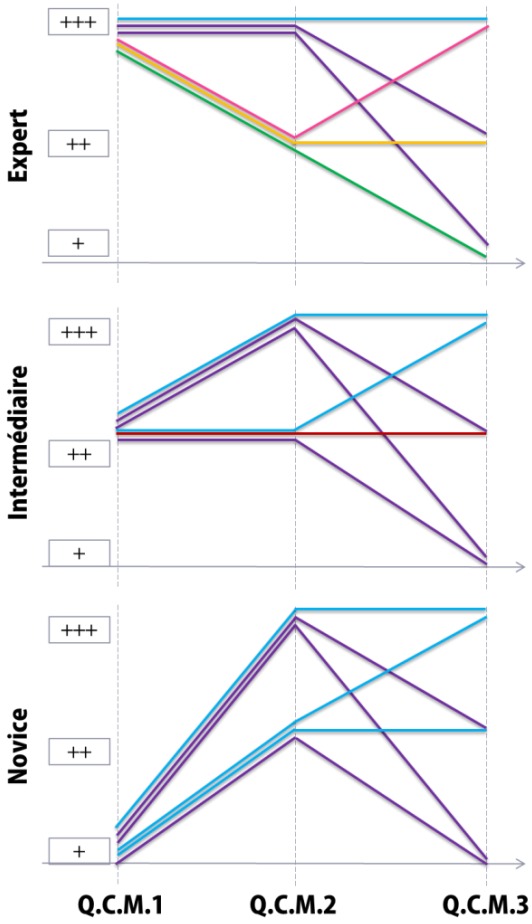


Figure 108 : Illustration des différentes évolutions de « connaissances » possibles par profil

La **Figure 108** illustre les différentes évolutions de connaissance en exploration et visualisation d'espace de conception possibles par profil d'expertise. En bleu sont représentés les évolutions attendues et en violets les cas où la phase Test a créé de la confusion. Nous relevons trois cas particuliers : l'intermédiaire constant en rouge, l'expérimentation créant une totale confusion chez l'expert en vert, les experts perturbés par la phase de Formation en rose et jaune.

3.4.3 Synthèse

Nous testons trois graphiques différents : le Scatter Plot Matrix (SPM), le Simple Scatter Plot (SSP) et le Parallel Coordinates Plot (PCP). Quarante-deux élèves-ingénieurs en 3ième année de l'ESTIA spécialisés en « conception de produits » ont participé à l'expérimentation. Chaque participant a résolu deux problèmes de conception avec un graphique : les problèmes « Vessel » et «Disk ».

Ainsi, la résolution des deux problèmes est pour nous le moyen d'identifier des graphiques pertinents pour les trois phases de l'exploration d'espace de conception et pour la prise de décision informée. Avant de présenter l'analyse de ces quatre dimensions, nous proposons une analyse préliminaire qui traite les données liées aux participants.

3.5 ANALYSE PRELIMINAIRE : GROUPES DE PARTICIPANTS

Dans cette partie nous analysons les données obtenues lors des tests afin de savoir si nous pouvons distinguer des profils de concepteur au sein des testeurs et/ou s'il existe une différence entre les groupes (i.e. trois groupes – trois graphiques). Comme spécifié dans la section précédente, nous utilisons les résultats des trois Q.C.M. Nous utilisons également le niveau de confiance du participant en sa décision (questionnaire post-problème).

Connaissance en exploration et visualisation d'espace de conception

Pour « contrôler » la connaissance en exploration d'espace de conception nous avons utilisé trois Q.C.M. de dix questions. Le premier Q.C.M. est donné avant que le test ne commence afin d'avoir le profil du participant (Expert, Intermédiaire ou Novice). Le deuxième Q.C.M. est donné après la phase de training et avant de commencer la résolution de problèmes afin de contrôler une évolution chez le participant. Enfin le troisième Q.C.M. est donné à la fin du test, une fois que les problèmes de conception ont été résolus afin de contrôler que l'exercice (test) n'ait pas perturbé le participant (ex : une chute de son niveau de connaissance). Il s'agit donc pour nous de variables qualitatives ordinales.

Niveau de confiance

A la fin de chacun des deux problèmes, trois questions ont permis de relever le niveau de confiance du participant dans sa décision. Les réponses à ces trois questions étaient données sur une échelle comprise entre 1 et 10 :

1. Sur une échelle de 1 à 10, pourriez-vous noter à quel niveau êtes-vous certain d'avoir choisi une solution optimale (10= très certain)?

   1 5 10

2. Sur une échelle de 1 à 10, pourriez-vous noter à quel niveau estimez-vous avoir choisi une solution qui reflète vos préférences (1= aucunes préférences, 10= toutes mes préférences)?

   1 5 10

3. Sur une échelle de 1 à 10, pourriez-vous donner votre niveau de certitude? (1= je ne sais pas, 5= j'ai fait un compromis, 10=j'ai sélectionné la meilleure solution)

   1 5 10

Il s'agit donc pour nous de variables qualitatives ordinales.

3.5.1 Connaissance en exploration d'espace de conception

3.5.1.1 Analyse Intra-Graphiques

Pour cette analyse nous utilisons le test de Friedman car la variable est qualitative ordinale et que l'étude est composée de trois mesures appariées (n=14 pour chaque mesure). Notons qu'un test de Friedman donne comme résultat une version statistique du Khi-carré noté csqr, le degré de liberté noté df et la valeur de p notée p.

Les analyses sont réalisées pour les trois graphiques indépendamment. Ainsi, nous posons les hypothèses statistiques pour chacun des graphiques :

- H0 : il n'existe pas de différence entre les réponses aux trois Q.C.M.
- H1 : il existe une différence entre les réponses données aux trois Q.C.M. (QCM1<QCM2<QCM3)

3.5.1.1.1 SSP

Les moyennes de rang sont égales à 1 pour le QCM1, 2.2 pour le QCM2 et 2.8 pour le QCM3. Le test nous donne csqr=23.89, df=2 et p<0.0001.

Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence entre les trois mesures. Nous opérons donc un test de de Wilcoxon (rang signé) par paires :

- QCM1 vs. QCM2 : W=-105, Z=-3.28 et p=0.0005.
- QCM2 vs. QCM3 : W=-76, Z=-2.64 et p=0.0041.
- (QCM1 vs. QCM3 n'a pas d'intérêt pour nous)

Les résultats des tests de Wilcoxon réalisés par paires nous indiquent qu'il existe une différence significative entre les trois réponses données aux QCM pour les participants ayant utilisé le SSP : QCM1<QCM2<QCM3.

3.5.1.1.2 PCP

Les moyennes de rang sont égales à 1 pour le QCM1, 2.1 pour le QCM2 et 2.9 pour le QCM3. Le test nous donne csqr=26.14, df=2 et p<0.0001.

Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence entre les trois mesures. Nous opérons donc un test de de Wilcoxon (rang signé) par paires :

- QCM1 vs. QCM2 : W=-105, Z=-3.28 et p=0.0005.
- QCM2 vs. QCM3 : W=-103, Z=-3.22 et p=0.0006.
- (QCM1 vs. QCM3 n'a pas d'intérêt pour nous)

Les résultats des tests de Wilcoxon réalisés par paires nous indiquent qu'il existe une différence significative entre les trois réponses données aux QCM pour les participants ayant utilisé le PCP : QCM1<QCM2<QCM3.

3.5.1.1.3 SPM

Les moyennes de rang sont égales à 1 pour le QCM1, 2.2 pour le QCM2 et 2.8 pour le QCM3. Le test nous donne csqr=21.09, df=2 et p<0.0001.

Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence entre les trois mesures. Nous opérons donc un test de de Wilcoxon (rang signé) par paires :

- QCM1 vs. QCM2 : W=-107.5, Z=-3.36 et p=0.0004.
- QCM2 vs. QCM3 : W=-53, Z=-1.83 et p=0.0336.
- (QCM1 vs. QCM3 n'a pas d'intérêt pour nous)

Les résultats des tests de Wilcoxon réalisés par paires nous indiquent qu'il existe une différence significative entre les 3 réponses données aux QCM pour les participants ayant utilisé le SPM : QCM1<QCM2<QCM3.

Les résultats pour les trois graphiques nous permettent de conclure que les participants ont tous un profil de novice au début du test et qu'ils ont tous gagné en connaissance (pour l'exploration d'espace de conception).

Il convient maintenant de vérifier s'il existe une différence d'évolution de cette connaissance entre les trois groupes (i.e. les trois graphiques).

3.5.1.2 Analyse Inter-Graphiques

Pour cette analyse nous utilisons le test de Kruskal-Wallis car la variable est qualitative ordinale et que l'étude est composée de trois mesures indépendantes (n=14 pour chaque mesure). Notons qu'un test de Kruskal-Wallis donne comme résultat la valeur du test notée H, le degré de liberté noté df et la valeur de p notée p.

Les analyses sont réalisées pour les trois QCM indépendamment. Ainsi, nous posons les hypothèses statistiques pour chacun des QCM :

- H0 : il n'existe pas de différence de réponses aux Q.C.M entre les trois graphiques.
- H1 : il existe une différence de réponses données aux Q.C.M. entre les trois graphiques (SPM<>PCP<>SSP)

3.5.1.2.1 QCM1

Les moyennes de rang sont égales à 23.1 pour le SPM, 19.5 pour le PCP et 21.9 pour le SSP. Le test nous donne H=0.61, df=2 et p=0.7371.

Nous devons donc rejeter l'hypothèse alternative. Les résultats nous indiquent qu'il n'existe pas de différence significative entre les trois graphiques pour les réponses données au QCM1.

3.5.1.2.2 QCM2

Les moyennes de rang sont égales à 22 pour le SPM, 18.4 pour le PCP et 24.1 pour le SSP. Le test nous donne H=1.55, df=2 et p=0.4607.

Nous devons donc rejeter l'hypothèse alternative. Les résultats nous indiquent qu'il n'existe pas de différence significative entre les trois graphiques pour les réponses données au QCM2.

3.5.1.2.3 QCM3

Les moyennes de rang sont égales à 16.8 pour le SPM, 26.3 pour le PCP et 21.5 pour le SSP. Le test nous donne H=4.2, df=2 et p=0.12.

Nous devons donc rejeter l'hypothèse alternative. Les résultats nous indiquent qu'il n'existe pas de différence significative entre les trois graphiques pour les réponses données au QCM3.

En conclusion, il n'y a pas de différence de réponses aux trois QCM entre les trois graphiques. De plus, le niveau de connaissance en exploration d'espace de conception a évolué (entre les QCM1, QCM2 et QCM3) pour les participants des trois groupes (i.e. les trois graphiques). **La phase de formation de l'expérimentation a été bénéfique pour les trois graphiques et, dans les trois cas, l'utilisation du graphique pour la résolution des problèmes de conception a contribué à l'évolution du niveau de connaissance des participants (voir Figure 109).**

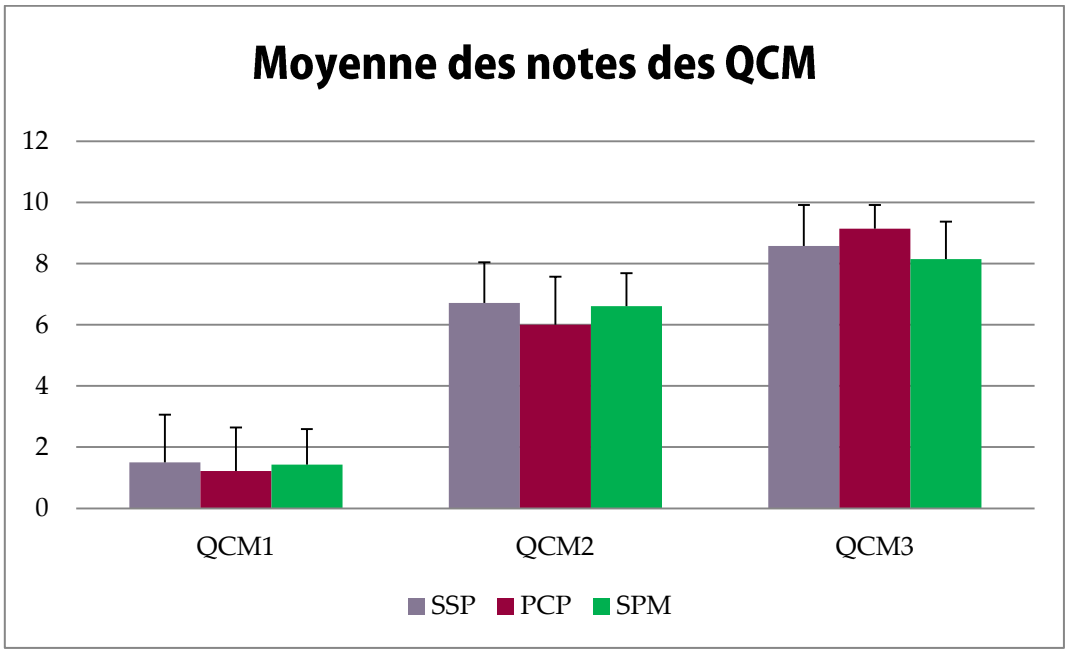


Figure 109 : Moyennes des notes obtenues aux QCM et écart-type

3.5.2 Confiance du Participant

Pour cette analyse nous utilisons le test de Kruskal-Wallis car la variable est qualitative ordinale (réponse sur une échelle) et que l'étude est composée de trois mesures indépendantes (i.e. les trois graphiques).

Les analyses sont réalisées pour les trois questions de confiance indépendamment. Ainsi, nous posons les hypothèses statistiques pour chacune des questions :

- H0 : il n'existe pas de différence de réponses aux questions de confiance entre les trois graphiques.
- H1 : il existe une différence de réponses données aux questions de confiance entre les trois graphiques (SPM<>PCP<>SSP)

3.5.2.1 Problème Vessel

Pour ce problème nous avons des tailles d'échantillon telles que n(SPM)=13, n(PCP)=12 et n(SSP)=13. Nous proposons ici de présenter les résultats des trois tests effectués dans le **Tableau 12**. Notons que dans les trois cas, le degré de liberté df=2. Nous observons que pour les trois questions de confiance, il n'existe pas de différence entre les trois graphiques.

En effet, nous obtenons, pour la question 1, une valeur H=3.3 et p=0.192 ; pour la question 2, une valeur de H=0.17 et p=0.9185 ; et pour la question 3, une valeur de H=0.42 et p=0.8106. Nous devons donc rejeter l'hypothèse alternative dans les trois cas et inférer qu'il n'existe pas de différence.

Tableau 12 : Résultats des tests de Kruskal-Wallis pour les réponses aux trois questions de confiance (Vessel)

	Graphique	Moyenne de rang	Valeur du H	Valeur de p
Question 1	SPM	24	3.3	0.192
	PCP	16.9		
	SSP	17.3		
Question 2	SPM	19.1	0.17	0.9185
	PCP	20.6		
	SSP	18.9		
Question 3	SPM	21.1	0.42	0.8106
	PCP	18.5		
	SSP	18.8		

3.5.2.2 Problème Disk

Pour ce problème nous avons des tailles d'échantillon telles que n(SPM)=13, n(PCP)=14 et n(SSP)=13. De la même manière que pour le problème Vessel, nous proposons de présenter les résultats des trois tests effectués dans le **Tableau 13**. Notons que dans les trois cas, le degré de liberté df=2. Nous observons que pour les trois questions de confiance, il n'existe pas de différence entre les trois graphiques.

En effet, nous obtenons, pour la question 1, une valeur H=1.38 et p=0.5016; pour la question 2, une valeur de H=3.02 et p=0.2209 ; et pour la question 3, une valeur de H=0.6 et p=0.7408. Nous devons donc rejeter l'hypothèse alternative dans les trois cas et inférer qu'il n'existe pas de différence.

Tableau 13 : Résultats des tests de Kruskal-Wallis pour les réponses aux trois questions de confiance (Disk)

	Graphique	Moyenne de rang	Valeur du H	Valeur de p
Question 1	SPM	23.2	1.38	0.5016
	PCP	20.5		
	SSP	17.8		
Question 2	SPM	23.1	3.02	0.2209
	PCP	2.4		
	SSP	15.9		
Question 3	SPM	19.1	0.6	0.7408
	PCP	22.4		
	SSP	19.9		

3.5.3 Conclusion pour les données des participants

Les résultats obtenus pour les données des participants nous permettent de conclure que les participants ont tous un profil de concepteur novice au début du test et qu'ils ont tous gagné en connaissance pour l'exploration d'espace de conception (de manière significative selon les analyses statistiques).

Comme nous montrons qu'il n'existe pas de différence de réponses aux trois QCM entre les trois graphiques, les résultats nous permettent d'affirmer que le niveau de connaissance en exploration d'espace de conception a évolué (entre les QCM1, QCM2 et QCM3) pour les participants des trois groupes (i.e. les trois graphiques). La phase de formation de l'expérimentation a été bénéfique pour les trois graphiques et, dans les trois cas, l'utilisation du graphique pour la résolution des problèmes de conception a contribué à faire évoluer le niveau de connaissance des participants.

Nous pouvons donc nous affranchir de diviser les groupes par niveaux de connaissance pour la suite des analyses.

Enfin, il n'existe pas de différence entre les trois graphiques pour les réponses aux questions nous permettant de mesurer la confiance du sujet. Nos résultats indiquent que l'utilisation de chacun des trois graphiques n'influence pas la confiance que porte le participant à sa décision.

Les trois graphiques sont donc potentiellement des graphiques que nous pouvons identifier sans craintes comme performants pour une décision informée (telle que nous l'avons

définie). En effet, nous ne pouvons pas aboutir à un résultat où un graphique non pertinent induit une grande confiance chez le participant (et inversement).

Comme précisé dans l'introduction de la troisième partie, nous poursuivons la présentation de l'expérimentation n°2 en quatre sous-parties :

- Expérimentation 2.1 qui traite la phase de sélection
- Expérimentation 2.2 qui traite la phase d'optimisation
- Expérimentation 2.3 qui traite la phase de découverte
- Expérimentation 2.4 qui traite la décision informée

Pour chacune d'entre elles, nous décrivons les mesures utilisées et le traitement des résultats. Nous proposons une conclusion intermédiaire à la fin de chaque sous-partie.

Pour l'ensemble des analyses statistiques, nous choisissons de prendre un risque de première espèce à 10%.

3.6 EXPERIMENTATION 2.1 : PHASE SELECTION

3.6.1 Mesures liées à la phase de sélection

Dans cette partie nous définissons les trois indicateurs utilisés dans la phase de sélection.

Solution « standardisable »

Pour chacun des deux problèmes de conception, il y avait deux paramètres de conception « standardisables ». Nous appelons paramètre « standardisable » un paramètre pour lequel les solutions dites Pareto-Optimales tendent vers une valeur fixe de ce paramètre. Ainsi, pour cet indicateur nous relevons la recherche / sélection (ou non) d'une solution comportant les deux paramètres « standardisables ».

Pour le problème Vessel, le rayon (R) et l'épaisseur (T) sont standardisables. Et pour le problème Disk, la force d'action (F) et l'épaisseur des disques (t) sont standardisables.

Le relevé a été effectué par analyse vidéo de l'écran combiné avec l'eye-tracking.

Solution Non-dominée au sens de Pareto

Nous relevons pour cet indicateur la sélection d'une solution non-dominée au sens de Pareto dans l'ensemble final des solutions possibles générées par le participant.

Qualité de la solution sélectionnée

Enfin, pour la phase de sélection, nous mesurons la qualité de la solution sélectionnée à l'aide de l'Indice Basique dans l'ensemble final de solutions disponibles. Il est important de noter que nous mesurons la qualité de la solution sélectionnée dans cet espace final et donc en tenant compte des bornes qualités (défini pour la phase d'optimisation).

3.6.2 Traitement des résultats 2.1

Pour l'analyse des données de cette partie, nous posons les hypothèses statistiques globales suivantes :

- H0 : il n'existe pas de différence en les trois groupes
- H1 : il existe une différence entre les trois groupes (le PCP est le graphique le plus performant pour la sélection)

3.6.2.1 Solution « standardisable »

Pour cet indicateur nous analysons les données des deux problèmes conjointement. Nous pouvons, en effet, analyser les données ensembles puisqu'ici nous relevons simplement si le participant a sélectionné une solution standardisable ou non. Ainsi, notre étude comporte deux variables qualitatives : le groupe (SPM, PCP et SSP) et le standard (oui ou non). Nous appliquons un test du Khi2 (le tableau de contingence est présenté **Tableau 14**).

Tableau 14 : Tableau de contingence des variables groupes et standard

	Standard	Non standard	Total
SPM	7	19	26
PCP	16	10	26
SSP	8	18	26
Total (freq)	31 (0.40)	47 (0.60)	78 (1)

Nous réalisons donc le test du Khi2 pour répondre à nos hypothèses statistiques: X^2=7.816, ddl=2 et p=0.02. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite des tests du Khi2 par paires :

- SSP vs PCP : X^2=4.952, ddl=1 et p=0.026
- SSP vs SPM : X^2=0.094, ddl=1 et p=0.759
- PCP vs SPM : X^2=6.315, ddl=1 et p=0.012

Il existe une différence significative entre les graphiques PCP et SSP ; et SPM et PCP. **Nous pouvons conclure que le PCP est le graphique avec lequel les participants ont sélectionné le plus de solutions standardisables**.

3.6.2.2 Solution Non-dominée au sens de Pareto

Pour cet indicateur nous analysons les données des deux problèmes conjointement. Nous pouvons, en effet, analyser les données ensembles puisqu'ici nous relevons simplement si le participant a sélectionné une solution dominée ou non-dominée. Ainsi, notre étude comporte deux variables qualitatives : le groupe (SPM, PCP et SSP) et le « Pareto » (oui ou non).

Tableau 15 : Tableau de contingence des variables groupes et « Pareto »

	Non-dominée	Dominée	Total
SPM (freq attendue)	23 (24.7)	3 (1.3)	26
PCP (freq attendue)	26 (24.7)	0 (1.3)	26
SSP (freq attendue)	25 (24.7)	1 (1.3)	26
Total (freq)	74 (0.95)	4 (0.05)	78 (1)

Nous observons dans le tableau de contingence (**Tableau 15**) que plus de 20% de nos résultats (la moitié) ont une fréquence attendue inférieure à 5. Nous réalisons donc le test du Khi2 avec la correction de Yates pour répondre à nos hypothèses statistiques: X^2=1.647, ddl=2 et p=0.44. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes.

Bien que **nous ne pouvons pas inférer une différence entre les trois graphiques pour la sélection d'une solution non-dominée**, il est important de noter qu'aucunes solutions dominées n'a été sélectionnées avec le graphique PCP (3/26 pour le SPM et 1/26 pour le SSP).

3.6.2.3 Qualité de la solution sélectionnée

3.6.2.3.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 0.58 pour le SSP (n=13), 0.73 pour le PCP (n=12) et 0.80 pour le SPM (n=13).

Nous réalisons une analyse de variance (ANOVA-*between*) pour répondre à nos hypothèses statistiques car notre étude comporte une variable quantitative et trois groupes indépendants: F(2,35)= 2.46 et p= 0.10. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : t(23)= 1.3 p= 0.1
- SSP vs SPM : t(24)= 2.12 p = 0.02
- PCP vs SPM : t(23)= 0.86 p = 0.20

Il existe une différence significative entre les graphiques SPM et SSP ; et SSP et PCP. Nous pouvons conclure que pour le problème Vessel, **le SSP est le graphique avec lequel les participants ont sélectionné des solutions de moins bonnes qualités (voir Figure 110)**.

3.6.2.3.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 0.74 pour le SSP (n=13), 0.79 pour le PCP (n=14) et 0.77 pour le SPM (n=13).

Nous réalisons une analyse de variance (ANOVA-between) pour répondre à nos hypothèses statistiques car notre étude comporte une variable quantitative et trois groupes indépendants: F(2,37)= 0.12 et p= 0.89. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes (voir **Figure 110**).

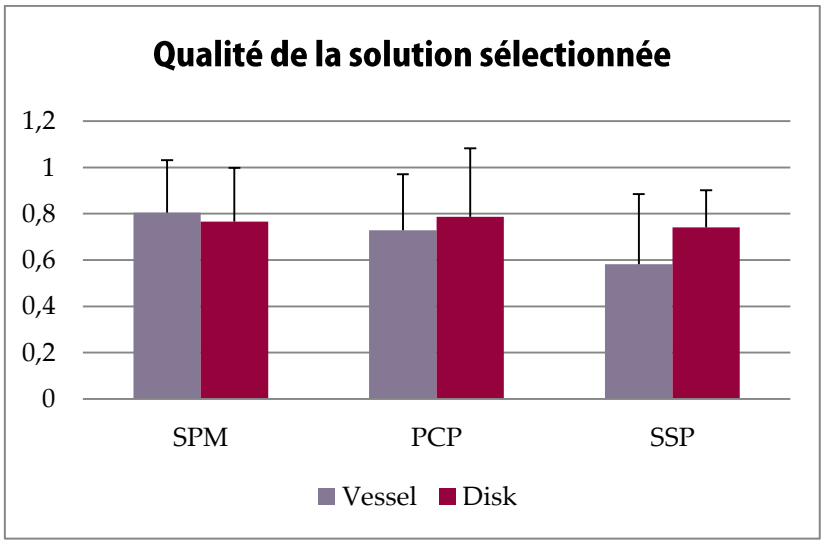


Figure 110 : Moyennes et écart-types des résultats pour la qualité de la solution sélectionnée

3.6.3 Conclusion pour la phase de sélection

Le Parallel Coordinates Plot (PCP) est le graphique le plus pertinent pour la phase de sélection. En acceptant un risque de première espèce α à 10%, le PCP est le graphique avec lequel les participants ont sélectionné le plus de solutions standardisables.

Nous devons également indiquer que le Simple Scatter Plot (SSP) est le graphique montrant les moins bons résultats pour la phase de sélection (voir le **Tableau 16**). Nous déconseillons donc l'utilisation de ce graphique pour la sélection car il est le graphique avec lequel les participants ont sélectionné des solutions de moins bonne qualité.

Tableau 16 : Synthèse des résultats pour la phase de Sélection

<table>
<tr><td></td><td>Standard</td><td>Pareto</td><td colspan="2">Qualité</td></tr>
<tr><td></td><td>2 problèmes</td><td>2 problèmes</td><td>Vessel</td><td>Disk</td></tr>
<tr><td>Meilleur</td><td>PCP</td><td rowspan="3">PCP≡SPM≡SSP</td><td rowspan="2">PCP≡SPM</td><td rowspan="3">PCP≡SPM≡SSP</td></tr>
<tr><td>Intermédiaire</td><td rowspan="2">SPM≡SSP</td></tr>
<tr><td>Pire</td><td>SSP</td></tr>
</table>

Nous démontrons donc que le Parallel Coordinates Plot est le graphique le plus pertinent pour la sélection en exploration d'espace de conception que le processus soit multi-attribut ou multi-objectif.

3.7 EXPERIMENTATION 2.2 : PHASE OPTIMISATION

3.7.1 Mesures liées à la phase d'optimisation

Dans cette partie, nous utilisons l'Indice Basique (déjà défini auparavant) pour mesurer la qualité des solutions dans l'espace de conception. Lors de l'expérimentation, les participants avaient à disposition une fonction permettant de colorer les solutions possibles en fonction de l'Indice Basique. Il convient donc d'utiliser le même indice pour analyser les données.

Nous définissons les bornes qualités de l'espace de conception dans le temps par l'indice de qualité de la meilleure solution possible (i_b) de l'espace et l'indice de qualité de la pire solution de l'espace (i_w). Les données sont temporelles car à chaque génération de points de conception ou à chaque réduction de l'espace, les données sont enregistrées (t=0 correspond à la première génération de points de conception).

Indicateur « sens »

Nous relevons ici une erreur (ou non) durant l'optimisation (voir une illustration en **Figure 111 b. et c**.).

Si $i_b(t_{fin}) < i_b(t)$ ou $i_w(t_{fin}) < i_w(t)$ alors le participant à fait une erreur . L'objectif est de réduire et donc d'avancer dans l'espace de conception vers une zone plus performante.

Indicateur « différence de l'étendue de qualité »

Considérons à l'instant t, $Ei = i_b - i_w$ soit l'étendue de qualité que couvre l'ensemble des solutions.

Si $Ei(t=0) > Ei(t_{fin})$ alors l'étendue de qualité que couvre l'ensemble des solutions a été réduite (sensé être vers une zone performante de l'espace de conception).

Par contre si $Ei(t=0) < Ei(t_{fin})$ alors l'étendue de qualité s'est agrandie et l'optimisation n'est pas optimale (l'étendue de qualité des solutions possibles diverge). Evidement l'étendue peut diverger alors que la qualité augmente (entre t_0 et t_{fin}). Cela veut dire que le participant a peuplé l'espace de conception avec des solutions plus performantes sans écarter (i.e. réduire l'espace) les solutions moins performantes (i.e. des solutions qui ne sont pas de bonnes réponses au problème).

Une illustration est proposée en **Figure 111 d. et e**.

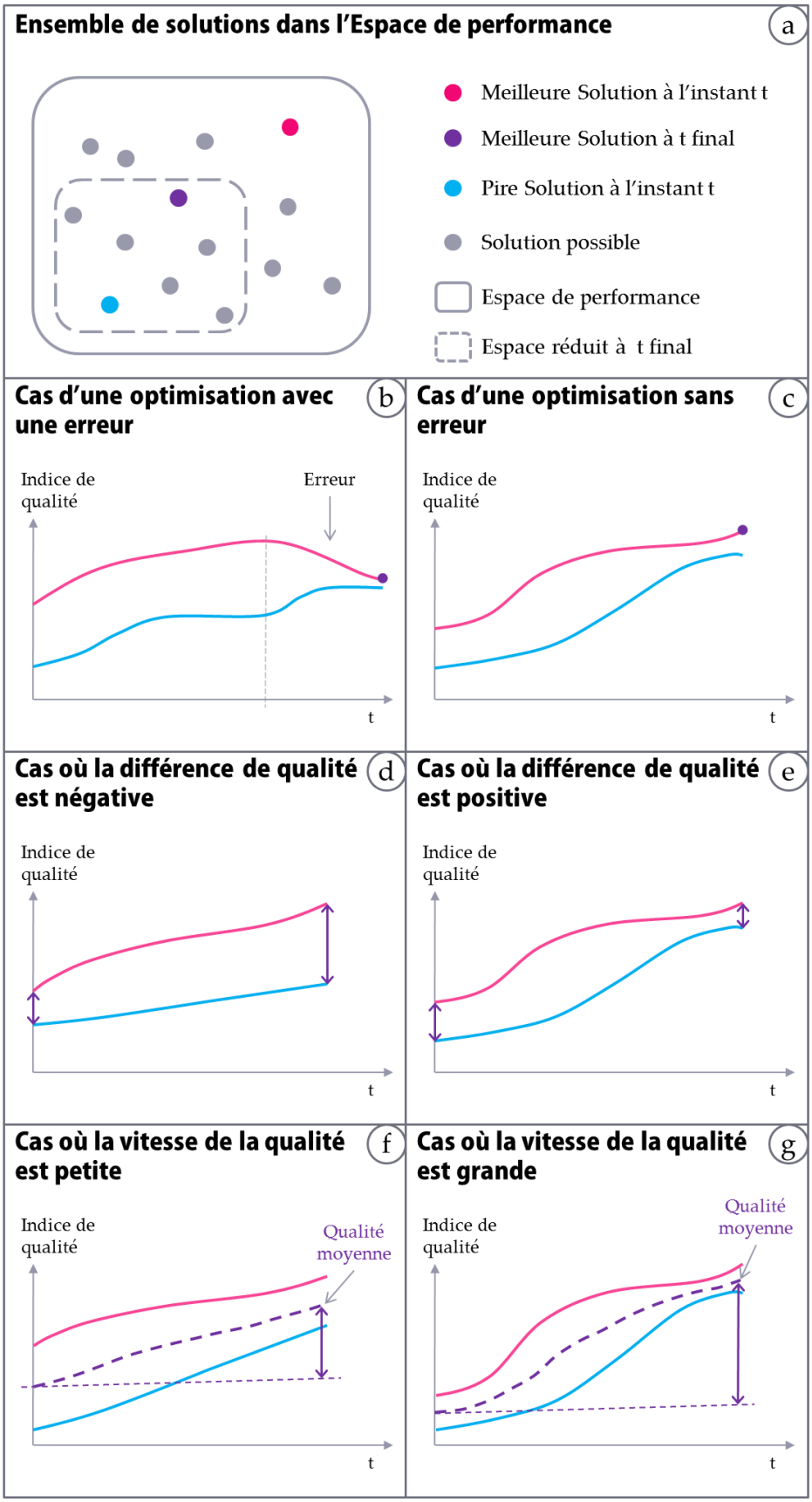


Figure 111 : Illustrations des indicateurs (b et c) : « sens », (d et e) : différence de l'étendue de qualité et (f et g) : différence de la qualité moyenne

Indicateur « différence de la qualité moyenne »

Nous mesurons ici la différence de la qualité moyenne de l'espace de conception. Pour cela nous calculons premièrement la qualité moyenne de l'espace (voir une illustration en **Figure 111 f. et g.**) :

$$i_{moy} = \frac{i_b + i_w}{2}$$

Puis la différence de qualité :

$$d_i = \frac{i_{moy}(t_{fin}) - i_{moy}(t_0)}{t_{fin}}$$

Ainsi, plus le participant peuple l'espace de bonnes solutions plus la différence de qualité (de son ensemble de solutions) entre t_0 et t_{fin} sera grande.

Indicateur « convergence »

Enfin, nous prenons comme indicateur pour la phase d'optimisation la convergence. Pour cela nous considérons conjointement l'étendue de la qualité (à chaque instant) et la taille de l'espace en pourcentage :

$$C = (i_b - i_w) * S$$

Où S est la taille de l'espace de conception à l'instant t. Notons qu'à t=0, S=100%.

Puis nous calculons la convergence telle que :

$$d_c = \frac{C(t_0) - C(t_{fin})}{t_{fin}}$$

Ainsi, plus le participant converge rapidement (i.e. un espace de plus en plus réduit ayant un une étendue de la qualité de plus en plus petite) et meilleur est son indicateur de « convergence ».

3.7.2 Traitement des résultats 2.2

Pour l'analyse des données de cette partie, nous posons les hypothèses statistiques globales suivantes :

- H0 : il n'existe pas de différence en les trois groupes
- H1 : il existe une différence entre les trois groupes (le SSP est le graphique le plus performant pour l'optimisation)

3.7.2.1 Le « sens »

Pour cet indicateur nous analysons les données des deux problèmes conjointement. Nous pouvons, en effet, analyser les données ensembles puisqu'ici nous relevons simplement si le

participant a fait une erreur ou non. Ainsi, notre étude comporte deux variables qualitatives : le groupe (SPM, PCP et SSP) et l'erreur (oui ou non). Nous appliquons un test du Khi2 (le tableau de contingence est présenté **Tableau 17**).

Tableau 17 : Tableau de contingence des variables groupes et erreur

	Pas d'erreurs	Avec erreurs	Total
SPM	30	22	52
PCP	42	10	52
SSP	42	10	52
Total (freq)	114 (0.73)	42 (0.27)	156 (1)

Nous réalisons donc le test du Khi2 pour répondre à nos hypothèses statistiques: X^2=9.383, ddl=2 et p=0.009. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite des tests du Khi2 par paires :

- SSP vs PCP : X^2=0, ddl=1 et p=1
- SSP vs SPM : X^2=6.5, ddl=1 et p=0.01
- PCP vs SPM : X^2=6.5, ddl=1 et p=0.01

Il existe une différence significative entre les graphiques SPM et SSP ; et SPM et PCP. **Nous pouvons conclure que le SPM est le graphique avec lequel les participants ont commis le plus d'erreurs**. (Le même nombre d'erreurs a été commis avec SSP et le PCP).

Pour la suite, les variables sont quantitatives et nous avons trois groupes indépendants. Nous poursuivons donc avec l'utilisation d'ANOVA-between.

3.7.2.2 Différence de l'étendue de la qualité

3.7.2.2.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 0.46 pour le SSP (n=13), 0.42 pour le PCP (n=12) et 0.45 pour le SPM (n=13). L'ANOVA nous donne comme résultats : F(2,35)= 0.05 et p= 0.95. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes.

3.7.2.2.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 0.55 pour le SSP (n=13), 0.45 pour le PCP (n=14) et 0.55 pour le SPM (n=13). L'ANOVA nous donne comme résultats : F(2,37)= 0.75 et p= 0.48. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes.

Nous devons conclure qu'il n'existe pas de différence entre les trois groupes pour la différence entre les bornes qualité en entrée et sortie. **La réduction de l'étendue de la qualité est similaire avec les trois graphiques.**

3.7.2.3 Différence de la qualité moyenne

3.7.2.3.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 4.3 pour le SSP (n=13), 4.2 pour le PCP (n=12) et 3.6 pour le SPM (n=13). L'ANOVA nous donne comme résultats : $F(2,35)= 0.16$ et $p= 0.83$. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes.

3.7.2.3.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 6.0 pour le SSP (n=13), 4.6 pour le PCP (n=14) et 5.6 pour le SPM (n=13). L'ANOVA nous donne comme résultats : $F(2,37)= 0.62$ et $p= 0.54$. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes.

Nous devons conclure qu'il n'existe pas de différence entre les 3 groupes pour la différence de la qualité moyenne de l'espace de conception.

3.7.2.4 La vitesse de convergence

3.7.2.4.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 5.2 pour le SSP (n=13), 7.4 pour le PCP (n=12) et 5.9 pour le SPM (n=13). L'ANOVA nous donne comme résultats : $F(2,35)= 0.54$ et $p= 0.58$. Nous devons rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les trois groupes.

3.7.2.4.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 12.5 pour le SSP (n=13), 7.5 pour le PCP (n=14) et 9.5 pour le SPM (n=13). L'ANOVA nous donne comme résultats : $F(2,37)= 3.02$ et $p= 0.06$. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : $t(25)= 2.38$ $p= 0.01$
- SSP vs SPM : $t(24)= 1.35$ $p = 0.09$
- PCP vs SPM : $t(25)= 1.08$ $p = 0.14$

Il existe une différence significative entre les graphiques SPM et SSP ; et SSP et PCP. Nous pouvons conclure que pour le problème Disk, **le SSP est le graphique avec lequel les participants ont mieux convergé vers une solution**.

3.7.3 Conclusion pour la phase d'optimisation

Le Simple Scatter Plot (SSP) est le graphique le plus pertinent pour la phase d'optimisation. En acceptant un risque de première espèce α à 10%, le SSP est le graphique permettant aux participants de converger le plus rapidement vers une « zone » de performance.

Nous devons également indiquer que le Scatter Plot Matrix (SPM) est le graphique montrant les moins bons résultats pour la phase d'optimisation (voir le **Tableau 18**). Nous déconseillons donc l'utilisation de ce graphique pour l'optimisation car il est le graphique avec lequel les participants ont commis le plus d'erreurs.

Tableau 18 : Synthèse des résultats pour la phase d'optimisation

<table>
<tr><td></td><td>« Sens »</td><td>Différence de l'étendue de qualité</td><td>Différence de la qualité moyenne</td><td colspan="2">Convergence</td></tr>
<tr><td></td><td>2 problèmes</td><td>2 problèmes</td><td>2 problèmes</td><td>Vessel</td><td>Disk</td></tr>
<tr><td>Meilleur</td><td rowspan="2">SSP≡PCP</td><td rowspan="3">SSP≡PCP≡SPM</td><td rowspan="3">SSP≡PCP≡SPM</td><td rowspan="3">SSP≡PCP≡SPM</td><td>SSP</td></tr>
<tr><td>Intermédiaire</td><td rowspan="2">SPM≡PCP</td></tr>
<tr><td>Pire</td><td>SPM</td></tr>
</table>

3.8 EXPERIMENTATION 2.3 : PHASE DECOUVERTE

3.8.1 Mesures liées à la phase de découverte

Dans cette partie nous définissons les trois indicateurs utilisés dans la phase de découverte.

Nombre de découvertes réalisées

Nous mesurons avec cet indicateur le nombre de découvertes réalisées par le participant sachant que pour chaque problème il y avait un maximum de sept découvertes à réaliser. L'objectif pour le participant est de réaliser l'ensemble des découvertes (ne sachant pas qu'il y en a sept).

Pour le problème Vessel, nous avons trois découvertes au niveau global et quatre découvertes au niveau local (i.e. sur les solutions Pareto-Optimales) :

- Découvertes d'un point de vu global :
  - Corrélation « grossière » positive entre R et V
  - Corrélation « grossière » positive entre R et T
  - Corrélation « grossière » positive entre T et V
- Découvertes d'un point de vu local :
  - Corrélation positive entre W et L
  - Corrélation positive entre V et L
  - Les solutions tendent vers T=4
  - Les solutions tendent vers R=36

Notons qu'en réalisant toutes ces découvertes, il est aisé d'identifier la variable de décision (L : la longueur de la cuve).

Pour le problème Disk, nous avons une découverte au niveau global et six découvertes au niveau local (i.e. sur les solutions Pareto-Optimales) :

- Découverte d'un point de vu global :
  - Corrélation « grossière » négative entre S et Z
- Découvertes d'un point de vu local :
  - Corrélation « grossière » positive entre Ri & Ro
  - Un point de transition pour Ro=100
  - Un point de transition pour M=1
  - Les solutions tendent vers t=1
  - Les solutions tendent vers F=1000
  - Corrélation positive entre M et Z

Notons qu'en réalisant toutes ces découvertes, il est aisé d'identifier la variable de décision (Z : le nombre de disque).

Le relevé du nombre de découvertes a été effectué par analyse vidéo de l'écran combiné à l'eye-tracking. Nous proposons d'illustrer des découvertes réalisées avec le SSP (**Figure 112**), le PCP (**Figure 113**) et le SPM (**Figure 114**).

Sur la **Figure 112**, nous pouvons remarquer que le participant observe une corrélation négative entre les variables S et Z dans le problème Disk. Les ronds jaunes numérotés sont l'illustration du parcours de son regard sur l'interface. La corrélation est observée (ronds n°3, 4 et 5) et le nom des variables est vérifié (ronds n°2, 6 et 7).

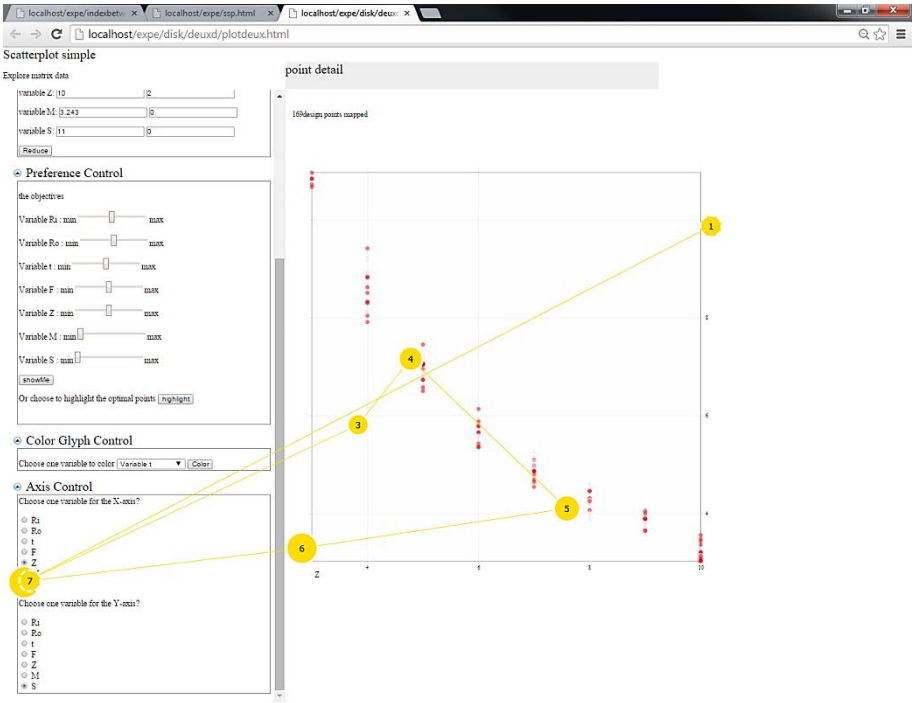


Figure 112 : Export de l'eye-tracking du participant n°11 pour le problème Disk

Sur la **Figure 113**, nous pouvons remarquer que le participant observe une corrélation « grossière » positive entre les variables R et V dans le problème Vessel. Les ronds bleus numérotés sont l'illustration du parcours de son regard sur l'interface.

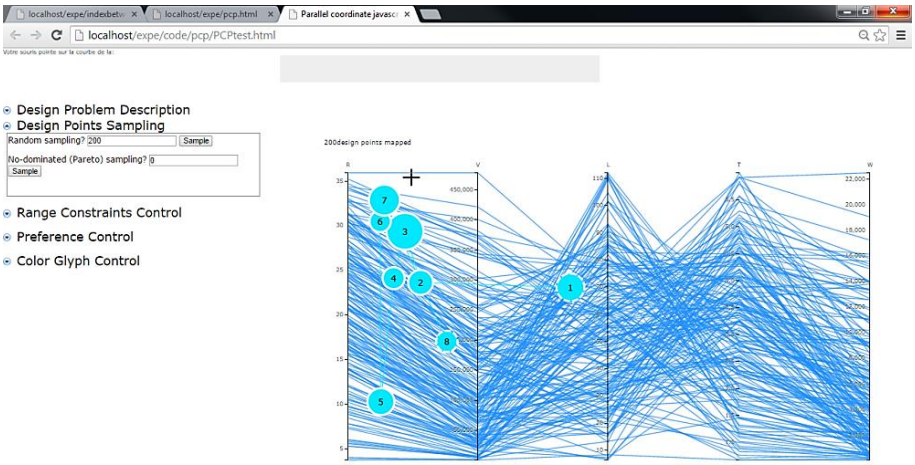


Figure 113 : Export de l'eye-tracking du participant n°9 pour le problème Vessel

Sur la **Figure 114**, nous pouvons remarquer que le participant observe un paramètre standardisable dans le problème Disk : les solutions Pareto-Optimales ont des valeurs qui tendent vers 1000 pour la variable F. Les ronds oranges numérotés sont l'illustration du parcours de son regard sur l'interface.

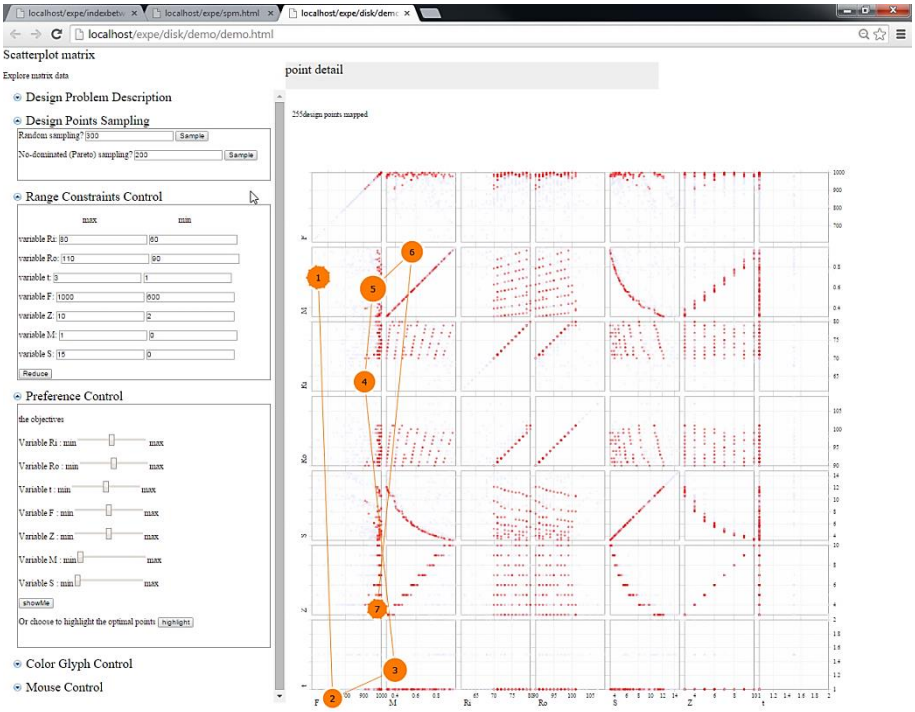


Figure 114 : Export de l'eye-tracking du participant n°26 pour le problème Disk

Le temps avant la première découverte

Pour cet indicateur, nous mesurons le temps écoulé avant que le participant réalise sa première découverte. Ce temps ne dépend pas du temps écoulé pour s'approprier le problème car le chronomètre se déclenche lorsque le participant a lu la description du problème et qu'il effectue une première génération de points de conception.

Cette mesure est, pour nous, un premier indice sur la rapidité de découvertes réalisées avec les trois graphiques.

Temps moyen par découverte

Enfin, nous calculons ici le temps moyen pour réaliser une découverte i.e. le temps total pour effectuer toutes les découvertes divisé par le nombre de découvertes réalisées. Nous ne pouvons pas simplement mesurer le temps total que le participant a mis pour effectuer toutes les observations (i.e. découvertes) puisque tous les participants n'ont pas effectué toutes les observations. Nous préférons donc mesurer un temps moyen « nécessaire » pour réaliser une découverte.

Cette mesure est, pour nous, un deuxième indice sur la rapidité de découvertes réalisées avec les trois graphiques.

3.8.2 Traitement des résultats 2.3

Pour l'analyse des données de cette partie, nous posons les hypothèses statistiques globales suivantes :

- H0 : il n'existe pas de différence en les trois groupes
- H1 : il existe une différence entre les trois groupes (le SPM est le graphique le plus performant pour la découverte)

Aussi, les variables sont quantitatives et nous avons trois groupes indépendants. Nous réalisons donc l'analyse statistique avec l'utilisation d'ANOVA-*between*.

3.8.2.1 Nombre de découvertes réalisées

3.8.2.1.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 2.66 découvertes pour le SSP (n=12), 3.25 découvertes pour le PCP (n=12) et 6.3 découvertes pour le SPM (n=13). Le nombre total de découvertes possibles est de 7. L'ANOVA nous donne comme résultats : $F(2,34)= 24.27$ et $p<0.0001$. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : $t(22)=-0.92$ $p=0.18$
- SSP vs SPM : $t(23)=6.62$ $p<0.0001$
- PCP vs SPM : $t(22)=5.86$ $p<0.0001$

Il existe une différence significative entre les graphiques SPM et SSP ; et SPM et PCP. Nous pouvons conclure que **pour le problème Vessel, le SPM permet de réaliser un plus grand nombre de découvertes.**

3.8.2.1.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 2.92 découvertes pour le SSP (n=13), 3 découvertes pour le PCP (n=14) et 5.4 découvertes pour le SPM (n=13). Le nombre total de découvertes possibles est de 7. L'ANOVA nous donne comme résultats : F(2,37)= 17.01 et p<0.0001. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : t(25)= 0.16 p= 0.4370835
- SSP vs SPM : t(24)= 4.99 p <.0001
- PCP vs SPM : t(25)= 5.08 p <.0001

Il existe une différence significative entre les graphiques SPM et SSP ; et SPM et PCP.

Nous pouvons conclure que le SPM est le graphique qui permet de réaliser un plus grand nombre de découvertes (voir Figure 115).

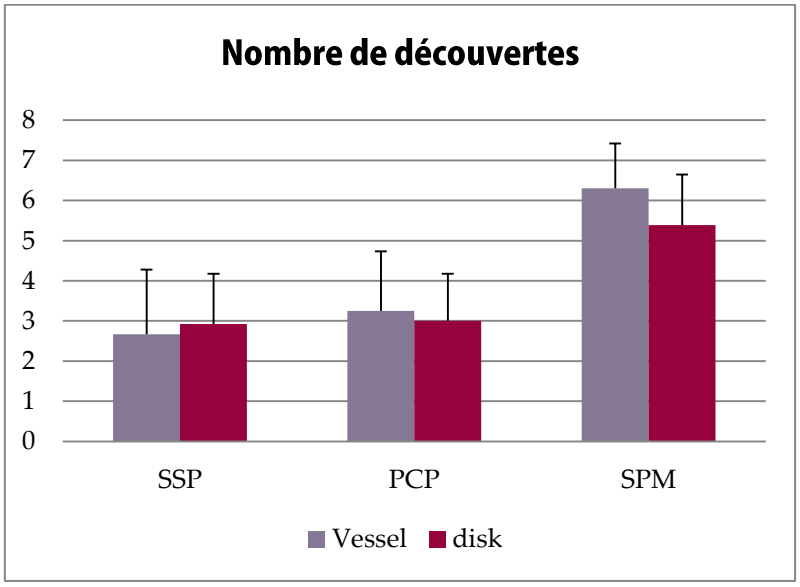


Figure 115 : Moyennes et écart-types des résultats pour le nombre de découvertes réalisées

3.8.2.2 Temps pour la première découverte

3.8.2.2.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 209.6 secondes écoulées avant la première découverte pour le SSP (n=12), 115.8 secondes écoulées avant la première découverte pour le PCP (n=12) et 14.8 secondes écoulées avant la première découverte pour le SPM (n=13). L'ANOVA nous donne comme résultats : F(2,33)= 8.65 et p= 0.000954. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de student) par paires :

- SSP vs PCP : t(21)=-1.57 p= 0.0656805

- SSP vs SPM : t(22)= 3.69 p= 0.0006405
- PCP vs SPM : t(23)= 4.55 p<0.0001

Il existe une différence significative entre les trois graphiques. Nous pouvons conclure que **pour le problème Vessel, les participants ont mis moins de temps pour réaliser une première découverte avec le SPM et plus de temps avec le SSP**.

3.8.2.2.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 175.2 secondes écoulées avant la première découverte pour le SSP (n=13), 107.2 secondes écoulées avant la première découverte pour le PCP (n=14) et 52.4 secondes écoulées avant la première découverte pour le SPM (n=13). L'ANOVA nous donne comme résultats : F(2,37)= 9.17 et p= 0.000583. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : t(25)= 2.39 p= 0.0123525
- SSP vs SPM : t(24)= 4.44 p <.0001
- PCP vs SPM : t(25)= 1.89 p = 0.035203

Il existe une différence significative entre les trois graphiques SPM, SSP et PCP. Nous pouvons conclure que **pour le problème Disk, les participants ont mis moins de temps pour réaliser une première découverte avec le SPM et que le SSP est le graphique avec lequel les participants ont mis le plus de temps pour réaliser une première découverte.**

Nous pouvons conclure que le graphique SPM est le graphique avec lequel les participants ont réalisé une première découverte avec un temps le plus court et avec le SSP le temps le plus long (voir Figure 116).

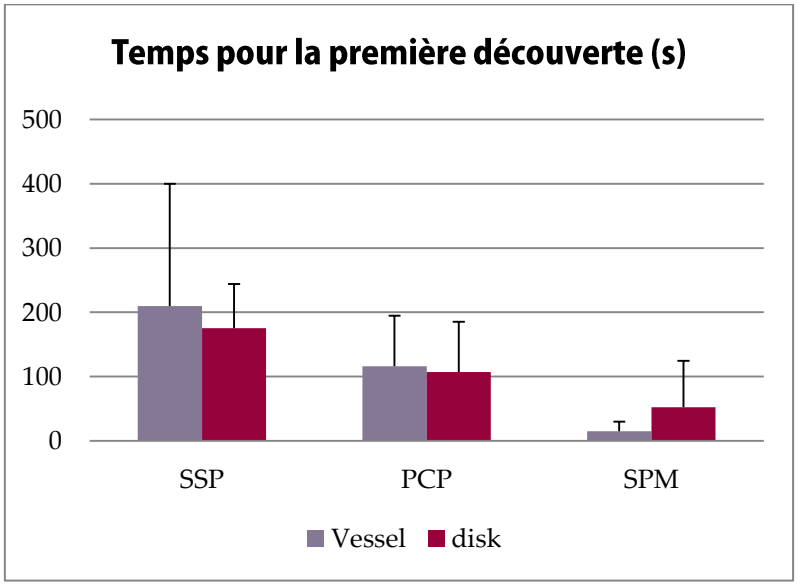


Figure 116 : Moyennes et écart-types des résultats pour le temps pour la première découverte

3.8.2.3 Temps moyen pour réaliser une découverte

3.8.2.3.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 198.4 secondes de temps moyen par découverte pour le SSP (n=12), 131.1 secondes de temps moyen par découverte pour le PCP (n=12) et 65.1 secondes de temps moyen par découverte pour le SPM (n=13). L'ANOVA nous donne comme résultats : $F(2,33)= 6.45$ et $p= 0.004321$. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : $t(21)= 1.44$ $p= 0.082305$
- SSP vs SPM : $t(22)= 3.19$ $p= 0.0021155$
- PCP vs SPM : $t(23)= 3.52$ $p= 0.000918$

Il existe une différence significative entre les trois graphiques. Nous pouvons conclure que **pour le problème Vessel, le SPM permet de réaliser des découvertes le plus rapidement et le SSP le moins rapidement.**

3.8.2.3.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 129.1 secondes de temps moyen par découverte pour le SSP (n=13), 93.3 secondes de temps moyen par découverte pour le PCP (n=14) et 75.3 secondes de temps moyen par découverte pour le SPM (n=13). L'ANOVA nous donne comme résultats : $F(2,37)= 5.71$ et $p= 0.006899$. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : $t(25)= 1.94$ $p= 0.03187$
- SSP vs SPM : $t(24)= 3.24$ $p = 0.001743$
- PCP vs SPM : $t(25)= 1.44$ $p = 0.081137$

Il existe une différence significative entre les trois graphiques. Nous pouvons conclure que **pour le problème Disk, le SPM permet de réaliser des découvertes le plus rapidement et le SSP le moins rapidement.**

Nous pouvons conclure que le graphique SPM permet de faire des découvertes le plus rapidement, le SSP étant le graphique avec lequel les participants ont été les moins rapides (voir Figure 117).

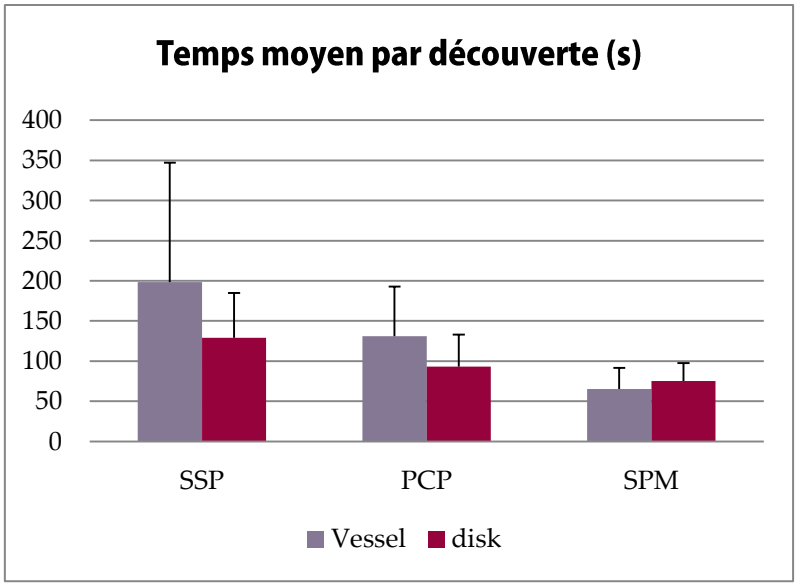


Figure 117 : Moyennes et écart-types des résultats pour le temps moyen par découverte

3.8.3 Conclusion pour la phase de découverte

Le Scatter Plot Matrix (SPM) est le graphique le plus pertinent pour la phase de découverte. En acceptant un risque de première espèce α à 10%, le SPM est le graphique permettant un plus grand nombre de découvertes, une première découverte dans un temps le plus court et un temps moyen par découverte le plus court.

Tableau 19 : Synthèse des résultats obtenus pour la phase de Découverte

	Nombre de découvertes réalisées		**Temps pour la première découverte**		**Temps moyen pour réaliser une découverte**	
	Vessel	**Disk**	**Vessel**	**Disk**	**Vessel**	**Disk**
Meilleur	SPM	SPM	SPM	SPM	SPM	SPM
Intermédiaire	PCP ≡ SSP	PCP ≡ SSP	PCP	PCP	PCP	PCP
Moins bon			SSP	SSP	SSP	SSP

Nous devons également indiquer que le Simple Scatter Plot (SSP) est le graphique montrant les moins bons résultats pour la phase de découverte (voir le **Tableau 19**). Nous déconseillons donc l'utilisation de ce graphique pour la découverte.

3.9 PERFORMANCE DES GRAPHIQUES DANS L'EXPLORATION D'ESPACE DE CONCEPTION

Nous retenons trois graphiques pertinents pour l'exploration d'espace de conception. Nous conseillons donc l'utilisation du Scatter Plot Matrix pour la phase de découverte, le Scatter Plot Simple pour la phase d'optimisation et le Parallel Coordinates Plot pour la phase de sélection (voir **Tableau 20**).

Pour la phase de découverte, le Scatter Plot Matrix est le plus adapté car c'est un graphique permettant une vue d'ensemble avec une représentation cartésienne. Ce type de représentation est efficace pour interpréter facilement les interactions entre les variables en un coup d'œil. En comparaison, le Simple Scatter Plot offre une représentation cartésienne mais en deux dimensions. Il est donc nécessaire de réaliser plusieurs permutations pour visualiser l'ensemble des interactions entre les variables. Le Parallel Coordinates Plot propose une vue d'ensemble de l'ensemble des données mais nécessite plusieurs permutations des axes pour visualiser l'ensemble des interactions entre les variables. De plus, ça représentation n'est pas cartésienne et donc difficile à interpréter.

Pour la phase d'optimisation, le Simple Scatter Plot est plus adapté car c'est le graphique qui présente le plus de détail sur les données. De plus, il a l'avantage d'afficher les variables dans une représentation cartésienne.

Enfin, pour la phase de sélection c'est le Parallel Coordinates Plot qui est le plus adapté. En effet, c'est un graphique qui offre une vue d'ensemble des solutions possibles. Dans cette phase il n'est plus nécessaire de chercher des interactions entre les variables (i.e. comparer les variables deux à deux). En comparaison au Scatter Plot Matrix, le Parallel Coordinates Plot à l'avantage de présenter une solution possible en un seul item (une poly-ligne) alors que dans le Scatter Plot Matrix une solution est représentée par plusieurs items (au tant qu'il y a de nuages de points dans la matrice). Aussi, il est plus facile d'identifier des paramètres standardisables car ils sont représentés par un point sur un axe (« métaphore » d'une valeur fixe du paramètre) alors qu'ils sont représentés par une ligne dans le Scatter Plot Matrix.

Tableau 20 : Synthèse des résultats de l'expérimentation n°2

		Phase		
		Découverte	Optimisation	Sélection
Graphique	Conseillé	SPM	SSP	PCP
	Déconseillé	SSP	SPM	SSP

De plus, afin de compléter nos recommandations, il est important de noter que :

- le SPM est le graphique permettant d'être le plus performant pour la phase découverte, cependant nous le déconseillons fortement pour la phase d'optimisation car c'est le graphique affichant les moins bons résultats de cette phase.
- le SSP est le graphique permettant d'être le plus performant pour la phase d'optimisation, cependant nous le déconseillons fortement pour les phases de découverte et de sélection car c'est le graphique affichant les moins bons résultats de cette phase.

Il convient de vérifier que l'utilisation de ces graphiques permet au concepteur d'aboutir à une décision informée.

3.10 EXPERIMENTATION 2.4 : DECISION INFORMEE

3.10.1 Mesures liées à la décision informée

Nous avons sélectionné deux indicateurs nécessaires pour vérifier l'influence du graphique sur la décision informée : Le nombre d'actions réalisées et l'information acquise relative au problème de conception.

Le nombre d'actions réalisées

Cet indicateur est le reflet du nombre d'allers-retours entre la phase de découverte et la phase d'optimisation. L'aller-retour entre les deux phases est un premier indice sur la décision informée : pour nous il représente le nombre d'observations-réductions réalisées qui correspond donc à une décision informée.

L'information acquise relative au problème de conception

Il s'agit de la réponse des participants à une question post-test :

1. « Quelles sont les informations qui vous ont permis de construire votre cheminement jusqu'à la solution proposée?

   D'un point de vu global ou local (les optimums): Interactions entre des variables? Découvertes d'indices tout au long de l'exploration?

   - *Corrélation?*
   - *Tendance?*
   - *Point de transition?*
   - *Valeur « fixe »? »*

Les réponses à cette question nous permettent de savoir si le participant peut « justifier » sa décision à partir des informations acquises durant son exploration. Notons que pour l'analyse de cet indicateur, nous ne regarderons que si le participant justifie sa décision ou non (variable qualitative nominale). Nous ne pouvons pas ici analyser la quantité d'information que le participant utilise pour justifier sa réponse car cette quantité dépend de la phase de découverte mais aussi du participant. De la même manière, les formulaires blancs ne sont pas considérés (car pas de réponses n'implique pas que le participant ne sache pas justifier sa décision). Vous trouverez un exemple d'une réponse pour une décision informée et d'une réponse pour une décision non-informée **en annexes p.241**.

3.10.2 Traitement des résultats 2.4

Pour l'analyse des données de cette partie, nous posons les hypothèses statistiques globales suivantes :

- H0 : il n'existe pas de différence en les trois groupes
- H1 : il existe une différence entre les trois groupes

3.10.2.1 Nombre d'actions

3.10.2.1.1 Problème Vessel

Pour ce problème nous obtenons une moyenne de 6.31 actions pour le SSP (n=13), 6.58 actions pour le PCP (n=12) et 9.31 actions pour le SPM (n=13).

Nous réalisons une analyse de variance (ANOVA-*between*) pour répondre à nos hypothèses statistiques car notre étude comporte une variable quantitative et trois groupes indépendants. Nous prenons un risque de première espèce égale à 10% : $F(2,35)= 3.64$ et $p=0.037$. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite un test-t (de Student) par paires :

- SSP vs PCP : $t(23)=-0.26$ $p=0.399$
- SSP vs SPM : $t(24)=-2.31$ $p=0.015$
- PCP vs SPM : $t(23)=2.06$ $p=0.025$

Il existe une différence significative entre les graphiques SPM et SSP ; et SPM et PCP. Nous pouvons conclure que pour le problème Vessel, **le SPM permet de réaliser un plus grand nombre de d'actions (et donc d'allers-retours entre les phases de découverte et d'optimisation)**.

3.10.2.1.2 Problème Disk

Pour ce problème nous obtenons une moyenne de 7.54 actions pour le SSP (n=13), 6.5 actions pour le PCP (n=14) et 8.46 actions pour le SPM (n=13).

Nous réalisons une analyse de variance (ANOVA between) pour répondre à nos hypothèses statistiques car notre étude comporte une variable quantitative et trois groupes indépendants: $F(2,37)= 1.18$ et $p=0.319$. Nous devons rejeter l'hypothèse alternative et inférer qu'**il n'existe pas de différences entre les trois groupes pour le problème Disk**.

3.10.2.2 Question post-test

Pour cet indicateur nous analysons les données des deux problèmes conjointement. Nous pouvons, en effet, analyser les données ensembles puisqu'ici nous relevons simplement si le participant a fait une erreur ou non. Ainsi, notre étude comporte deux variables qualitatives : le groupe (SPM, PCP et SSP) et la réponse à la question (décision justifiée ou non). Nous appliquons un test du Khi2 (le tableau de contingence est présenté **Tableau 21**).

Tableau 21 : Tableau de contingence des variables groupes et réponse à la question

	Décision justifiée	Décision non-justifiée	Total
SPM	18	4	22
PCP	12	11	23
SSP	14	10	24
Total (freq)	44 (0.64)	25 (0.36)	69 (1)

Nous réalisons donc le test du Khi2 pour répondre à nos hypothèses statistiques: X^2=4.747, ddl=2 et p=0.093. Nous devons rejeter l'hypothèse nulle et inférer qu'il existe une différence. Nous opérons ensuite des tests du Khi2 par paires :

- SSP vs PCP : X^2=0.18, ddl=1 et p=0.671
- SSP vs SPM : X^2=2.99, ddl=1 et p=0.084
- PCP vs SPM : X^2=4.447, ddl=1 et p=0.035

Il existe une différence significative entre les graphiques SPM et SSP ; et SPM et PCP. **Nous pouvons conclure que le SPM est le graphique avec lequel les participants ont réalisé le plus de décisions informées**.

Afin de consolider l'analyse de données pour la prise de décision informée nous proposons ici de comparer deux groupes: les participants ayant abouti à une décision informée et les participants ayant abouti à une décision non-informée (pour les graphiques PCP et SSP confondus).

3.10.2.3 Comparaison des décisions informées et non-informées

Nous décidons de comparer ces deux groupes (informée/non-informée) pour la phase de découverte. Nous utilisons donc les mesures utilisées pour cette phase : le nombre de découvertes réalisées, le temps écoulé avant de réaliser une première découverte et le temps moyen pour effectuer une découverte.

Nous sommes dans les trois cas en présence de deux groupes et de variables quantitatives. Nous opérons donc trois Test-t.

Aussi, pour cette partie nous proposons de poser des hypothèses statistiques globales :

- H0 : il n'y a pas de différence entre les participants ayant abouti à une décision informée et ceux aboutissant à une décision non-informée.
- H1 : il existe une différence entre les participants, ceux ayant abouti à une décision informée ont été plus performant en phase de découverte.

3.10.2.3.1 Nombre de découvertes réalisées

Pour cet indicateur, nous obtenons une moyenne de 2.38 découvertes pour le groupe « non-informée » (n=21) et de 3.17 découvertes pour le groupe « informée » (n=24). Les résultats du

Test-t nous donne t=-2.06 et p=0.023 avec ddl=43. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les 2 groupes**. Les participants ayant abouti à une décision informée ont réalisé plus de découvertes**.

3.10.2.3.2 Temps pour la première découverte

Pour cet indicateur, nous obtenons une moyenne de 192.05 secondes pour le groupe « non-informée » (n=20) et de 142.83 secondes pour le groupe « informée » (n=24). Nous remarquons qu'il y a un individu de moins pour le groupe « non-informée » par rapport à l'indicateur précédent. Il ne s'agit pas d'une erreur. Un participant a effectué zéro découverte, il nous est donc impossible de mesurer un temps. Les résultats du Test-t nous donne t=1.44 et p=0.079 avec ddl=42. Nous devons donc rejeter l'hypothèse nulle et inférer qu'il existe une différence significative entre les deux groupes**. Les participants ayant abouti à une décision informée ont réalisé une première découverte en un temps plus court**.

3.10.2.3.3 Temps moyen pour réaliser une découverte

Pour cet indicateur, nous obtenons une moyenne de 158.11 secondes pour réaliser une découverte pour le groupe « non-informée » (n=20) et de 130.08 secondes pour réaliser une découverte pour le groupe « informée » (n=24). Les résultats du Test-t nous donne t=1.01 et p=0.159 avec ddl=42. Nous devons donc rejeter l'hypothèse alternative et inférer qu'il n'existe pas de différence entre les deux groupes.

Nous pouvons conclure que les participants ayant abouti à une décision informée ont été plus performants dans la phase de découverte que ceux ayant abouti à une décision non-informée.

3.10.3 Conclusion sur la décision informée

Le SPM est le graphique avec lequel les participants ont réalisé le plus d'actions (pour le problème Vessel). De plus, il s'agit du graphique utilisé pour lequel les participants montrent un plus grand nombre de décisions justifiées par les informations acquises durant l'exploration. Les résultats de ces deux indicateurs nous permettent de conclure que le SPM est le graphique le plus adapté pour une décision informée.

Il n'existe pas de différence significative entre le SSP et PCP. Nous notons que pour ces deux graphiques les participant ont, dans la majorité, exprimé une décision informée :

- SSP : 14 décisions sur 24 sont informées (58%)
- PCP : 12 décisions sur 23 sont informées (52%)

De plus, les participants ayant exprimés une décision informée sont plus performant dans la phase de découverte. La performance durant cette phase de découverte semble déterminante pour la prise de décision informée en exploration d'espace de conception.

CONCLUSION DE LA TROISIEME PARTIE

Nous retenons trois graphiques permettant d'être performant pour l'exploration d'espace de conception. Nous conseillons donc l'utilisation du Scatter Plot Matrix pour la phase de découverte, le Scatter Plot Simple pour la phase d'optimisation et le Parallel Coordinates Plot pour la phase de sélection (voir **Tableau 22**).

Tableau 22 : Synthèse des résultats des expérimentations

		Phase		
		Découverte	Optimisation	Sélection
Graphique	Conseillé	SPM	SSP	PCP
	Déconseillé	SSP	SPM	SSP

De plus, afin de compléter nos recommandations, il est important de noter que :

- le SPM est le graphique le plus performant pour la phase découverte, cependant nous le déconseillons fortement pour la phase d'optimisation car c'est le graphique affichant les moins bons résultats de cette phase.
- le SSP est le graphique le plus performant pour la phase d'optimisation, cependant nous le déconseillons fortement pour les phases de découverte et de sélection car c'est le graphique affichant les moins bons résultats de cette phase.

Il est important de souligner ces deux dernières spécifications car, comme déjà mentionné, les phases de découverte et d'optimisation s'exécutent de concert jusqu'à la phase de sélection d'une solution optimale (voir une illustration **Figure 118**).

Enfin, nous montrons que le Scatter Plot Matrix est le graphique permettant d'aboutir à une décision informée. Par ailleurs, nous montrons que les participants ayant abouti à une décision informée sont ceux qui ont été performants durant la phase de découverte. Nous encourageons donc les concepteurs à exploiter les capacités du Scatter Plot Matrix qui est le graphique montrant les meilleurs résultats dans la phase de découverte.

Ainsi, la conclusion que nous portons pour l'exploration d'espace de conception en vue d'aboutir à une décision informée est de se former à l'utilisation du Scatter Plot Matrix en phase de découverte, du Simple Scatter Plot en phase d'optimisation et du Parallel Coordinates Plot en phase de sélection.

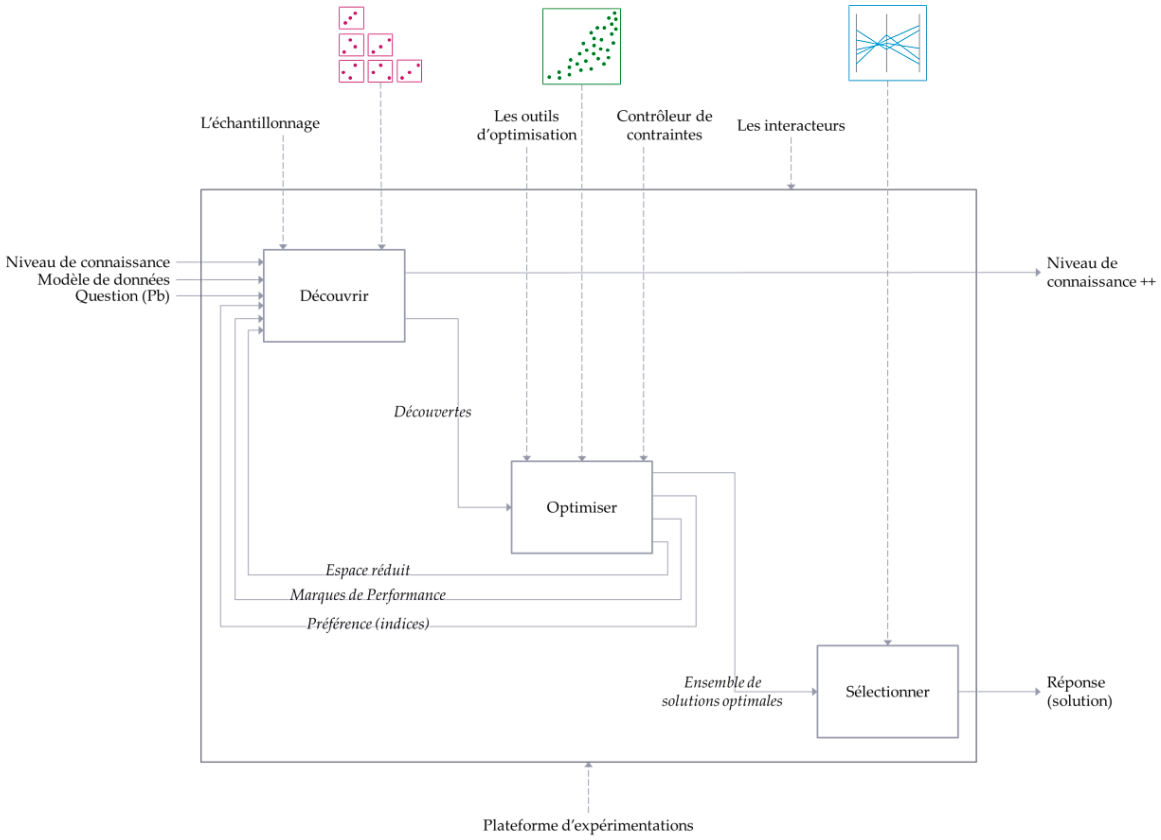


Figure 118 : SADT de l'exploration d'espace de conception avec nos recommandations

Quatrième partie

Apports, Limites et Perspectives

4 QUATRIEME PARTIE : Apports, limites et perspectives

Dans cette quatrième partie, nous présentons les apports, limites et perspectives de nos travaux de recherche. Elle est divisée en deux sections distinctes.

La première section établit les apports de nos travaux de thèse. Nous retenons sept apports principaux que nous présentons sous trois thématiques. Nous décrivons deux apports qui sont des préconisations de visualisation d'information pour la prise de décision, deux apports orientés méthodologie et trois apports que nous définissons comme à destination de la communauté scientifique en génie industriel.

La deuxième section décrit les limites actuelles de nos travaux et de notre approche et propose les perspectives envisagées. Nous illustrons le cadre actuel de nos apports afin de le délimiter dans un cadre dit idéal. Cette illustration nous permet de dresser les limites actuelles de nos travaux qui nous servent de socle pour proposer des perspectives.

4.1 Apports

Nos travaux de recherche ont permis d'apporter des recommandations concernant la visualisation d'information pour l'exploration d'espace de conception en vue d'aboutir à une décision informée. Nous retenons sept apports principaux de nos travaux de recherche que nous scindons sous trois thématiques.

4.1.1 Préconisations de visualisation d'information pour la prise de décision

Apport 1 : Une identification des graphiques pertinents pour chacune des phases du processus.

Les résultats de nos travaux nous permettent d'identifier des graphiques pertinents pour les trois phases principales de l'exploration d'espace de conception. Nous recommandons d'utiliser le Scatter Plot Matrix pour la phase de découverte, le Simple Scatter Plot pour la phase d'optimisation et le Parallel Coordinates Plot pour la phase de sélection. De plus, nous montrons la pertinence d'utiliser le Parallel Coordinates Plot pour la sélection dans une situation de prise de décision multi-attribut et multi-objectif.

Apport 2 : Une identification d'un graphique performant pour aboutir à une décision informée.

Les résultats de nos travaux nous permettent d'identifier un graphique pertinent pour aboutir à une décision informée en exploration d'espace de conception. Nous montrons l'importance de la phase de découverte pour aboutir à une décision informée et recommandons l'utilisation du Scatter Plot Matrix pour aboutir à une décision informée.

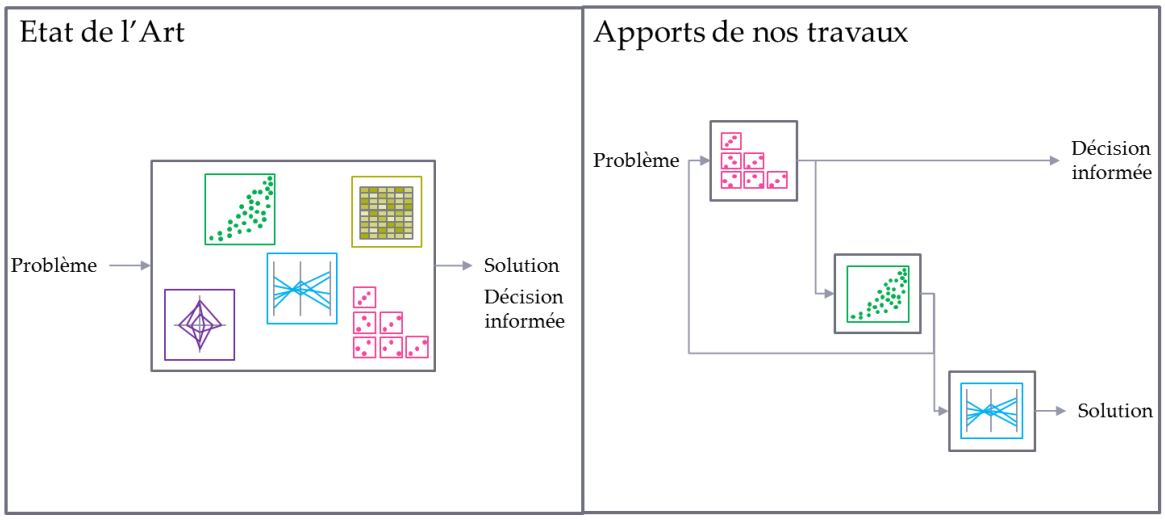


Figure 119 : Illustration des premiers apports de nos travaux

4.1.2 Apports orientés méthodologie

Apport 3 : Une plateforme en ligne proposant 4 problèmes de conception.

La plateforme développée et utilisée pour l'expérimentation n°2 (pour les étapes de formation et de test) est accessible en ligne[8]. Elle peut être un support utile à la formation des concepteurs à l'exploration d'espace de conception et/ou l'utilisation des graphiques pour l'optimisation en conception.

Apport 4 : Une formalisation du processus qui structure la méthode.

Nous proposons une modélisation du processus d'exploration d'espace de conception sous forme d'une SADT. Notre modélisation se concentre sur l'activité du concepteur et elle décrit les micro-situations auxquelles doit faire face le concepteur. En ce sens, elle structure la méthode d'exploration d'espace de conception et permet d'être un support solide pour se former à la méthode.

4.1.3 Apports à destination de la communauté scientifique en génie industriel

Apport 5 : Trois indices de qualité et leurs constructions.

Nous proposons la construction de 3 indices de qualité constitués à partir de la dominance au sens de Pareto et le « *goal programming* ». Ces indices peuvent être utiles pour l'évaluation des solutions possibles pour l'optimisation multicritère en conception et pour évaluer le *Design* d'un concepteur en conception basée sur les ensembles. Nous tenons à souligner que ces indices sont utiles pour évaluer des solutions et non les concepteurs. Notons que ces indices doivent être adaptés au contexte d'utilisation et plus particulièrement au type de représentation des données.

Apport 6 : Une définition de la décision informée.

Nos travaux nous permettent de mettre en lumière les entrés/soties (mécanisme) et les éléments clés d'une décision informée. En effet, c'est un terme employé dans la littérature et, à partir d'un état de l'art, nous en proposons une définition claire.

Apport 7 : Un protocole expérimental détaillé et reproductible.

Nous détaillons dans nos travaux le protocole expérimental de l'expérimentation n°2. Il décrit la procédure à mettre en place et les mesures à utiliser pour tester la pertinence de supports graphiques lors de l'exploration d'espace de conception. Plus largement, le protocole peut être réappliqué pour tout type de graphiques utilisés en conception et permettre d'identifier les supports graphiques pertinents pour la découverte, l'optimisation, la sélection et la prise de décision informée.

[8] http://these.aaa.alwaysdata.net/expe2/

4.2 Limites et perspectives

Actuellement, nos résultats sont principalement validés suite à notre expérimentation n°2 qui s'est attachée à tester l'utilisation de représentations graphiques pour l'exploration d'espace de conception avec des étudiants engagés individuellement dans le processus de conception.

Dans la deuxième expérimentation, nous avons testé les trois graphiques indépendamment i.e. un participant a utilisé un seul et même graphique pour les trois phases du processus. Cette approche est suffisante et nécessaire pour répondre à notre problématique. Néanmoins, nos résultats nous poussent à nous questionner sur la combinaison des trois graphiques. Nous pensons qu'il serait pertinent de tester l'utilisation combinée des trois graphiques pour résoudre un problème de conception en exploration d'espace de conception.

Le modèle a été testé principalement avec des concepteurs novices dans un cadre « pédagogique ». Il est indispensable d'évaluer la combinaison des trois graphiques pour l'exploration d'espace de conception avec des experts engagés dans un processus de conception collaboratif. En effet, ce type de méthode s'emploie lors de grands projets dans un processus multi-acteur (concepteurs / ingénieurs) et multi-expertise. Ainsi, l'adoption de la méthode et des outils par des experts serait un cadre idéal.

La **Figure 120 a.** illustre l'exploration d'espace de conception par un concepteur (un novice). Elle représente, dans une illustration simplifiée, l'apport actuel de nos travaux.

La **Figure 120 c.** illustre, dans un contexte industriel, l'exploration d'espace de conception dans une situation multi-expertise où les concepteurs ont des variables de performance différentes qui se retrouvent souvent en conflit (Canbaz, 2013 ; 2014). Ce cadre représente un apport terrain. Pour nous, cette situation est la plus réaliste et c'est dans cette situation de conception où il faut trouver des compromis que les outils de représentations graphiques sont les plus utiles et pertinents.

Par ailleurs, Simpson et ses collègues avancent que la visualisation d'information en exploration d'espace de conception est un support de communication important pour la collaboration (Simpson *et al*., 2008). En poursuivant cette idée, il serait pertinent de tester les trois graphiques afin d'identifier le plus performant pour le rôle de support à la communication (voire tester les cinq graphiques identifiés dans le **chapitre 3**).

Ensuite, nos travaux se concentrent sur l'exploration de l'espace de conception (divisée en trois phases). Evidemment, l'exploration nécessite la construction du modèle au préalable. Dans nos expérimentations, le modèle de données, i.e. la définition du problème de conception, a été réalisée par nos soins et la description et les objectifs donnés aux participants. La construction du modèle est une étape qui peut influencer la représentation du problème (par le concepteur) et donc influencer l'exploration d'espace de conception. Il serait intéressant de tester l'exploration suite à la construction du modèle par le concepteur afin de vérifier la validité de nos résultats dans un contexte plus réaliste et plus global.

La **Figure 120 b**. illustre les trois phases de l'exploration d'espace de conception : D pour découvrir, O pour optimiser et S pour sélectionner.

La **Figure 120 d.** illustre les trois phases du processus d'exploration d'espace de conception précédées d'une phase supplémentaire noté M qui est la définition du problème de conception (le Modèle). Ce cadre représente l'idéal car avant de résoudre un problème de conception il est nécessaire de définir le problème (même si tous les critères ne sont pas encore définis ou fixés).

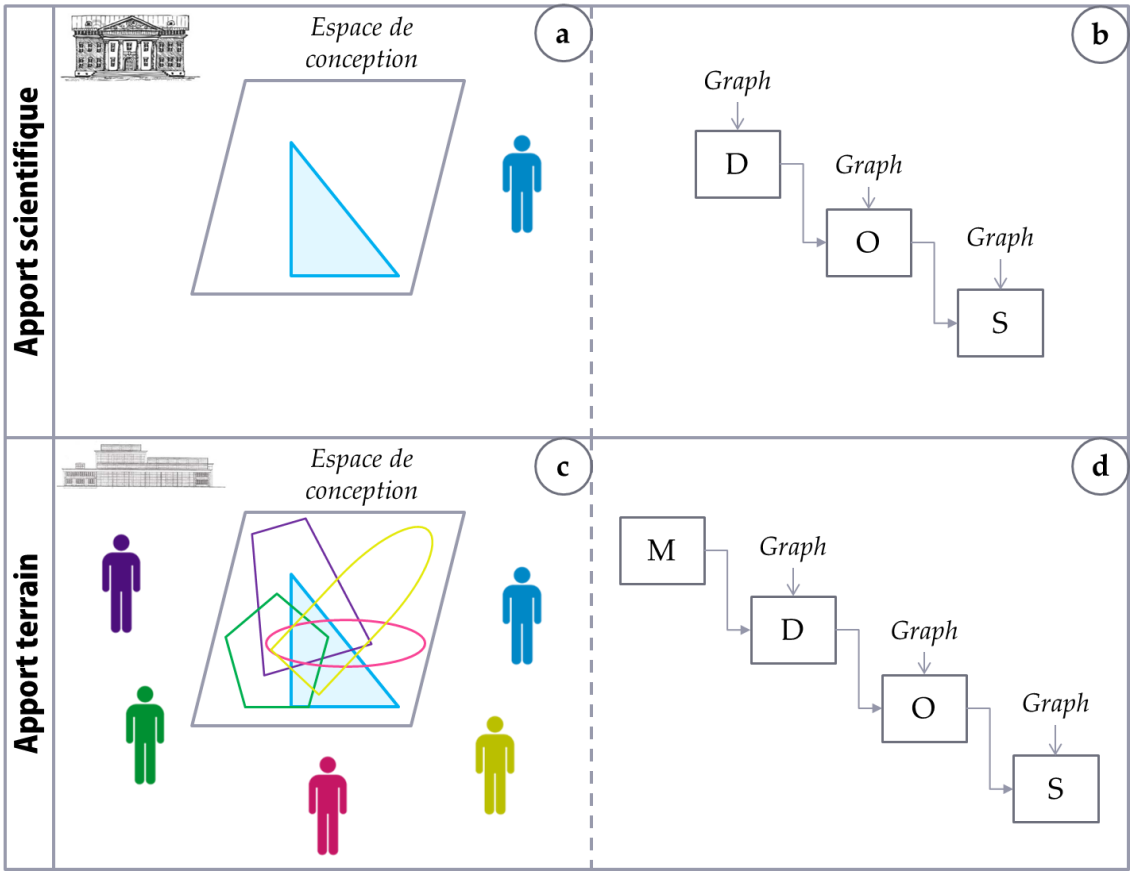


Figure 120 : Illustration de l'apport scientifique (a et b) et d'un apport terrain (c et d). M pour construire le Modèle, D pour Découvrir, O pour Optimiser et S pour Sélectionner.

Enfin, nous illustrons nos perspectives à travers la **Figure 121 a**. **b**. **c**. et **d**. :

La **Figure 121 a**. illustre l'utilisation de l'indicateur de bien-être (*wellbeing*) proposé par Canbaz (voir Canbaz, 2013). Ce type d'indicateur a été proposé pour la conception en *Set-Based-Design* (Sobek *et al*., 1999) dans un paradigme de réduction de l'incertitude. Il a pour nous sa place en *Design by Shopping* qui est un paradigme qui met en avant la satisfaction du concepteur lors d'optimisation en conception. Il serait pertinent, pour la conception en *Design by Shopping*, d'intégrer cet indicateur dans le processus d'exploration d'espace de conception afin de garantir la satisfaction des concepteurs.

La **Figure 121 b**. illustre l'intégration, dans le processus, de la phase de définition de l'indicateur de bien-être par chacun des concepteurs notée WB (i.e. *Wellbeing*). Par ailleurs,

nous proposons d'utiliser des représentations graphiques comme entrées à la définition de ces indicateurs.

La **Figure 121 c**. illustre la visualisation des indicateurs de bien-être par les différents concepteurs. Outre l'intégration d'un indicateur de bien-être, il serait pertinent de visualiser les indicateurs de l'ensemble des concepteurs au cours du processus d'exploration d'espace de conception. De cette manière, chaque concepteur est conscient du niveau de satisfaction de ces collaborateurs et peut « avancer » dans son processus tout en étant informé.

La **Figure 121 d**. illustre l'utilisation de l'indicateur de bien-être comme support aux trois phases de l'exploration d'espace de conception. Par la voie de la visualisation d'information, cet indicateur peut être un support à l'exploration d'espace de conception. En effet, il peut être perçu comme un indice supplémentaire lors de la phase de découverte et favoriser la compréhension du problème à résoudre. En effet, dans cette idée, un concepteur peut interagir avec le modèle, à travers un graphique, et observer l'impact de certaines modifications sur le niveau de bien-être de ses collaborateurs.

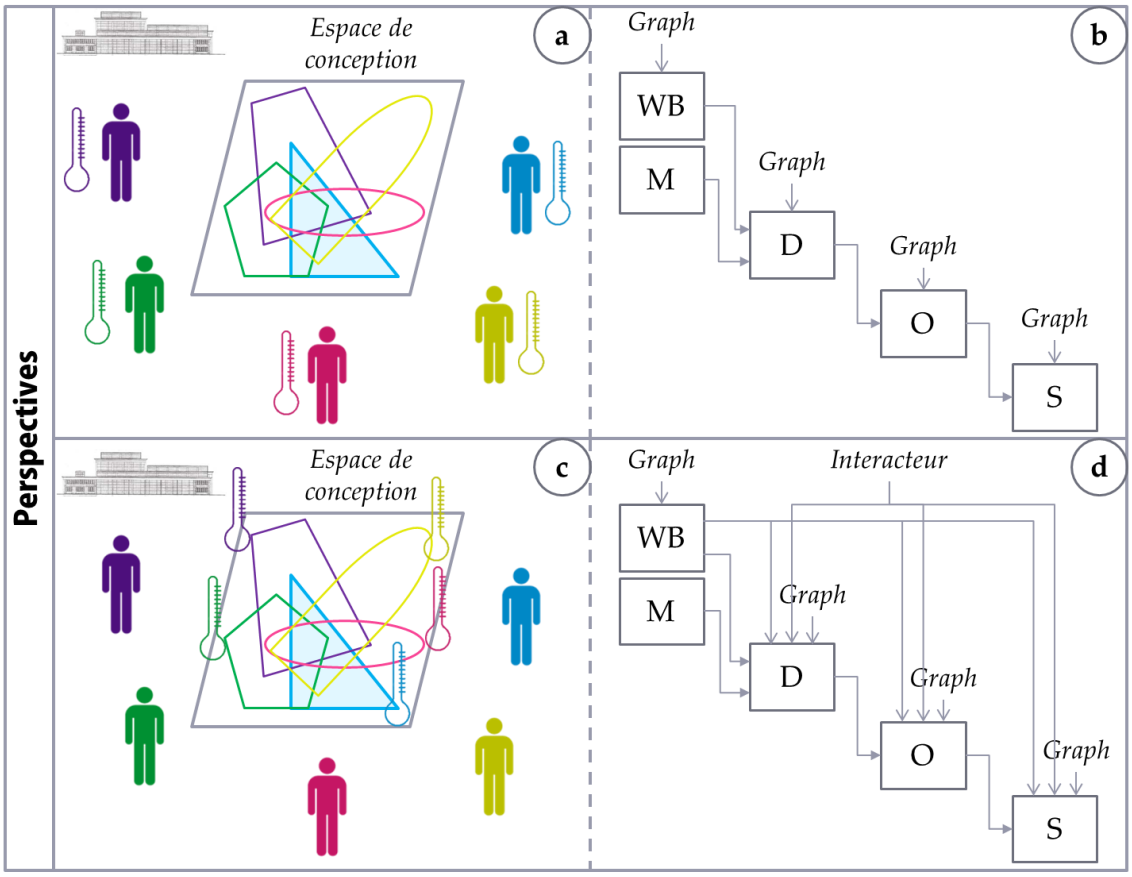


Figure 121 : Illustration des perspectives. M pour construire le Modèle, WB pour la définition du *WellBeing*, D pour Découvrir, O pour Optimiser et S pour Sélectionner

La **Figure 121 d**. illustre également l'utilisation d'interacteurs qui représentent pour nous les fonctionnalités à identifier et utiles aux concepteurs. Le principal défi est de préconiser et de mettre au point des fonctionnalités d'outils graphiques afin de manipuler et interagir avec les

solutions faisables et infaisables de l'espace de conception. Nous pouvons citer comme exemple pertinent le « visual steering attractor sampler » de Stump *et al.* (Stump *et al.*, 2009) qui permet notamment de générer des solutions possibles autour d'un point (l'attractor) défini par le concepteur dans l'hypercube formé par l'espace de conception.

Poursuivant la même idée, il serait pertinent de faire évoluer notre plateforme d'expérimentations vers l'usage collaboratif et d'y intégrer la phase de construction de modèle, de définition des indicateurs de bien-être et la visualisation de ces indicateurs au cours du processus d'exploration. Actuellement, la plateforme propose de résoudre quatre problèmes de conception. Ces intégrations nous permettraient donc d'élever la plateforme au statut d'outil à part entière car les concepteurs pourront modéliser leurs problèmes de conception et les résoudre.

Conclusion Générale

CONCLUSION GENERALE

Dans la première partie nous avons présenté trois chapitres décrivant notre positionnement scientifique et notre état de l'art.

Le premier chapitre positionne nos travaux de recherche en exploration d'espace de conception. L'exploration d'espace de conception est une méthode pour l'optimisation multicritère en conception (en situation multi-attribut ou multi-objectif) en Design by Shopping. Le processus de conception, avec cette méthode, est décomposé en trois phases principales : la découverte, l'optimisation et la sélection. Nous constatons que la phase de sélection peut être dans une situation de prise de décision multi-objectif (intégrée dans le processus complet) ou dans une situation multi-attribut.

Le deuxième chapitre confirme la formulation d'une décision informée pour la conception en Design by Shopping. Nous en proposons une définition :

Une réponse à une question, l'élection d'une alternative ou un ensemble d'alternatives parmi plusieurs qui permettra d'obtenir les avantages optimaux et les inconvénients minimaux suite à un processus de traitement et d'analyse itératif et interactif durant lequel l'acteur gagne en compréhension, connaissance et « insights » grâce à la visualisation et la manipulation de grands ensembles et/ou modèle de données.

Nous soulignons la nécessité d'inclure l'humain dans le processus, l'activité d'analyse et de traitement de données et le besoin de transformation de données en information et par la suite en connaissance. Nous constatons que l'un des outils utile à la décision informée est la visualisation d'information et l'interaction (entre l'utilisateur et les données et les modèles).

Le troisième chapitre propose un état de l'art des représentations graphiques d'ensembles de données multidimensionnels. Nous réalisons un comparatif de ces graphiques afin d'identifier ceux qui semblent pertinents dans notre contexte. Nous nous appuyons notamment sur : des caractéristiques physiques, fonctionnelles et d'apprentissage. Ce chapitre établit les graphiques qui permettent potentiellement d'être performant pour l'optimisation multicritère en Design by Shopping. Cinq graphiques sont retenus et une fiche identité est proposée pour chacun d'entre eux.

C'est dans la deuxième partie du manuscrit que nous décrivons notre problématique de recherche qui est :

« Quel(s) graphique(s) permet(tent) aux concepteurs d'aboutir à une décision informée lors de l'exploration de l'espace de conception ? »

Cette problématique est constituée de deux composantes :

- La première composante est liée à la performance du processus d'optimisation multicritère en conception lors de l'exploration d'espace de conception. Le premier objectif est la définition et l'identification de graphiques pertinents pour la résolution de problèmes multicritères de conception dans une approche de Design by Shopping.
- La deuxième composante est liée au concepteur. Elle concerne le concepteur comme décisionnaire en exploration d'espace de conception. Le deuxième objectif est l'identification de graphiques permettant d'aboutir à une décision informée i.e. une décision prise principalement à partir de l'information acquise durant l'exploration d'espace de conception.

De ces deux composantes de notre problématique découlent nos quatre hypothèses de recherche.

La troisième partie du manuscrit présente les expérimentations réalisées dans nos travaux pour démontrer nos hypothèses de recherche.

L'expérimentation n°1 a comme objectif de démontrer quels sont les graphiques performants pour la phase de sélection dans une situation de décision multi-attribut.

L'expérimentation n°2 a pour enjeux de valider :

- les graphiques performants pour la résolution de problèmes en exploration d'espace de conception dans une situation de décision multi-objectifs.
- les graphiques pour la décision informée en exploration d'espace de conception.

De plus, nous avons construit un indice de qualité relié à un scénario de conception. Nous proposons de calculer de manière automatique cet indice pour déterminer quel graphique permet aux concepteurs de prendre les « meilleures » décisions (utilisé pour l'expérimentation n°1).

Nous démontrons que suivant les différentes phases du processus, il est optimal d'associer certains types de graphique pour une décision informée, à savoir l'utilisation du :

- **Scatter Plot Matrix en phase de découverte**
- **Simple Scatter Plot en phase d'optimisation**
- **Parallel Coordinates Plot en phase de sélection.**

Enfin, dans la quatrième partie, nous présentons les apports, limites et perspectives de nos travaux de recherche en deux sections distinctes.

La première section établit les apports de nos travaux de thèse. Nous retenons sept apports principaux que nous présentons sous trois thématiques. Nous décrivons deux apports qui

sont des préconisations de visualisation d'information pour la prise de décision, deux apports orientés méthodologie et trois apports que nous définissons comme à destination de la communauté scientifique en génie industriel.

La deuxième section décrit les limites actuelles de nos travaux et de notre approche et propose les perspectives envisagées. Nous illustrons le cadre actuel de nos apports afin de le délimiter dans un cadre dit idéal. Cette illustration nous permet de dresser les limites actuelles de nos travaux qui nous servent de socle pour proposer des perspectives.

Nous sommes bien décidés à poursuivre nos travaux dans cette thématique de visualisation d'information pour la prise de décision en conception. Nous comptons continuer notamment en développant et en proposant des éléments industrialisables pour les entreprises de demain.

Liste des publications

Revue international avec comité de lecture

- **Abi Akle, A.**, Minel, S., & Yannou, B. (soumis). Information Visualization for Selection in Design by Shopping. Research in Engineering Design.

Chapitre d'ouvrage

- **Abi Akle, A.**, Borgiel, K., Minel, S., & Merlo, C. (soumis). Design of Interfaces for Seniors in the Context of Health Care. Ergonomics and Usability in Design.

Communication en congrès internationaux avec actes et comité de lecture

- **Abi Akle, A.**, Minel, S., & Yannou, B. (2015). Graphical support adapted to designers for the selection of an optimal solution in Design by Shopping. In International Conference on Engineering Design (ICED15) Design for Life. Milan, Italy, 27-30.07.2015
- **Abi Akle, A.**, Borgiel, K., Minel, S., & Merlo, C. (2014, July). Recommendations for User Centered Design of Interfaces for Seniors in the Context of Health Care. In Advances in Human Factors and Ergonomics 2014 (No. Section 13, p. 647). AHFE Conference© 2014.
- **Abi Akle, A.**, Bertoluci, G., & Minel, S. (2013). From environmental assessment to usage centered eco-design: taking into account the real impact of container-content system for the liquid laundry detergent. In DS 75-5: Proceedings of the 19th International Conference on Engineering Design (ICED13) Design For Harmonies, Vol. 5: Design for X, Design to X, Seoul, Korea 19-22.08. 2013.
- Chapotot, E., **Abi Akle, A.**, Minel, S., & Yannou, B. (2011). Comparative study of theoretical and real uses of eco-designed laundry detergents. In DS 68-6: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 6: Design Information and Knowledge, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011.

Communication en congrès avec comité de lecture

- **Abi Akle, A.**, Minel, S., & Yannou, B. (2013). CHANGE THE VISUAL ALERT DISPLAY TO IMPROVE THE DECISION MAKING TIME: INDUSTRIAL SYSTEMS SUPERVISION CASE STUDY. In CONFERE2013, Biarritz, France.
- **Abi Akle, A.**, Bertoluci, G., & Minel, S. (2012). FACTEURS INFLUENTS SUR LE CONSENTEMENT À TRIER DES CONSOMMATEURS FRANÇAIS. CAS DES EMBALLAGES DE LESSIVES LIQUIDES EN FRANCE. In CONFERE2012, Venise, Italie.

BIBLIOGRAPHIE

1. Adla, A. (2010). Aide à la Facilitation pour une prise de Décision Collective: Proposition d'un Modèle et d'un Outil (Doctoral dissertation, Université de Toulouse, Université Toulouse III-Paul Sabatier).
2. Aloysius,J., Davis,F.,Wilson,D.,Taylor,A., Kottemann,J., "User acceptance of multi-criteria decision support systems: The impact of preference elicitation techniques", European Journal of Operational Research 169 (2006) 273–285.
3. Andrienko, G., & Andrienko, N. (2001, March). Constructing parallel coordinates plot for problem solving. In 1st International Symposium on Smart Graphics (pp. 9-14).
4. Ankerst, M. (2001). Visual data mining with pixel-oriented visualization techniques. In Proceedings of the ACM SIGKDD Workshop on Visual Data Mining.
5. Ankerst, M., Keim, D. A., & Kriegel, H. P. (2011). 'Circle Segments': A Technique for Visually Exploring Large Multidimensional Data Sets. Dim, 1501, 1.
6. Anscombe, F.J. (1973), « Graphs in statistical analysis », The american statistician, 27, 17-21.
7. Balling, R. (1999, May). Design by shopping: A new paradigm?. In Proceedings of the Third World Congress of structural and multidisciplinary optimization (WCSMO-3) (Vol. 1, pp. 295-297).
8. Barabel, M. (1996). Un style de décision à la française?. Revue française de gestion, (111), 159-170.
9. Barron, K., Simpson, T. W., Rothrock, L., Frecker, M., Barton, R. R., & Ligetti, C. (2004, January). Graphical user interfaces for engineering design: impact of response delay and training on user performance. In ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 11-20). American Society of Mechanical Engineers.
10. Bass, T. (2000). Intrusion detection systems and multisensor data fusion. Communications of the ACM, 43(4), 99-105.
11. Bellut, S. (2002). Les processus de la décision: Démarches, méthodes et outils. AFNOR.
12. Berchet, C. (2000). Modélisation pour la simulation d'un système d'aide au pilotage industriel (Doctoral dissertation).
13. Bertin, J. (1973). Sémiologie graphique: Les diagrammes-Les réseaux-Les cartes.
14. Bertini, E., & Lalanne, D. (2009, June). Surveying the complementary role of automatic data analysis and visualization in knowledge discovery. In Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration (pp. 12-20). ACM.
15. Canbaz B., Yannou B., Yvars P.-A. (2013) 'Preventing design conflicts in distributed design systems composed of heterogeneous design agents', Engineering Application of Artificial Intelligence, vol. 28, pp. 142-154, doi: 10.1016/j.engappai.2013.11.017

16. Canbaz B., Yannou B., Yvars P.-A. (2014) 'Improving Process Performance of Distributed Set based Design Systems by Controlling Wellbeing Indicators of Design Actors', Journal of Mechanical Design, vol. 136, no. 2, doi: 10.1115/1.4026034
17. Canbaz B., Yannou B., Yvars P.-A. (2014) 'Resolving design conflicts and promoting solidarity in distributed design', IEEE Transactions on Systems, Man and Cybernetics: Systems, vol. 44, no. 8, pp. 1044-1055, doi: 10.1109/TSMC.2013.2296275
18. Canbaz, B. (2013). Preventing and resolving design conflicts for a collaborative convergence in distributed set-based design (Doctoral dissertation, Ecole Centrale Paris).
19. Carlsen, D., Malone, M., Kollat, J., & Simpson, T. W. (2008, January). Evaluating the performance of visual steering commands for user-guided pareto frontier sampling during trade space exploration. In ASME 2008 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 499-509). American Society of Mechanical Engineers.
20. Chambers JM, Cleveland WS, Kleiner B, Tukey PA (1983) Graphical methods for data analysis.Wadsworth International Group
21. Chandrasegaran, S. K., Ramani, K., Sriram, R. D., Horváth, I., Bernard, A., Harik, R. F., & Gao, W. (2013). The evolution, challenges, and future of knowledge representation in product design systems. Computer-Aided Design, 45(2), 204-228.
22. Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. Journal of the American Statistical Association, 68(342), 361-368.
23. Cook, D., Hofmann, H., Lee, E. K., Yang, H., Nikolau, B., & Wurtele, E. (2007). Exploring gene expression data, using plots. Journal of Data Science, 5(2), 151-182.
24. Cui, Q., Ward, M. O., & Rundensteiner, E. A. (2006, January). Enhancing scatterplot matrices for data with ordering or spatial attributes. In Electronic Imaging 2006 (pp. 60600R-60600R). International Society for Optics and Photonics.
25. Deb, K., & Srinivasan, A. (2006, July). Innovization: Innovating design principles through optimization. In Proceedings of the 8th annual conference on Genetic and evolutionary computation (pp. 1629-1636). ACM.
26. Donndelinger, J., Ferguson, S., & Lewis, K. (2006). Exploring mass trade-offs in preliminary vehicle design using Pareto sets. In Proceedings of the 11th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Portsmouth, VA, AIAA Paper No. AIAA-2006-7056.
27. Dumont d'Ayot, G. (2005). Coopération et évaluation cognitive d'agents artificiels pour la supervision (Doctoral dissertation, Toulouse, INSA).
28. Fathianathan, M., & Panchal, J. H. (2009). Modelling an ongoing design process utilizing top-down and bottom-up design strategies. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 223(5), 547-560.
29. Feather, M., Cornford, S., Hicks, K., Kiper, J., & Menzies, T. (2008). Application of a broad-spectrum quantitative requirements model to early-lifecycle decision making. IEEE Software, 1673-1680.

30. Ferguson, S., Gurnani, A., Donndelinger, J., & Lewis, K. (2005). A study of convergence and mapping in preliminary vehicle design. International Journal of Vehicle Systems Modelling and Testing, 1(1), 192-215.
31. Fernagu Oudet, S. (2012). Concevoir des environnements de travail capacitants: l'exemple d'un réseau réciproque d'échanges des savoirs. Formation emploi, 119(3), 7-27.
32. Ferringer, M. P., Spencer, D. B., & Reed, P. (2009, May). Many-objective reconfiguration of operational satellite constellations with the large-cluster epsilon non-dominated sorting genetic algorithm-II. In Evolutionary Computation, 2009. CEC'09. IEEE Congress on (pp. 340-349). IEEE.
33. Few, S. (2005). Keep Radar Graphs Below the Radar–Far Below. Perceptual Edge.
34. Flury, B., & Riedwyl, H. (1981). Graphical representation of multivariate data by means of asymmetrical faces. Journal of the American Statistical Association, 76(376), 757-765.
35. Gettinger, J., Kiesling, E., Stummer, C., & Vetschera, R. (2013). A comparison of representations for discrete multi-criteria decision problems. Decision support systems, 54(2), 976-985.
36. Ghoniem M., fekete J.-D., castagliola P., "A comparison of the readability of graphs using node-link and matrix-based representations", Proceedings of IEEE Symposium on Information Visualization, 2004, pp.17–24.
37. Glaser, S. D., & Tolman, A. (2008). Sense of sensing: from data to informed decisions for the built environment. Journal of infrastructure systems, 14(1), 4-14.
38. Hubka, V., & Eder, W. E. (1987). A scientific approach to engineering design. Design studies, 8(3), 123-137.
39. Hunault Gilles. Cours en ligne de de l'Université d'Angers (consulté en février 2015) : http://www.info.univ-angers.fr/~gh/wstat/statgen.htm
40. Inselberg, A. (1985). The plane with parallel coordinates. The Visual Computer, 1(2), 69-91.
41. Ireson, N. (2009, June). Local community situational awareness during an emergency. In Digital Ecosystems and Technologies, 2009. DEST'09. 3rd IEEE International Conference on (pp. 49-54). IEEE.
42. Jordan, D. D., Spencer, D. B., Simpson, T. W., Yukish, M. A., & Stump, G. M. (2008). Optimal spacecraft trajectories via visual trade space exploration. In AAS/AIAA Space Flight Mechanics Meeting (pp. 27-31).
43. Karandikar, H., & Mistree, F. (1992). Designing a composite material pressure vessel for manufacture: a case study in concurrent engineering. Engineering Optimization, 18(4), 235-262.
44. Kast, R. (1993). La théorie de la décision (pp. 61-64). Paris: La Découverte.
45. Keim, D. A. (2000). Designing pixel-oriented visualization techniques: Theory and applications. Visualization and Computer Graphics, IEEE Transactions on, 6(1), 59-78.

46. Keim, D. A. (2001). Visual exploration of large data sets. Communications of the ACM, 44(8), 38-44.
47. Keim, D. A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006, July). Challenges in visual data analysis. In Information Visualization, 2006. IV 2006. Tenth International Conference on (pp. 9-16). IEEE.
48. Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008b). Visual analytics: Scope and challenges (pp. 76-90). Springer Berlin Heidelberg.
49. Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008a). Visual analytics: Definition, process, and challenges (pp. 154-175). Springer Berlin Heidelberg.
50. Kollat, J. B., & Reed, P. (2007). A framework for visually interactive decision-making and design using evolutionary multi-objective optimization (VIDEO). Environmental Modelling & Software, 22(12), 1691-1704.
51. Le Moigne J. L. (1974). Les systèmes de décision dans les organisations. Presses Universitaires de France, Paris, ISBN 2-13-032739-7.
52. Levine, P. & Pomerol, J. (1989). Systèmes interactifs d'aide à la décision et systèmes experts. Editions Hermès.
53. Lewis, K., & Mistree, F. (1998). Collaborative, sequential, and isolated decisions in design. Journal of Mechanical Design, 120(4), 643-652.
54. Ligetti, C., Simpson, T. W., Frecker, M., Barton, R. R., & Stump, G. (2003). Assessing the impact of graphical design interfaces on design efficiency and effectiveness. Journal of Computing and Information Science in Engineering, 3(2), 144-154.
55. Longueville, B. (2003). Capitalisation des processus de décision dans les projets d'innovation: Application à l'automobile (Doctoral dissertation, Ecole Centrale Paris).
56. Lotov, A. V., & Miettinen, K. (2008). Visualizing the Pareto frontier. In Multiobjective optimization (pp. 213-243). Springer Berlin Heidelberg.
57. Lurie, N. H., & Mason, C. H. (2007). Visual representation: Implications for decision making. Journal of Marketing, 71(1), 160-177.
58. Malak Jr, R. J., Aughenbaugh, J. M., & Paredis, C. J. (2009). Multi-attribute utility analysis in set-based conceptual design. Computer-Aided Design, 41(3), 214-227.
59. Mavris, D. N., Pinon, O. J., & Fullmer Jr, D. (2010). Systems design and modeling: A visual analytics approach. In Proceedings of the 27th International Congress of the Aeronautical Sciences (ICAS), Nice, France.
60. Mennis, J. L., Peuquet, D. J., & Qian, L. (2000). A conceptual framework for incorporating cognitive principles into geographical database representation. International Journal of Geographical Information Science, 14(6), 501-520.
61. Meyer, J., Thomas, J., Diehl, S., Fisher, B. D., Keim, D. A., Laidlaw, D., ... & Ynnerman, A. (2010). From Visualization to Visually Enabled Reasoning. Scientific visualization: Advanced concepts, 1, 227-245.

62. Mezzich, J. E., & Worthington, D. R. (1978). A comparison of graphical representations of multidimensional psychiatric diagnostic data. Graphical representation of multivariate data, 123-141.
63. Miettinen, K. (2014). Survey of methods to visualize alternatives in multiple criteria decision making problems. OR Spectrum, 36(1), 3-37.
64. Mintzberg, H. (1979): "The structuring of Organization", Prentice Hall.
65. Mistree, F., Smith, W. F., Bras, B., Allen, J. K., & Muster, D. (1990). Decision-based design: a contemporary paradigm for ship design. Transactions, Society of Naval Architects and Marine Engineers, 98, 565-597.
66. O'Hara, J. J., Stump, G. M., Yukish, M. A., Harris, E. N., Hanowski, G. J., & Carty, A. (2007, April). Advanced visualization techniques for trade space exploration. In Collection of Technical Papers—AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference (Vol. 2, pp. 2137-2142).
67. Pahl, G. & Beitz, W. (1996). Engineering Design: a Systematic Approach. Springer, London.
68. Paredis, C., Malak Jr, R. J., & Aughenbaugh, J. M. (2007). Multi-attribute utility analysis in set-based conceptual design.
69. Pauzié, A. (2008). A method to assess the driver mental workload: The driving activity load index (DALI). IET Intelligent Transport Systems, Special Issue – selected papers from HCD 2008.
70. Petersen, S., & Svendsen, S. (2010). Method and simulation program informed decisions in the early stages of building design. Energy and buildings, 42(7), 1113-1119.
71. Pryke, A., Mostaghim, S., & Nazemi, A. (2007, January). Heatmap visualization of population based multi objective algorithms. In Evolutionary multi-criterion optimization (pp. 361-375). Springer Berlin Heidelberg.
72. Rangavajhala,S. (2007). Decision Making in Design Under uncertainty with multiobjective Robust Design Optimization. 205p, Rensselaer Polytechnic Institute, Troy, NY.
73. Reymen, I. (2001). Improving design processes through structured reflection: A domain-independent approach. Technische Universiteit Eindhoven.
74. Riveiro, M., Falkman, G., & Ziemke, T. (2008, July). Visual analytics for the detection of anomalous maritime behavior. In Information Visualisation, 2008. IV'08. 12th International Conference (pp. 273-279). IEEE.
75. Ross, A. M. (2003). Multi-attribute tradespace exploration with concurrent design as a value-centric framework for space system architecture and design (Doctoral dissertation, Massachusetts Institute of Technology).
76. Ross, A. M., & Hastings, D. E. (2005). The Tradespace Exploration Paradigm. In in INCOSE International Symposium.

77. Ross, A. M., Hastings, D. E., Warmkessel, J. M., & Diller, N. P. (2004). Multi-attribute tradespace exploration as front end for effective space system design. Journal of Spacecraft and Rockets, 41(1), 20-28.
78. Ross, A. M., McManus, H. L., Rhodes, D. H., & Hastings, D. E. (2010, August). Revisiting the tradespace exploration paradigm: structuring the exploration process. In AIAA Space.
79. Rothrock, L., & Narayanan, S. (2011). Human-in-the-Loop Simulations: Methods and Practice. Springer Science & Business Media.
80. Roy, B. et Bouyssou D., 1993, Aide multicritère à la décision: Méthodes et cas. Economica, Paris.
81. Ruault, J. R., Kolski, C., & Vanderhaegen, F. (2012, October). Persona pour la conception de systèmes complexes résilients. In Proceedings of the 2012 Conference on Ergonomie et Interaction homme-machine (p. 17). ACM.
82. Russell, A. D., Chiu, C. Y., & Korde, T. (2009). Visual representation of construction management data. Automation in Construction, 18(8), 1045-1062.
83. Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. Visualization and Computer Graphics, IEEE Transactions on, 11(4), 443-456.
84. Scaravetti, D. (2004). Formulation préalable d'un problème de conception, pour l'aide à la décision en conception préliminaire (Doctoral dissertation, Ecole nationale supérieure d'arts et métiers-ENSAM).
85. Schneider, D. K. (1994). Modélisation de la démarche du décideur politique dans la perspective de l'intelligence artificielle. Thèse de l'Université de Genève, Suisse.
86. Scott Morton M. (1971): “Management decision systems, computer based support for decision making”, Harvard University, Boston (MA, USA).
87. Scott, M. J., & Antonsson, E. K. (2000, September). Using indifference points in engineering decisions. In Proceedings of ASME Design Engineering Technical Conferences, Baltimore, USA.
88. Scott, M. J., & Antonsson, E. K. (2005). Compensation and weights for trade-offs in engineering design: beyond the weighted sum. Journal of Mechanical Design, 127(6), 1045-1055.
89. Seguy, A. (2008). Décision collaborative dans les systèmes distribués: application à la e-maintenance (Doctoral dissertation).
90. Shanteau J (1992), Competence in experts: the role of task characteristics. Organ Behavior and Human Decision 53(2):252-266.
91. Shneiderman, B. (1992). Tree visualization with tree-maps: 2-d space-filling approach. ACM Transactions on graphics (TOG), 11(1), 92-99.
92. Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In Visual Languages, 1996. Proceedings., IEEE Symposium on (pp. 336-343). IEEE.

93. Shneiderman, B., & Wattenberg, M. (2001, October). Ordered treemap layouts. In Information Visualization, IEEE Symposium on (pp. 73-73). IEEE Computer Society.
94. Siirtola, H. (2007). Interactive visualization of multidimensional data. Tampereen yliopisto.
95. Simon, H. A. (1992). What is an "explanation" of behavior?. Psychological Science, 3(3), 150-161.
96. Simpson, T. W., Carlsen, D. E., Congdon, C. D., Stump, G., & Yukish, M. A. (2008). Trade space exploration of a wing design problem using visual steering and multi-dimensional data visualization. In 4th AIAA Multidisciplinary Design Optimization Specialist Conference (pp. 7-10).
97. Simpson, T. W., Spencer, D. B., Yukish, M. A., & Stump, G. (2008, September). Visual steering commands and test problems to support research in trade space exploration. In 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (pp. 10-12).
98. Sobek, D. K., Ward, A. C., & Liker, J. K. (1999). Toyota's principles of set-based concurrent engineering. Sloan management review, 40(2), 67-84.
99. Stadler, W. (Ed.). (1988). Multicriteria Optimization in Engineering and in the Sciences (Vol. 37). Springer Science & Business Media.
100. Stal-Le Cardinal, J. (2000). Etude des dysfonctionnements dans la prise de décision. Application au choix d'acteur (Doctoral dissertation, Ecole Centrale Paris).
101. Stump, G. M., Yukish, M., Simpson, T. W., & O'Hara, J. J. (2004, March). Trade space exploration of satellite datasets using a design by shopping paradigm. In Aerospace Conference, 2004. Proceedings. 2004 IEEE (Vol. 6, pp. 3885-3895). IEEE.
102. Stump, G., Lego, S., Yukish, M., Simpson, T. W., & Donndelinger, J. A. (2009). Visual steering commands for trade space exploration: User-guided sampling with example. Journal of Computing and Information Science in Engineering, 9(4), 044501.
103. Stump, G., Simpson, T. W., Donndelinger, J. A., Lego, S., & Yukish, M. (2009). Visual steering commands for trade space exploration: User-guided sampling with example. Journal of Computing and Information Science in Engineering, 9(4), 044501.
104. Stump, G., Yukish, M., Martin, J. D., & Simpson, T. W. (2004, August). The ARL trade space visualizer: An engineering decision-making tool. In 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (Vol. 30).
105. Sullivan, K. J., Griswold, W. G., Cai, Y., & Hallen, B. (2001, September). The structure and value of modularity in software design. In ACM SIGSOFT Software Engineering Notes (Vol. 26, No. 5, pp. 99-108). ACM.
106. Wang, G. G., & Shan, S. (2007). Review of metamodeling techniques in support of engineering design optimization. Journal of Mechanical Design, 129(4), 370-380.
107. Ware, C. (2004). Information Visualization: Perception for Design (Interactive Technologies). Morgan Kaufmann 2nd edition.

108. Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. Journal of the American Statistical Association, 85(411), 664-675

109. Wickens, C. D. (1992). Engineering psychology and human performance. HarperCollins Publishers.

110. Wiederhold, G. (1992). Mediators in the architecture of future information systems. Computer, 25(3), 38-49.

111. Wilkins, B. (2003). MELD: a pattern supported methodology for visualisation design (Doctoral dissertation, University of Birmingham).

112. WilsonTD. Schooler JW. (1991). Thinking too much: introspection can reduce the quality of preferences and decisions. J Pers Soc Psychol 60(2):181-192.

113. Wolf, D., Hyland, J., Simpson, T. W., & Zhang, X. L. (2011). The Importance of Training for Interactive Trade Space Exploration: A Study of Novice and Expert Users. Journal of Computing and Information Science in Engineering, 11(3), 031009.

114. Wolf, D., Simpson, T. W., & Zhang, X. L. (2009, January). A preliminary study of novice and expert users' decision-making procedures during visual trade space exploration. In ASME 2009 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 1361-1371). American Society of Mechanical Engineers.

115. Womack, J. P., & Jones, D. T. (2010). Lean thinking: banish waste and create wealth in your corporation. Simon and Schuster.

116. Wood, K. L., Otto, K. N., & Antonsson, E. K. (1992). Engineering design calculations with fuzzy parameters. Fuzzy Sets and Systems, 52(1), 1-20.

117. Yan, X., Qiao, M., Li, J., Simpson, T. W., Stump, G. M., & Zhang, X. (2012, January). A Work-Centered Visual Analytics Model to Support Engineering Design with Interactive Visualization and Data-Mining. In System Science (HICSS), 2012 45th Hawaii International Conference on (pp. 1845-1854). IEEE.

118. Yan, X., Qiao, M., Simpson, T. W., Li, J., & Zhang, X. L. (2011, January). LIVE: A Work-Centered Approach to Support Visual Analytics of Multi-Dimensional Engineering Design Data With Interactive Visualization and Data-Mining. In ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 825-834). American Society of Mechanical Engineers.

119. Yannou, B., & Hamdi, A. (2004). Truss dimensioning with an uncertainty reduction paradigm. In DS 32: Proceedings of DESIGN 2004, the 8th International Design Conference, Dubrovnik, Croatia.

120. Yannou, B., Moreno, F., Thevenot, H. J., & Simpson, T. W. (2005, January). Faster generation of feasible design points. In ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 355-363). American Society of Mechanical Engineers.

121. Yannou, B., Troussier, N., Chateauneuf, A., & Boudaoud, N. (2009). Dimensioning a product in preliminary design through different exploration techniques. International Journal of Product Development, 9(1), 140-163.

122. Yannou, B., Yvars, P. A., Hoyle, C., & Chen, W. (2013). Set-based design by simulation of usage scenario coverage. Journal of Engineering Design, 24(8), 575-603.

123. Yoon, K. P., & Hwang, C. L. (1995). Multiple attribute decision making: an introduction (Vol. 104). Sage Publications.

124. Zanakis, S. H., Solomon, A., Wishart, N., & Dublish, S. (1998). Multi-attribute decision making: A simulation comparison of select methods. European journal of operational research, 107(3), 507-529.

125. Zhang, X., Simpson, T., Frecker, M., & Lesieutre, G. (2012). Supporting knowledge exploration and discovery in multi-dimensional data with interactive multiscale visualisation. Journal of Engineering Design, 23(1), 23-47.

ANNEXES

A.1. Encodage Visuel

Bertin (1973) a listé 7 variables visuelles et a présenté les effets de la variation des propriétés perceptives des variables visuelles afin d'en tirer des représentations pertinentes. Il y a 2 variables « planar » (les positions x et y) et 5 dites « retinal » (la taille, la couleur, la teinte, la forme et l'orientation) traduit par certains chercheurs comme « pré-attentive » (Tidwell par exemple).

Bertin distingue *selective, associative, ordered* et *quantitative* variables visuelles. Une variable visuelle est *selective* si tous les symboles peuvent être facilement isolés (sélection perceptuelle). Une variable visuelle est *associative* si elle permet de grouper de façon perceptuelle toutes catégories de symboles en se basant sur leur caractéristique visuelle particulière. Pour Bertin, seule les variables taille et teinte ont des caractéristiques perceptuelles dissociatives. Les variables dissociative sont soit *ordered* soit *quantitative*. Une variable est dite *ordered* s'il est possible de ranger de façon perceptuelle les symboles en se basant sur leur caractéristique visuelle particulière. S'il est possible de quantifier de façon perceptuelle le degré de variation du symbole visuel, la variable visuelle est définie comme *quantitative*. (Voir des illustrations dans la **Figure 122**, la **Figure 123** et la **Figure 124**)

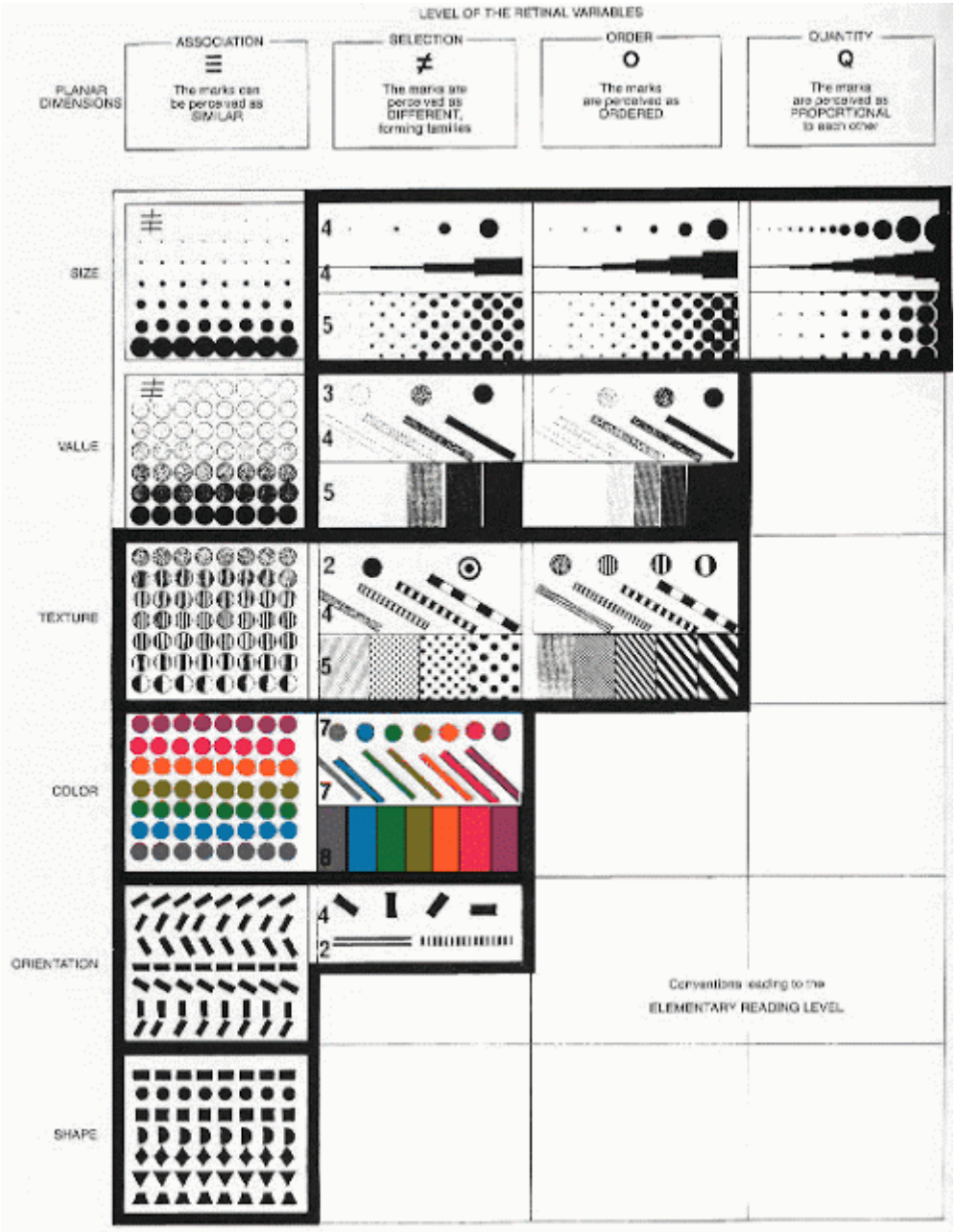


Figure 122 : Variables visuelles de Bertin 1/3

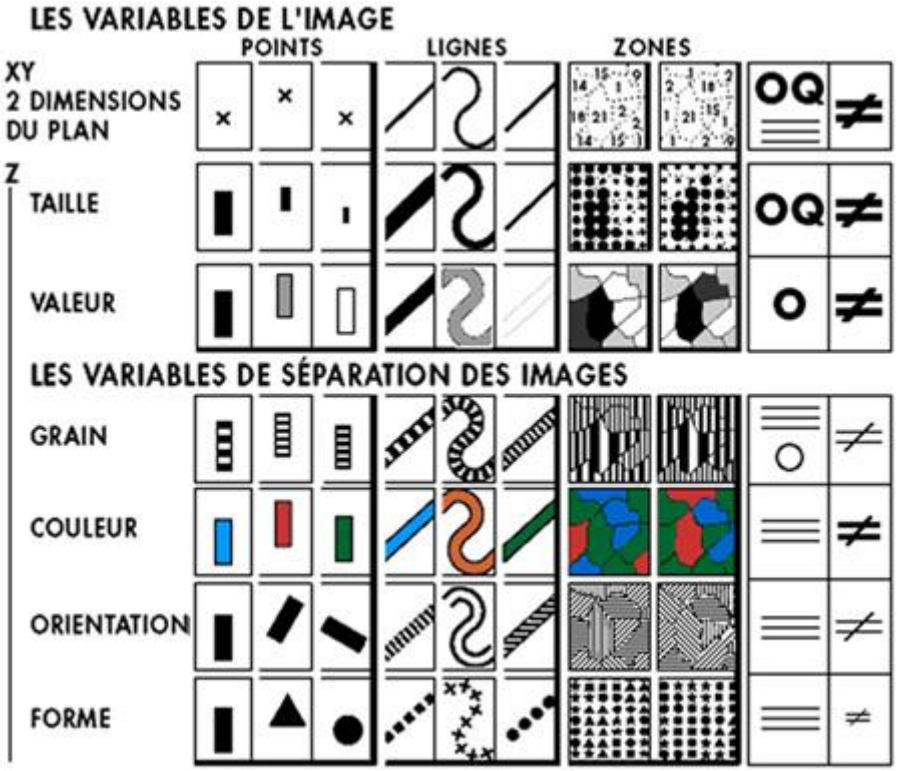


Figure 123 : Variables visuelles de Bertin 2/3

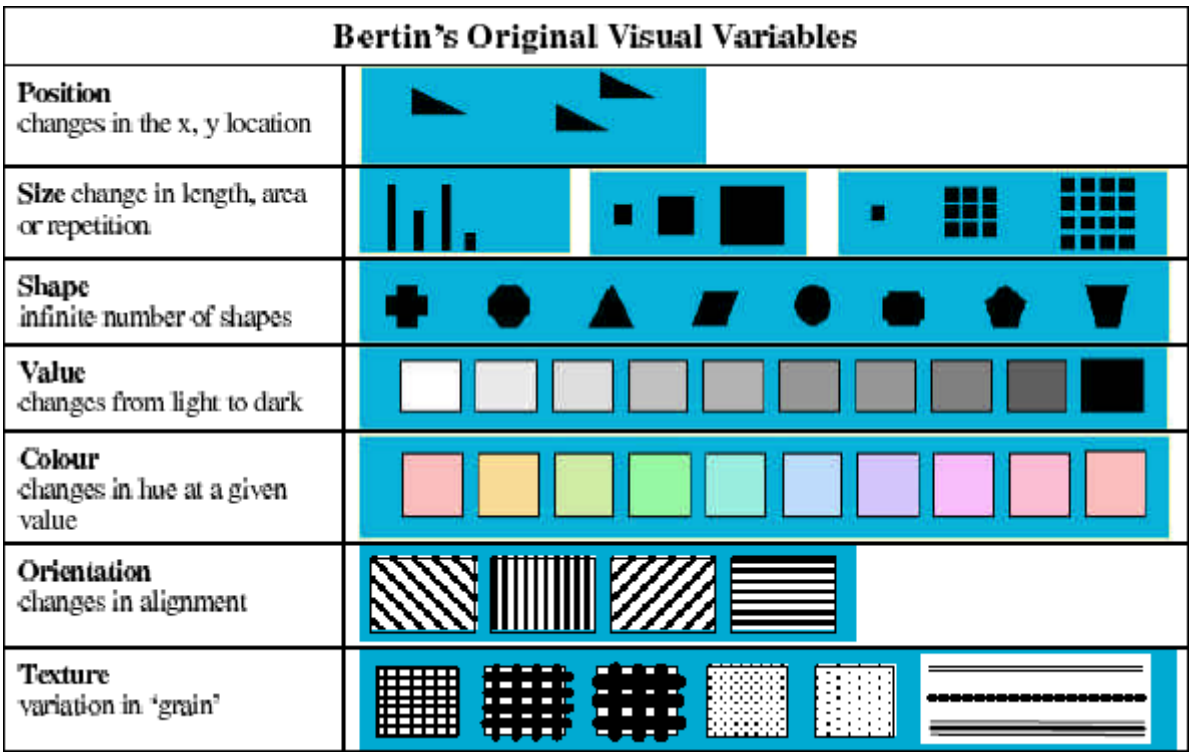


Figure 124 : Variables visuelles de Bertin 3/3

A.2. Capture d'écran de la table de données des voitures

	Info générale				Mécanique					Mesures								Performances				
											Dimension				Volume coffre (l)				Temps			
	Marque	Modèle	Année	Prix	Cylindrée	Nombre de cylindre	Puissance		Couple	Poids à vide (kg)	L	l	h	Réservoir	min	max	Diamètre des jantes	Vitesse max	0 à 100	Consommation	CO2	Niveau
1	Fiat	500 Abarth	2012	18 950€	1.4L	4	135ch	5500tr/min	206 Nm	1035	3.66	1.63	1.48	351	185	610	195	205 km/h	7.9 s	6.6 l/100	155	D
2	Volkswagen	Polo 1.6 TDI 90 DSG	2012	16 150€	1.6L	4	90ch	4200tr/min	229 Nm	1157	3.97	1.69	1.49	451	280	952	185	180 km/h	11.5 s	4.2 l/100	112	B
3	Dacia	Sandero 1.2 16V GPL	2012	8 500€	1.4L	4	75ch	5500tr/min	107 Nm	1067	4.02	1.75	1.54	841	320	1200	185	161 km/h	14.5 s	6 l/100	139	C
4	Renault	Twingo Gordini	2012	15 250€	1.2L	4	100ch	5500tr/min	155 Nm	980	3.68	1.65	1.47	401	165	959	185	189 km/h	9.8 s	5.7 l/100	130	C
5	Citroen	C1 1.0i	2012	9 450€	1.0L	3	68ch	6000tr/min	93 Nm	790	3.43	1.63	1.47	351	130	300	155	157 km/h	14.2 s	4.5 l/100	103	B
6	Skoda	Fabia 1.2 TSI 105 DSG7	2012	17 730€	1.2L	4	105ch	5000tr/min	252 Nm	1080	4.00	1.65	1.50	451	712	1180	195	189 km/h	9.9 s	5.3 l/100	124	C
7	Mitsubishi	Colt Ralliart	2011	17 890€	1.5L	4	150ch	6000tr/min	209 Nm	1060	3.88	1.70	1.52	471	160	568	205	210 km/h	7.4 s	6.9 l/100	161	E
8	Toyota	iQ 100 VVT-i	2011	15 800€	1.3L	3	98ch	6000tr/min	122 Nm	930	2.99	1.68	1.50	321	132	242	175	170 km/h	11.8 s	4.8 l/100	113	B
9	Alfa-Romeo	MiTo Quadrifoglio Verde	2012	22 400€	1.4L	4	170ch	5000tr/min	230 Nm	1145	4.07	1.72	1.45	451	270	687	215	219 km/h	7.5 s	6 l/100	139	C
10	Seat	Ibiza 1.4i 16V	2012	13 515€	1.4L	4	85ch	5000tr/min	132 Nm	1000	4.04	1.70	1.43	451	284	284	185	177 km/h	11.8 s	5.9 l/100	139	C
11	Opel	Agila 1.2	2011	12 890€	1.2L	4	86ch	5500tr/min	113 Nm	1120	3.74	1.68	1.59	451	225	1050	185	175 km/h	12.3 s	5.5 l/100	130	C
12	Ford	Fiesta 1.4 TDCi	2011	14 600€	1.4L	4	68ch	4000tr/min	159 Nm	1086	3.95	1.73	1.44	421	295	979	175	162 km/h	14.9 s	4.2 l/100	110	B
13	Volvo	V40 D3	2012	26 600€	2.0L	5	150ch	3500tr/min	350 Nm	1404	4.37	1.80	1.45	601	335	1032	205	210 km/h	9.6 s	4.3 l/100	114	B
14	Kia	Pro Ceed 1.6 CRDi	2012	22 590€	1.6L	4	128ch	4000tr/min	260 Nm	1358	4.25	1.79	1.45	531	340	1130	205	197 km/h	11 s	4.5 l/100	117	B
15	Nissan	Note 1.5 dCi 90	2012	17 400€	1.5L	4	90ch	4000tr/min	200 Nm	1180	4.09	1.69	1.55	461	280	1332	175	168 km/h	12.8 s	4.2 l/100	110	B
16	Opel	Corsa CDTI ecoFLEX 95	2012	17 990€	1.4L	4	95ch	4000tr/min	190 Nm	1163	4.00	1.72	1.49	451	285	1100	185	177 km/h	12.3 s	3.6 l/100	94	A
17	Lancia	Delta 1.9 Multijet 190	2012	30 500€	1.9L	5	190ch	4000tr/min	400 Nm	1430	4.52	1.80	1.50	581	380	1190	205	222 km/h	7.9 s	5.7 l/100	149	D
18	Seat	Leon 1.6 TDI 105 Ecomotive	2012	21 605€	1.6L	4	105ch	4400tr/min	250 Nm	1325	4.32	1.77	1.46	551	341	1166	195	185 km/h	11.7 s	4.2 l/100	109	B
19	Honda	Jazz 1.4 i-VTEC	2012	15 600€	1.4L	4	100ch	6000tr/min	126 Nm	975	3.90	1.70	1.53	421	379	863	175	182 km/h	11.4 s	5.4 l/100	128	C
20	Skoda	Citigo 60	2012	9 090€	1.0L	3	60ch	5000tr/min	105 Nm	929	3.56	1.64	1.48	351	251	951	165	160 km/h	14.4 s	4.5 l/100	105	B
21	Audi	A1 Quattro	2012	51 190€	2.0L	4	256ch	6000tr/min	350 Nm	1390	3.99	1.74	1.41	451	210	860	225	245 km/h	5.7 s	8.6 l/100	199	E
22	Chevrolet	Spark 1.0 16V	2012	8 990€	1.0L	4	68ch	6100tr/min	93 Nm	939	3.64	1.60	1.53	351	170	980	155	154 km/h	15.5 s	5.1 l/100	119	B
23	Mini	Mini Cooper S	2012	24 400€	1.6L	4	184ch	5500tr/min	260 Nm	1205	3.73	1.69	1.41	501	160	680	195	228 km/h	7 s	5.8 l/100	136	C
24	Smart	Fortwo mhd 71	2012	10 850€	1.0L	3	71ch	5800tr/min	92 Nm	750	2.70	1.56	1.55	331	220	340	155	145 km/h	13.7 s	4.2 l/100	97	A
25	Bmw	Serie 1 116d	2012	25 200€	2.0L	4	116ch	4000tr/min	260 Nm	1385	4.33	1.77	1.43	521	360	1200	195	200 km/h	10.3 s	4.5 l/100	117	B
26	Ford	Focus 2.0 TDCi 140	2012	24 850€	2.0L	4	140ch	3750tr/min	313 Nm	1421	4.36	1.86	1.48	551	277	1148	205	207 km/h	8.9 s	5 l/100	129	C
27	Mercedes	Classe A 160 BlueEFFICIENCY	2011	18 950€	1.5L	4	95ch	5200tr/min	140 Nm	1215	3.89	1.77	1.6	541	435	1370	185	175 km/h	12.6 s	6 l/100	139	C
28	Dacia	Logan 1.5 dCi 90	2011	11 750€	1.5L	4	90ch	3750tr/min	196 Nm	1075	4.29	1.74	1.53	501	510	510	185	173 km/h	12.8 s	4.1 l/100	108	B
29	Peugeot	207 1.6 HDi 92	2011	18 100€	1.6L	4	92ch	4000tr/min	225 Nm	1214	4.03	1.72	1.48	501	270	923	185	183 km/h	11.9 s	4.2 l/100	110	B
30	Nissan	Micra DIG-S	2011	10 490€	1.2L	3	98ch	6000tr/min	137 Nm	980	3.78	1.66	1.52	451	265	1132	165	180 km/h	11 s	4.1 l/100	95	A
31	Peugeot	107 1.0	2011	9 450€	1.0L	3	68ch	6000tr/min	93 Nm	790	3.43	1.63	1.47	351	130	712	155	157 km/h	13.7 s	4.5 l/100	103	B
32	Lancia	Ypsilon 1.3 Multijet 95 DPF	2011	15 500€	1.3L	4	95ch	4000tr/min	145 Nm	1050	3.84	1.68	1.52	401	245	245	185	183 km/h	11.4 s	3.8 l/100	99	A
33	Volkswagen	Golf 6 GTD	2012	31 240€	2.0L	4	170ch	4200tr/min	350 Nm	1329	4.22	1.79	1.51	551	350	1305	225	222 km/h	8.1 s	5.3 l/100	139	C
34	Renault	Megane 3 Energy dCi 110	2012	24 350€	1.5L	4	110ch	4000tr/min	260 Nm	1205	4.30	1.81	1.48	601	405	1162	195	195 km/h	9.8 s	3.5 l/100	90	A
35	Renault	Clio 3 1.5 dCi 90	2011	17 250€	1.5L	4	90ch	4000tr/min	196 Nm	1165	4.03	1.71	1.50	551	288	1038	185	175 km/h	11 s	4 l/100	106	B
36	Daihatsu	Sirion 1.0 12S	2011	11 990€	1.0L	3	69ch	6000tr/min	92 Nm	890	3.60	1.66	1.55	401	225	630	175	160 km/h	14 s	5.1 l/100	118	B
37	Fiat	Punto Evo 1.2 8V S&S	2011	12 390€	1.2L	4	69ch	5500tr/min	101 Nm	1015	4.07	1.69	1.49	451	275	1020	175	156 km/h	14.4 s	5.2 l/100	123	C
38	Toyota	Yaris 90 D4D	2011	14 620€	1.4L	4	90ch	4000tr/min	190 Nm	1055	3.75	1.70	1.53	421	272	737	185	175 km/h	10.7 s	4.1 l/100	109	B
39	Fiat	Panda 1.2 8V	2011	9 350€	1.2L	4	69ch	5500tr/min	101 Nm	860	3.54	1.59	1.54	351	206	775	155	162 km/h	13.4 s	4.9 l/100	110	B
40	Chevrolet	Aveo 1.2 16V	2011	10 090€	1.2L	4	84ch	6000tr/min	113 Nm	1062	3.92	1.68	1.51	451	220	980	185	170 km/h	12.8 s	5.5 l/100	132	C

A.3. Liste des solutions avec le nombre de dominations et la distance au front de Pareto pour le scénario TAILLE

Numéro de la solution	Nbr de solutions que la solution domine	Distance au front de Pareto
27	7	0
30	4	0
11	3	0
22	3	0
39	3	0
33	3	0
2	3	0
6	2	0
20	2	0
19	2	0
38	2	0
35	2	0
16	2	0
24	1	0
15	1	0
3	1	0
31	1	0
37	1	0
12	1	0

Numéro de la solution	Nbr de solutions que la solution domine	Distance au front de Pareto
28	0	0
40		0,16768418
4		0,19505755
29		0,12032592
1		0,22348018
7		0,21872578
17		0,22957942
34		0,27506728
18		0,30368599
14		0,29337979
25		0,31401358
5		0,39163498
9		0,27292719
21		0,22113228
8		0,52687205
26		0,32772033
32		0,44289464
13		0,33962521
23		0,36800537

36	0	0	10		0,66205728

A.4. Liste des solutions avec le nombre de dominations et la distance au front de Pareto pour le scénario PUISSANCE

Numéro de la solution	Nbr de solutions que la solution domine	Distance au front de Pareto	Numéro de la solution	Nbr de solutions que la solution domine	Distance au front de Pareto
1	11	0	32		0,12930975
6	9	0	10		0,10841606
7	8	0	38		0,23604043
4	8	0	2		0,24292207
23	6	0	36		0,08335862
9	6	0	35		0,29451435
33	4	0	29		0,31010544
19	4	0	14		0,4478622
8	4	0	28		0,32309509
39	3	0	25		0,46772747
21	2	0	40		0,23248631
31	1	0	11		0,25629439
5	1	0	37		0,14096797
17	0	0	18		0,40643568
24	0	0	16		0,36221312
20	0	0	22		0,13999684
13		0,245774	3		0,21745883
30		0,05915802	12		0,29784972
34		0,19702953	27		0,383638
26		0,29210839	15		0,4331915

A.5. Liste des solutions avec le nombre de dominations et la distance au front de Pareto pour le scénario TRAJET

Numéro de la solution	Nbr de solutions que la solution domine	Distance au front de Pareto	Numéro de la solution	Nbr de solutions que la solution domine	Distance au front de Pareto
34	27	0	10		0,38999862
28	20	0	15		0,45897921
13	15	0	32		0,53643699
14	15	0	37		0,4708063
33	9	0	23		0,54357684
26	9	0	11		0,50966833
17	5	0	40		0,51578615
6	2	0	30		0,5824075
3	0	0	7		0,59507323
27	0	0	21		0,73573729

18		0,172612449
25		0,162491023
35		0,406360702
9		0,386740486
16		0,41564367
29		0,396650681
2		0,439979722
12		0,42666195
38		0,456808507
19		0,388852351

36		0,62343739
4		0,67168809
20		0,69292539
8		0,7825575
1		0,75501569
5		0,91311841
31		0,91311841
39		0,90239966
24		0,94827284
22		0,92972516

A.6. Support pour la phase de formation de l'expérimentation 2

Exploration et Visualisation d'un espace de conception

Training

EXPLORATION

Introduction

Exploration de l'espace de conception (Design Space Exploration):

- **L'espace de conception** (design space): Ensemble de toutes les alternatives de conception possibles (l'Espace des possibles)
- **L'exploration** fait référence à l'activité systématique de découverte et d'évaluation des éléments de l'espace de conception

Vocabulaire:

Les alternatives:
- Solutions possibles
- Candidats
- Points de conception (design points)

L'espace de conception (design space) est aussi appelé « TradeSpace ». Trade fait référence au « trade-off » (compromis) que le concepteur doit souvent réaliser.

Le but:

Rechercher dans un grand espace de conception des alternatives qui puissent satisfaire l'ensemble des contraintes et qui soient optimales au regard des objectifs

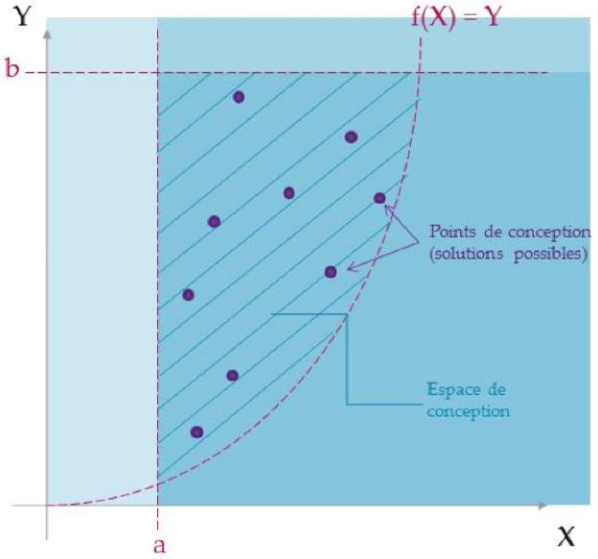
Illustration

Du problème à l'espace de conception

Imaginons un problème de conception composé de 2 variables X et Y tel que:

- C1: X > a
- C2: Y < b
- C3: Y > f(X)

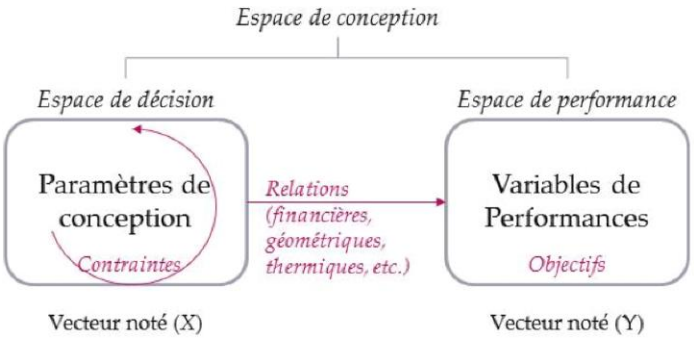
jouer



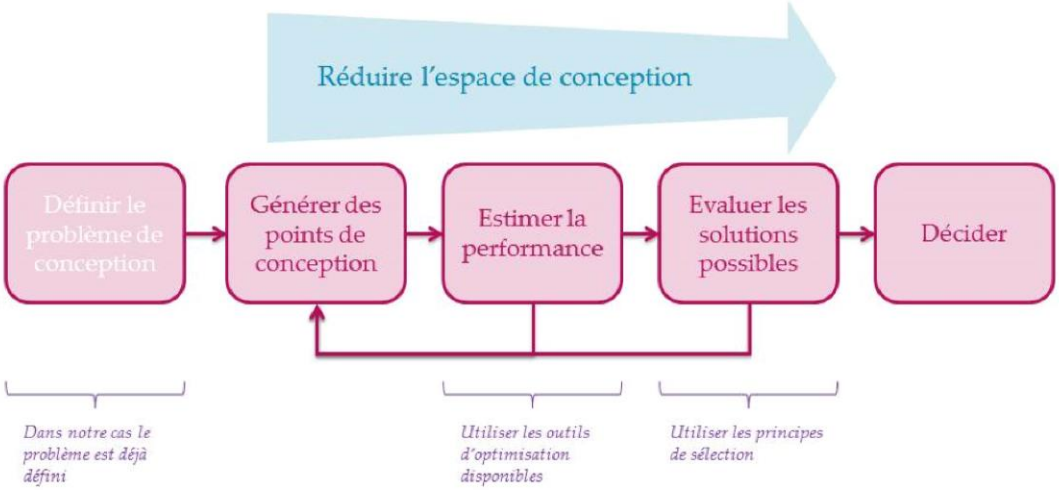
Le problème de conception

Le problème est composé de:

- **Paramètres de conception** (des dimensions par exemple). Ces paramètres sont liés entre eux par des contraintes (égalité ou inégalité). Les paramètres de conception forment « l'espace de décision » et le vecteur des paramètres de conception qui composent la solution est souvent noté (X).
- **Variables de performances** (une puissance, une taille ou un temps par exemple). Les variables de performances sont liées aux paramètres de conception par des relations/équations qui peuvent être financières, géométriques, thermiques, etc. (tout dépend de la variable de performance). Les objectifs du problème de conception se trouvent sur les variables de performances et le vecteur des performances de conception des solutions faisables est souvent noté (Y).



Les étapes



Optimiser

Les objectifs du problème de conception

Les **objectifs** du problème de conception sont de **minimiser** et/ou **maximiser** les variables de performance.

Dans certains problèmes de conception, nous pouvons trouver des « poids » sur les variables de performance. Dans notre cas nous considérons des poids égaux pour toutes les variables de performance (poids = 1). C'est-à-dire que toutes les variables de performance ont la même importance.

Par ailleurs, nous ne considérerons que 2 « outils » pour l'optimisation:

- L'indice de qualité (vecteur de préférence)
- La dominance au sens de Pareto

Optimiser

Indice de qualité (vecteur de préférence)

L'indice de qualité est un indice calculé pour chaque solution possible. Il s'agit d'une estimation (mathématique) de la performance. L'indice de qualité est compris entre 0 et 1, 1 étant la solution idéale et 0 la solution « anti-idéale » ou « nadir » (en général l'idéal et le nadir n'existent pas).

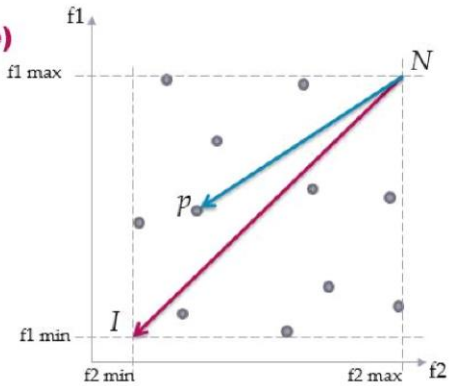


Illustration où f1 et f2 sont à minimiser

Comment obtient-on l'indice de qualité?

Soit $\overrightarrow{NI}$ un vecteur de préférence et *fi* variable de performance où $i=\{1, \ldots, n\}$ et n est le nombre de variables de performance.

Si *fi* est à minimiser, $fi(N) = fi_{max}$ et $fi(I) = fi_{min}$

Si *fi* est à maximiser, $fi(N) = fi_{min}$ et $fi(I) = fi_{max}$

Ainsi, nous calculons un indice de qualité Ip pour chaque point P (par rapport au vecteur $\overrightarrow{NI}$) tel que:

$$Ip = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{fi(p)-fi(N)}{fi(I)-fi(N)}\right| \qquad \text{Notons que} \quad Ip \in [0;1]$$

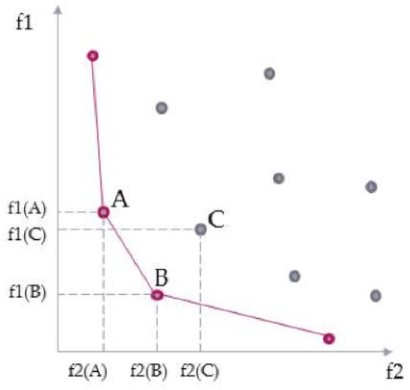
Optimiser

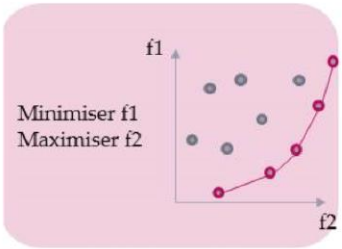
Dominance au sens de Pareto

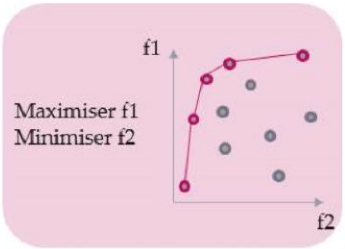
Considérons un problème où il faut minimiser f1 et f2.

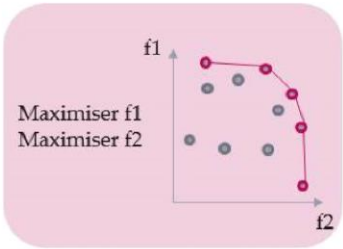
Les points A et B sont dits Pareto-optimaux ou non-dominés au sens de Pareto si:
[f1(A) ≥ f1(B)] & [f2(A) ≤ f2(B)]

Ainsi, l'ensemble des points non-dominés forme la frontière (ou « front ») de Pareto (en rouge sur le graphique). Pour la prise de décision, nous considérons que chaque point sur le front de Pareto est optimal (A n'est pas meilleur que B et B n'est pas meilleur que A).









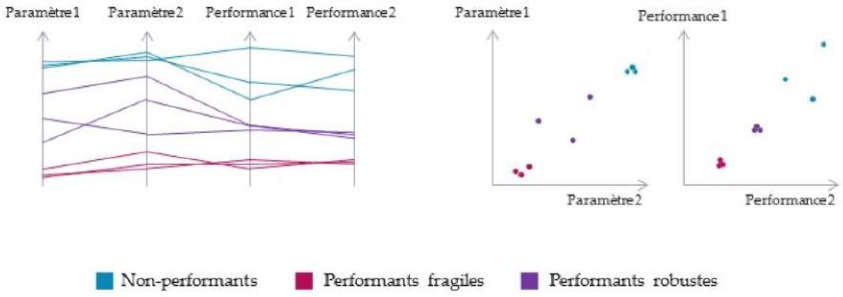
La sélection

Choix d'une solution optimale

Robustesse

La robustesse d'un système se définit comme la « stabilité de sa performance ».

Globalement, il s'agit d'une analyse de sensibilité. C'est-à-dire quantifier la part de variabilité apportée par chacune des variables d'entrée (= les paramètres de conception) sur la réponse (= les variables de performance).



La sélection

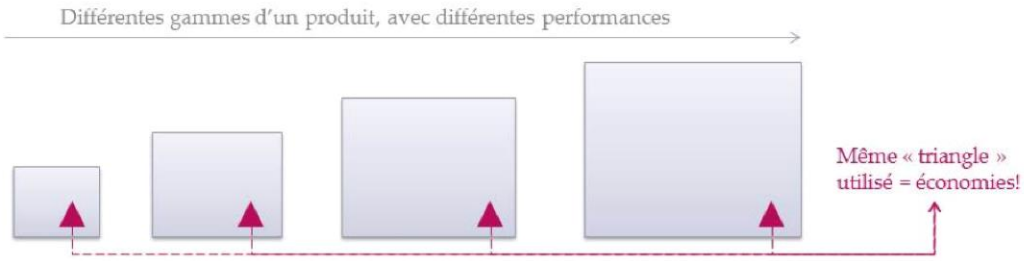
Choix d'une solution optimale

Dimension Industrielle

On appelle ici, « dimension industrielle » la préférence des concepteurs à choisir des solutions comportant des paramètres de conception « standardisable » (parfois même au détriment de solutions performantes-robustes pour des solutions performantes-fragiles).

Imaginons la conception d'un moteur électrique, vous cherchez à maximiser la puissance tout en minimisant la taille du produit. Et vous observez que les solutions Pareto-optimales tendent vers une valeur fixe d'un paramètre (l'armature par exemple). Vous chercherez alors une solution qui a cette valeur fixe.

Cette valeur fixe pourra être standardisée au sein de l'entreprise. Car si vous devez plus tard concevoir d'autres moteurs avec une plus petite puissance , vous pourrez réutiliser la même technologie et la même valeur fixe du paramètre de conception (armature dans l'exemple). Cela entrainera des économies (couts, maintenance, « ingrédients », etc.)



Les interactions entre les variables

VISUALISATION

Introduction

La visualisation de qui ? De quoi?

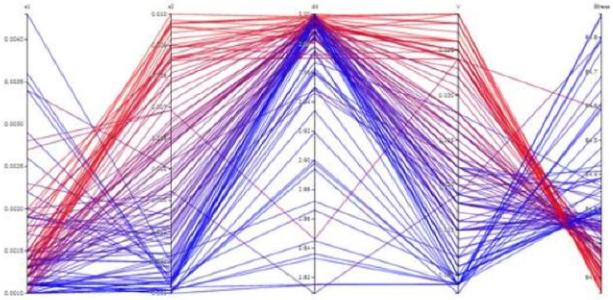
Pour l'activité d'exploration d'espace de conception, les concepteurs utilisent des outils graphiques. Il est pertinent d'employer des moyens visuels pour représenter l'espace de conception et les points de conception. L'utilisation des graphiques est, en effet, une bonne manière de découvrir l'espace de conception (les interactions entre les variables) et d'évaluer les points de conception. Car un concepteur manipule une centaine voire des milliers de solutions possibles dans un espace de conception composé de 5 jusqu'à plus de 20 variables.

Lorsqu'on parle d'interactions entre les variables, on fait référence à un pattern reconnaissable entre les variables (une corrélation, un cluster, une tendance…)

En plus des interactions entre les variables qui permettent de gagner en connaissance et compréhension du problème de conception, les graphiques permettent la découverte « d'indices » guidant le concepteur dans sa démarche.

Introduction

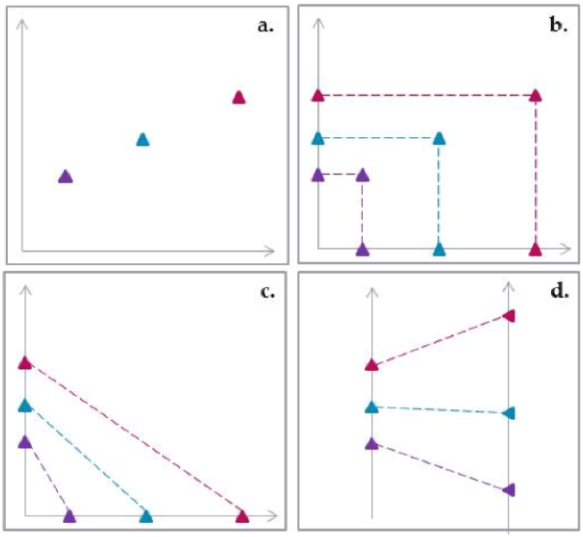
Le graphique qui nous intéresse



« Parallel Coordinate Plot »

La construction des graphiques

Le Parallel Coordinate Plot



Avant d'apprendre à lire le graphique « Parallel Coordinate Plot » (PCP) il est intéressant de comprendre comment il se construit:

1. Partons d'un nuage de points simple avec 3 points de conception (vignette a.)
2. Nous opérons une projection des points sur les axes en ordonnées et en abscisse (vignette b.)
3. Puis nous relions les coordonnées des points projetées sur les 2 axes par des lignes (vignette c.)
4. Enfin nous plaçons les axes en parallèle.

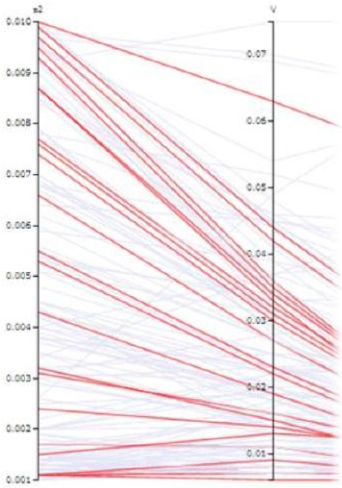
Ainsi, les axes sont les variables et les lignes qui coupent les axes sont les points de conception (ou solutions possibles).

Corrélation

Positive entre 2 variables (A et B)

Plus les valeurs de la variable (A) augmentent et plus les valeurs de la variable (B) augmentent

Les lignes ne s'entrecoupent pas. Elles peuvent aller vers le haut où le bas

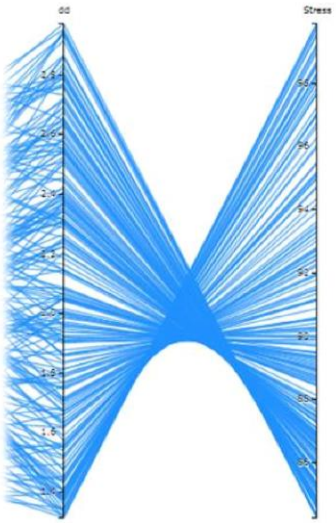


Corrélation

Négative entre 2 variables (A et B)

Plus les valeurs de la variable (A) augmentent et plus les valeurs de la variable (B) diminuent

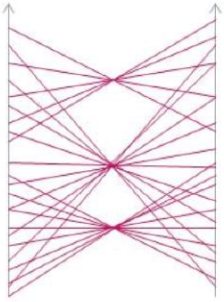
Les lignes s'entrecoupent en un point



Corrélation

3 corrélations négatives entre 2 variables

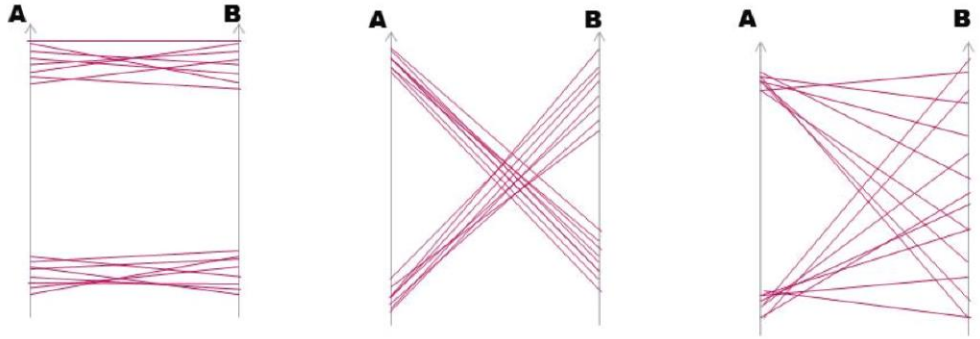
On observe 3 « lots » sur le graphique où les lignes s'entrecoupent en un point



Clusters

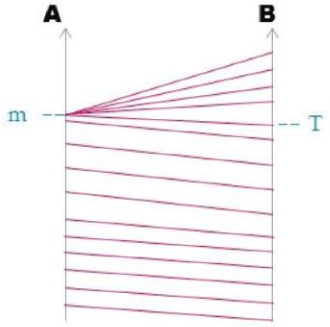
Des exemples

Un cluster est un groupement de solutions

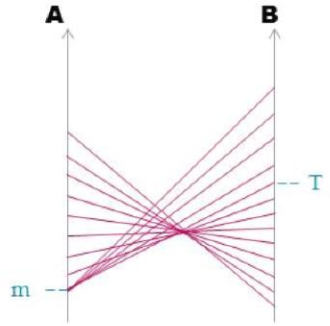


Point de transition

La variable A atteint un maximum (m) à partir d'une certaine valeur de la variable B (T)



La variable A atteint un minimum (m) à partir d'une certaine valeur de la variable B (T)



A.7. Storyboard pour le PCP de l'expérimentation 2

StoryBoard : Parallel Coordinate Plot

Nous proposons un exemple d'utilisation des graphiques et des fonctionnalités à l'aide d'un storyboard illustrant le cas du problème « two-members truss design ».

François, ingénieur, accepte de résoudre le problème intitulé « two-members truss design » à l'aide de notre Plateforme d'expérimentations de l'explorateur de points de conception.

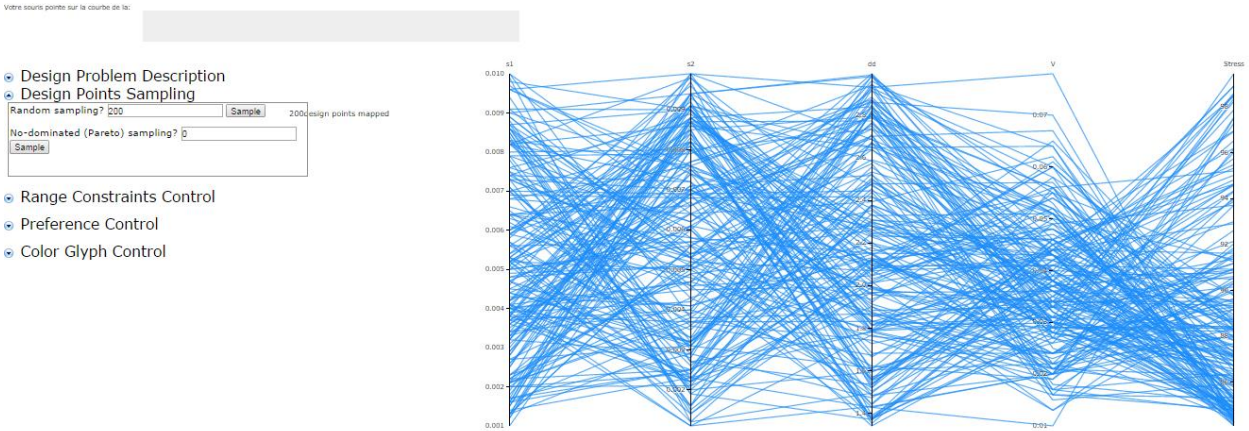
Il démarre avec la description du problème: le 2-members truss. Il s'agit d'un problème de dimensionnement.

Il y a 2 variables de performance qu'il faut minimiser: le volume total des 2 poutres (V) et l'effort maximum développé dans les 2 poutres (Stress).

Il y a 3 variables de décision (ou variables de conception): les sections des 2 poutres (en m²) notées s1 et s2 et la longueur verticale notée dd (en m).

Vous voulez faire comme François ?

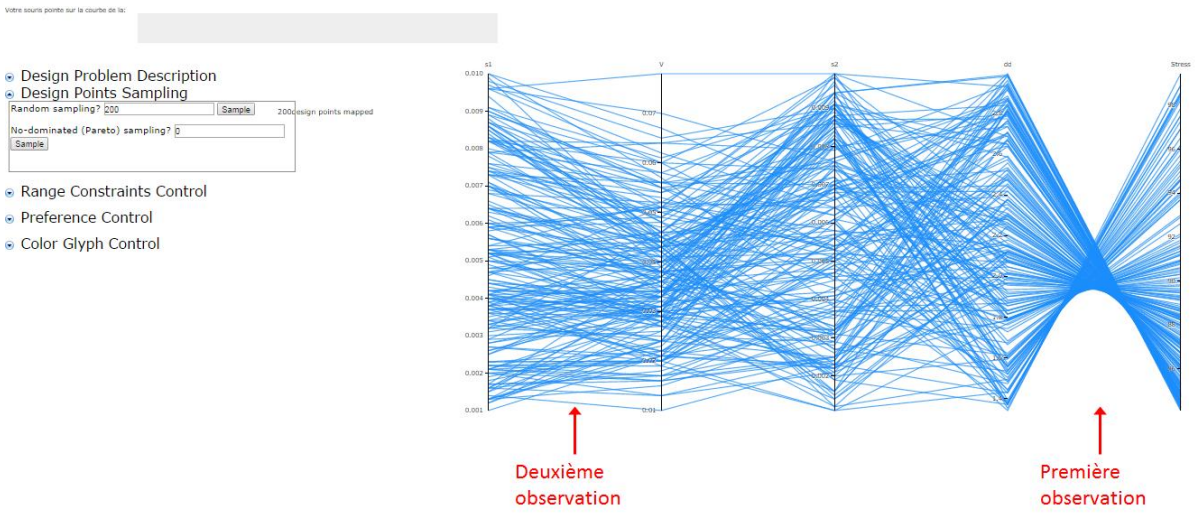
A cette étape, il n'y a rien à faire ! Au lancement de la page, la description du problème s'affiche automatiquement.



François commence par générer 200 points aléatoires dans l'espace de conception afin d'observer les interactions entre les variables d'un point de vue global.

Vous voulez faire comme François ?

Lors de cette étape, commencez par cliquer sur l'onglet « Design Problem Description » pour le « rétracter ». Le bloc « Design Points Sampling » est déjà ouvert. Entrez le nombre 200 dans le champ dédié au « Random Sampling » et cliquez sur le bouton « sample » correspondant.



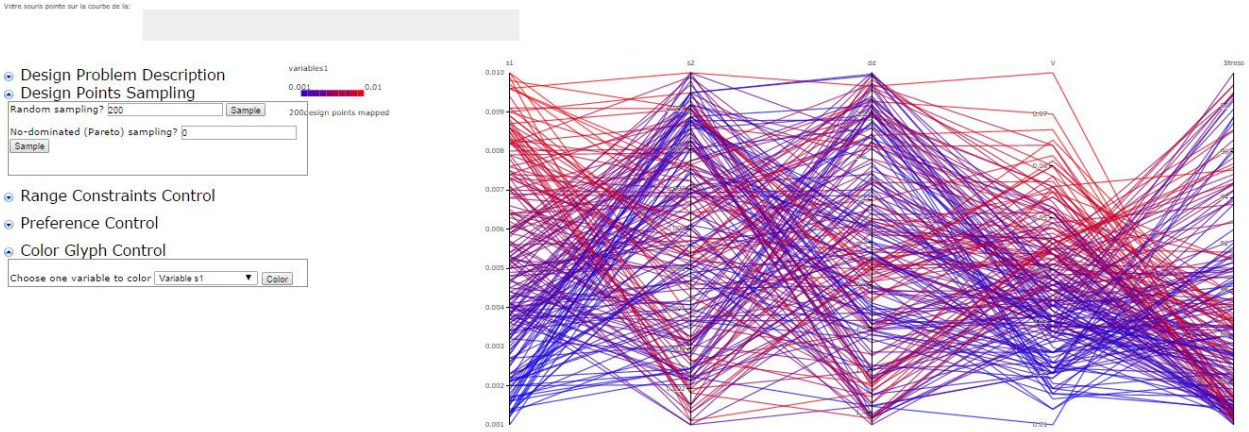
François déplace les axes afin de trouver des interactions possibles entre les variables.

La première observation est une corrélation négative entre dd et la variable de performance « stress »: ainsi plus la longueur des poutres est grande et plus l'effort développé est petit.

Pour la deuxième observation, il semble y avoir une corrélation positive « grossière » entre s1 et V.

Vous voulez faire comme François ?

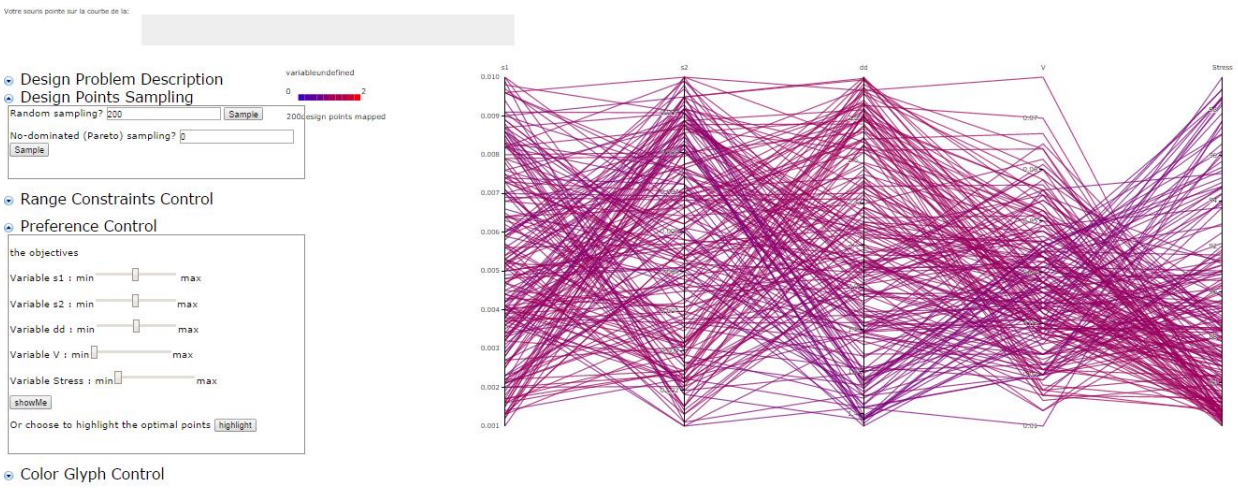
Lors de cette étape, avec votre souris allez au-dessus du label de l'axe V, cliquez, maintenez enfoncez et glissez l'axe pour le déposer entre les axes de s1 et s2.



François décide de mettre un marqueur de couleur sur la variable s1 et s'aperçoit qu'il y a bien une tendance : plus les valeurs de s1 sont grande et plus les valeurs de V sont grandes (bleu pour les petites valeurs et rouge pour les plus grandes).

Vous voulez faire comme François ?

Lors de cette étape, cliquez sur l'onglet « Color Glyph Control » et choisissez la variable s1 dans la liste déroulante et cliquez sur le bouton « color ». (attention les axes se sont replacés à leur position d'origine)



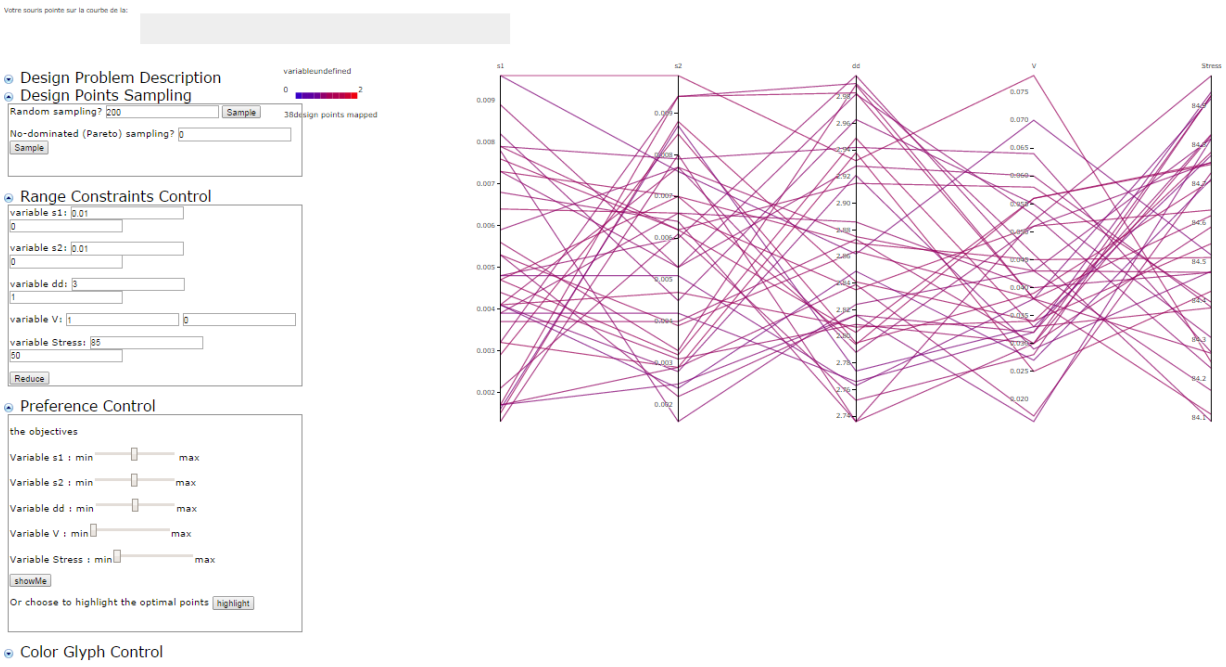
L'ingénieur décide d'observer la "qualité" des points affichés. Il va donc dans le menu préférence et indique qu'il souhaite minimiser les variables Stress et V.

Il se rend compte en observant les couleurs que les solutions possibles ont un indice de qualité compris environ entre 30 et 50% (voir l'échelle de couleur entre 0 et 2 : la couleur violette observée semble être entre le premier tiers et la moitié).

On peut observer un petit mieux pour les points ayant des valeurs de Stress inférieures à 85. Les meilleures solutions semble donc être en dessous de Stress=85.

Vous voulez faire comme François ?

Cliquez sur l'onglet « Preference Control » pour faire apparaitre les fonctionnalités de préférence. Faites glisser les sliders des variables de performance V et Stress vers le minimum et cliquez sur le bouton « showMe ».



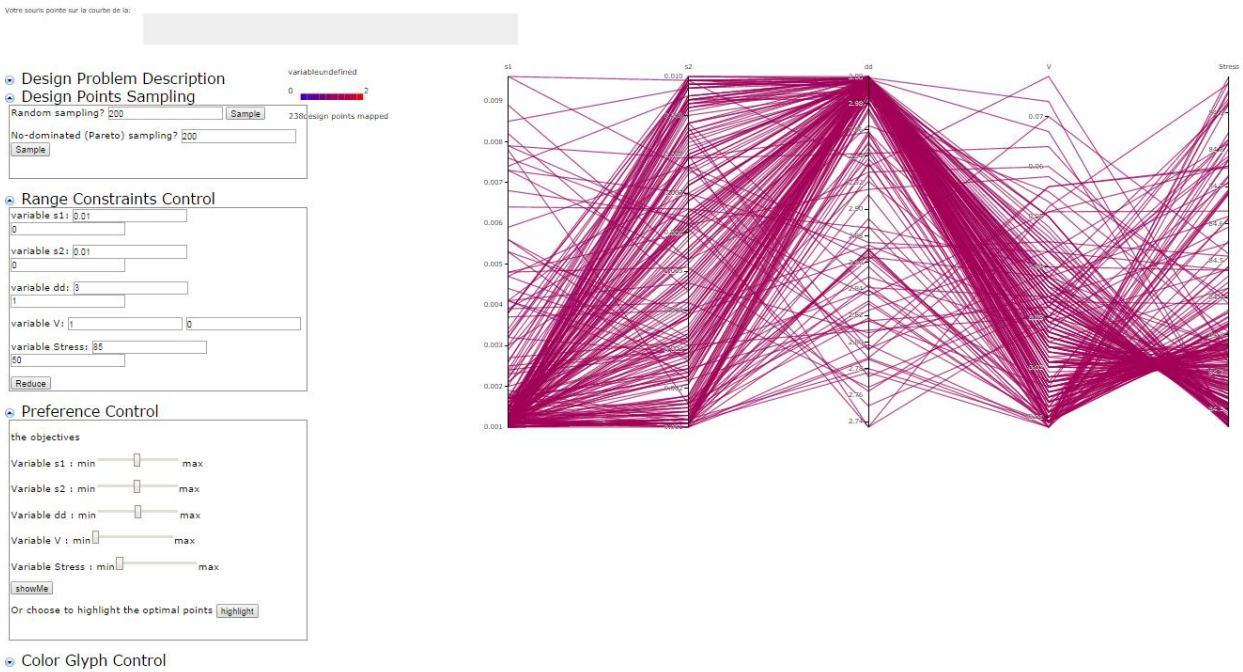
Il décide donc de réduire l'espace de conception en passant de Stress_max = 100 à Stress_max =85.

Il ne reste plus que 38 points dans l'espace de conception.

Vous voulez faire comme François ?

L'idée pour réduire l'espace de conception est de modifier les bornes supérieures et/ou inférieures des variables. Ainsi, cliquez sur l'onglet « Range Constraints Control » puis entrez le nombre 85 dans le champ dédié à la valeur maximum de la variable Stress (qui est à 100 par défaut). Enfin cliquez sur le bouton « Reduce ».

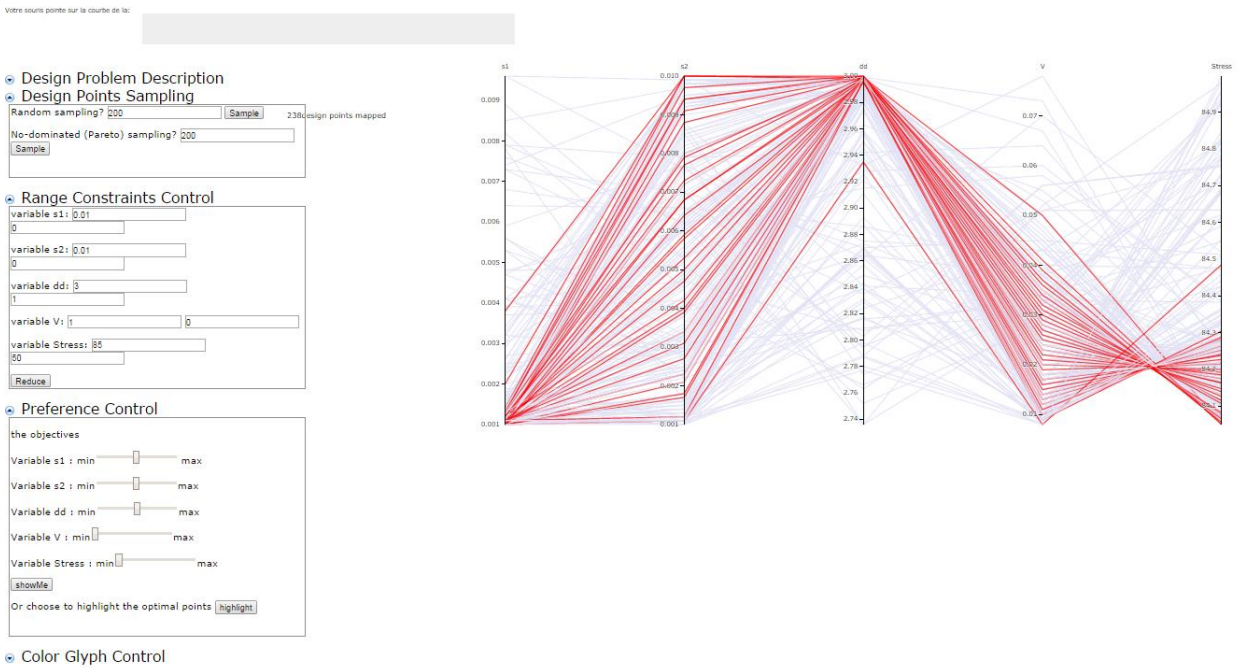
Le nombre de points affichés dans les graphiques est indiqué sous l'échelle de couleur.



38 points n'est pas suffisant. François décide donc de repeupler l'espace de conception (qui a été réduit) avec 200 nouveaux points dits non-dominés.

Vous voulez faire comme François ?

Sous l'onglet « Design Points Sampling » si vous ne l'avez pas rétracté entre temps, entrez le nombre 200 dans le champ dédié au « No-dominated (Pareto) Sampling » et cliquez sur le bouton « sample » correspondant.



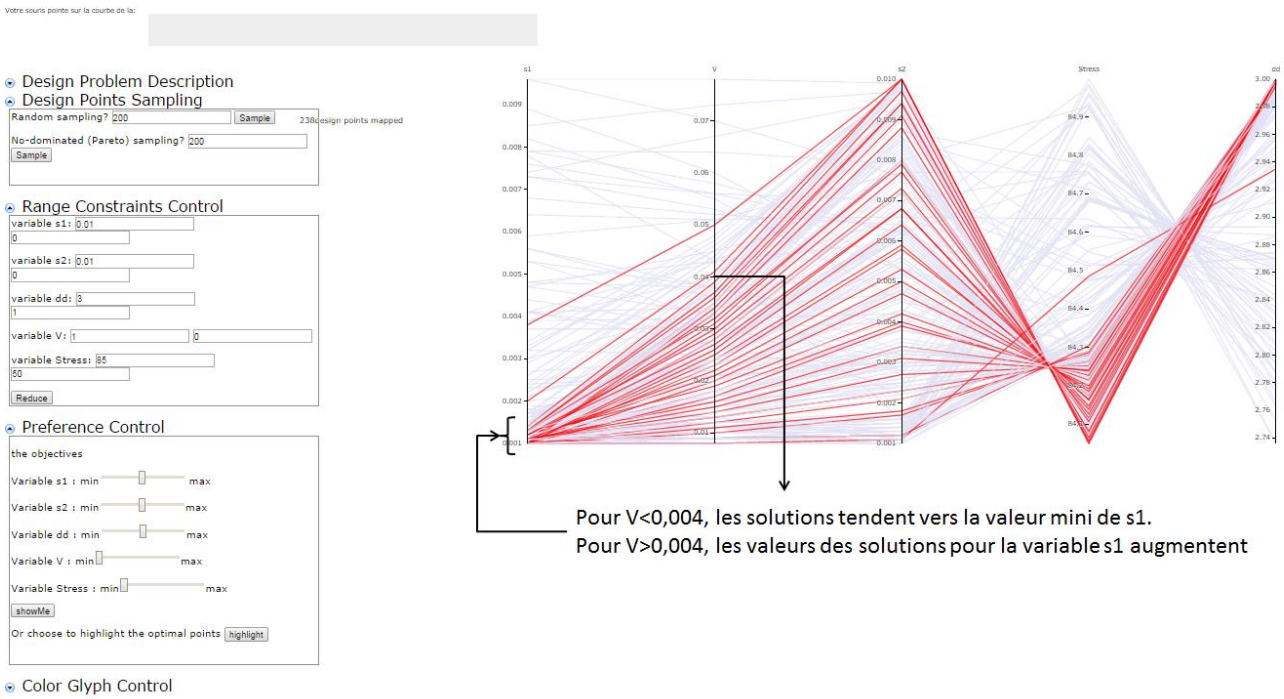
Et décide de mettre en lumière les points dits optimums (au sens de Pareto) pour tenter d'observer de nouvelles interactions entre les variables d'un point de vue plus local cette fois-ci.

Il peut enfin réaliser des observations « locales » (sur l'ensemble des optimums):

Les optimums semblent tendre vers dd=3 et s1=0,001: la solution finale devra surement prendre ces valeurs.

Vous voulez faire comme François ?

La fonctionnalité de mise en lumière des points dits optimums se trouve sous l'onglet « Preference Control ». Pour activer la fonction, cliquez sur le bouton « highlight ».



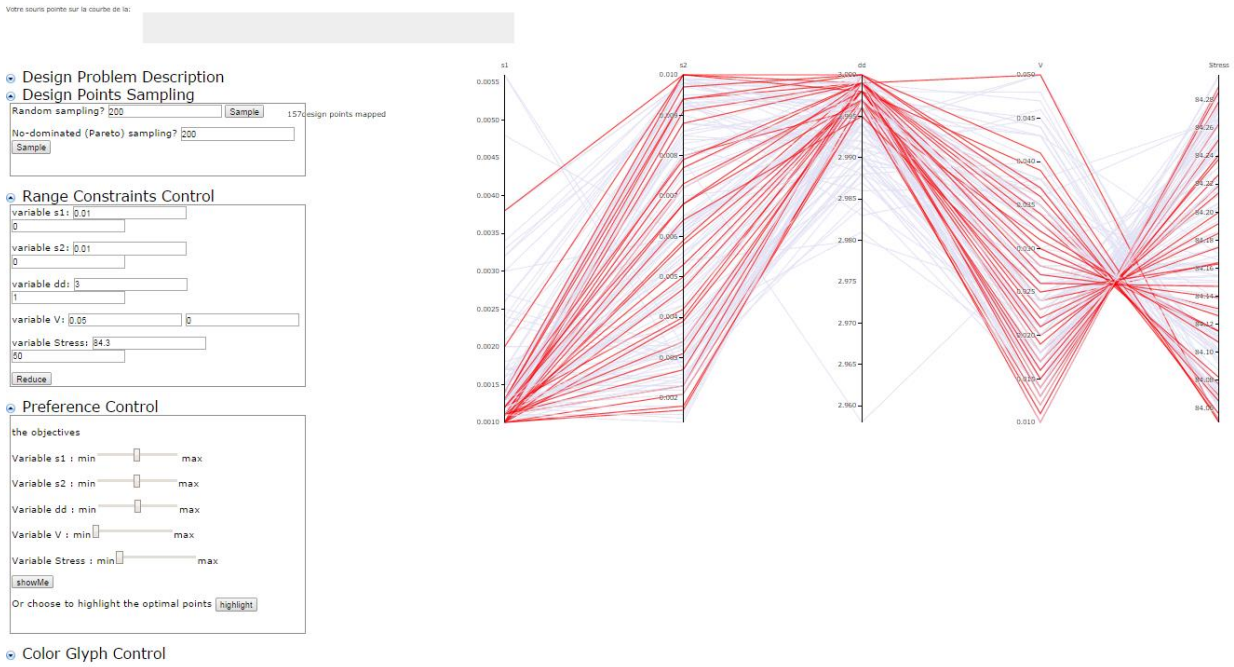
François déplace les axes des variables V et Stress pour découvrir des indices sur les données.

Nous pouvons observer qu'il semble y avoir un point de transition pour V=0,04. En effet, s2 évolue jusqu'à ce point tandis que s1 commence à évoluer à partir de celui-ci.

Nous relevons également une corrélation négative entre s2 et Stress

Vous voulez faire comme François ?

Réalisez le « glisser / déposer » pour l'axe de la variable V (entre s1 et s2) et pour l'axe de la variable Stress (entre s2 et dd).

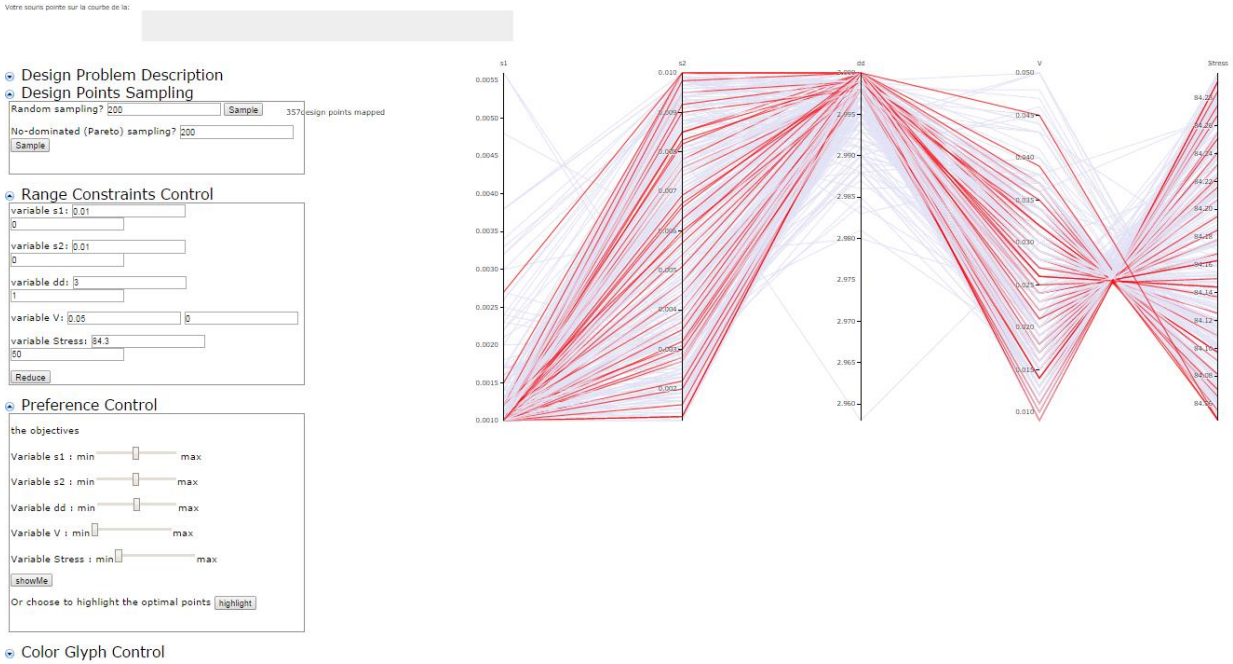


Visuellement il s'aperçoit qu'il peut limiter l'espace de conception à Stress_max=84.3

Il réduit donc l'espace de conception à Stress_max = 84,3 et en profite pour réduire également V_max à 0,06.

Vous voulez faire comme François ?

Modifiez le champ de la valeur maximum de Stress par la valeur qui vous semble la plus grande pour la variable Stress (elle peut être différente que celle de François car les données sont créés à la volées) et la valeur maximum de V par 0.06 puis cliquez sur le bouton « Reduce ».



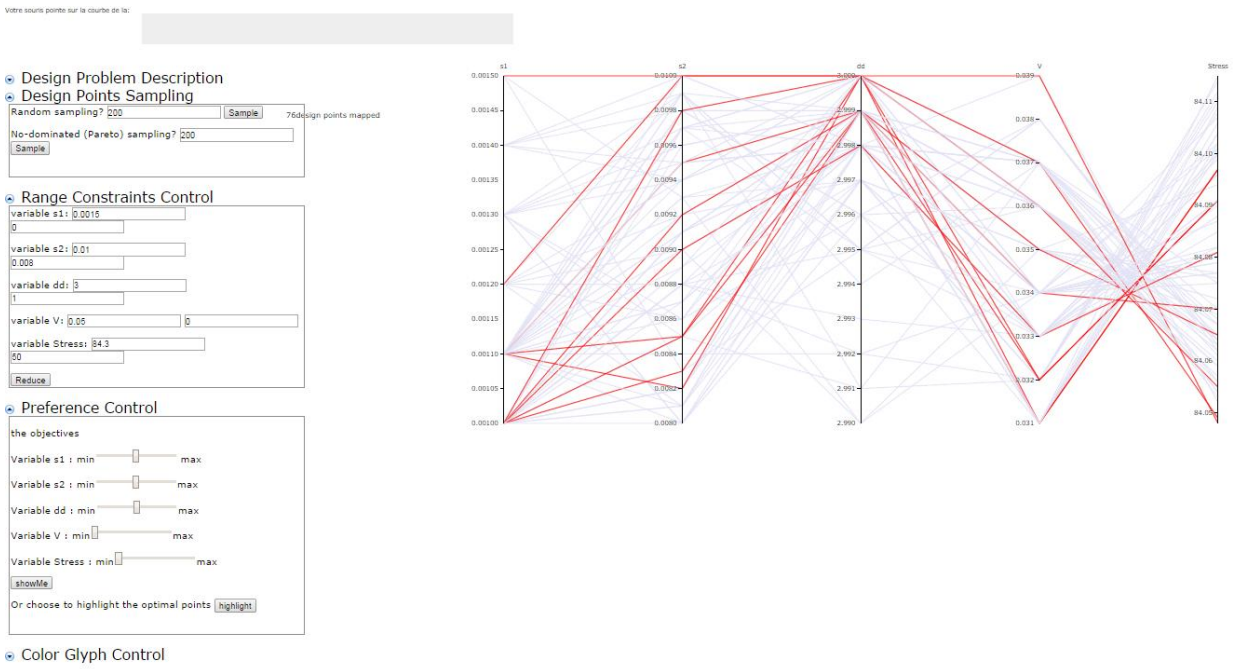
Après réduction, il relance 200 nouveaux points dits non-dominés.

L'espace des performances semble assez réduit pour l'ingénieur. Il décide maintenant de se pencher sur les variables de conception (ou variables de décisions).

Comme déjà vu, les valeurs de dd (=3) et s1 (=0,001) semblent déjà fixées. Ainsi la variable s2 est la réelle variable de décision.

Vous voulez faire comme François ?

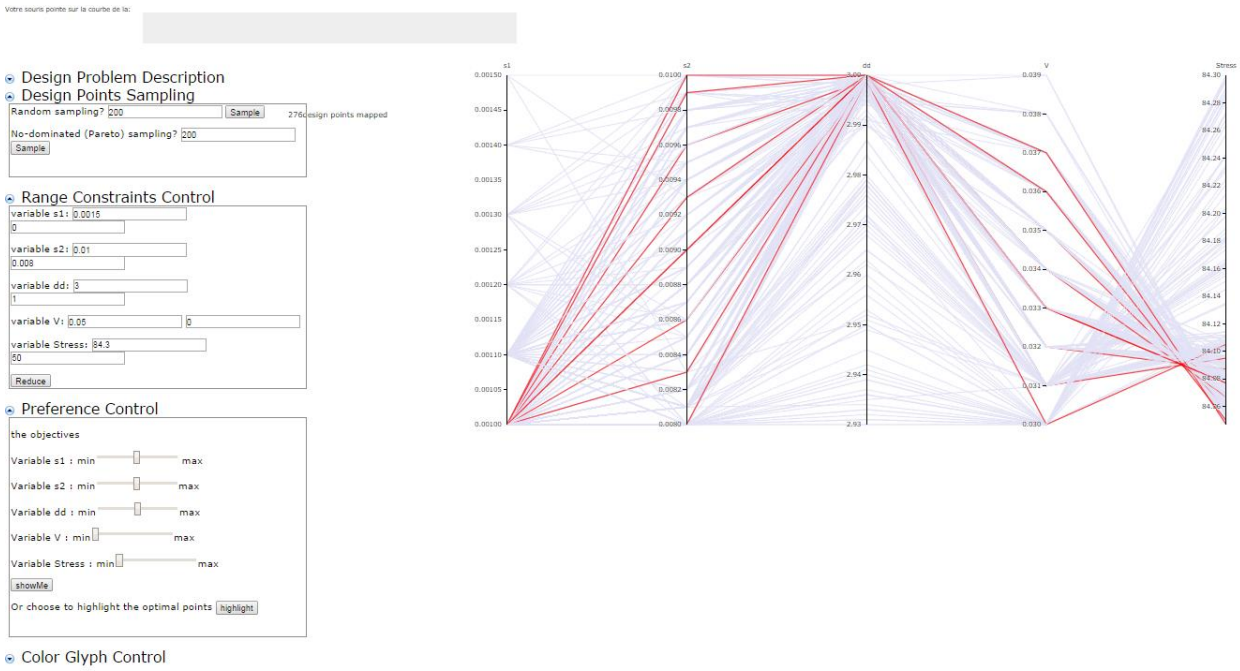
Cliquez sur le bouton « sample » dédié à générer des points non-dominés.



Ayant donc des valeurs de V inférieur à 0.06 (ce qui est correct au regard des objectifs), François réduit l'espace de conception en passant la valeur s2_mini=0.008 et s1_max =0.0015.

Vous voulez faire comme François ?

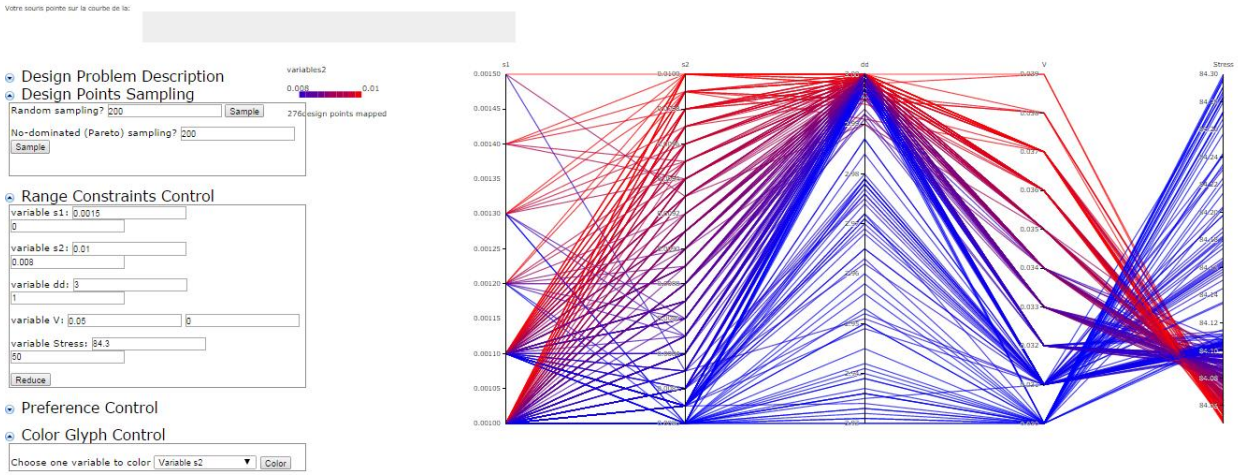
Dans le bloc sous l'onglet « Range Constraints Control » modifier le champ de la valeur minimum de la variable s2 par 0.008 et le champ de la valeur maximum de la variable s1 par 0.0015. Cliquez sur le bouton « Reduce ».



Comme après chaque réduction de l'espace de conception, l'ingénieur génère de nouveaux points dits non-dominés, il relance 200 nouveaux points.

Vous voulez faire comme François ?

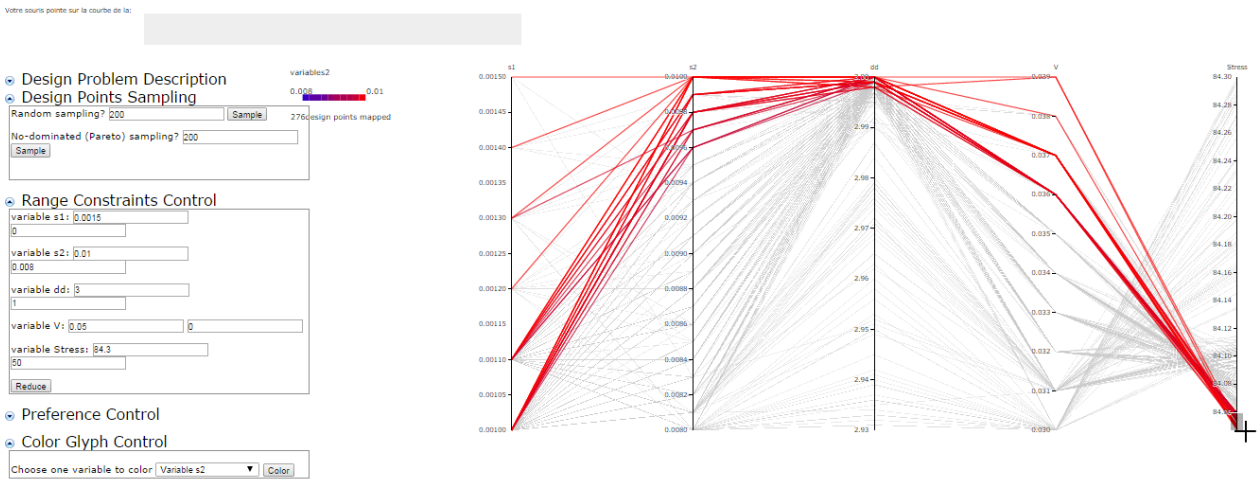
Dans le bloc de « sampling » cliquez sur son bouton « sample » dédié aux points non-dominés.



François utilise le marqueur de couleur sur la variable s2 pour identifier où se trouvent les valeurs les plus hautes de cette variable. En effet, la corrélation négative des optimums entre s2 et Stress nous pousse à chercher des points avec une grande section de la poutre 2. Mais il est également nécessaire d'observer la variable s2 conjointement avec les variable s1 (que nous voulons à 0.001) et dd (que nous voulons à 3).

Vous voulez faire comme François ?

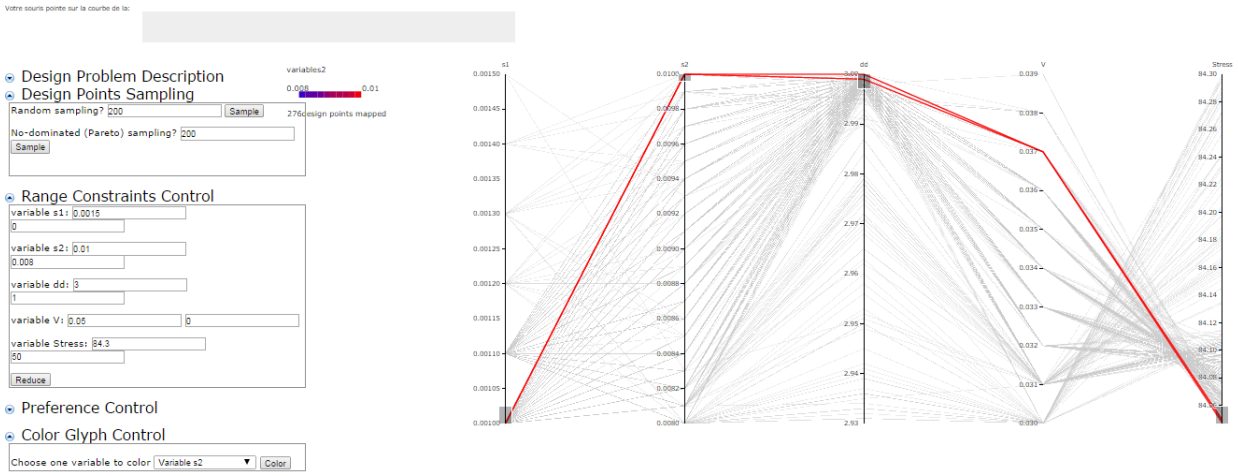
Cliquez sur l'onglet « Color Glyph Control » (vous pouvez cliquer sur l'onglet « Preference Control » pour le rétracter et gagner de la place). Puis dans la liste déroulante sélectionnez la variable s2. Cliquez sur le bouton « Color ».



L'ingénieur utilise la fonction de « brush » afin d'isoler des points de conception dans une petite zone où il lui apparait y avoir des solutions optimales.

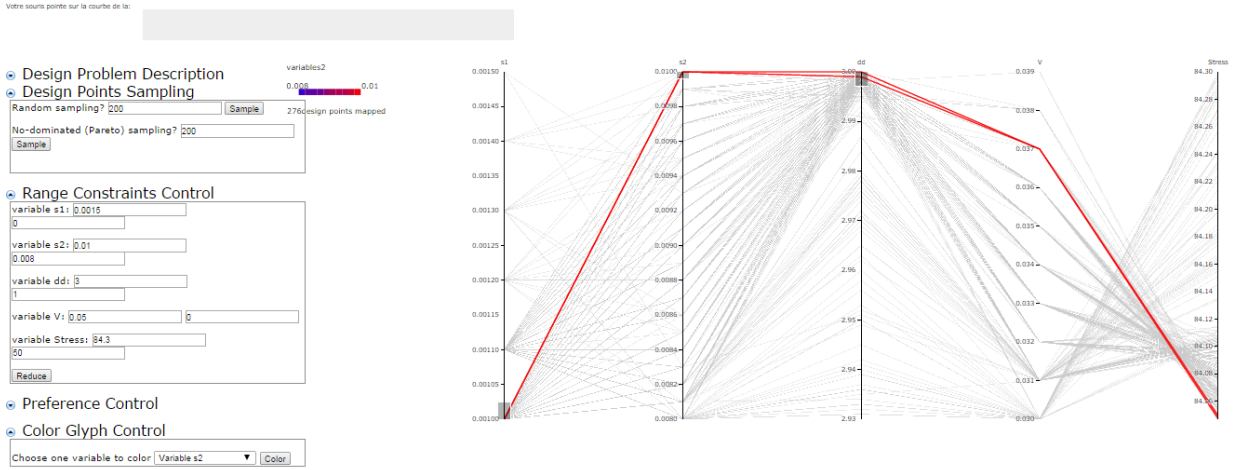
Vous voulez faire comme François ?

Avec votre souris allez sur l'axe de la variable Stress (vous verrez votre souris se transforme en croix noire). Cliquez et maintenez enfoncé puis tirez vers le bas, vous verrez un rectangle gris se dessiner.



François ajoute des « brush » sur les axes des variables s1 (pour une valeur égale à 0.001), s2 (pour une valeur égale à 0.01) et dd (pour une valeur égale à 3).

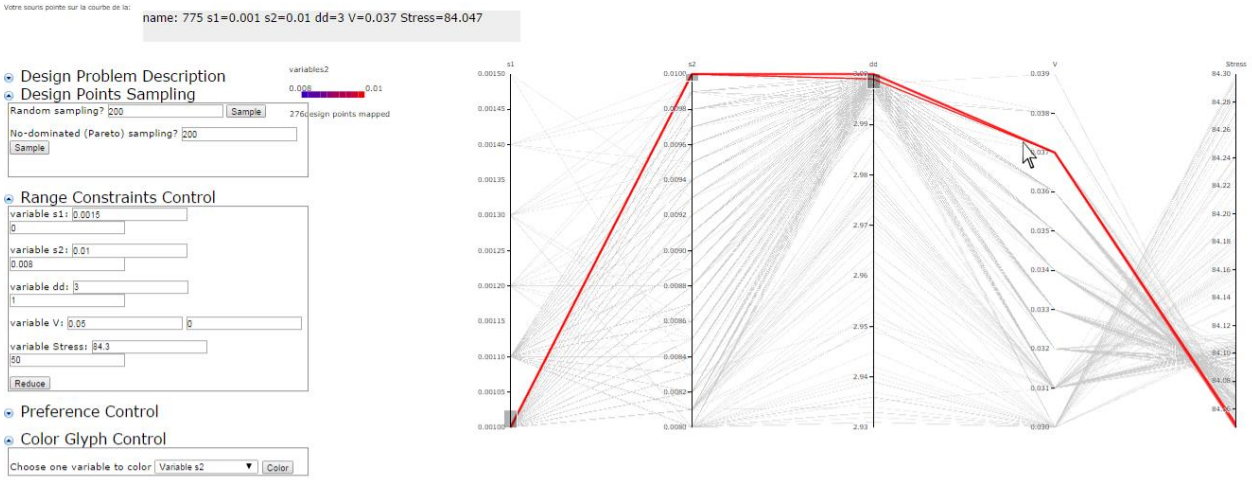
Puis il retire le brush placé sur l'axe de la variable Stress. Il n'y a pas plus de solutions possibles : il a trouvé un espace réduit avec deux solutions robustes.



Vous voulez faire comme François ?

Comme à l'étape précédente, utilisez votre souris pour isoler les solutions qui vous intéressent.

Pour retirer le brush, placer votre souris au-dessus du brush à retirer, la souris est en forme de croix, cliquez.



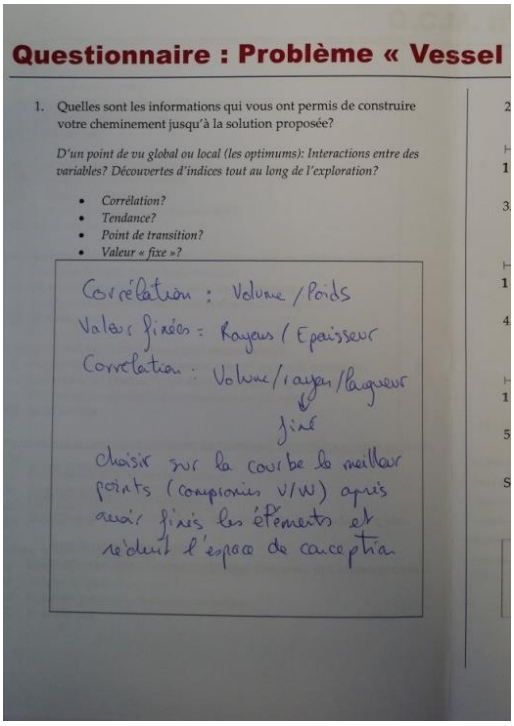
François utilise la fonction « selection » pour sélectionner un point et obtenir du détail dessus.

Il choisit le point numéro 775 avec s1=0.001, s2=0.01 et dd=3.

Vous voulez faire comme François ?

Cherchez un point qui convient en tenant compte des objectifs et des observations faites durant l'activité d'exploration de l'espace de conception (attention, dans votre session ce point ne sera peut-être pas le 775). Pour obtenir le détail, placez simplement votre souris sur la ligne correspondant à votre solution optimale.

A.8. Exemples de réponses informées et non-informées à l'expérimentation 2.4



Questionnaire : Problème « Vessel

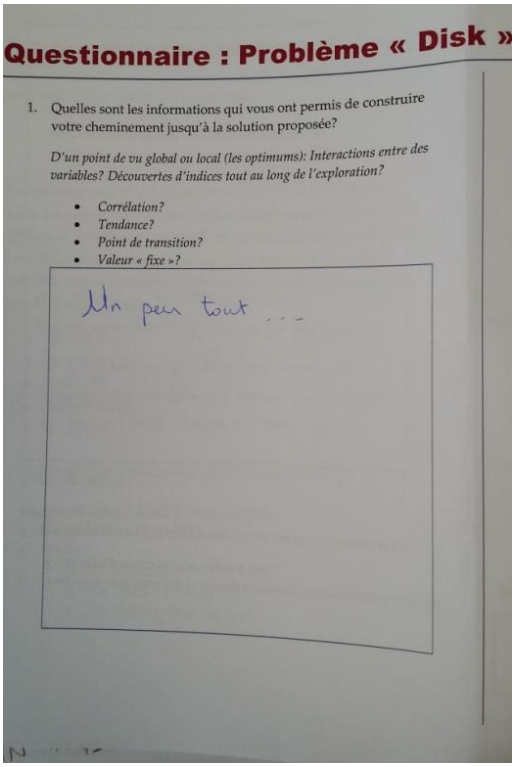
1. Quelles sont les informations qui vous ont permis de construire votre cheminement jusqu'à la solution proposée?

D'un point de vu global ou local (les optimums): Interactions entre des variables? Découvertes d'indices tout au long de l'exploration?

- *Corrélation?*
- *Tendance?*
- *Point de transition?*
- *Valeur « fixe »?*

Corrélation : Volume / Poids
Valeur fixées : Rayons / Epaisseur
Correlation : Volume / rayon / largueur
fixé
choisir sur la courbe le meilleur points (comparaison V/W) après avoir fixés les éléments et réduit l'espace de conception

Figure 125 : Exemple d'une réponse pour une décision informée



Questionnaire : Problème « Disk »

1. Quelles sont les informations qui vous ont permis de construire votre cheminement jusqu'à la solution proposée?

D'un point de vu global ou local (les optimums): Interactions entre des variables? Découvertes d'indices tout au long de l'exploration?

- *Corrélation?*
- *Tendance?*
- *Point de transition?*
- *Valeur « fixe »?*

Un peu tout ...

Figure 126 : Exemple d'une réponse pour une décision non-informée

Résumé / Abstract

Résumé

Lors de l'exploration d'espace de conception, les données résultantes de la simulation d'un grand nombre d'alternatives de conception peuvent conduire à la surcharge d'information quand il s'agit de choisir une bonne solution de conception. Cette exploration d'espace de conception s'apparente à une méthode d'optimisation en conception multicritère mais en mode manuel pour lequel des outils appropriés à la visualisation de données multidimensionnelle sont employés. Pour le concepteur, un processus en trois phases – découverte, optimisation, sélection – est suivi selon un paradigme dit de Design by Shopping. Le fait de « parcourir » l'espace de conception permet de gagner en intuition sur les sous-espaces de solutions faisables et infaisables et sur les solutions offrant de bons compromis. Le concepteur apprend au cours de ces manipulations graphiques de données. La sélection d'une solution optimale se fait donc sur la base d'une décision dite informée. L'objectif de cette recherche est la performance des représentations graphiques pour l'exploration d'espace de conception, pour les trois phases du processus en Design by Shopping. Pour cela, cinq représentations graphiques, identifiées comme potentiellement performantes, sont testées à travers deux expérimentations. Dans la première, trente participants ont testé trois graphiques, pour la phase de sélection dans une situation multi-attribut, à travers trois scénarios de conception où une voiture doit être choisie parmi quarante selon des préférences énoncées. Pour cela, un indice de qualité est proposé pour calculer la qualité de la solution du concepteur pour un des trois scénarios définis, la solution optimale selon cet indice étant comparée à celles obtenues après manipulation des graphiques. Dans la deuxième expérimentation, quarante-deux concepteurs novices ont résolu deux problèmes de conception à l'aide de trois graphiques. Dans ce cas, la performance des graphiques est testée pour la prise de décision informée et pour les trois phases du processus dans une situation multi-objectif. Les résultats révèlent qu'un graphique est adapté à chacune des trois phases du Design by Shopping :: le graphique Scatter Plot Matrix pour la phase de découverte et pour la prise de décision informée, le graphique Simple Scatter pour la phase d'optimisation et le graphique Parallel Coordinate Plot pour la phase de sélection aussi bien dans une situation multi-attribut que multi-objectif.

Mots-clés : Exploration d'espace de conception, Design by Shopping, Décision informée, Optimisation en conception multicritère, Visualisation d'information, Représentation graphique multidimensionnelle.

Abstract

In Design space exploration, the resulting data, from simulation of large amount of new design alternatives, can lead to information overload when one good design solution must be chosen. The design space exploration relates to a multi-criteria optimization method in design but in manual mode, for which appropriate tools to support multi-dimensional data visualization are employed. For the designer, a three-phase process - discovery, optimization, selection - is followed according to a paradigm called Design by Shopping. Exploring the design space helps to gain insight into both feasible and infeasible solutions subspaces, and into solutions presenting good trade-offs. Designers learn during these graphical data manipulations and the selection of an optimal solution is based on a so-called informed decision. The objective of this research is the performance of graphs for design space exploration according to the three phases of the Design by Shopping process. In consequence, five graphs, identified as potentially efficient, are tested through two experiments. In the first, thirty participants tested three graphs, in three design scenarios where one car must be chosen out of a total of forty, for the selection phase in a multi-attribute situation where preferences are enounced. A response quality index is proposed to compute the choice quality for each of the three given scenarios, the optimal solutions being compared to the ones resulting from the graphical manipulations. In the second experiment, forty-two novice designers solved two design problems with three graphs. In this case, the performance of graphs is tested for informed decision-making and for the three phases of the process in a multi-objective situation. The results reveal three efficient graphs for the design space exploration: the Scatter Plot Matrix for the discovery phase and for informed decision-making, the Simple Scatter Plot for the optimization phase and the Parallel Coordinate Plot for the selection phase in a multi-attribute as well as multi-objective situation.

Keywords: Design space exploration, Design by Shopping, Informed decision, Multi-criteria design optimization, Information visualization, Multidimensional data visualization.