Synthèse Bibliographique: Étude de Variantes du Problème de sac à dos (knapsack) dans le Cadre des Jeux de Fantasy Cyclisme

Yuxiang ZHANG
Louiza CHENAOUI
Kenan ALSAFADI
Master AI2D - Sorbonne Université

24 novembre 2025

Résumé

Cette synthèse bibliographique examine les approches d'optimisation pour les jeux de fantasy cyclisme, en se concentrant sur les modélisations basées sur le problème du sac à dos (knapsack) et ses variantes multi-étapes. L'analyse couvre les modèles de programmation linéaire en nombres entiers, les méthodes de résolution exactes et approchées, ainsi que les perspectives d'analyse post-optimale. La revue identifie les principaux défis computationnels et propose des directions de recherche futures pour ce domaine émergent situé à l'intersection de la recherche opérationnelle et des sports analytics. Notre analyse critique met en lumière les lacunes actuelles de la recherche, notamment l'absence d'approches intégrant optimisation multi-périodes et analyse de robustesse dans le contexte spécifique du cyclisme fantasy.

Mots-clés : fantasy cyclisme, sac à dos (knapsack) multi-étapes, optimisation combinatoire, PLNE, analyse post-optimale, sports analytics

Table des Matières

1 Introduction **4**

1.1 Contexte et problématique 4

1.2 Délimitation du champ d'étude 4

2 État de l'art **5**

2.1 Cartographie du champ scientifique 5

2.2 Les fondements de la modélisation optimisationnelle 5

2.3 Le contexte sportif et l'impact des règles 6

2.4 L'analyse de robustesse et stabilité des solutions 7

2.5 Tensions Méthodologiques, Synthèse et Positionnement 7

3 Modélisation du problème **8**

3.1 Présentation du modèle de Beliën et al. (2017) 8

3.2 Ensembles et indices 8

3.3 Paramètres du modèle 8

3.4 Variables de décision 9

3.5 La fonction objectif 9

3.6 Les contraintes du modèle détaillées 10

3.6.1 Contraintes de composition d'équipe 10

3.6.2 Contraintes budgétaires initiales 10

3.6.3 Contraintes budgétaires dynamiques et de transfert 11

3.6.4 Contraintes d'attribution des points 11

3.6.5 Contraintes des joueurs remplaçants 12

3.6.6 Définition des domaines de variables 12

3.7 Extensions du modèle pour objectifs alternatifs 13
3.7.1 Scénario avec pool de joueurs 13
3.7.2 Maximisation des gains monétaires 13
3.8 Inégalités valides pour le renforcement 14
3.9 Synthèse de l'application à Gigabike 15

4 Analyse post-optimale pour le problème de sac à dos (knapsack) 15
4.1 Introduction à l'analyse de tolérance 15
4.2 Définitions formelles des limites de tolérance 16
4.3 Méthodes exactes pour l'analyse de tolérance 16
4.4 Méthodes approximatives 16
4.5 Résultats expérimentaux et comparaisons 17
4.6 Synthèse des contributions méthodologiques 17

5 Conclusion et Perspectives 17
5.1 Synthèse du cadre méthodologique et lacunes de la littérature 17
5.2 Perspectives de recherche, défis d'adaptation et implications pratiques . . . 18

Bibliographie 19

1 Introduction

1.1 Contexte et problématique

Le fantasy cyclisme est une catégorie émergente des jeux de fantasy sports. Le principe pour le participant, qui endosse le rôle de manager sportif, est de composer une équipe de coureurs pour une compétition multi-étapes (comme le Tour de France) afin de maximiser les points accumulés. Cette sélection est soumise à une contrainte budgétaire stricte, chaque coureur possédant un coût et un rendement estimé (les points potentiels). Le défi principal consiste à choisir l'ensemble optimal de coureurs dont la somme des coûts ne dépasse pas le budget, tout en maximisant la valeur totale. L'environnement est rendu dynamique par la possibilité d'effectuer des transferts entre les étapes pour ajuster la formation en fonction des performances, des abandons ou des changements stratégiques. Ces transferts sont généralement limités en nombre et peuvent impliquer des coûts additionnels, ajoutant une dimension stratégique supplémentaire.

D'un point de vue académique, cette problématique s'inscrit dans le cadre des problèmes d'optimisation combinatoire et présente des similarités avec le problème classique du sac à dos (knapsack). Cependant, elle s'en distingue par plusieurs aspects fondamentaux : la dimension multi-périodes introduite par les différentes étapes, la possibilité de modifier la composition de l'équipe via des transferts, et les contraintes spécifiques liées aux règles du cyclisme (nombre de coureurs par type, contraintes par équipe professionnelle, etc.). La complexité de ce problème réside dans l'arbitrage entre l'optimisation instantanée (le choix initial) et l'optimisation globale (les opportunités de transferts ultérieurs).

Pour comprendre la dynamique de ces jeux et valider les stratégies, l'analyse post-optimale offre un cadre privilégié. Elle permet d'étudier a posteriori les décisions optimales en connaissant toutes les performances réelles des coureurs, d'identifier les coureurs "incontournables", les combinaisons synergiques, et surtout, d'évaluer l'impact des règles du jeu sur les stratégies optimales. C'est pourquoi cet article de synthèse vise à cartographier les outils méthodologiques nécessaires à l'intégration de la modélisation dynamique et de l'analyse de robustesse.

1.2 Délimitation du champ d'étude

Cette synthèse se concentre spécifiquement sur les applications du problème de sac à dos (knapsack) au fantasy cyclisme, en excluant les autres sports fantasy. Nous nous limitons aux approches d'optimisation déterministes et post-optimales, laissant de côté les méthodes purement heuristiques ou les approches basées sur l'apprentissage automatique.

Le périmètre couvre les modèles de programmation linéaire en nombres entiers pour la sélection d'équipe et la gestion des transferts, ainsi que les méthodes d'analyse de robustesse pour l'évaluation de la stabilité des solutions.

Ce travail de synthèse bibliographique vise à cartographier les approches existantes pour modéliser et résoudre cette problématique complexe, en identifiant les liens avec les variantes du problème de sac à dos (knapsack) et en mettant en lumière les défis computationnels et méthodologiques qui demeurent ouverts.

2 État de l'art

Cette revue de littérature s'appuie sur une sélection stratégique de sept articles fondamentaux qui couvrent l'ensemble du spectre de recherche autour du fantasy cyclisme et des problèmes de sac à dos (knapsack) multi-étapes. L'analyse révèle un paysage scientifique structuré autour de trois axes complémentaires : la modélisation optimisationnelle, le contexte sportif spécifique, et l'analyse de robustesse des solutions.

2.1 Cartographie du champ scientifique

Notre analyse révèle un paysage structuré autour de trois écoles principales. L'école belge, menée par Beliën et ses collaborateurs, se concentre sur la modélisation MIP appliquée aux jeux fantasy. L'école française, représentée par Bampis et al., apporte des contributions théoriques sur le sac à dos (knapsack) multi-étapes. Enfin, l'école scandinave, avec Pisinger, développe des méthodes avancées d'analyse post-optimale. Ces différentes approches reflètent la diversité méthodologique du domaine et la complémentarité des perspectives entre recherche théorique et applications pratiques.

2.2 Les fondements de la modélisation optimisationnelle

Les travaux de Beliën et ses collaborateurs constituent la pierre angulaire de notre étude. Leur premier article de 2011 a marqué une étape importante en démontrant l'applicabilité pratique des méthodes de programmation linéaire en nombres entiers (MIP) à travers l'étude du jeu Gigabike. Cette application concrète, structurée en trois étapes - problème de sac à dos (knapsack) par période, gestion multi-périodes avec budget et transferts, contraintes logiques sur les transferts - a permis d'obtenir des solutions optimales significativement supérieures aux performances des meilleurs joueurs humains. Ces résultats ont validé l'approche méthodologique et souligné le potentiel d'amélioration offert par l'optimisation mathématique dans le domaine des jeux fantasy.

Fort de ces résultats prometteurs, Beliën et ses collaborateurs ont développé en 2017 un cadre MIP générique pour l'analyse ex-post des jeux de fantasy sports. Ce modèle flexible intègre l'ensemble des contraintes caractéristiques du domaine : limitations budgétaires, composition d'équipe équilibrée, et gestion stratégique des transferts sur plusieurs périodes. La méthodologie développée permet de maximiser les points totaux tout en considérant les coûts de transaction associés aux transferts, offrant ainsi un cadre formel adaptable à divers sports. Cette généralisation représente une contribution majeure, étendant l'applicabilité des approches d'optimisation au-delà du cyclisme vers l'ensemble des sports fantasy.

Sur le plan théorique, l'article de Bampis et al. (2022) fournit les fondements algorithmiques les plus récents pour aborder la dimension temporelle de notre problème. En formalisant le problème du sac à dos (knapsack) multi-étapes avec transitions entre périodes, les auteurs développent un PTAS pour un nombre constant d'étapes, établissant ainsi des garanties de performance théoriques pour cette classe de problèmes. L'analyse de la relaxation linéaire et la modélisation ILP proposée offrent des pistes méthodologiques précieuses pour notre étude des transitions d'équipe dans le fantasy cyclisme.

2.3 Le contexte sportif et l'impact des règles

La revue exhaustive de Durán (2021) sur les sports analytics fournit le contexte général nécessaire pour situer notre problématique dans le paysage plus large de la recherche en optimisation sportive. Cette synthèse bibliographique couvrant la planification, l'assignation et l'analyse de performances démontre la maturité croissante des méthodes d'optimisation dans le domaine sportif. L'identification des défis computationnels récurrents et des tendances émergentes offre des perspectives précieuses pour notre travail sur le fantasy cyclisme.

Dans une perspective plus spécifique, l'article récent d'Ausloos (2024) apporte une critique essentielle sur l'influence des règles sportives dans les stratégies d'optimisation. Par une analyse statistique comparative entre la méthode UCI classique et des indicateurs alternatifs de performance, l'étude démontre comment la définition des contraintes et des métriques de scoring modifie fondamentalement la hiérarchie des coureurs et les décisions optimales. Ces résultats soulignent l'importance d'une modélisation fine des règles spécifiques du jeu dans l'élaboration des stratégies.

2.4 L'analyse de robustesse et stabilité des solutions

Les travaux de Pisinger et Saidi (2017) sur l'analyse de tolérance pour le sac à dos (knapsack) 0-1 fournissent le cadre théorique complet pour l'analyse post-optimale qui est au cœur de notre problématique. Leurs algorithmes de programmation dynamique pour l'analyse de tolérance exacte, complétés par des méthodes approximatives basées sur des bornes supérieures, permettent d'évaluer la robustesse de la solution optimale face aux perturbations des coefficients. Cette méthodologie est directement applicable à l'étude de la stabilité des équipes optimales dans le fantasy cyclisme.

Plus récemment, les travaux de Kumabe et Yoshida (2022) sur la sensibilité moyenne du problème de sac à dos (knapsack) apportent une contribution méthodologique cruciale pour l'étude de la stabilité des solutions. Leur algorithme stable, combinant classification des objets, approche gloutonne modifiée et mécanisme exponentiel, permet d'évaluer la robustesse des solutions face aux perturbations des données d'entrée. Cette approche est particulièrement pertinente dans notre contexte où les performances des coureurs sont sujettes à variabilité.

2.5 Tensions Méthodologiques, Synthèse et Positionnement

Le champ de l'optimisation pour les jeux fantasy est marqué par une série de tensions méthodologiques. Si le modèle de Beliën offre une grande généralité, sa complexité computationnelle peut limiter son applicabilité en temps réel. À l'inverse, les méthodes de Pisinger, bien que plus rapides, n'intègrent pas la dimension multi-périodes cruciale pour le fantasy cyclisme. Cette tension entre exhaustivité et efficacité représente un défi central du domaine. Les approches théoriques de Bampis et ses collaborateurs fournissent des garanties de performance mais peinent à capturer la complexité des règles réelles des jeux fantasy, tandis que les analyses contextuelles d'Ausloos et Durán rappellent l'importance d'ancrer les modèles dans la réalité sportive.

Malgré ces divergences, la complémentarité de ces approches permet d'appréhender le problème de fantasy cyclisme dans toute sa complexité. Les modèles de Beliën fournissent le cadre formel, les analyses contextuelles d'Ausloos et Durán ancrent le problème dans la réalité sportive, tandis que les travaux théoriques de Bampis, Kumabe et Pisinger offrent les outils d'analyse de robustesse nécessaires à l'étude post-optimale. Cet état de l'art révèle cependant une lacune importante : l'absence d'études intégrant simultanément ces trois dimensions dans le contexte spécifique du cyclisme fantasy avec analyse systématique de la stabilité des solutions optimales au cours d'une compétition multi-étapes. C'est sur cette intersection critique que se positionne l'originalité de notre recherche.

3 Modélisation du problème

3.1 Présentation du modèle de Beliën et al. (2017)

Le modèle proposé par Beliën, Goossens et Van Reeth (2017) constitue le cadre formel le plus complet pour l'analyse ex-post des jeux de fantasy sports. Cette modélisation en programmation linéaire en nombres entiers (PLNE) permet de déterminer la stratégie optimale de sélection d'équipe et de gestion des transferts lorsque toutes les performances des joueurs sont connues a posteriori.

3.2 Ensembles et indices

Le modèle fait intervenir les ensembles fondamentaux suivants :

$p, p' \in P = \{1, 2, \ldots, \|P\|\}$	Joueurs (coureurs cyclistes)
$\pi \in \Pi = \{1, 2, \ldots, \|\Pi\|\}$	Types de joueurs
$P_\pi \subseteq P$	Ensemble des joueurs de type π
$t \in T = \{1, 2, \ldots, \|T\|\}$	Périodes de transfert
$\tau \in \Theta = \{1, 2, \ldots, \|\Theta\|\}$	Événements (courses cyclistes)
$G \subseteq \Theta$	Ensembles de jeux pour l'attribution des prix
$t_G \in T$	Période de référence pour le budget restant dans l'ensemble G
$e \in E_G = \{1, 2, \ldots, \|E_G\|\}$	Rangs pour lesquels un prix est attribué

3.3 Paramètres du modèle

Les paramètres principaux qui caractérisent le problème sont :

$v_{p\tau}$	Points obtenus par le joueur p dans l'événement τ
c_{pt}	Coût du joueur p durant la période t
B	Budget total disponible
A_t	Nombre maximal de transferts autorisés en période t
D_τ	Nombre maximal de joueurs marquant des points dans l'événement τ
L	Pénalité par transfert utilisé
ε	Poids du critère de départage (budget restant)
$n_{\pi t}$	Nombre minimal de joueurs de type π en période t
$N_{\pi t}$	Nombre maximal de joueurs de type π en période t
$a_{p\tau} = 1$	Si le joueur p a participé à l'événement τ, 0 sinon
ω_{eG}	Valeur (prix) pour l'obtention du rang e dans l'ensemble G
H_{eG}	Score de l'adversaire ayant obtenu le rang e dans l'ensemble G
R_{eG}	Budget restant du participant ayant obtenu le rang e dans l'ensemble G
Q_π	Nombre maximal de joueurs de type π dans le pool de joueurs

3.4 Variables de décision

Le modèle utilise les variables de décision suivantes :

$x_{pt} \in \{0,1\}$	1 si le joueur p est dans l'équipe en période t, 0 sinon
$y_{p\tau} \in \{0,1\}$	1 si le joueur p marque des points dans l'événement τ, 0 sinon
$r_t \geq 0$	Budget restant en période t
$z_{pt} \in \{0,1\}$	1 si le joueur p est transféré dans l'équipe en période t, 0 sinon
$s_{pt} \in \{0,1\}$	1 si le joueur p est remplaçant en période t, 0 sinon
$w_{eG} \in \{0,1\}$	1 si l'équipe obtient le rang e dans l'ensemble G, 0 sinon
$q_p \in \{0,1\}$	1 si le joueur p est inclus dans le pool de joueurs

3.5 La fonction objectif

La fonction objectif principale est structurée autour de trois composantes essentielles qui reflètent les différents aspects stratégiques du jeu :

$$\text{Maximiser} \quad \underbrace{\sum_{p\in P}\sum_{\tau\in\Theta} v_{p\tau} y_{p\tau}}_{\text{Maximisation des points}} \quad - \quad \underbrace{L\sum_{p\in P}\sum_{t\in T} z_{pt}}_{\text{Pénalisation des transferts}} \quad + \quad \underbrace{\varepsilon r_{|T|}}_{\text{Valorisation du budget restant}} \tag{1}$$

La première composante maximise la somme des points marqués par l'ensemble des joueurs sélectionnés sur tous les événements, où $v_{p\tau}$ représente les points du joueur p dans l'événement τ et $y_{p\tau}$ est une variable binaire indiquant si le joueur p contribue aux points dans l'événement τ. La deuxième composante introduit une pénalité proportionnelle au nombre total de transferts effectués, avec L représentant le coût en points par transfert utilisé et z_{pt} identifiant les transferts. La troisième composante valorise le budget restant en fin de saison comme critère de départage, où ε est un poids positif faible assurant que ce critère n'intervient qu'en cas d'égalité.

3.6 Les contraintes du modèle détaillées

3.6.1 Contraintes de composition d'équipe

$$\sum_{p\in P_\pi} x_{pt} \geq n_{\pi t} \quad \forall\pi, \forall t \tag{2a}$$

$$\sum_{p\in P_\pi} x_{pt} \leq N_{\pi t} \quad \forall\pi, \forall t \tag{2b}$$

Explication : Ces contraintes imposent un nombre minimal $n_{\pi t}$ et maximal $N_{\pi t}$ de joueurs de chaque type π dans chaque période t. Elles garantissent une composition d'équipe équilibrée selon les règles spécifiques du sport.

Application à Gigabike : Dans Gigabike, il n'y a pas de types de joueurs ($\Pi = \emptyset$), donc ces contraintes ne s'appliquent pas. L'équipe doit simplement contenir exactement 30 coureurs.

3.6.2 Contraintes budgétaires initiales

$$\sum_{p\in P} c_{p1} x_{p1} + r_1 = B \tag{3}$$

Explication : Cette contrainte établit le budget initial. La somme des coûts des joueurs sélectionnés dans la première période plus le budget restant r_1 doit égaler le budget total B.

Application à Gigabike : Le budget CQ initial est de 20,000 points. La sélection des 30 coureurs doit respecter cette limite.

3.6.3 Contraintes budgétaires dynamiques et de transfert

$$\sum_{p \in P} c_{pt}(x_{pt} - x_{p,t-1}) = r_{t-1} - r_t \quad \forall t > 1 \tag{4a}$$

$$z_{pt} \geq x_{pt} - x_{p,t-1} \quad \forall p, \forall t > 1 \tag{4b}$$

$$\sum_{p \in P} z_{pt} \leq A_t \quad \forall t > 1 \tag{4c}$$

Explication :

- **(4a) :** Gère l'évolution du budget entre les périodes. Le coût net des transferts (nouveaux joueurs moins joueurs vendus) est financé par le budget restant de la période précédente.

- **(4b) :** Identifie les transferts entrants. La variable z_{pt} vaut 1 si un joueur rejoint l'équipe à la période t.

- **(4c) :** Limite le nombre total de transferts par période à A_t.

Application à Gigabike : Gigabike permet 5 transferts maximum par période. Le budget restant se reporte d'une période à l'autre.

3.6.4 Contraintes d'attribution des points

$$y_{p\tau} \leq x_{pt} \quad \forall p, \forall t, \forall \tau \in T_t \tag{5}$$

$$\sum_{p \in P} y_{p\tau} \leq D_\tau \quad \forall \tau \tag{6}$$

Explication :

— **(5) :** Un joueur ne peut marquer des points que s'il fait partie de l'équipe durant la période contenant l'événement τ.

— **(6) :** Seuls les D_τ meilleurs joueurs de l'équipe contribuent aux points dans chaque événement τ.

Application à Gigabike : Contrainte CRUCIALE. Dans chaque course, seuls les 8 meilleurs coureurs ($D_\tau = 8$) marquent des points pour l'équipe.

3.6.5 Contraintes des joueurs remplaçants

$$\sum_{p \in P_\pi} s_{pt} = 1 \quad \forall \pi, \forall t \tag{7a}$$

$$s_{pt} \leq x_{pt} \quad \forall p, \forall t \tag{7b}$$

$$y_{p\tau} \leq (1 - s_{pt}) + \sum_{p' \in P_\pi \setminus p} (1 - a_{p'\tau}) x_{p't} \quad \forall \pi, \forall p \in P_\pi, \forall t, \forall \tau \in T_t \tag{7c}$$

Explication : Ces contraintes gèrent un système avec remplaçants :

— **(7a) :** Exactement un remplaçant par type de joueur.

— **(7b) :** Le remplaçant doit faire partie de l'équipe.

— **(7c) :** Le remplaçant ne marque des points que si au moins un titulaire de son type n'a pas participé à l'événement.

Application à Gigabike : NON APPLICABLE. Gigabike n'utilise pas de système de remplaçants.

3.6.6 Définition des domaines de variables

$$x_{pt} \in \{0, 1\} \quad \forall p, \forall t \tag{8}$$

$$y_{p\tau} \in \{0, 1\} \quad \forall p, \forall \tau \tag{9}$$

$$z_{pt} \in \{0, 1\} \quad \forall p, \forall t > 1 \tag{10}$$

$$s_{pt} \in \{0, 1\} \quad \forall p, \forall t \tag{11}$$

$$r_t \geq 0 \quad \forall t \tag{12}$$

Explication : Ces contraintes définissent les domaines des variables : binaires pour les décisions discrètes, continues non-négatives pour le budget.

3.7 Extensions du modèle pour objectifs alternatifs

3.7.1 Scénario avec pool de joueurs

$$x_{pt} \leq q_p \quad \forall p, \forall t \tag{13}$$

$$\sum_{p \in P_\pi} q_p \leq Q_\pi \quad \forall \pi \tag{14}$$

$$q_p \in \{0, 1\} \quad \forall p \tag{15}$$

Explication : Ce scénario modélise les jeux où les participants sélectionnent d'abord un pool de joueurs, puis forment des équipes à partir de ce pool.

— **(13) :** Seuls les joueurs du pool peuvent être sélectionnés.

— **(14) :** Limite la taille du pool par type de joueur.

— **(15) :** Variable binaire définissant l'appartenance au pool.

Application à Gigabike : NON APPLICABLE.

3.7.2 Maximisation des gains monétaires

$$\text{Maximiser} \quad \sum_{G} \sum_{e \in E_G} \omega_{eG} w_{eG} \tag{16}$$

$$\text{s.t.} \quad \sum_{p \in P} \sum_{\tau \in G} v_{p\tau} y_{p\tau} + \varepsilon r_{t_G} - w_{eG}(H_{eG} + \varepsilon R_{eG}) \geq 0 \quad \forall G, \forall e \in E_G \tag{17}$$

$$\sum_{e \in E_G} w_{eG} \leq 1 \quad \forall G \tag{18}$$

$$w_{eG} \in \{0, 1\} \quad \forall e, \forall G \tag{19}$$

Explication : Cette formulation maximise les gains monétaires plutôt que le score total.

— **(16) :** Maximise la somme des prix gagnés.

— **(17) :** Un rang e n'est obtenu que si le score dépasse celui du participant réel ayant ce rang.

— **(18) :** Au plus un rang par ensemble de jeux.

— **(19) :** Variable binaire pour l'obtention d'un rang.

3.8 Inégalités valides pour le renforcement

Pour renforcer la formulation du modèle, particulièrement dans le scénario de maximisation des gains monétaires, Beliën et al. proposent des inégalités valides de la forme :

$$w_{eG} \leq \sum_{(p,t)\in Q} y_{pt} \quad \forall G, \forall e \in E_G \tag{20}$$

Explication : Cette inégalité valide stipule que pour obtenir un rang e dans un ensemble de jeux G, il est nécessaire d'avoir sélectionné au moins un des joueurs de l'ensemble Q durant la période correspondante.

Mécanisme de construction de Q :

— L'ensemble Q est construit à l'aide d'un arbre de recherche binaire

— Chaque niveau de l'arbre correspond à la décision d'inclure ou non un joueur particulier

— On calcule une borne supérieure du score réalisable compte tenu des décisions prises

— Si cette borne supérieure devient inférieure au score H_{eG} du participant réel ayant le rang e, alors on sait qu'il faut inclure au moins un des joueurs exclus pour espérer atteindre ce rang

— L'ensemble Q est formé par ces paires (joueur, période) critiques

Avantage computationnel : Ces inégalités réduisent l'espace de recherche en éliminant précocement des solutions non prometteuses, accélérant ainsi la résolution du modèle.

Application à Gigabike : Ces inégalités sont particulièrement utiles pour les calculs de maximisation des gains dans les différentes catégories de prix.

3.9 Synthèse de l'application à Gigabike

Pour Gigabike, le modèle se simplifie considérablement :

— **Contraintes actives :** (1), (3), (4a-4c), (5), (6), (8-10), (12)

— **Contraintes inactives :** (2a-2b), (7a-7c), (11), (13-15)

— **Extensions optionnelles :** (16-19) pour l'analyse des prix, (20) pour le renforcement

— **Paramètres spécifiques :** $L = 0$, $D_\tau = 8$, $B = 20,000$, $A_t = 5$, équipe de 30 coureurs

4 Analyse post-optimale pour le problème de sac à dos (knapsack)

4.1 Introduction à l'analyse de tolérance

L'analyse post-optimale constitue une étape cruciale dans l'étude des problèmes d'optimisation combinatoire, permettant d'évaluer la robustesse des solutions optimales face aux incertitudes sur les données.

La modélisation du problème de sélection d'équipe (Section 3) repose sur des pronostics de performance qui sont, par nature, très incertains. Cette instabilité est amplifiée dans le contexte des courses cyclistes multi-étapes par des facteurs imprévisibles comme les abandons, les disqualifications ou les blessures. L'article d'Ausloos (2024) illustre concrètement ce défi : en analysant l'impact des remplacements et des retraits sur les classements finaux des équipes, il démontre que les résultats sont structurellement instables. Cette observation concrète nous rappelle que la meilleure équipe sur le papier n'a que peu de valeur si elle est fragilisée par le premier imprévu. Il devient donc essentiel, non seulement de calculer la sélection optimale, mais aussi de quantifier sa robustesse face aux variations de données, ce qui est l'objet de l'analyse post-optimale.

Contrairement à la programmation linéaire où l'analyse de sensibilité est bien établie, l'analyse post-optimale pour les problèmes en nombres entiers présente des défis computationnels significatifs. Pour le problème du sac à dos (knapsack) en nombres entiers (0-1), le cadre méthodologique de l'analyse post-optimale est celui de l'analyse de tolérance, formalisée par Pisinger et Saidi (2017). Leur approche s'intéresse plus spécifiquement à déterminer les intervalles dans lesquels les coefficients du problème peuvent varier sans que le vecteur solution optimal ne change. Cette approche diffère de l'analyse de sensibilité classique en se focalisant sur la robustesse de la solution optimale plutôt que sur les variations des solutions de base.

4.2 Définitions formelles des limites de tolérance

Soit KP une instance du problème de sac à dos (knapsack) 0-1 avec une solution optimale x^* et une valeur optimale z^*. Les problèmes perturbés sont définis pour les profits et les poids. Pour la perturbation des profits, on substitue p_k par $p_k + \Delta p_k$ dans la fonction objectif. Pour la perturbation des poids, on substitue w_k par $w_k + \Delta w_k$ dans la contrainte de capacité. Les limites de tolérance exactes sont définies comme les intervalles les plus larges pour lesquels la solution optimale x^* reste optimale, avec des définitions formelles pour les bornes inférieures et supérieures des profits et des poids.

4.3 Méthodes exactes pour l'analyse de tolérance

Pisinger et Saidi proposent un algorithme exact basé sur la programmation dynamique avec une complexité amortie de $O(c \log n)$ par item. L'idée centrale repose sur la réutilisation des sous-problèmes dans une structure arborescente. L'algorithme calcule les limites de tolérance en résolvant des sous-problèmes où chaque item est successivement retiré, utilisant des relations précises entre la solution originale et les solutions des sous-problèmes pour déterminer les intervalles de tolérance.

La structure arborescente exploitant les sous-problèmes chevauchants permet de réduire la complexité computationnelle. Chaque niveau de l'arbre ajoute $n/2^i$ items à 2^i sous-problèmes, conduisant à une complexité totale de $O(nc \log n)$. Cette approche innovante permet de réutiliser les calculs intermédiaires et d'éviter les redondances dans la résolution des sous-problèmes similaires.

4.4 Méthodes approximatives

L'utilisation de la borne supérieure de Dantzig permet d'obtenir des limites de tolérance approximatives en temps $O(\log n)$ par item. Cette borne, calculée à partir de la solution du problème relaxé, fournit une estimation rapide mais potentiellement sous-optimale des intervalles de tolérance. Une alternative plus rapide mais moins précise utilise la borne de Dembo-Hammer, calculable en temps constant mais avec une qualité de solution généralement inférieure.

4.5 Résultats expérimentaux et comparaisons

Les tests expérimentaux sur différentes catégories d'instances montrent que l'algorithme exact résout toutes les instances en moins d'une seconde, tandis que les approches approximatives sont extrêmement rapides mais moins précises. La qualité des limites de tolérance varie significativement selon le type d'instance, avec des intervalles plus larges pour les instances non corrélées et plus étroits pour les instances fortement corrélées. Ces résultats soulignent l'importance du compromis entre précision et temps de calcul dans le choix de la méthode d'analyse.

4.6 Synthèse des contributions méthodologiques

L'analyse de tolérance pour le problème de sac à dos (knapsack) 0-1, telle que développée par Pisinger et Saidi (2017), représente un outil puissant pour l'analyse de robustesse des solutions optimales. Les travaux futurs pourraient s'orienter vers l'extension de ces résultats à d'autres problèmes combinatoires, l'intégration de l'analyse de tolérance dans les solveurs commerciaux, et le développement de méthodes pour l'analyse de tolérance multi-paramètres. Cette approche trouve des applications naturelles dans les problèmes de pricing pour la génération de colonnes et la séparation de coupes, où la connaissance des limites de tolérance peut accélérer la résolution des sous-problèmes.

5 Conclusion et Perspectives

5.1 Synthèse du cadre méthodologique et lacunes de la littérature

Le présent travail a réalisé une synthèse critique et détaillée des outils d'optimisation mathématique pertinents pour l'analyse des jeux de fantasy cyclisme. Notre principal apport réside dans la cartographie complète et intégrée de la chaîne méthodologique nécessaire pour transformer la solution optimale théorique en une solution quantifiablement robuste.

Cette revue s'est concentrée sur deux axes : l'analyse approfondie du cadre MIP générique de Beliën et al. (2017) (Section 3), l'outil formel le plus complet identifié pour la modélisation multi-étapes et la gestion des joueurs remplaçants ; et la synthèse de l'arsenal méthodologique de l'Analyse Post-Optimale (Section 4), notamment les travaux de Pisinger et Saidi (2017) sur l'analyse de tolérance. La contribution essentielle de cette étude réside dans la proposition de convergence entre ces deux axes, démontrant que la modélisation dynamique de type sac à dos (knapsack) (knapsack) doit être systématiquement complétée par l'analyse de tolérance.

Cette synthèse met en évidence plusieurs lacunes significatives dans la littérature actuelle. La principale lacune est l'absence d'études intégrant simultanément l'optimisation multi-périodes et l'analyse de robustesse dans le contexte spécifique du cyclisme fantasy. De plus, le manque de modèles stochastiques adaptés à l'incertitude des performances sportives, ainsi que la rareté des travaux sur l'optimisation robuste multi-objectifs, représentent des défis majeurs pour l'applicabilité pratique des modèles.

5.2 Perspectives de recherche, défis d'adaptation et implications pratiques

L'analyse des lacunes ouvre directement la voie aux futures recherches. L'analyse de tolérance étudiée (Pisinger et Saidi, 2017) doit être la base pour de nouvelles applications dans le fantasy cyclisme, notamment l'identification des coureurs critiques (ceux dont la performance affecte le plus la solution) et la gestion robuste des transferts par l'évaluation de la stabilité de l'équipe.

Cependant, l'adaptation de ces méthodes au modèle dynamique de Beliën pose plusieurs défis méthodologiques. La gestion de la dimension temporelle et des transferts introduit une complexité accrue, nécessitant le développement d'approches novatrices pour l'analyse de sensibilité dans un contexte dynamique. Cette complexité substantielle appelle à l'élaboration d'algorithmes optimisés et de méthodes d'approximation efficaces.

Au-delà des contributions théoriques, l'intégration de ces analyses a des implications pratiques directes. Ces outils pourraient être intégrés dans des plateformes d'aide à la décision pour les participants, des systèmes de recommandation de transferts robustes et des analyses de risque personnalisées. Pour les organisateurs de jeux, ces analyses pourraient guider la conception de règles équilibrées et stimulantes. En définitive, l'étude des variantes du problème de sac à dos dans le cadre des jeux de fantasy cyclisme se positionne à la frontière entre la recherche opérationnelle et l'analyse sportive, un champ dont l'importance croissante est confirmée par des revues récentes (Durán, 2021). Ce domaine offre un terrain fertile pour l'innovation méthodologique.

Bibliographie

Références

[Ausloos, 2024] Ausloos, M. (2024). *Should one (be allowed to) replace the Cipollini's ?*. Annals of Operations Research, 1-19. DOI : 10.1007/s10479-024-06206-y.

[Bampis et al., 2022] Bampis, E., Escoffier, B., & Teiller, A. (2022). *Multistage knapsack*. Journal of Computer and System Sciences, 126, 106-118. DOI : 10.1016/j.jcss.2022.01.002.

[Beliën et al., 2017] Beliën, J., Goossens, D., & Van Reeth, D. (2017). *Optimization modelling for analyzing fantasy sport games*. INFOR : Information Systems and Operational Research, 55(4), 275-294. DOI : 10.1080/03155986.2017.1279899.

[Beliën et al., 2011] Beliën, J., Goossens, D., Van Reeth, D., & De Boeck, L. (2011). *Using mixed-integer programming to win a cycling game*. INFORMS Transactions on Education, 11(3), 93-99. DOI : 10.1287/ited.1110.0062.

[Durán, 2021] Durán, G. (2021). *Sports scheduling and other topics in sports analytics : a survey with special reference to Latin America*. TOP, 29(1), 125-155.

[Kumabe & Yoshida, 2024] Kumabe, S., & Yoshida, Y. (2024). *Average sensitivity of the knapsack problem*. arXiv preprint arXiv :2405.13343.

[Pisinger & Saidi, 2017] Pisinger, D., & Saidi, A. (2017). *Tolerance analysis for 0-1 knapsack problems*. European Journal of Operational Research, 258(3), 866-876. DOI : 10.1016/j.ejor.2016.10.054.