



HAL
open science

Gestion robuste de la production électrique à horizon court terme

Sinda Ben Salem

► **To cite this version:**

Sinda Ben Salem. Gestion robuste de la production électrique à horizon court terme. Autre. Ecole Centrale Paris, 2011. Français. NNT : 2011ECAP0014 . tel-00594242

HAL Id: tel-00594242

<https://theses.hal.science/tel-00594242>

Submitted on 19 May 2011

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

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



ÉCOLE CENTRALE DES ARTS
ET MANUFACTURES
« ÉCOLE CENTRALE PARIS »

THÈSE
présentée par

Sinda BEN SALEM

pour l'obtention du
GRADE DE DOCTEUR

Spécialité : **Génie Industriel**
Laboratoire d'accueil : **Laboratoire de Génie industriel**

Gestion robuste de la production électrique à horizon court-terme

soutenue le **11 mars 2011** devant un jury composé de :

| | | |
|-----------------------------|---|--------------------|
| Pr Alexandre Dolgui, | <i>Ecole des Mines de Saint-Etienne</i> | Président |
| Pr Abdel Lisser, | <i>Université de Paris Sud</i> | Rapporteur |
| Pr Adam Ouorou, | <i>Orange R&D</i> | Rapporteur |
| Pr Vincent Mousseau, | <i>Ecole Centrale Paris</i> | Examineur |
| M. Ala Ben Abbes, | <i>EDF R&D</i> | Examineur |
| Pr. Michel Minoux, | <i>Université de Paris VI</i> | Directeur de thèse |

Cette thèse a été effectuée conjointement au Laboratoire Génie Industriel de l'Ecole Centrale Paris et dans le département **OSIRIS d'EDF R&D** sous la forme d'un contrat CIFRE



N° 2011 ECAP 0014

Remerciements

Je souhaite en premier lieu adresser mes sincères remerciements à l'ensemble des membres de mon jury de thèse : Pr. Adam Ouorou et Pr. Abdel Lisser, qui ont accepté d'être les rapporteurs de ma thèse, Pr. Vincent Mousseau et M. Ala Ben Abbes, examinateurs et Pr. Alexandre Dolgui qui a accepté de présider le jury de thèse.

Ce projet n'aurait pas vu le jour sans la confiance que m'a accordée mon directeur de thèse, Pr. Michel Minoux. J'ai eu l'honneur de travailler avec un éminent chercheur et un grand professeur, auprès duquel j'ai beaucoup appris. Un encadrant à l'écoute, trouvant toujours les bons mots pour m'encourager dans mes avancées, pour m'aider dans mes réflexions, et pour m'éclairer dans mes interrogations. Merci infiniment pour votre confiance.

Mes travaux de thèse se sont déroulés dans le cadre d'un partenariat entre le laboratoire Génie Industriel de l'Ecole Centrale Paris et le département OSIRIS (Optimisation, SIMulation, RISques & Statistiques) à EDF R&D. D'où la double chance de faire de belles rencontres dans les deux milieux.

Je remercie donc tout d'abord Jérôme Pigat, ancien chef de groupe et Yannick Jacquemart, chef du département OSIRIS, pour leur confiance en me proposant ce projet de thèse.

J'ai vécu nombre de bons souvenirs au sein de mon équipe "R37", et pour tous ces moments je remercie Jean-Marc Luel et tous mes collègues du groupe qui m'ont imprégnée de leur savoir, qui allait des contraintes sur la Durance aux actualités brûlantes. Merci Jean-Claude, Laurent, Xavier, Bernard, Seb, Constantin, Félix.. Merci tout R37.

J'ai également eu la chance de voir la création à OSIRIS d'une équipe dédiée aux méthodes d'optimisation. Cette nouvelle structure m'a permis de travailler et de discuter avec des personnes intéressantes, captivantes, voire attachantes :) Merci donc à Anes, Véronique, Pierre, Thomas... Merci à toute l'équipe R36.

Et je ne peux aborder mes quatre années passées à EDF sans remercier Jérôme Quenu et Ala Ben Abbes. J'ai beaucoup appris à leurs côtés, autant humainement que scientifiquement. C'était une véritable chance que de travailler avec vous. Et je suis fier d'avoir eu la chance de croiser vos chemins. Merci encore.

Même si j'ai passé la majeure partie de mon temps à EDF R&D, ceci ne m'a pas empêchée de vivre d'excellents moments dans mon laboratoire d'accueil, le LGI. Un grand Merci à Emna, Asma, Céline, Anne, Sylvie, Myriam, Adel et toute l'équipe du LGI pour votre accueil constamment chaleureux malgré mes passages toujours trop courts. C'était un vrai plaisir pour moi que de vous avoir connus.

Enfin, un grand merci à mes supportrices et amies les plus fidèles : Linda, Neila, Amel, Salma ! A ma merveilleuse famille qui a cru en moi, et qui m'a donné la force et le courage pour réaliser ce projet. Et une mention particulière à Lina, mon adorable nièce qui a insisté pour assister à ma soutenance (elle avait 15 mois :)).

Enfin, les mots et les pages ne me suffiront jamais pour remercier l'être qui m'a tant donné pour réussir ce travail. Son soutien et ses encouragements m'ont bercée durant toutes ces années. Son plus beau soutien était sa fierté que je lisais dans ses yeux, et qui est ma plus belle récompense... Merci Rami.

Résumé

Dans un marché électrique concurrentiel, EDF a adapté ses outils de gestion de production pour permettre une gestion optimale de son portefeuille, particulièrement sur les horizons journalier et infra-journalier, derniers leviers pour une gestion optimisée de la production. Et plus l'horizon d'optimisation s'approche du temps réel, plus les décisions prises aux instants précédents deviennent structurantes voire limitantes en terme d'actions. Ces décisions sont aujourd'hui prises sans tenir compte du caractère aléatoire de certaines entrées du modèle.

En effet, pour les décisions à court-terme, la finesse et la complexité des modèles déjà dans le cas déterministe ont souvent été un frein à des travaux sur des modèles tenant compte de l'incertitude.

Pour se prémunir face à ces aléas, des techniques d'optimisation en contexte incertain ont fait l'objet des travaux de cette thèse. Nous avons ainsi proposé un modèle robuste de placement de la production tenant compte des incertitudes sur la demande en puissance. Nous avons construit pour cette fin un ensemble d'incertitude permettant une description fine de l'aléa sur les prévisions de demande en puissance. Le choix d'indicateurs fonctionnels et statistiques a permis d'écrire cet ensemble comme un polyèdre d'incertitude.

L'approche robuste prend en compte la notion de coût d'ajustement face à l'aléa. Le modèle a pour objectif de minimiser les coûts de production et les pires coûts induits par l'incertitude. Ces coûts d'ajustement peuvent décrire différents contextes opérationnels. Une application du modèle robuste à deux contextes métier est menée avec un calcul du coût d'ajustement approprié à chaque contexte.

Enfin, le présent travail de recherche se situe, à notre connaissance, comme l'un des premiers dans le domaine de la gestion optimisée de la production électrique à court terme avec prise en compte de l'incertitude. Les résultats sont par ailleurs susceptibles d'ouvrir la voie vers de nouvelles approches du problème.

Mots-clés

Optimisation robuste, programmation linéaire mixte, gestion de production électrique, contexte incertain

Abstract

Robust Optimization is an approach typically offered as a counterpoint to Stochastic Programming to deal with uncertainty, especially because it doesn't require any precise information on stochastic distributions of data. In the present work, we deal with challenging unit-commitment problem for the French daily electricity production under demand uncertainty. Our contributions concern both uncertainty modelling and original robust formulation of unit-commitment problem. We worked on a polyhedral set to describe demand uncertainty, using statistical tools and operational indicators. In terms of modelling, we proposed robust solutions that minimize production and worst adjustment costs due to uncertainty observation. We study robust solutions under two different operational contexts. Encouraging results to the convex unit-commitment problems under uncertainty are thus obtained, with interesting research topics for future work.

Keywords

Robust optimization, mixed integer programming, unit-commitment problem, approach under uncertainty

Table des matières

| | |
|---|-----------|
| Avant propos | 7 |
| 1 Introduction à l'optimisation sous incertitude | 15 |
| 1.1 Programmation avec recours | 16 |
| 1.1.1 Modèles avec recours à deux niveaux | 16 |
| 1.1.2 Résolution dans le cas linéaire : Méthode <i>L-shaped</i> | 17 |
| 1.1.3 Modèles "Multi-étapes" : Modélisation de l'espace des scénarios | 19 |
| 1.1.4 Résolution par programmation dynamique stochastique | 21 |
| 1.2 Modèles stochastiques sous contraintes en probabilités | 22 |
| 1.2.1 Théorie | 22 |
| 1.2.2 Contraintes probabilistes indépendantes | 24 |
| 1.2.3 Contraintes probabilistes jointes | 24 |
| 1.3 Conclusions & Applications | 26 |
| 2 Optimisation robuste | 29 |
| 2.1 Problèmes linéaires | 29 |
| 2.1.1 Incertitude par colonnes | 30 |
| 2.1.2 Incertitude par lignes | 32 |
| 2.2 Problèmes discrets | 37 |
| 2.2.1 Travaux de Bertsimas & Sim dans le cas discret | 37 |
| 2.2.2 Incertitude sous forme de scénarios | 38 |
| 2.3 Problèmes robustes dynamiques | 40 |
| 2.3.1 Règles de décision linéaires | 41 |
| 2.3.2 Règles de décision linéaires par morceaux | 41 |
| 3 Introduction à la gestion de production électrique à EDF | 45 |
| 3.1 EDF : 1 ^{er} producteur français | 45 |
| 3.1.1 EDF en quelques chiffres | 45 |
| 3.1.2 Missions & Objectifs | 45 |
| 3.1.3 Le parc électrique EDF | 46 |
| 3.1.4 Management d'énergies | 48 |
| 3.2 Le processus journalier | 50 |
| 3.2.1 Contexte et réglementation | 50 |

| | | |
|----------|--|-----------|
| 3.2.2 | Outils d'optimisation | 52 |
| 3.3 | Le processus infra-journalier | 56 |
| 3.3.1 | Le Mécanisme d'Ajustement | 56 |
| 3.3.2 | Les re-déclarations | 57 |
| 3.3.3 | Vers une gestion optimisée des re-déclarations | 57 |
| 3.4 | Aléas physiques et financiers : Limites d'une vision déterministe | 58 |
| 4 | Formulation robuste du problème de gestion de production électrique à court terme | 61 |
| 4.1 | Introduction | 61 |
| 4.2 | Modèle robuste proposé | 62 |
| 4.2.1 | Notations | 63 |
| 4.2.2 | Fonction objectif | 63 |
| 4.3 | Etude de la fonction de coût d'ajustement $\Phi(P)$ | 63 |
| 4.3.1 | Modélisation d'un contexte industriel | 63 |
| 4.3.2 | Formulation mathématique | 65 |
| 4.3.3 | Convexité de $\Phi(P)$ | 66 |
| 4.4 | Modèle robuste : Exemples en faible dimension | 67 |
| 4.4.1 | Exemples en une dimension | 67 |
| 4.4.2 | Exemple en dimension 2 de la vision Pénalités | 72 |
| 4.5 | Résolution du problème général | 74 |
| 4.5.1 | Algorithme proximal conceptuel | 74 |
| 4.5.2 | Algorithme proximal effectif | 76 |
| 4.5.3 | Tests d'arrêt | 78 |
| 4.5.4 | Implémentation | 79 |
| 4.6 | Comparatif avec les travaux de Zhang et Guan [53] | 79 |
| 5 | Construction de l'ensemble d'incertitude | 83 |
| 5.1 | Prévision de la consommation électrique | 83 |
| 5.1.1 | Principe | 83 |
| 5.1.2 | Outils | 83 |
| 5.1.3 | Hypothèses de travail | 85 |
| 5.2 | Construction du polyèdre | 85 |
| 5.2.1 | Première construction à partir de scénarios d'erreurs | 85 |
| 5.2.2 | Consolidation de la construction | 86 |
| 5.2.3 | Polyèdre final | 89 |
| 5.3 | Validation basée sur un polyèdre | 89 |
| 5.3.1 | A partir d'un générateur | 90 |
| 5.3.2 | Validation avec Tirage uniforme | 90 |
| 5.3.3 | Validation avec tirage et changement de base | 92 |

| | | |
|----------|--|------------|
| 6 | Le modèle robuste avec pénalités : Analyse des résultats | 97 |
| 6.1 | Hypothèses de travail | 97 |
| 6.1.1 | Modèle déterministe | 97 |
| 6.1.2 | Modèle robuste | 101 |
| 6.1.3 | Développement | 102 |
| 6.1.4 | Données générales | 102 |
| 6.1.5 | Parc de production | 102 |
| 6.2 | Simulations : Mode opératoire | 103 |
| 6.2.1 | Objectif | 103 |
| 6.2.2 | Valorisation des écarts en puissance | 103 |
| 6.3 | Distribution des coûts : impact des pénalités | 104 |
| 6.3.1 | Pénalisation plus marquée pour la sous-production | 104 |
| 6.3.2 | Pénalisation équivalente de la sous-production et sur-production | 107 |
| 6.3.3 | Pénalisation plus marquée pour la sur-production | 109 |
| 6.3.4 | Interprétation et analyse | 111 |
| 7 | Le modèle robuste avec coûts de réorganisation | 115 |
| 7.1 | Introduction | 115 |
| 7.2 | Coût de réorganisation : définitions & propriétés | 116 |
| 7.3 | Approximations de la fonction initiale | 118 |
| 7.3.1 | Algorithme glouton | 119 |
| 7.4 | Approximation convexe | 122 |
| 7.4.1 | Evolution sur la modélisation du parc électrique | 122 |
| 7.4.2 | Etude expérimentale de la fonction φ | 122 |
| 7.4.3 | Distributions des coûts | 125 |
| 7.5 | Impact du polyèdre sur la distribution des coûts | 127 |
| 7.5.1 | Principe | 127 |
| 7.5.2 | Implémentation | 128 |
| 7.5.3 | Résultats | 128 |
| | Conclusions & Perspectives | 133 |
| | Bibliographie | 136 |

Avant propos

Issu du latin *Incertus* « *qui n'est pas fixé, déterminé à l'avance* », un événement est dit incertain dès lors que sa prévision exacte est impossible. Keynes disait

Notre seule certitude est que nous mourrons

Et la grande incertitude de l'humanité est simplement son avenir.

Si nos ancêtres ont répondu à ce problème par de la sagesse ou du fatalisme (le stoïcisme avait comme enseignement « *Ne demande point que les choses arrivent comme tu le désires, mais désire qu'elles arrivent comme elles arrivent, et tu prospéreras toujours* »), notre ère post moderne a cherché à mieux définir, comprendre et réduire l'incertitude, par des travaux en sociologie, en philosophie, et surtout en sciences. Ces travaux alimentent ainsi ce désir grandissant de maîtrise dans bien des domaines (à la météo de *maîtriser* le temps, aux instituts de sondage de *maîtriser* la réalité des résultats prédits et au scientifiques de *maîtriser* la "réalité"). Ainsi, afin de gérer l'incertain ou l'imprévisible, il a fallu d'abord l'accepter, c'est à dire la considérer comme une composante inhérente de la réalité.

Au départ, la programmation mathématique déterministe s'est développée comme outil d'aide aux décideurs. Les décisions étaient prises à l'aide de modèles déterministes, néanmoins le contexte demeurait incertain voire imprévisible.

La notion d'incertitude dans la programmation mathématique est apparue pour la première fois dans les années 50 avec les travaux de Bellman, Dantzig [20], et elle a rencontré depuis un développement rapide.

Sa nouveauté était la prise en compte du phénomène aléatoire dans les modèles d'optimisation eux-mêmes. Des données telles que les coûts de production, les prix des marchés, les délais de livraison ou la demande des clients sont alors des données incertaines ou aléatoires.

Dans le cas de la gestion de production électrique à EDF, l'incertitude fut modélisée dans les outils optimisant les décisions à moyen-terme, c'est à dire sur un horizon d'une à cinq années. Pour les décisions à court-terme, la finesse et la complexité des modèles déjà dans le cas déterministe ont souvent été un frein à des travaux sur des modèles tenant compte de l'incertitude.

Le présent travail de recherche se situe, à notre connaissance, comme l'un des *premiers* sur un *modèle d'optimisation de la production électrique à court terme avec une prise en compte de l'incertitude*, dont les résultats sont susceptibles d'ouvrir la voie vers de nouvelles approches du problème de gestion de production dans l'incertain.

La première partie (chapitres 1 à 3) est une introduction au contexte *scientifique et industriel* ayant motivé ce travail.

Le premier chapitre de la thèse est *une introduction à l'optimisation sous incertitude*, généralement connue sous l'appellation de *programmation stochastique*. On présente en particulier deux familles de problèmes souvent employées dans le cadre de la gestion de production électrique, que sont *la programmation stochastique avec recours*, et *la programmation sous contraintes en probabilités*.

Dans la littérature, une autre grande famille de modèles que sont les approches robustes, est souvent distinguée de la programmation stochastique, même si elle traite également de problèmes en contexte incertain. Le deuxième chapitre se consacre à une introduction spécifique à *la programmation robuste*.

Née du besoin de proposer des solutions à des problèmes d'optimisation sous incertitude pour lesquels on ne dispose pas d'une connaissance précise des lois de probabilités de leurs paramètres aléatoires, l'approche robuste est un élément dans nos travaux avec la proposition d'un *modèle de gestion de production robuste face aux aléas*.

Le troisième chapitre est *une introduction à la gestion de production électrique à EDF*. Après une brève présentation du contexte d'ensemble, nous nous intéressons à la *Gestion d'énergies* à horizon court-terme, et plus particulièrement la gestion journalière et infra-journalière de la production.

La deuxième partie (chapitres 4 et 5) de la thèse décrit *l'approche robuste proposée* pour répondre à cette problématique :

Dans le chapitre 4, nous présentons une formulation robuste nouvelle pour répondre à la problématique de gestion de production en contexte incertain. A travers cette formulation, on introduit la notion de *coût d'ajustement*. L'objectif du modèle étant de *minimiser les coûts de production et les pires coûts induits par l'observation de l'incertitude*. Les coûts induits par la réalisation de l'incertitude peuvent ainsi *décrire différents contextes opérationnels*, selon les besoins du métier.

Nous montrons que ce problème peut se reformuler comme la minimisation d'une fonction convexe non partout différentiable. Enfin nous présentons l'algorithme proximal utilisé pour la résolution du problème.

Dans le chapitre 5, après une brève présentation du processus de prévision de la consommation électrique, nous exposons notre démarche pour construire *un polyèdre décrivant*

l'incertitude sur la demande en puissance. Nous avons fait le choix d'indicateurs basés sur un ensemble de scénarios de demande en puissance, et permettant une bonne description de l'erreur sur la demande. Cet ensemble d'incertitude peut également être un outil de validation des modèles, à travers un tirage uniforme de scénarios de demande dans le polyèdre.

La troisième partie (chapitres 6 et 7) des travaux comporte une discussion approfondie des *principaux résultats obtenus* ainsi que les conclusions et ouvertures au-delà de ces travaux.

Le chapitre 6 décrit les hypothèses de travail, le modèle déterministe de départ, et le modèle robuste proposé. Nous présentons également le mode opératoire des simulations, qui a permis *une étude comparative rigoureuse* entre l'approche déterministe et l'approche robuste.

Le chapitre 7 propose *une nouvelle modélisation du coût d'ajustement*, permettant de valoriser, de manière approchée, la réorganisation des programmes de production suite à la réalisation d'une incertitude.

Cette modélisation nouvelle est intégrée dans le modèle robuste, et validée selon le mode opératoire décrit dans le chapitre 6. On étudie la stabilité des solutions robustes obtenues à travers *une analyse de sensibilité* du modèle robuste vis à vis du *polyèdre d'incertitude*. Des résultats encourageants sont observés, confirmant l'arbitrage possible entre la stabilité du modèle ou la réduction du coût de gestion.

Le dernier chapitre est une conclusion générale des travaux présentés. Nous présentons un bilan des contributions apportées sur les différents axes de recherche abordés. Nous ouvrons enfin la voie sur nombre de perspectives et pistes de réflexion, comme l'extension du modèle robuste au cadre non convexe.

Chapitre 1

Introduction à l'optimisation sous incertitude

Introduction

La programmation stochastique se positionne comme un complément aux modèles classiques d'optimisation (tels que la programmation linéaire, dynamique ou non-linéaire) lorsque ces derniers comportent des paramètres aléatoires ou incertains.

Néanmoins, du fait de la prise en compte du caractère incertain de certains paramètres, les problèmes deviennent plus difficiles à résoudre, tant sur le plan théorique que numérique. Et la difficulté apparaît dès la phase de modélisation de l'aléa, avec un équilibre délicat à atteindre entre un modèle décrivant finement la réalité et un modèle traitable.

Dans ce chapitre, nous nous intéressons à deux familles des problèmes dits stochastiques. Notre choix s'est porté sur les modèles avec recours et les modèles avec contraintes en probabilités.

Ces deux approches ont pour qualité d'avoir proposé une modélisation intéressante pour de nombreuses problématiques métier. De nombreux travaux de recherche (notamment à EDF R&D et à France Télécom) ont servi à décrire des besoins du métier à l'aide de problèmes avec recours ou à l'aide de problèmes sous contraintes en probabilités.

Nous présenterons donc, dans une première section, la programmation avec recours, et la modélisation de l'ensemble d'incertitude spécifique à cette approche. En deuxième partie, nous introduisons les problèmes d'optimisation sous contraintes en probabilités. Nous concluons par quelques applications de ces approches pour les problèmes de gestion de production électrique à EDF.

1.1 Programmation avec recours

1.1.1 Modèles avec recours à deux niveaux

La programmation stochastique avec recours s'intéresse particulièrement à des problèmes où on doit prendre une décision sur le champ sous présence d'incertitude, sans attendre la réalisation de certaines variables aléatoires. Des décisions ou actions correctives peuvent être prises ensuite après réalisation de l'aléa.

Le mot "*recours*" révèle la possibilité dont on dispose de remédier à une décision éventuellement non pertinente à t_0 , en mettant en oeuvre des actions correctives, pour satisfaire les contraintes concernant les instants ultérieurs $t_i, i > 0$.

L'ensemble des variables de décision est formé de deux ensembles :

- Un certain nombre de décisions doit être pris avant réalisation de l'incertitude. Ces décisions sont appelées "*decisions de premier niveau*" ou décisions de la première étape.
- Un certain nombre de décisions est pris après observation de l'aléa. Ce sont les "*décisions de seconde étape*" ou décisions de la seconde étape.

Le problème classique de la programmation stochastique linéaire avec recours à 2 étapes s'écrit de la manière la plus générale (voir par exemple Kall & Wallace [32]) sous la forme suivante :

$$\begin{aligned}
 & \min_x \quad c^T x + \mathbb{E}_\xi \left[\min_y q(\xi)^T y(\xi) \right] \\
 \text{s.c : } & Ax = b, \\
 & T(\xi)x + W(\xi)y(\xi) = h(\xi) \\
 & x \geq 0, y(\xi) \geq 0
 \end{aligned} \tag{1.1}$$

Avec la matrice de contraintes A ($A : m_1 \times n_1$) et les vecteurs c ($c \in \mathbb{R}^{n_1}$) et b ($b \in \mathbb{R}^{m_1}$).

Les "*décisions de premier niveau*" sont représentées par le vecteur x , de dimension n_1 .

Le coût des décisions pour la seconde étape ou "*décisions de second niveau*" $y(\xi)$ ($y \in \mathbb{R}^{n_2}$) est minimisé en espérance, par rapport à la variable aléatoire ξ , de même pour le vecteur h ($h \in \mathbb{R}^{m_2}$), q ($q \in \mathbb{R}^{n_2}$) ou les matrices de contraintes T et W ($T : m_2 \times n_1$ et $W : m_2 \times n_2$)

Ce type de problème est connu également sous le nom de "*here and now problems*", puisqu'il est nécessaire de déterminer à un moment t_0 certaines variables de décision qui ne peuvent plus changer par la suite. Des extensions de ce modèles à des processus de décision à plus de deux étapes ont été étudiés en particulier par Birge [14], avec un aléa très souvent modélisé sous la forme d'un arbre de scénarios.

On cherche ainsi à minimiser les coûts des variables de la première étape x , mais on veille en même temps à ce que les coûts au second niveau, après observation de ξ soient minimaux en moyenne.

Cette double minimisation est la deuxième famille de contraintes couplant les variables de

décision fixes x et les variables de décision $y(\xi)$ dépendant de l'aléa. Les variables $y(\xi)$ sont calculées de manière à minimiser les conséquences des décisions prises à la première étape. Enfin, le fait que le vecteur de décisions de première étape x ne dépende pas de ξ est caractéristique de la *non-anticipativité*. Les décisions de première étape seront donc appliquées indépendamment de la réalisation de l'aléa. Si cette propriété n'était pas respectée, alors on aurait pour toute réalisation ξ_i un vecteur de décisions x_i pour la première étape qui minimiserait la fonction de coût.

Dans de nombreuses applications, on se restreint à un espace de réalisations discret avec des scénarios comme suit :

Notons ξ la variable aléatoire décrivant l'aléa. Notons Ξ l'espace dans lequel la variable ξ prend ses valeurs et $\mathbb{E}_\xi(\cdot)$ l'opérateur espérance par rapport à cette variable. Les hypothèses de discrétisation sont :

- On suppose l'ensemble des réalisations possibles comme étant un ensemble fini $\Xi = [\xi^1, \xi^2, \dots, \xi^S]$, les réalisations ξ^i seront appelées scénarios, S étant le nombre de scénarios,
- On suppose qu'il existe une mesure de probabilité \mathbb{P} telle que $\mathbb{P}\{\xi = \xi^s\} = p_s$, p_s est donc la probabilité que le scénario ξ^s se réalise

Sous ces hypothèses, le modèle avec recours à deux étapes s'écrit de la façon suivante :

$$\begin{aligned}
 & \min_{x, y^i, i \in S} \quad c^T x + \sum_{i \in S} p_i q_i^T y^i \\
 \text{s.c :} \quad & Ax = b, \\
 \text{avec} \quad & T^i x + W^i y^i = h^i, \quad i = 1 \dots S \\
 & x \geq 0, y^i \geq 0, \quad i = 1 \dots S
 \end{aligned} \tag{1.2}$$

Avec p_i la probabilité d'occurrence du scénario i , et q_i^T le coût de la décision y^i .

Dans le cas de variables continues, le modèle avec recours ci-dessus se présente comme un programme linéaire. Sa résolution est rapide lorsque la taille du problème reste raisonnable. Dans le cas où le nombre de scénarios est très élevé, il est préférable de s'orienter vers des méthodes de résolution spécifiques à cette famille de problèmes. Parmi les approches possibles, nous citons la méthode L-shaped détaillée ci-dessous.

1.1.2 Résolution dans le cas linéaire : Méthode *L-shaped*

En s'inspirant de la méthode de décomposition de Benders de 1962 [9] qui consiste à partitionner les variables d'un programme linéaire en deux catégories, Van Slyke et Wets ont développé en 1969 [52] la méthode du *L-shaped*. L'idée de l'algorithme *L-shaped* consiste à décomposer le problème en deux parties distinctes, un problème *maître* et des problèmes *esclaves* ou *satellites*, tels que :

- Le problème maître calcule les variables x du premier niveau avec une représentation approchée des problèmes du deuxième niveau,

- Un problème satellite $Prob^s$ pour chaque scénario s . Ce problème calcule les variables y_s de recours pour un scénario s donné. Il a en entrée les valeurs des variables x du premier niveau (il y a autant de problèmes satellites que de scénarios représentés dans le problème)
- La résolution de chaque sous-problème permet d'obtenir deux informations : la valeur de la fonction objectif ainsi qu'un sous-gradient. Ces informations permettent de définir une coupe qui approche la valeur de la fonction de recours (ou encore la fonction de deuxième étape $\sum_s p_s q_s^T y_s$). La coupe s'ajoute au problème maître et améliore par là l'approximation des sous-problèmes.
- L'algorithme s'arrête lorsque la fonction de recours est approchée, d'après l'utilisateur, de manière suffisamment précise.

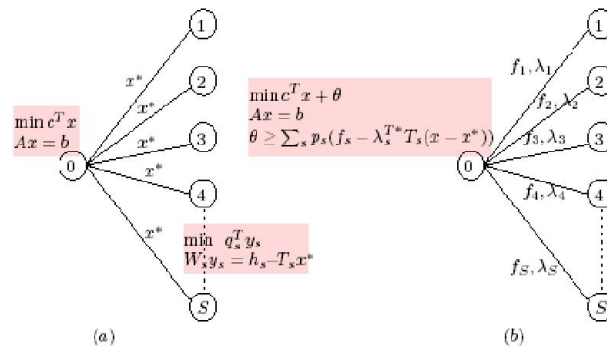


FIGURE 1.1 – Flux d'information du maître aux sous-problèmes pendant la première itération : (a) les solutions x^* sont transférées du maître à tout sous-problème, (b) la valeur de la fonction f_s et le vecteur dual optimal λ_s sont envoyés de tout sous-problème au maître ; à partir de ces informations une nouvelle coupe se rajoute au problème maître

La figure 1.1 illustre l'interaction entre problème maître et problèmes esclaves.

La convergence de la méthode a été prouvée par Van Slyke et Wets [52]. Une variante de l'algorithme *L-shaped* a été proposée par Birge et Louveaux en 1988 [15] et remplace la coupe d'optimalité agrégée par plusieurs coupes par scénarios. D'où l'appellation *L-shaped* version *multi-coupes*.

Extension au cas non-linéaire

Dans le cas où les variables du premier niveau sont purement binaires, Laporte et Louveaux [34] ont proposé en 1993 une variante du *L-shaped*.

Leur algorithme calcule les variables du premier niveau par "branch & bound", et utilise les coupes d'optimalité et réalisabilité pour approcher la valeur de la fonction de recours. Cette variante fut très utile pour résoudre les problèmes de "location-routing". Néanmoins, cette méthode ne fonctionne plus si une seule des variables du premier niveau est continue.

Caroe et Tind ont proposé en 1998 [17] un algorithme de résolution pour des variables de premier niveau mixtes en utilisant la programmation non-linéaire pour approcher la fonction de recours.

Mais selon l'étude bibliographique de Sahinidis datant de 2004 [48], aucune résolution pratique n'a encore abouti avec cet algorithme.

Toujours d'après Sahinidis, la programmation stochastique en nombre entiers suscite néanmoins de plus en plus d'intérêt et certains algorithmes commencent à présenter des résultats encourageants.

En effet, cet intérêt pour la programmation stochastique en nombres entiers est limité principalement lorsque les variables de niveaux supérieurs sont en nombres entiers.

Les modèles avec recours à deux étapes "Two-Stage problems" décrivent la possibilité d'apporter des décisions correctives après observation de l'aléa.

Le modèle multi-étapes ("Multi-Stage problems") est l'extension du modèle en deux étapes dans le cas où le problème nécessite plusieurs prises de décision, avec une observation partielle de l'aléa à chaque période de décision. Le modèle multi-étapes peut être schématiquement présenté comme la succession de modèles à deux étapes, et ce sur la période totale d'étude.

Les modèles "multi-stage" sont généralement étudiés avec un espace des scénarios d'aléas représentés sous forme d'*arbre de scénarios*. Nous développons dans la section suivante cette modélisation, puis nous introduirons une méthode de résolution - la programmation dynamique stochastique - appropriée à ces modèles.

1.1.3 Modèles "Multi-étapes" : Modélisation de l'espace des scénarios

La modélisation de l'espace des scénarios sous forme d'un arbre est souvent appliquée dans le cas d'un modèle avec recours multi-étapes.

Soit l'arbre de la figure 1.2 suivante :

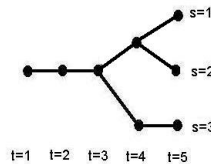


FIGURE 1.2 – Arbre de scénarios classiques

matrice), et les décisions à l'étape $i + 1$. Le problème formulé en 1.3 peut être écrit comme un grand problème linéaire. Mais la dimension augmentant exponentiellement en fonction du nombre de scénarios et des branchements (ou étapes), la résolution des modèles avec recours à deux ou plusieurs étapes atteint rapidement ses limites.

1.1.4 Résolution par programmation dynamique stochastique

La programmation dynamique fut introduite pour la première fois dans les années '50 par Bellman [3]. Cette technique s'intéresse aux problèmes écrits sous la forme d'un système dynamique. Ce système est défini par :

$$x_{i+1} = f_i(x_i, u_i, w_i) \quad (1.4)$$

avec x_i l'état du système à la $i^{\text{ème}}$ période, u_i la décision à la période i , et w_i une variable aléatoire modélisant une perturbation de l'état obtenue à la période $i + 1$. Ces variables expriment facilement le coût d'une décision en fonction de l'état du système. Soit g_i le coût instantané et $g_{N+1}(x_{N+1})$ le coût du système dans son état final x_{N+1} . Ce coût étant supposé additif, il s'écrit sous la forme :

$$g_{N+1}(x_{N+1}) + \sum_{i=1}^N g_i(x_i, u_i, w_i) \quad (1.5)$$

En règle générale, une fonction coût différente est écrite pour la dernière période afin de mieux valoriser l'état en fin de période.

Le critère de minimisation des coûts s'écrit alors comme une minimisation en espérance :

$$\min_{u_i} \mathbb{E}_{w_i} \left\{ g_{N+1}(x_{N+1}) + \sum_{i=1}^N g_i(x_i, u_i, w_i) \right\} \quad (1.6)$$

Principe d'optimalité

La programmation dynamique repose sur le *principe d'optimalité de Bellman*, qui s'énonce comme suit :

Soit une suite de décisions $u = (u_i)_{i=1}^N$ associées chacune à l'état du système au pas i . Si u est une suite de décisions optimale, alors toute sous-suite $u^k = (u_i)_{i=k}^N$ est elle aussi optimale pour la période allant de k à N . Cette sous suite de décision débutant de l'état x_k (issu de la mise en oeuvre de la suite de décisions optimale sur la période allant de 1 à $k - 1$) et conduisant au même état final x_{N+1} .

Ce principe d'optimalité permet de résoudre le problème de manière rétrograde. La fonction coût (ou fonction de Bellman) peut alors s'écrire sous la forme suivante : on initialise la fonction à l'état final $V^{N+1}(x_{N+1})$, et pour tous les états allant de 1 à N , on calcule :

$$V^i(x_i) = \min_{u_i} \mathbb{E}_{w_i} (g_i(x_i, u_i, w_i) + V^{i+1}(x^{i+1})) \quad (1.7)$$

La programmation dynamique présente comme avantage la possibilité d'une résolution exacte du problème traité, sans exiger aucune hypothèse de convexité ou de différentiabilité des fonctions de Bellman et de la fonction coût. Cette méthode souffre néanmoins de *la limite des dimensions*, l'effort de calcul et la mémoire nécessaire pour le stockage croît en effet de manière exponentielle avec la dimension du vecteur d'états.

Application à un modèle multi-étapes

La programmation dynamique stochastique permet de résoudre un modèle avec recours multi-étapes dont l'espace de scénarios est modélisé sous forme d'arbre. La fonction de Bellman s'écrit alors sous la forme suivante :

$$V^i(n) = \mathbb{E}_{m|m \in \text{fils}(n)} \left(\min_{j \in \text{Tr}(i)} \{f(i, j, n, m) + V^j(m)\} \right) \quad (1.8)$$

où $V^i(n)$ désigne la valeur de Bellman associée à l'état i au noeud n , $\text{fils}(n)$ est l'ensemble des noeuds m fils de n , $\text{Tr}(i)$ désigne l'ensemble des états j joignables à partir de l'état i , et $f(i, j, n, m)$ le coût instantané lié à cet état.

On obtient alors une résolution par programmation dynamique sur un espace d'états représenté par l'ensemble des noeuds d'un arbre de scénarios modélisé en 1.3.

L'intérêt de la résolution par programmation dynamique réside dans sa capacité à résoudre des problèmes sans aucune contraintes sur la régularité des fonctions (convexité, différentiabilité, variables mixtes ou entières). Néanmoins, dans le cas où la fonction économique possède des propriétés intéressantes, d'autres algorithmes de résolution sont exploitables tels que la méthode "L-shaped".

1.2 Modèles stochastiques sous contraintes en probabilités

1.2.1 Théorie

Dans une approche avec recours, le décideur cherche les actions correctives réalisables après observation de l'aléa et minimise en moyenne le coût de ces corrections. L'approche sous contraintes en probabilités s'intéresse quant à elle à fiabiliser jusqu'à une certaine mesure le système de manière à ce que les décisions obtenues soient réalisables en environnement incertain, et ce avec une probabilité suffisamment forte.

La classe des modèles d'optimisation sous contraintes en probabilités a été introduite par Charnes, Cooper et Symonds [19] sous le nom de "*chance constrained programming*". elle constitue actuellement une branche importante du domaine de la programmation stochastique. Soit un problème linéaire sous la forme classique suivante :

$$\begin{aligned} & \min \quad c^T x \\ \text{s.c : } & Ax \geq b, \\ & x \geq 0 \end{aligned} \quad (1.9)$$

où c et $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ et A est une matrice $m \times n$. Dans un contexte incertain, les coefficients de la matrice A ainsi que le vecteur b peuvent être fonctions d'un paramètre aléatoire ξ scalaire ou vectoriel, avec un système exigeant de satisfaire ces contraintes avec une probabilité $p \in [0, 1]$. Pour tenir compte du caractère probabiliste de ce problème, la contrainte $Ax \geq b$ sera prise en compte de la façon suivante : une solution x sera déclarée acceptable si la probabilité que le système $Ax \geq b$ soit satisfait est suffisamment grande :

$$\mathbb{P}(A(\xi)x \geq b(\xi)) \geq p \tag{1.10}$$

où \mathbb{P} désigne une mesure de probabilité, et p une probabilité requise par le décideur et reflétant le niveau de fiabilité exigé de la solution. Ce type d'exigences se retrouve dans des domaines tels que la finance, les télécoms et l'industrie.

Il existe une multitude de modèles écrits à l'aide de contraintes en probabilité. Néanmoins, une majorité des problèmes étudiés présentent la matrice A et/ou le second membre b comme fonctions linéaires en ξ . Ainsi, deux classes principales de contraintes en probabilités se distinguent :

1. Contraintes de la forme : $\mathbb{P}(Ax - b\xi \geq c) \geq p$ avec A est une matrice à coefficients déterministes et le vecteur b linéaire en ξ : $b(\xi) = c + b\xi$
2. Contraintes de la forme : $\mathbb{P}(A(\xi)x \geq c) \geq p$ où $A(\xi)$ est une matrice à coefficients aléatoires

On considérera par la suite les problèmes écrits avec des contraintes du 1^{er} type, c'est à dire où la contrainte en probabilités admet le caractère linéaire en (x, ξ) . Une étude plus détaillée des autres formulations est proposée dans les articles [1] et [43].

Soit la contrainte en probabilité suivante :

$$\mathbb{P}(Ax - b\xi \geq c) \geq p \tag{1.11}$$

En toute généralité, la probabilité 1.11 lie plusieurs inégalités.

Exemple 1 :

On souhaite que la probabilité que la machine A produise plus de 100 produits a soit supérieure à 0.99 et que la probabilité que la machine B produise plus de 500 produits b soit supérieure à 0.99.

Exemple 2 :

Ou alors, on peut modéliser la volonté que la probabilité que la machine A produise plus de 100 produits a et que la machine B produise plus de 500 produits b soit supérieure à 0.99.

Ces deux exemples simples illustrent la différence conceptuelle et mathématique entre des contraintes en probabilités dites individuelles ou indépendantes (exemple 1), et des contraintes en probabilités jointes (exemple 2), qui sont souvent plus réalistes.

1.2.2 Contraintes probabilistes indépendantes

Le cas de contraintes en probabilités individuelles présente le cadre le plus favorable pour cette classe de problèmes. La contrainte 1.11 s'écrit pour chaque inégalité sous la forme suivante :

$$\mathbb{P}(A_i x \geq b_i(\xi_i)) \geq p_i \quad \forall i = 1 \dots n \quad (1.12)$$

où n représente le nombre d'inégalités, $b_i(\xi) = b_i \xi_i + c_i$ le second membre linéaire en ξ ($b(\xi) = c - b\xi$), et A_i désigne la i -ème ligne de la matrice A .

ξ_i étant uni-dimensionnel, la contrainte probabiliste 1.12 peut s'écrire sous la forme :

$$\mathbb{P}(A_i x \geq b_i(\xi_i)) \geq p_i \quad \Leftrightarrow \quad F_{\xi_i}(A_i x) \geq p_i \quad (1.13)$$

avec F_{ξ_i} la fonction de répartition de ξ_i .

En posant q_{p_i} le quantile d'ordre p_i de ξ_i , on obtient alors :

$$A_i x \geq q_{p_i} \quad \text{avec} \quad q_{p_i} := \inf \{t \in \mathbb{R} | F_{\xi_i}(t) \geq p_i\} \quad (1.14)$$

La résolution du problème devient plus simple, la connaissance de la fonction de répartition de ξ_i permet d'obtenir le quantile d'ordre p_i souhaité.

Ainsi, dans le cas de contraintes indépendantes, la résolution du problème est simplifiée. Ce contexte permet ainsi la résolution d'un grand nombre de problèmes appliqués. Néanmoins, ce sont généralement les contraintes en probabilités jointes qui sont les plus réalistes.

1.2.3 Contraintes probabilistes jointes

Dans le cadre de contraintes probabilistes jointes, toute solution admissible doit respecter l'ensemble de ces contraintes avec une probabilité supérieure ou égale à un niveau p donné :

$$\mathbb{P}(A_i x \geq b_i(\xi_i) \quad \forall (i = 1 \dots m)) \geq p \quad (1.15)$$

La formulation d'un problème à l'aide de contraintes probabilistes jointes est souvent plus appropriée dans un contexte industriel (par exemple un ensemble d'usines électriques dont les productions à chaque pas de temps dépendent des pas antérieurs et postérieurs). Par ailleurs, une solution admissible pour un problème avec contraintes jointes l'est aussi pour un problème avec contraintes indépendantes, la réciproque n'est pas nécessairement vérifiée :

$$\begin{aligned}
 (a) \quad \max \quad & c^T x \quad s.c : \quad \mathbb{P}(x_1 \geq b_1) \geq 0.5 \quad x_1 \geq 0, x_2 \geq 0 \\
 & \mathbb{P}(x_2 \geq b_2) \geq 0.5 \\
 (b) \quad \max \quad & c^T x \quad s.c : \quad \mathbb{P} \left(\begin{array}{l} x_1 \geq b_1 \\ x_2 \geq b_2 \end{array} \right) \geq 0.5 \quad x_1 \geq 0, x_2 \geq 0,
 \end{aligned}
 \tag{1.16}$$

Dans l'exemple (a), on garantit avec une probabilité de 0.5 que les variables x_1, x_2 ne dépassent pas une valeur critique. Or si on choisit comme scénarios possibles le couple $(b_1, -b_2), (-b_1, b_2)$, ces scénarios respectent les contraintes de (a), mais violent les contraintes de (b) puisque aucun de ces deux scénarios n'est admissible. Plus généralement l'inclusion suivante ([44]) est toujours vérifiée :

$$\begin{aligned}
 & M_p \subseteq \bar{M}_p \\
 \text{avec :} & \\
 & \bar{M}_p := \{x \in \mathbb{R}^n | \mathbb{P}(A_i x \geq b_i(\xi_i)) \geq p_i \quad \forall i = 1 \dots m\} \\
 & M_p := \{x \in \mathbb{R}^n | \mathbb{P}(A_i x \geq b_i(\xi_i) \quad \forall i = 1 \dots m) \geq p\}
 \end{aligned}
 \tag{1.17}$$

où \bar{M}_p et M_p désignent l'ensemble des solutions réalisables pour une écriture avec contraintes individuelles et une écriture avec contraintes jointes.

Parmi les premiers travaux sur les contraintes en probabilités jointes, on peut citer ceux de Prékopa en 1970 ([43]) (ces travaux ont établi la convexité de l'ensemble des solutions réalisables dans le cas de fonctions de répartition multi-variées log-concave).

D'autres cas standards avec des lois continues sont résumés dans [44].

Par ailleurs, Prékopa, Dentcheva et Ruszczyński [21] ont explicité des bornes inférieures et supérieures de la solution optimale dans le cas de fonctions de répartition discrètes.

Il est à noter que de manière générale, l'ensemble des solutions réalisables n'est pas convexe comme dans l'exemple suivant :

$$\begin{aligned}
 & \max \quad c^T x \\
 s.c : & \quad \mathbb{P} \left(\begin{array}{l} x_1 + x_2 \geq b_1 \\ x_1 + 3x_2 \geq b_2 \end{array} \right) \geq 0.5 \\
 & \quad x_1 \geq 0 \\
 & \quad x_2 \geq 0
 \end{aligned}
 \tag{1.18}$$

avec b_1 et b_2 des variables aléatoires dépendantes avec :

$$\mathbb{P}(b_1 = 2, b_2 = 4) = 0.5 \quad \mathbb{P}(b_1 = 3, b_2 = 0) = 0.5
 \tag{1.19}$$

L'ensemble (S) de solutions réalisables de 1.18 inclut :

- les solutions satisfaisant $x_1 + x_2 \geq 2$ et $x_1 + 3x_2 \geq 4$ représenté par (S_1)
- les solutions satisfaisant $x_1 + x_2 \geq 3$ et $x_1 + 3x_2 \geq 0$ représenté par (S_2)

Même si (S_1) et (S_2) sont convexes, leur union $S = S_1 \cup S_2$ ne l'est pas comme illustré dans la figure 1.3.

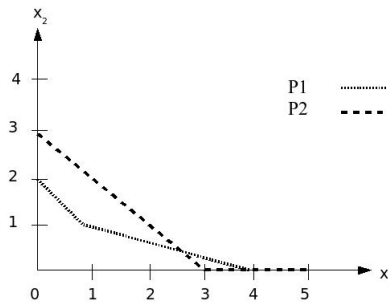


FIGURE 1.3 – Ensemble (S) des solutions réalisables

1.3 Conclusions & Applications

L'évolution des problèmes d'optimisation déterministes vers la prise en compte de facteurs aléatoires a donné lieu à une multitude de travaux.

Les modèles avec recours ont introduit la notion de mesures correctives que l'on peut apporter une fois l'aléa observé. Les modèles avec deux étapes s'étendent à des modèles multi-étapes mais les méthodes de résolution correspondantes sont plus complexes.

Un prototype de modèle avec recours résolu sur arbre de scénarios a été développé pour les besoins de la gestion hebdomadaire de la production électrique. L'étude a montré l'inexistence d'enjeu à utiliser cette modélisation en lieu et place d'une modélisation avec vision moyenne des aléas.

Parmi les modèles les plus souvent rencontrés, on retrouve les modèles avec contraintes probabilistes. La probabilité représentant un indicateur de la fiabilité du système rend ces modèles présents dans les problèmes de contrôle et de sûreté.

Une modélisation avec contraintes en probabilités jointes a permis de résoudre un problème de production d'une vallée hydraulique comprenant deux réservoirs d'eau connectés et alimentés par des apports indépendants pour chaque réservoir [1]. Ces apports suivaient des lois normales multi-variées et étaient corrélés dans le temps pour chaque réservoir. L'objectif était de maximiser le gain de la vallée hydraulique tout en respectant les bornes des volumes des réservoirs. Cette approche fut comparée à une modélisation avec contraintes en probabilités individuelles et une modélisation avec contraintes en espérance. Les comparaisons ont confirmé l'intérêt d'une telle approche.

Hormis les modèles avec recours et les modèles avec contraintes probabilistes, dont il a été question ci-dessus, d'autres voies sont explorées dans la littérature (comme la programmation floue introduite par Zadeh (1970)).

Dans la littérature, une autre grande famille de modèles, l'approche robuste, est souvent distinguée des problèmes stochastiques, même s'ils traitent de problèmes en contexte incertain. Le terme *robuste* synonyme dans la langue française de *Résistant*, *solide*, ou encore *inébranlable*.

L'optimisation robuste approche et traite les problèmes avec incertitude de manière différente de ceux cités jusque-là. Cette approche propose ainsi d'optimiser un problème vis à vis du pire cas possible dans un ensemble d'incertitude donné. On présentera dans la suite les principaux modèles robustes étudiés dans la littérature.

Chapitre 2

Optimisation robuste

En programmation stochastique, définir la réalisabilité d'une solution dans un contexte incertain nécessite une description probabiliste de l'incertitude. Or cette modélisation souffre de deux difficultés majeures. D'une part, il est difficile de connaître les réelles distributions des variables aléatoires dans le modèle. D'autre part, l'écriture d'un problème avec des contraintes en probabilités peut détruire ses propriétés de convexité et complexifier le nouveau problème.

L'optimisation robuste est née du besoin de proposer des solutions à des problèmes d'optimisation sous incertitude pour lesquels on ne dispose pas des lois de probabilités de leurs paramètres aléatoires. On cherche des solutions insensibles aux variations de certaines données du problème, en cherchant à se prémunir contre toutes ou partie des réalisations possibles. Le modèle d'incertitude n'est pas stochastique mais déterministe et ensembliste. Le principal défi de la programmation robuste reste la conservation de la solvabilité du problème initial. Ce défi dépend souvent de la famille de problèmes qu'on traite.

Dans ce chapitre, nous présentons les travaux qui ont marqué la littérature robuste, depuis le début des années 70, jusqu'à nos jours. Notre regard fait la distinction entre **les problèmes robustes linéaires** et **les problèmes robustes discrets**. En troisième partie, nous nous intéressons aux **modèles robustes dynamiques**, et plus précisément ceux avec recours.

2.1 Problèmes linéaires

L'approche robuste s'est intéressée au départ aux problèmes linéaires. Selon "*l'emplacement*" de l'incertitude dans le problème nominal, les approches proposées étaient différentes.

Nous avons scindé ces approches en deux parties : les premiers travaux s'intéressant à des problèmes linéaires affectés par **une incertitude par colonne** sur la matrice des contraintes. Le cas particulier d'**une incertitude affectant le second membre** est également abordé.

La deuxième famille d'approches s'intéresse au cas où l'incertitude affecte les lignes de la

matrice des contraintes.

2.1.1 Incertitude par colonnes

Principe

L'incertitude par colonnes fut introduite par Soyster [50] en 1973. Selon ce modèle, pour une ligne donnée i , l'incertitude apparaît sur certains coefficients désignés par les colonnes J de la matrice de contraintes \mathbf{A} (matrice $m \times n$). Chaque colonne $A_j, j \in J$ appartient à un ensemble d'incertitude $\mathcal{K}_j \subset \mathbb{R}^m$ supposé convexe. Soyster s'est intéressé au problème suivant :

$$\begin{aligned} & \max_x c^T x \\ \text{s.c. : } & \sum_{j=1}^n \mathbf{A}_j x_j \leq b \quad \forall \mathbf{A}_j \in \mathcal{K}_j, j = 1, \dots, n \\ & 0 \leq x \leq u \end{aligned} \quad (2.1)$$

Soyster a montré que ce problème était équivalent à :

$$\begin{aligned} & \max_x c^T x \\ \text{s.c. : } & \sum_j \bar{\mathbf{A}}_j x_j \leq b \quad j = 1, \dots, n \\ & 0 \leq x \leq u \end{aligned} \quad (2.2)$$

avec $\bar{a}_{ij} = \sup_{\mathbf{A}_j \in \mathcal{K}_j} (A_{ij})$. Prenons l'exemple suivant : Soit \tilde{a}_{ij} un coefficient incertain de la matrice A , avec $j \in J$. En posant $\tilde{a}_{ij} \in \mathcal{K}_j$, avec $\mathcal{K}_j = [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$, la formulation robuste selon Soyster est la suivante :

$$\begin{aligned} & \max_x c^T x \\ \text{s.c. : } & \sum_j a_{ij} x_j + \sum_{j \in J} \hat{a}_{ij} x_j \leq b_i \quad \forall i \\ & 0 \leq x \leq u \end{aligned} \quad (2.3)$$

La solution robuste optimale se ramène à résoudre le problème linéaire initial en posant pour les coefficients aléatoires de la matrice A les valeurs suivantes :

$$\begin{aligned} \tilde{a}_{ij} &= a_{ij} + \hat{a}_{ij} \text{ si la } i\text{-ème contrainte est de la forme } \leq \\ \tilde{a}_{ij} &= a_{ij} - \hat{a}_{ij} \text{ si la } i\text{-ème contrainte est de la forme } \geq \end{aligned}$$

Le modèle de Soyster est celui qui se prémunit le plus contre l'incertitude, mais cette solution conservatrice est beaucoup plus chère que le problème nominal. Le surcoût élevé des solutions robustes fournies par cette approche a orienté les travaux vers d'autres modèles.

Incertitude sur le second membre

L'incertitude sur le second membre est un cas particulier de l'incertitude par colonnes. Le vecteur \mathbf{b} peut être intégré à la matrice de contraintes \mathbf{A} comme étant sa $n + 1$ -ième colonne. On introduit également une $n + 1$ -ième variable de décision x_{n+1} imposée à avoir la valeur 1. Ainsi le problème avec incertitude sur le second membre s'écrit sous la forme :

$$\begin{aligned} \max_x \quad & c^T x \\ \text{s.c.} : \quad & \sum_{j=1}^{n+1} \mathbf{A}_j x_j \leq 0 \quad \mathbf{A}_{n+1} \in \mathcal{K}_{n+1}, \\ & 0 \leq x \leq u \\ & x_{n+1} = 1 \end{aligned} \tag{2.4}$$

Ce cas particulier a été étudié par Minoux [39] [40] et [41], un exemple d'application typique est celle du problème d'ordonnancement PERT robuste (avec durée des tâches incertaines). Ce problème se formule de la manière suivante :

$$\max_{\pi \in P} \max_{d \in D} \left\{ \sum_{i \in \pi} d_i \right\}$$

P étant l'ensemble de chemins possibles dans le graphe associé, d_i la durée de la tâche i appartenant au chemin π et D l'ensemble d'incertitude sur la durée des tâches.

Ce problème a été étudié avec deux modèles d'incertitude :

- D est considéré comme un boule relative à la norme L_∞ . D s'écrit donc pour ce problème comme un produit cartésien d'intervalles,
- D est considéré comme une boule de rayon maximum fixé relativement à la distance de Hamming. En associant à chaque tâche i une variable binaire u_i , l'ensemble d'incertitude s'écrit alors :

$$D = \left\{ \theta = (\theta)_{i=1, \dots, n} \mid \theta_i = d_i + \Delta_i u_i \quad (i = 1, \dots, n) \text{ tel que : } \sum_{i=1}^n u_i \leq \Gamma \right\}$$

La résolution étant obtenue par programmation dynamique dans le second cas. Ce problème d'optimisation robuste de l'ordonnancement PERT se résout efficacement en un temps polynomial. L'auteur de [39] explique dans son article l'efficacité de cette approche pour résoudre ce problème. Les différences avec l'approche proposée par Bertsimas & Sim sont également mises en évidence (voir [10] et [11]).

2.1.2 Incertitude par lignes

Les travaux de Ben-Tal et Nemirovski (1998-2000) [5],[6],[7] d'une part, et de Bertsimas et Sim (2003-2004) [10], [11] d'autre part ont contribué à l'essor récent de l'optimisation robuste linéaire. Leur intérêt s'est porté à des modèles d'incertitude affectant les lignes de la matrices des contraintes, une approche susceptible de fournir des solutions robustes moins conservatives que celle de Soyster. Pour la suite de cette section, on s'intéressera à des problèmes d'optimisation linéaires robustes du type :

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.c :} \quad & \mathbf{a}_i^T x \leq b_i \quad \forall \mathbf{a}_i \in \mathcal{U}_i, \quad i = 1, \dots, m \end{aligned} \quad (2.5)$$

où $x = (x_1, \dots, x_n)$ est le vecteur de variables de décision, \mathbf{a}_i la i -ème ligne de la matrice incertaine \mathbf{A} , pouvant prendre toutes les valeurs possibles dans l'ensemble d'incertitude $\mathcal{U}_i \subseteq \mathbb{R}^n$.

Sans perte de généralité, on suppose que l'incertitude n'affecte ni les coefficients c^T de la fonction objectif ni le second membre b .

En effet, supposons, dans le premier cas, que l'incertitude affecte le vecteur c^T de la fonction objectif, avec $c \in \mathcal{U}_c$.

En ajoutant une variable supplémentaire z , le problème 2.5 peut s'écrire :

$$\begin{aligned} \min_{x,z} \quad & z \\ \text{s.c :} \quad & \mathbf{a}_i^T x \leq b_i \quad \forall \mathbf{a}_i \in \mathcal{U}_i, \quad i = 1, \dots, m \\ & c^T x \leq z \quad \forall c \in \mathcal{U}_c \end{aligned} \quad (2.6)$$

Dans le second cas, où l'incertitude apparaît sur les coefficients de la matrice *ainsi que sur le second membre*, il suffit d'introduire une variable supplémentaire x_{n+1} **imposée** à -1 . On écrit alors le vecteur $b(\xi)$ comme la $(n+1)$ -ième colonne de la matrice (A) :

$$(A_{n+1} = (\mathbf{b}_1, \dots, \mathbf{b}_m))$$

Le problème écrit en 2.5 introduit un nouveau type de contraintes. En effet, pour chaque contrainte i , l'inégalité $\mathbf{a}_i^T x \leq b_i$ doit être respectée pour toutes les valeurs que peut prendre le vecteur a_i dans l'ensemble d'incertitude \mathcal{U}_i . La contrainte $\mathbf{a}_i^T x \leq b_i, \forall a_i \in \mathcal{U}_i$ est donc respectée si et seulement si :

$$\max_{\mathbf{a}_i \in \mathcal{U}_i} \mathbf{a}_i^T x \leq b_i \quad \forall i \quad (2.7)$$

Le problème 2.5 s'écrit alors :

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.c :} \quad & \max_{\mathbf{a}_i \in \mathcal{U}_i} \mathbf{a}_i^T x \leq b_i \quad \forall i = 1, \dots, m \end{aligned} \quad (2.8)$$

Formulation robuste de Ben-Tal & Nemirovski

Le choix de Ben-Tal et Nemirovski de considérer les ellipsoïdes comme ensembles d'incertitude est discuté dans [6]. Ce choix permet d'obtenir une contrepartie robuste traitable dans le cas des programmes linéaires.

En définissant l'ellipsoïde \mathcal{U}_i de la manière suivante :

$$\mathcal{U}_i = \{ \mathbf{a}_i = a_i^0 + \Delta_i * u_i, \|u_i\|_2 \leq \rho \}$$

avec $a_i^0 = (a_{i1}^0, \dots, a_{in}^0)$ la ligne i de la matrice du problème nominal, $u_i \in \mathbb{R}^n$ appartient à la sphère de rayon ρ .

$\Delta_i \in \mathbb{R}^n$ tel que $\Delta_i * u_i = (\Delta_{ij} u_{ij})_{j=0}^n \in \mathbb{R}^n$ appartient à l'ellipsoïde défini par le couple (Δ_i, ρ) .

La contrepartie robuste du problème 2.7 s'écrit d'après [6] :

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.c :} \quad & a_i^0 x \leq b_i - \rho \|\Delta_i x\|_2 \quad \forall i \in 1 \dots m \end{aligned} \quad (2.9)$$

La contrepartie robuste d'un problème linéaire est donc un problème conique quadratique, et se résout en temps polynomial par la méthode des points intérieurs. La conservation du caractère traitable du problème revient au choix de l'ellipsoïde comme ensemble d'incertitude.

Partant des mêmes ensembles d'incertitude, cette formulation est moins conservative que celle proposée par Soyster. En effet, soit le cas où les coefficients incertains \tilde{a}_{ij}

s'écrivent :

$$\tilde{a}_{ij} = (1 + \epsilon \xi_{ij}) a_{ij}$$

avec $\xi_i = (\xi_{ij})_{j \in J}$ des variables aléatoires indépendantes à valeurs dans $[-1, 1]$ de moyenne nulle.

Les coefficients \tilde{a}_{ij} sont donc compris dans l'intervalle $[a_{ij} - \epsilon a_{ij}, a_{ij} + \epsilon a_{ij}]$.

D'après [7], le vecteur de décisions x solution de :

$$\begin{aligned} & \min_{x,y,z} c^T x \\ \text{s.c. : } & \sum_j a_{ij} x_j + \sum_{j \in J_i} \hat{a}_{ij} x_j + \kappa_i \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 z_{ij}^2} \leq b_i \quad \forall i \\ & -y_{ij} \leq x_j - z_{ij} \leq y_{ij} \quad \forall i, j \in J_i \\ & 0 \leq x_j \leq u \\ & 0 \leq y_{ij} \end{aligned} \tag{2.10}$$

est solution réalisable du problème nominal et la probabilité de violation des contraintes incertaines est bornée par $\exp(-\kappa_i^2/2)$ (κ étant un paramètre ≥ 0). D'où un modèle moins conservatif que le modèle de Soyster pour une probabilité de violation de contraintes assez faible.

Le problème robuste 2.10 est néanmoins un problème non linéaire (ensemble d'incertitude d'ordre 2) avec un nombre de variables supérieur au modèle de Soyster.

L'étude d'un cas particulier *et* traitable de cette approche fut proposée par les mêmes auteurs, dans le cas où l'ensemble d'incertitude est un polyèdre. D'après [6], lorsque \mathcal{U}_i est un polyèdre, la contrepartie robuste du problème nominal est, comme nous allons le voir ci-dessous, un problème linéaire.

Soit le problème robuste où les contraintes $i = 1, \dots, m$ est affectée par des coefficients incertains :

$$\begin{aligned} & \min_x c^T x \\ \text{s.c. : } & \max_{\mathbf{A}_i} \mathbf{A}_i^T x \leq b_i \quad i = 1, \dots, m \end{aligned} \tag{2.11}$$

L'ensemble d'incertitude est décrit par un polyèdre, c'est à dire sous la forme :

$$Q_i \mathbf{A}_i \leq t_i \quad \forall i = 1, \dots, m$$

Intéressons nous alors à une contrainte i donnée du problème 2.11. La contrainte i peut être assimilée à un **sous-problème** de maximisation sous la forme suivante :

$$\begin{aligned} \max_{\mathbf{A}_i} \quad & \mathbf{A}_i^T x \\ \text{s.c :} \quad & Q_i \mathbf{A}_i \leq t_i \end{aligned} \tag{2.12}$$

Le paramètre x étant fixe, le **dual** de ce sous-problème s'écrit :

$$\begin{aligned} \min_{y_i} \quad & t_i^T y_i \\ \text{s.c :} \quad & y_i^T Q_i = x \\ & y_i \geq 0 \end{aligned} \tag{2.13}$$

y_i étant la variable duale associée la contrainte i .

En appliquant le dual à *toutes les lignes* de la matrice \mathbf{A} affectées par l'incertitude, le problème d'optimisation devient :

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.c :} \quad & y_i^T t_i \leq b_i \quad i = 1, \dots, m \\ & y_i^T Q_i = x \quad i = 1, \dots, m \\ & y_i \geq 0 \quad i = 1, \dots, m \end{aligned} \tag{2.14}$$

Ainsi, dans le cas particulier d'un polyèdre comme ensemble d'incertitude, la contrepartie robuste du problème linéaire reste un problème linéaire avec une augmentation polynomiale de sa taille. Malheureusement ceci est le seul cas où la contrepartie robuste conserve la structure du problème initial.

Plus généralement, la version robuste a une complexité supérieure à celle du problème nominal. Exemple : lorsque le problème nominal a une contrainte quadratique, alors le problème robuste équivalent est un problème semi-défini qui admet une solution exacte si l'ensemble d'incertitude est réduit à un unique et même ellipsoïde pour tous les coefficients incertains. Le problème devient NP-difficile dans le cas d'une intersection d'ellipsoïdes ou dans le cas d'un polyèdre.

Le caractère non-linéaire des contreparties robustes selon l'approche de Ben-Tal et al. rend inattractives pour les problèmes d'optimisation robuste discrets, voire inapplicable pour les problèmes de grande taille.

Formulation robuste de Bertsimas & Sim

Bertsimas & Sim [10] ont proposé une prise en compte de l'incertitude par ligne qui permet de réécrire le problème robuste comme un problème linéaire.

Un paramètre Γ_i , associé à chaque contrainte i indique le nombre maximum de coefficients pouvant varier autour de leur valeur moyenne sur la ligne i . En posant $\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. Alors, au plus Γ_i coefficients \tilde{a}_{ij} ont une valeur différente de a_{ij} , et ce pour la ligne i . Le problème à résoudre s'écrit :

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.c :} \quad & \\ & \sum_j a_{ij} x_j + \max_{\{S_i \mid |S_i| = \Gamma_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} x_j \right\} \leq b_i \quad \forall i \\ & 0 \leq x \leq u \end{aligned} \tag{2.15}$$

Le paramètre Γ_i permet, en variant dans l'intervalle $[0, J_i]$ (J_i étant le nombre de coefficients aléatoires sur la ligne i), de décider du degré de conservatisme que l'on souhaite donner à la solution recherchée.

La formulation linéaire équivalente à ce problème s'obtient par passage au dual du sous-problème de maximisation (le sous-problème correspond au terme de maximisation présent sur les contraintes du problème ci-dessus). Elle s'écrit sous la forme :

$$\begin{aligned} \min \quad & c^T x \\ \text{s.c :} \quad & \sum_j a_{ij} x_j + z_i \Gamma_i + \sum_{j \in J_i} p_{ij} \leq b_i \quad \forall i \\ & z_i + p_{ij} \geq \hat{a}_{ij} x_j \quad \forall i, j \in J_i \\ & 0 \leq x_j \leq u_j \quad \forall j \\ & p_{ij} \geq 0 \quad \forall i, j \in J_i \\ & z_i \geq 0 \quad \forall i \end{aligned} \tag{2.16}$$

Cette formulation linéaire augmente raisonnablement la taille du problème, et la conservation de la structure linéaire du problème nominal rend cette approche attractive même pour des problèmes de grande taille. D'un point de vue probabiliste, il a été démontré que la probabilité de violation des contraintes est de l'ordre de $\exp(-\frac{\Gamma_i^2}{2J_i})$. La qualité de cette borne est relativement comparable à la borne trouvée par Ben-Tal et Nemirovski pour leur modèle. Enfin, on notera le rôle du paramètre Γ_i , qui permet d'arbitrer entre robustesse de la solution (probabilité de satisfaction de la contrainte) et le surcoût induit (impact sur le coût de la décision optimale).

2.2 Problèmes discrets

La littérature s'est également penchée sur l'utilisation des approches robustes dans le cadre de problèmes discrets, c'est à dire dont toutes ou partie des variables de décisions appartiennent à des ensembles discrets.

En premier, nous citons ci-dessous une extension des travaux de Bertsimas et Sim [11]. Avec une démarche analogue au cas linéaire, ils proposent un modèle pour l'incertitude avec des résultats plus encourageants.

Ensuite, nous présentons succinctement les principales idées qu'ont présentées Kouvelis et Yu dans leur livre [33].

Dans cet ouvrage, ils ont parcouru l'ensemble des problèmes classiques de l'optimisation discrète, en proposant à chaque fois une contrepartie robuste. Malheureusement, la contrepartie robuste d'un grand nombre de problèmes discrets (même polynomiaux) devient NP-difficile.

2.2.1 Travaux de Bertsimas & Sim dans le cas discret

Dans [11], Bertsimas et Sim ont étendu avec succès leur approche au cas de problèmes discrets voire combinatoires.

Le modèle d'incertitude pour les coefficients de la matrice \mathbf{A} conserve les mêmes propriétés que dans le cas de problèmes linéaires. Une incertitude sur le vecteur coût de la fonction objectif est par ailleurs introduite. Le vecteur \mathbf{c}_j est donc une variable aléatoire variant dans l'intervalle $[c_j, c_j + d_j]$. Enfin, parmi les n variables de décision $x_j, j = 1 \dots n$, les k premières variables sont entières.

Le problème robuste équivalent proposé s'écrit alors sous la forme :

$$\begin{aligned} \min_x \quad & c^T x + \max_{\{S_0 | S_0 \subseteq J_0 \mid |S_0| \leq \Gamma_0\}} \left\{ \sum_{j \in S_0} d_j |x_j| \right\} \\ \text{s.c :} \quad & \sum_j a_{ij} x_j + \max_{\{S_i \mid |S_i| = \Gamma_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} x_j \right\} \leq b_i \quad \forall i \\ & 0 \leq x \leq u \end{aligned}$$

Avec $\Gamma_i, i = 0 \dots n$ supposés entiers (mais cette approche reste valable dans le cas où $\Gamma_i, i = 1 \dots n$ sont de simples réels positifs, avec $\sum S_i = \Gamma_i$).

Ce modèle robuste non linéaire peut s'écrire néanmoins sous forme d'un problème linéaire. Cette réécriture est obtenue par passage au dual du terme de maximisation.

Dans le cas où l'incertitude (telle que décrite ci-dessus) affecte uniquement le vecteur coût, avec des variables de décision *entièrement binaires*, la version robuste se résout avec au plus $n + 1$ résolutions du problème nominal. Ainsi, une résolution en temps polynomial est possible pour les problèmes combinatoires s'ils admettent une solution en temps polynomial dans un cadre déterministe.

Ce résultat ne s'étend pas aux problèmes mixtes.

Dans le cas où le problème nominal n'est pas traitable en un temps polynomial, mais qu'il admet une solution α -approchée en temps polynomial, alors sa version robuste admet également une solution α -approchée en temps polynomial.

Pour exemple, la version robuste du problème du sac à dos avec 200 variables binaires dégrade la solution optimale de près de 3% alors que près de 50% des coefficients sont sujets à de l'incertitude ([33]).

Ainsi, pour les problèmes combinatoires avec un temps de calcul raisonnable, cette approche pour résoudre la version robuste est à l'heure actuelle la plus intéressante proposée dans la littérature.

2.2.2 Incertitude sous forme de scénarios

Dans le cas de problèmes discrets, Kouvelis et Yu ont défini trois critères de robustesse, l'incertitude étant modélisée par un ensemble de scénarios indépendants.

Soient les notations suivantes :

| | |
|----------|--|
| S | l'ensemble des scénarios possibles |
| D | l'ensemble des données en entrée du modèle nominal |
| D^s | les données en entrée du modèle nominal spécifiques au scénario s |
| F_s | l'ensemble des décisions réalisables dans le cas où le scénario s est modélisé en entrée du modèle nominal |
| X | Ensemble des variables de décision du problème |
| X_s^* | solution optimale du problème nominal avec connaissance parfaite du futur (données D_s en entrée du modèle) |
| $f(.,.)$ | fonction coût du modèle |

Les trois critères de robustesse proposés pour minimiser la fonction coût $f()$ sont les suivants :

- Critère de **robustesse absolue** (ou critère du *Min-Max*) :

X_A est une solution robuste *absolue* optimale ssi elle minimise le **pire coût** de $f()$ selon les différents scénarios. X_A vérifie :

$$z_A = \max_{s \in S} f(X_A, D^s) = \min_{X \in \bigcap_{s \in S} F_s} \max_{s \in S} f(X, D^s)$$

- Critère de **Déviatiion robuste** (ou critère du *Min-Max Regret*) :

X_D est une solution avec *déviatiion robuste* optimale ssi les décisions X_D permettent de minimiser le **pire écart**, en terme de coût, entre le coûts de ces décisions et le coût des décisions parfaites (c'est à dire avec connaissance parfaite du futur) :

$$z_D = \max_{s \in \mathcal{S}} (f(X_D, D^s) - f(X_s^*, D^s)) = \min_{X \in \bigcap_{s \in \mathcal{S}} F_s} \max_{s \in \mathcal{S}} (f(X, D^s) - f(X_s^*, D^s))$$

- Critère de **robustesse relative** (ou critère du *Min-Max Relatif*) :

Dans le cas où $f(X_s^*, D^s)$ est non nulle, X_R est une solution robuste *relative* optimale ssi les décisions X_D permettent de minimiser le **pire écart**, en terme de coût, **relatif**, entre le coût de ces décisions et le coût des décisions parfaites (c'est à dire avec connaissance parfaite du futur) :

$$z_R = \max_{s \in \mathcal{S}} \frac{f(X_R, D^s) - f(X_s^*, D^s)}{f(X_s^*, D^s)} = \min_{X \in \bigcap_{s \in \mathcal{S}} F_s} \max_{s \in \mathcal{S}} \frac{f(X, D^s) - f(X_s^*, D^s)}{f(X_s^*, D^s)}$$

Le critère de robustesse absolu est un critère assez conservatif. Il présente néanmoins l'avantage d'une implémentation simple. Les deux critères restants sont plus proches des vrais besoins métier, ils sont néanmoins complexes a implémenter.

Kouvelis et Yu [33] ont étudié la version robuste des problèmes classiques d'optimisation combinatoire. On peut citer le problème du sac à dos, d'ordonnancement, d'allocation de ressource ou encore de gestion de production.

Les problèmes d'affectation et certains problèmes d'ordonnancement version robuste sont NP-difficiles. Est fortement NP-difficile la formulation robuste des problèmes du plus court chemin, de l'arbre couvrant, ou encore du sac à dos.

On observe ainsi la complexité due à l'aspect discret des variables combinée à une formulation robuste du problème nominal.

Enfin, certains problèmes (tels que le problème de *vendeur de journaux*) peuvent se résoudre en temps polynomial, et des résolutions approchées combinant méthodes de Branch & Bound et méthodes de relaxation sont proposées pour les problèmes NP-difficiles.

A travers ces deux premières sections, nous avons introduit, parfois de manière succincte, les grands travaux qui, selon notre connaissance, représentent le socle des avancées en optimisation robuste.

Néanmoins, ces approches se situent toutes dans un contexte décisionnel *statique*. On entend par là, que les décisions sont prises - et ce pour toutes les approches précédentes - avant réalisation de l'aléa. Cet environnement décisionnel ne traduit pas toutes les applications réelles, et plus particulièrement les applications qui ont la possibilité d'un certain

nombre de **décisions de recours**.

La programmation avec recours fut abondamment étudiée dans le cas de *la programmation stochastique*. Les travaux sur des approches **robustes dynamiques**, ou encore **avec recours** proposent actuellement des résultats intéressants. La section qui suit introduit ces modèles et les défis qu'ils relèvent.

2.3 Problèmes robustes dynamiques

Les modèles robustes étudiés jusque là sont résolus de manière statique, la décision robuste est obtenue en une fois, avant toute observation de l'aléa. Ce processus de décision peut être plus restrictif et contraignant que le contexte réel. En effet, on peut avoir à résoudre un problème où une partie de l'incertitude est dévoilée au fur et à mesure que l'on avance dans le temps. On s'intéressera donc dans ce paragraphe aux modèles robustes dans un contexte dynamique.

Les problèmes robustes avec décision séquentielle ont été étudiés dans le cadre de la programmation dynamique robuste (comme par Iyengar [31]), avec un aléa modélisé par un ensemble d'incertitude rectangulaire sur des problèmes de taille raisonnable afin d'éviter l'explosion de la dimension.

Le problème robuste linéaire à deux étapes sous sa forme la plus générale a été démontré comme NP-difficile par Ben-Tal et al. dans [8]. Les résultats récents de Minoux [40],[41] ont montré que la NP-difficulté est encore obtenue dans le cas très particulier des programmes linéaires correspondant à des problèmes de flots dans les réseaux et avec incertitude sur les seconds membres seulement. Une classe de problèmes d'optimisation dynamique robuste multi-étapes pouvant être résolue en temps pseudo-polynomial a été étudiée par Minoux [42].

En 2009, Zhang et Guan [53] ont appliqué l'approche robuste au problème de gestion de production électrique en contexte incertain. Les similitudes et les différences entre ces travaux et notre approche robuste seront discutées en détail dans le chapitre 4.

Afin d'aboutir à une version traitable du problème robuste avec recours, l'approche proposée [8] consiste à réduire l'espace fonctionnel des décisions de deuxième étape, en d'autres termes, les variables de seconde étape s'écrivant comme fonctions affines de l'incertitude.

Ce concept de règles de décision linéaires est couramment utilisé dans la programmation stochastique. Cette approximation dégrade certes la qualité de la solution mais rend traitable le problème à deux étapes.

Nous introduisons dans un premier temps cette approche qui fut appelée **AARC** pour *Affinely Adjustable Robust Counterpart*.

Nous présentons dans un second temps une extension aux problèmes mixtes étudiée par Bertsimas et Caramanis dans [12].

2.3.1 Règles de décision linéaires

Soit le problème robuste à deux étapes suivant :

$$\min_x \quad c^T x_1 + d^T x_2(\xi) \quad (2.17)$$

$$s.c : \quad (2.18)$$

$$\mathbf{A}_1^T(\xi)x_1 + \mathbf{A}_2^T(\xi)x_2(\xi) \leq b \quad \forall \xi \in \Xi \quad (2.19)$$

$$x_1 \in \chi_1, x_2 \in \chi_2 \quad (2.20)$$

x_1 sont les variables de première étape, $x_2(\xi)$ les variables de seconde étape, c et d étant leurs vecteurs coûts respectifs. \mathbf{A}_1 et \mathbf{A}_2 sont les matrices de contraintes associées, et Ξ est l'ensemble des réalisations possibles de la variable aléatoire ξ .

L'approche AARC consiste à écrire les variables de seconde étape comme **fonctions affines** de l'aléa ξ . Ainsi $x_2(\xi)$ s'écrit sous la forme :

$$x_2(\xi) = Q\xi + q$$

Toutefois, cette approximation 2.17 seule ne suffit pas pour obtenir un problème linéaire. Ceci étant dû à la dépendance en ξ de \mathbf{A}_2 et de x_2 .

En se plaçant dans le cas d'un recours dit *fixe* - qui fait l'hypothèse que la matrice \mathbf{A}_2 est constante, le problème de non-linéarité ne se pose plus. Dans le cas particulier d'un ensemble d'incertitude polyédral, la contrepartie robuste du problème robuste avec recours fixe conserve sa structure linéaire. Cette approximation est par ailleurs proposée dans l'article de Lemaréchal et al. [36] comme piste à étudier dans le cas d'un modèle robuste de réseau de télécommunications.

Dans le cas où la dépendance en ξ est non-linéaire, i.e. dans le cas où le recours est non fixe et où l'ensemble d'incertitude est un ellipsoïde, le problème AARC 2.17 s'écrit comme un problème semi-défini. On peut citer un certain nombre d'applications du modèle AARC tels que la gestion de portefeuille, la conception de circuits intégrés ou encore la gestion de trafic.

2.3.2 Règles de décision linéaires par morceaux

En considérant un problème robuste dont les variables de seconde étape sont entières, la démarche proposée par Ben-Tal et al n'est plus applicable. L'hypothèse de linéarité

$$x_2(\xi) = Q\xi + q$$

n'étant plus vérifiée.

L'approche proposée par Bertsimas et Caramanis dans [12] consiste à décrire les variables de deuxième étape comme fonctions linéaires par morceaux de l'aléa ξ . Cette approche est appelée **FA** pour *Finite Adaptability* en rapport avec le nombre *fini* de morceaux des fonctions utilisées.

Soit le problème robuste avec recours 2.17. Dans le cas où x_2 est un vecteur dont certaines variables sont entières, l'approche montre l'équivalence avec l'écriture suivante :

$$\min_{x=\bigcup_j \chi_j} \left[\begin{array}{l} \min_x \quad c^T x_1 + \max \{dx_2^1, \dots, dx_2^k\} \\ s.c : \quad \mathbf{A}_1^T(\xi)x_1 + \mathbf{A}_2^T(\xi)x_2^1 \leq b \quad \forall \xi^1 \in \chi \\ \quad \quad \quad \vdots \\ \quad \quad \quad \mathbf{A}_1^T(\xi)x_1 + \mathbf{A}_2^T(\xi)x_2^k \leq b \quad \forall \xi^k \in \chi \\ \quad \quad \quad x_1 \in \chi_1, x_2^j \in \chi_j \quad \quad \quad j = 1 \dots k \end{array} \right]$$

Il est important de noter que la partition de l'ensemble d'incertitude χ en k sous-ensembles (pas forcément disjoints) est à fixer par l'utilisateur. Cette formulation conserve la structure du problème robuste nominal, tout en augmentant d'un facteur k le nombre de variables de deuxième étape. Le problème 2.3.2 est traitable à partir du moment où le problème statique, qui correspond au cas particulier $k = 1$ est traitable.

Toutefois la réelle difficulté de cette approche consiste à trouver le partitionnement optimal de l'ensemble d'incertitude. Autrement dit, la partition qui permet d'obtenir la solution robuste optimale pour le problème nominal *et* 2.3.2 est une variable à déterminer.

Les auteurs de [12] ont montré que l'obtention de la partition optimale pour le simple cas $k = 2$ est un problème NP-difficile. Un résultat de complexité plus encourageant montre néanmoins que le problème pour $k = 2$ devient traitable si l'un des trois paramètres suivants est faible : la dimension de l'espace d'incertitude, le nombre de variables de décision ou encore le nombre de contraintes affectées par l'incertitude.

Conclusions

La principale application de l'optimisation robuste dans le cadre des problématiques de gestion de production d'EDF a été menée dans le cadre du travail de thèse de R. Aparigliato [2] pour la gestion des stocks hydrauliques.

Dans un contexte de gestion de production électrique optimisée, les réserves hydrauliques sont des moyens de production flexibles, sollicités généralement dans les périodes de pointe. Contrairement aux autres moyens de production classiques, la notion de stock existe et fait partie des variables d'optimisation du modèle. Les apports naturels en eau dans les réservoirs, considérés dans les modèles comme déterministe, sont sujets à de l'incertitude.

On peut citer deux contraintes majeures à ce problème : d'une part, le modèle de gestion

de production déterministe résout un problème de grande taille, et la prise en compte de l'incertitude ne doit pas augmenter de manière très forte la taille du problème. D'autre part, des contraintes dures sur le volume minimum et maximum des réservoirs sont à respecter et ce pour toute réalisation possible de l'aléa, ou du moins avec une probabilité de satisfaction élevée. Ces deux paramètres ont motivé le choix des techniques d'optimisation robuste pour résoudre le problème incertain, s'inspirant des travaux de Ben-Tal et Nemirovski sur les problèmes linéaires ainsi que les travaux sur les modèles robustes avec recours.

Cette approche a été testée sur une vallée comportant 3 réservoirs et 6 turbines sur une période de 7 jours (discretisée en 84 pas de temps). Les expériences numériques ont permis de comparer principalement une approche robuste à révisions périodiques (approche robuste statique relancée pour chaque jour optimisé après observation d'une partie de l'aléa), et une approche robuste à ajustement linéaires (utilisation d'un modèle robuste avec recours où les décisions de seconde étape suivent des règles de décision linéaires). Les résultats ont montré une bonne satisfaction des contraintes avec un surcoût estimé à 0.57% pour la première approche et 1.15% pour la deuxième comparée à une vision parfaite de l'aléa. La seconde approche a tout de même un handicap quant à la taille du problème équivalent qui est multiplié par un facteur oscillant entre 40 et 50 et ce pour une seule vallée. La première approche n'impose qu'une faible augmentation (facteur compris entre 1.2 et 3).

Les récentes avancées dans le domaine de l'optimisation robuste ont permis de donner un aperçu de la force qu'on peut tirer d'une approche robuste, en fonction de la modélisation de l'incertitude et de la structure du problème nominal. Certes, les modèles robustes augmentent la complexité numérique du problème, et certaines familles de problèmes voient leurs versions robustes généralement intraitables (problèmes combinatoires, problèmes non linéaires). Néanmoins, un choix pertinent du modèle d'incertitude combiné à un problème nominal traitable permet d'obtenir une version robuste traitable avec des résultats intéressants.

Chapitre 3

Introduction à la gestion de production électrique à *EDF*

3.1 EDF : 1^{er} producteur français

3.1.1 EDF en quelques chiffres

EDF est l'opérateur historique français. L'ouverture en 2005 du marché électrique français et européen à la concurrence a permis à EDF de s'affranchir du principe de spécialité et d'élargir son offre commerciale à l'international. Aujourd'hui, EDF est premier producteur nucléaire mondial, acteur principal des marchés français et britannique de l'électricité, bien implanté dans les grands pays d'Europe, ainsi qu'en Asie et aux États-Unis.

Fort de plus de 160 000 collaborateurs à travers le monde, le groupe satisfait les besoins énergétiques de plus de 35 millions de clients, soit une production mondiale totale d'environ 620 TWh. Cette activité dégage un chiffre d'affaires de plus de 66 Md€ (dont près de la moitié hors de France), offrant au groupe le statut d'un leader de l'énergie dans le monde.

3.1.2 Missions & Objectifs

La gestion optimisée de la production électrique a pour finalité de *maximiser le profit financier* du producteur par rapport à son portefeuille de clients.

RTE (Réseau de Transport d'Électricité) est le gestionnaire public du réseau de transport électrique français. Il est responsable de l'exploitation, la maintenance et le développement des réseaux, afin de maintenir au meilleur niveau la sûreté du système électrique. La séparation des activités entre les producteurs et le gestionnaire de réseau RTE permet aux producteurs de s'affranchir des problématiques liées au transport et acheminement de l'électricité, qui sont donc du ressort de RTE.

EDF, en tant que producteur, prend deux engagements vis-à-vis du gestionnaire de réseau : Il s'engage auprès de RTE à communiquer les informations nécessaires au bon

fonctionnement du système électrique, et surtout au maintien de son équilibre. En effet, L'équilibre du système électrique est d'autant plus important que, d'une part, l'électricité ne se stocke pas, et que d'autre part, tout déséquilibre physique peut avoir de graves répercussions économiques et sociétales. Cet engagement est d'autant plus important qu'EDF est le producteur historique en France.

Le second engagement est d'ordre financier : tout écart entre la production réalisée par EDF et la consommation réelle de son portefeuille de clients impliquera le règlement par le producteur d'une compensation financière.

Or l'équilibre entre offre et demande doit faire face à différents aléas qui peuvent affecter la demande en puissance (ex : fluctuation de la température, conjoncture économique), les coûts de production (ex : prix des combustibles), ou encore la capacité de production (indisponibilité imprévue d'unités de production).

Ces incertitudes introduisent des risques qu'il faut limiter, à long, moyen et court terme.

L'existence d'aléas oblige à ménager des marges de puissance supplémentaire par rapport au strict équilibrage de l'offre et de la demande prévisionnelle. Il existe trois types de réserves de puissance imposés par RTE :

- *la réserve primaire* est la régulation automatique de la puissance des tranches en fonction de la fréquence instantanée locale du réseau. La demande en réserve primaire pour la France est de l'ordre de 600 à 700 MW, et elle doit être mobilisable en moins de 30 secondes
- *la réserve secondaire* ou **télé réglage** peut être mobilisée en moins de quelques minutes par un signal national. Elle sert à reconstituer le volume de réserve primaire et compenser un nouvel écart éventuel du bilan offre-demande. Elle est de l'ordre de 600 à 800 MW.
- *la réserve tertiaire* correspond à une réserve manuelle, mise à disposition de RTE et dont le délai de mise en oeuvre n'excède pas 30 minutes. Elle sert d'une part à réaliser des ajustements d'écart entre production et demande et d'autre part à reconstituer le volume de réserve primaire et secondaire. 1000 MW sont actionnables en moins de 13 minutes, et 500 MW complémentaires en moins de 30 minutes.

3.1.3 Le parc électrique EDF

Le parc EDF en France se compose aujourd'hui de moyens de production assez diversifiés dans la mesure où il produit en proportions notables à partir de différents types d'énergie. Il est constitué de 58 tranches nucléaires, 47 centrales thermiques à flamme et près de 500 usines hydrauliques regroupées en 48 vallées.

En 2009, les puissances installées pour chaque type d'énergie et son utilisation pour la production électrique étaient les suivantes :

| Type d'énergie | Puissance disponible (en GW) | Participation à la production électrique (en %) |
|----------------|------------------------------|---|
| nucléaire | 63,1 | 87,1 |
| hydraulique | 20 | 9,3 |
| thermique | 13,9 | 3,6 |

TABLE 3.1 – Trois moyens de production majeurs constituant le parc EDF

Le nucléaire

Suite au choc pétrolier de 1973, un vaste programme nucléaire a permis de mettre en exploitation 58 tranches nucléaires, qui se répartissent en paliers standardisés (900MW, 1300MW et 1450 MW).

Le nucléaire occupe une place prépondérante dans le parc de production ; il correspond au mode de production le plus rentable (le coût de l'investissement ayant déjà été amorti). Nécessitant des arrêts périodiques (tous les ans environ) pour recharger le combustible, ces moyens de production font l'objet d'une optimisation spécifique de leurs plannings d'arrêts - généralement fréquents en été - pour une disponibilité maximale des centrales nucléaires en hiver.

Le thermique à flamme

Les moyens de production thermiques à flamme (ou encore appelés thermiques classiques) comprennent tous les types de centrales qui utilisent un combustible autre que nucléaire (ex : charbon, fioul, gaz naturel).

Plus cher et plus flexible que le nucléaire, ce type de production est utilisé essentiellement lors de pointes de consommation. Il a néanmoins un impact plus important sur l'environnement dû aux émissions d'oxyde de soufre, de carbone et d'azote. De plus, la nouvelle technologie utilisée dans les centrales CCG (Cycle Combiné Gaz) en construction va nécessiter une gestion des stocks de gaz couplée à la gestion de la production sur les différents horizons d'étude.

L'hydraulique

Le parc hydraulique est constitué d'usines situées en général dans les régions montagneuses et qui transforment l'énergie mécanique produite en énergie électrique, à partir des apports hydrauliques provenant de la pluie ou stockés suite à la fonte des neiges.

Une vallée hydraulique est constituée d'usines (plus de 500 usines au total) et de réservoirs connectés entre eux par des liens d'influencement. L'eau turbinée par une usine est prélevée de son réservoir *amont* et arrive dans son réservoir *aval* avec un certain délai de parcours. Certaines usines permettent d'effectuer du pompage d'eau du réservoir aval vers le réservoir amont, en consommant de l'électricité sur le réseau.

La figure 3.1 suivante illustre la vallée de Revin Saint-Nicolas, située dans le département des Ardennes. Elle est formée de deux usines de turbinage (REVI5H et BOCALH), une usine de pompage (REVI5P), du bassin supérieur des Marquisades (REVISR) et du bassin inférieur du Whitaker (REVIIR).

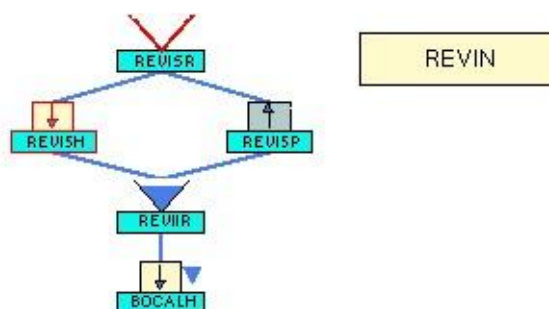


FIGURE 3.1 – Vallée du Revin

Une usine hydraulique est constituée d'un ensemble de groupes placés en parallèle et ayant des caractéristiques éventuellement différentes. Pour la gestion de production, ces groupes sont néanmoins agrégés et à une usine est associée un seul débit et une seule puissance. Les débits programmés sur une usine sont discrétisés en des points appelés *points de fonctionnement*. A chaque point de fonctionnement est associé 4 valeurs : le débit, la puissance générée, la réserve en primaire et la réserve secondaire.

La production hydraulique présente l'avantage d'utiliser un "combustible gratuit" qu'est l'eau mais disponible en quantité limitée. C'est un moyen de production flexible et réactif, mais en quantité limitée (selon les niveaux d'eau dans les réserves). Il est donc primordial de l'utiliser de façon optimale pour réajuster la production à la demande quand les coûts marginaux du système sont les plus élevés.

3.1.4 Management d'énergies

Le processus décisionnel pour une gestion optimisée de la production a été pensé comme une déclinaison de trois grandes périodes d'action. Des décisions prises à long-terme (entre 10 et 20 ans), des décisions à moyen-terme (entre 1 et 5 ans), et des décisions court-terme (entre quelques mois et quelques heures), comme illustré sur le schéma suivant 3.2 :

Les choix d'investissements en environnement concurrentiel et la valorisation de mix énergétiques sous des hypothèses fortes d'un point de vue réglementaires, économiques et environnementales font partie du processus de décisions à *long-terme*. Elles impactent la stratégie et les choix technologiques d'EDF dans les 10 à 20 prochaines années.

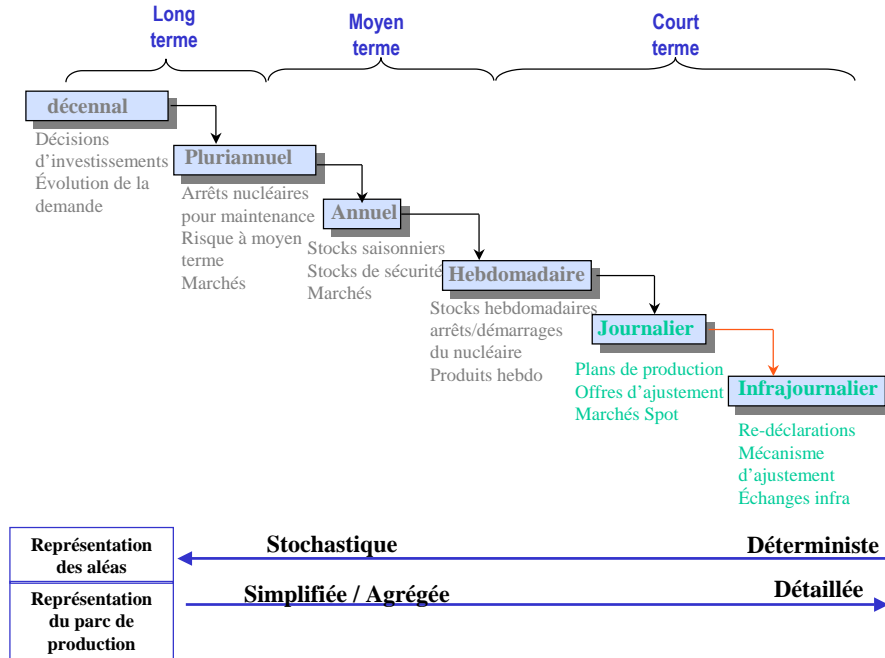


FIGURE 3.2 – Gestion de production & Processus décisionnel à EDF

A *moyen-terme*, les décisions portent sur un horizon allant de 1 à 5 années. Le placement des arrêts des centrales nucléaires, la logistique gazière et les placements sur les marchés font partie des décisions opérationnelles. Les indicateurs sur les coûts marginaux de la production électrique et la valorisation des stocks hydrauliques servent au processus décisionnel court-terme.

Le calcul des programmes de production journaliers, et la gestion fine des vallées hydrauliques et du parc thermique et nucléaire font partie des missions *court-terme*. Ce processus décisionnel est d'autant plus important qu'il représente le dernier levier qu'a le producteur pour maximiser son profit financier tout en respectant ses engagements vis à vis du gestionnaire de réseau.

On détaillera dans la suite la chaîne de décision court-terme qui se décompose en un processus journalier et un processus infra-journalier.

3.2 Le processus journalier

3.2.1 Contexte et réglementation

Le processus journalier consiste à assurer, la veille pour le lendemain, l'équilibre offre-demande prévisionnel d'EDF. Pour cela, un programme de production de tous les groupes nucléaires, thermiques à flamme et hydrauliques est établi avec respect des contraintes de fonctionnement du parc et des différents niveaux de réserves. Ce processus complexe peut être décomposé de manière simplifiée et très schématique en 3 phases : une phase d'estimation et d'échanges, une phase d'optimisation, et une phase de livraison

Estimation & échanges

La première partie du processus s'intéresse au calcul et collecte de données nécessaires à la phase d'optimisation du programme (deuxième phase du processus). La prévision de consommation des clients EDF, les échanges sur les marchés de gros (échanges de gré à gré ou sur les places de marché), les niveaux de réserve demandés par RTE, et le niveau d'effacement (possibilité de demander aux gros clients de baisser leur consommation) sont les informations obtenues en fin de première phase. On détaille de manière succincte les données collectées :

- i. Prévision de consommation** Historiquement, la prévision de consommation de la veille pour le lendemain se basait sur les niveaux de transit aux noeuds du réseau, qui étaient agrégés et traités pour obtenir une *Estimation Temps Réel* (ETR) de la consommation globale France. Aujourd'hui, l'ouverture du système électrique français à de nouveaux producteurs a entraîné un changement dans le calcul de cette prévision. Deux techniques sont aujourd'hui employées : La première consiste à défalquer de la prévision ETR la consommation *des clients perdus* par EDF. La deuxième alternative consiste à prévoir directement la consommation des clients EDF en se basant sur l'historique de production du parc EDF. Même si la première alternative alliait simplicité de calcul et bonne estimation, on assiste actuellement à un passage en douceur vers la prévision directe (démarche *bottom-up*), plus adaptée à un contexte dérégulé.
- ii. Echanges** L'arbitrage avec le marché de gros relève de la mission d' EDFTrading. En fonction des prix de marchés, et d'une première estimation des coûts de production, les traders effectuent leur placement sur le marché (principalement sur EPEXSPOT), profitant des écarts de prix entre le parc EDF et les marchés comme dans l'exemple simplifié de la figure qui suit.
- iii. Effacements** Le niveau d'effacement se base également sur une estimation des coûts de production/ prix de marchés. Valorisés par les équipes d'optimisation, ces contrats sont activés si les prix de marché sont plus élevés que leur valeur d'usage (prix de

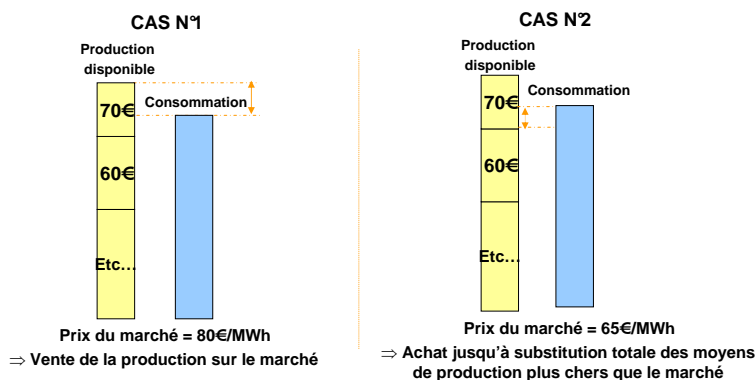


FIGURE 3.3 – Exemple d'arbitrage par EDF-T

l'électricité à partir duquel il est intéressant de les déclencher). Si un contrat d'effacement est activé, cela correspond à la baisse de consommation du client industriel ayant signé ce contrat, et ce sur une plage variant entre un et plusieurs jours.

iv. Réserves Enfin, pour faire face aux différents aléas pouvant affecter la sûreté du système, des marges de sécurité sont imposées par RTE et doivent être prises en compte dans le processus de planification du programme de production journalier.

Après obtention des données sur l'état physique du parc, la prévision de consommation, les arbitrages sur le marché et l'activation des contrats d'effacement, on effectue une optimisation du placement de la production sur le parc EDF. On revient de manière détaillée sur ce modèle dans la section suivante.

Optimisation

Le modèle d'optimisation ne prenant pas en compte toutes les contraintes techniques du parc de production, les opérateurs rectifient le programme optimal afin d'atteindre la réalisabilité. La version définitive est envoyée avant l'heure limite d'accès au réseau (fixée aujourd'hui à 16h) à RTE, et à toutes les unités de production et ce, de manière informatisée.

On prépare ensuite le passage au processus infra-journalier dont l'une des missions consiste à assurer le raccordement entre le programme d'appel décidé la veille qui se termine à 23h30 et le nouveau programme d'appel qui vient d'être envoyé à RTE et qui commence à minuit.

3.2.2 Outils d'optimisation

Le modèle Apogée

Aujourd'hui, le modèle Apogée - optimiseur développé au sein de la R&D d'EDF - détermine les programmes journaliers de production de l'ensemble des usines du parc. Ce modèle reçoit en entrée l'état de fonctionnement des unités de production, une estimation de la consommation électrique du lendemain, et le niveau de prescription pour les réserves. Il fournit en sortie des programmes journaliers pour chaque unité de production sur 48H avec un pas d'une demi-heure.

Plus concrètement, Apogée résout un problème d'optimisation dont la fonction objectif simplifiée s'écrit :

$$\sum_{i \in I} c_i(P_i) + c(dfP) + c(dfI) + c(dfB)$$

avec :

- I ensemble des moyens de production,
- c_i le coût du planning de l'unité de production i ¹,
- $c(dfP)$ le coût de la défaillance en puissance,
- $c(dfI)$ le coût de la défaillance en réserve primaire,
- $c(dfB)$ le coût de la défaillance en télé-régulation,

Les coûts à minimiser dans la fonction objectif sont les coûts d'exploitation des unités de production, le coût des pénalités lorsque la demande en puissance ou le niveau de réserve n'est pas satisfait. Le coût de l'écart en puissance et le coût de la défaillance en réserves sont fixés de manière à obliger le modèle à respecter les niveaux demandés.

Apogée permet de calculer cette fonction objectif tout en respectant les contraintes techniques citées ci-dessous :

Contraintes d'équilibre :

- équilibre offre-demande en puissance,
- équilibre offre-demande en réserve primaire,
- équilibre offre-demande en télé-régulation,

Contraintes physiques du parc thermique :

- durée minimale de marche ,
- durée minimale d'arrêt,
- limitations de production en puissance et en réserves par centrale,

1. le coût d'un planning inclut tous les coûts liés à l'exploitation des unités de production

- les contraintes de gradient ²,
- nombre maximum de démarrages par jour,
- la durée minimale entre deux modulations ³,
- la durée minimale entre deux passages en/hors réglage ⁴ ;

Contraintes physiques du parc hydraulique :

- les contraintes de flots - de volume sur les réservoirs,
- les contraintes de gradient (respect des valeurs autorisées pour les écarts de puissance entre deux pas de temps),
- interdiction de changement de sens pour une usine ;

En raison de la taille du problème (environ 1 million de variables et 2 millions de contraintes en formulation en nombres entiers MIP), et de sa structure (problème non convexe, variables discrètes), une résolution exacte en nombre entiers pourrait être envisagée. Néanmoins, cette possibilité est écartée à ce stade pour des contraintes de temps de calcul. Le choix de la résolution s'est donc porté sur une minimisation des coûts de production par relaxation lagrangienne en deux phases (voir [35], [25]).

En première phase, on dualise les contraintes de la demande, les multiplicateurs λ de Lagrange pénalisent ainsi la non satisfaction de l'équilibre offre-demande, avec une simplification de certaines contraintes du parc. Le problème dual à résoudre s'écrit :

$$\max_{\lambda} \min_P L(P, \lambda)$$

$$P_i \in \mathcal{X}_i, i \in I$$

avec $L(P, \lambda) = \sum_{i \in I} c_i(P_i) + \sum_{t \in T} \lambda_t (d^t - \sum_{i \in I} P_i^t)$

La fonction duale $\min_P L(P, \lambda)$ est concave non différentiable. Sa maximisation est résolue par la méthode des faisceaux [35]. Cette méthode nécessite de calculer la valeur de $\min_P L(P, \lambda)$ à chaque itération. Avec le passage au dual et grâce à la séparabilité par unité de production du problème, on peut effectuer ce calcul par décomposition.

La technique de décomposition utilisée peut être schématisée sous la forme suivante :

Le coordinateur envoie un signal de prix λ aux unités de production. A chaque unité correspond un sous-problème i qui calcule la production P_i . Cette commande est le résultat

2. les hausses et baisses de puissance doivent rester dans un intervalle prédéfini de valeurs
 3. c'est la durée minimale pour qu'une centrale change à nouveau de palier, généralement de 2h30
 4. c'est la durée minimale pour qu'une centrale change sa production de réserve, généralement d'1h30

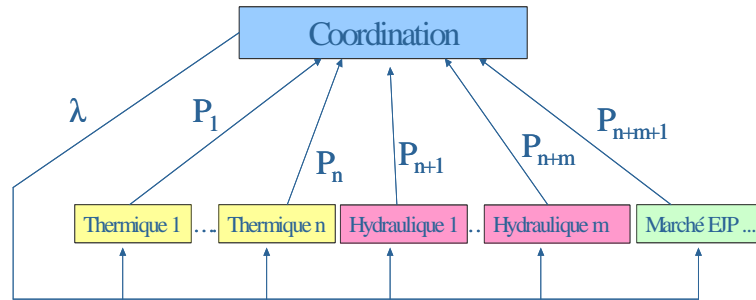


FIGURE 3.4 – Décomposition-Coordination du problème d’optimisation dans Apogée

d’un arbitrage entre la rémunération λ que propose le coordinateur et le coût de production c_i , dans le respect des contraintes physiques de cette unité. Le sous-problème s’écrit :

$$\min c_i(P_i) - \lambda P_i$$

Le coordinateur récupère les différents niveaux de production. Il ajuste à l’itération suivante la valeur des λ en fonction de l’écart positif ou négatif entre la demande à satisfaire et la somme des niveaux de production.

A un groupe thermique correspond un sous-problème, et à une vallée hydraulique un sous-problème. Les marchés sont également modélisés comme unité de production et sont assimilés à un sous-problème à optimiser. Les caractéristiques différentes de ces moyens de production implique une résolution différente. Le sous-problème thermique est résolu par programmation dynamique, et le sous-problème hydraulique se résout par programmation linéaire.

A la fin de cette phase, on obtient un optimum dual qui ne correspond pas à un optimum primal réalisable (saut de dualité pour non différentiabilité du problème dual). La satisfaction de la demande n’est donc toujours pas atteinte, et certaines contraintes physiques n’ont pas été prises en compte. Néanmoins le lagrangien simple a permis *schématiquement* de s’approcher de la région où se trouve l’optimum global, et d’obtenir des indicateurs λ qui approchent les coûts marginaux du système.

La seconde phase d'optimisation sert à obtenir des programmes réalisables, en pénalisant plus fortement l'écart entre l'offre et la demande, et en intégrant plus de contraintes techniques. La résolution du problème se fait avec la méthode du Lagrangien Augmenté [16], en initialisant les λ et les programmes P_i avec les solutions de la première phase. Le coût de la solution de la première phase est un minorant du coût de la solution finale. Le problème de la seconde phase s'écrit sous la forme suivante :

$$\max_{\mu} \min_{P, \tilde{P}} \begin{aligned} & \sum_{i \in I} c_i(P_i) + \sum_{i \in I} \mu(P_i - \tilde{P}_i) + \sum_{i \in I} \frac{c}{2} \|P_i - \tilde{P}_i\|^2 \\ & P_i \in \mathcal{X}_i, \\ & \sum_{i \in I} \tilde{P}_i = d, \end{aligned} \quad i \in I \quad (3.1)$$

Dans cette phase, on dédouble les variables de décision P en \tilde{P} . les variables de décision P doivent respecter les contraintes techniques du parc, les variables \tilde{P} satisfaire la demande prévue. La contrainte d'égalité $P_i^t = \tilde{P}_i^t$ est quant à elle dualisée avec la méthode du Lagrangien augmenté.

Pour minimiser les sous-problèmes en P et \tilde{P} de manière séparable, on applique le Principe du Problème Auxiliaire [45], qui consiste à linéariser le terme quadratique dans 3.1 autour de la solution de l'itération précédente, ce qui permet de traiter les sous-problèmes de manière analogue à ceux de la première phase.

A la fin de la seconde phase d'optimisation, on obtient un programme de production réalisable approchant la solution exacte du problème initial décrit en 3.2.2.

Évolutions futures

Concernant les outils d'optimisation à horizon journalier, on compte deux évolutions futures majeures en cours de développement et de validation.

La première évolution consiste à affiner la modélisation des groupes hydrauliques dans Apogée. Comme évoqué dans [24], la modélisation en continu de l'hydraulique dans Apogée ne permet pas la prise en compte de nombreuses contraintes discrètes caractérisant le parc hydraulique.

Face à la sous-optimalité de ce procédé, une version d'Apogée en nombres entiers est à l'étude (APOGENE), i.e. une version ayant une modélisation discrète des sous-problèmes hydrauliques.

Une évolution envisagée à horizon long (voire très long) terme est le passage d'une résolution duale du problème d'optimisation à une résolution frontale grâce à une combinaison de puissants solveurs commerciaux et une parallélisation massive du code de calcul.

Cette évolution est actuellement au stade embryonnaire. Néanmoins, la modélisation de l'hydraulique en nombre entiers et l'étude de l'opportunité d'une résolution frontale des sous-problèmes thermiques en PLNE sont actuellement à l'étude.

3.3 Le processus infra-journalier

Le processus infra-journalier a pour mission d'assurer le suivi du programme de production de la journée en cours, et d'effectuer toutes les actions nécessaires pour assurer l'équilibre offre-demande durant cette journée. Les tâches en infra-journalier consistent principalement à :

- suivre l'évolution des prévisions de consommation
- suivre l'état de fonctionnement du parc de production
- suivre les actions menées par EDF Trading

Et enfin à réajuster le programme d'appel en fonction de ces différentes observations.

3.3.1 Le Mécanisme d'Ajustement

Afin d'assurer la sûreté du système électrique, RTE a mis en place le Mécanisme d'Ajustement (voir à ce propos [37]). Ce dispositif mutualise l'ensemble des puissances disponibles pour un maintien constant de l'équilibre offre-demande.

Le Mécanisme fonctionne comme un appel d'offres continu ouvert aux producteurs raccordés au réseau électrique français, aux gros consommateurs industriels pouvant réguler (à la hausse comme à la baisse) leur consommation, et enfin aux acteurs étrangers interconnectés au réseau.

Une offre d'ajustement correspond généralement au niveau de régulation de la production/consommation que peut atteindre l'acteur sans mettre en danger son propre équilibre. A chaque offre correspond un prix (en MW/h), et ce pour tout l'horizon infra-journalier (24 heures au total découpées en 6 tranches).

Les offres d'ajustement sont mises à la disposition de RTE qui les rémunère si elles sont activées. Ainsi, lorsque le gestionnaire du réseau cherche à se procurer de l'énergie pour atteindre son équilibre, il va activer l'offre d'ajustement à la hausse dont le prix est le plus faible. Et lorsque le système connaît un surplus de production, le gestionnaire active une offre d'ajustement à la baisse, qui correspond au paiement par les acteurs de l'énergie excédentaire soutirée au prix proposé pour leur offre. Ainsi, le gestionnaire de réseau vend de l'énergie et choisit l'offre dont le prix de vente est le plus élevé possible.

Ce système de rémunération s'effectue a posteriori, sous l'appellation **Règlement des Ecartés**. Ce dispositif est totalement transparent pour les différents acteurs, et il est financièrement neutre pour RTE. Il permet ainsi de renvoyer aux acteurs du mécanisme la responsabilité financière des déséquilibres constatés a posteriori entre leur injection et leur soutirage.

3.3.2 Les re-déclarations

Pour palier aux différents aléas pouvant arriver au cours de la journée, il est prévu de pouvoir redéclarer au cours de la journée tout ou partie des programmes du producteur. Cette action permet par la même occasion de modifier les offres d'ajustement proposées.

Les modifications apportées au programme de marche à un guichet donné sont appelées "re-déclarations". Elles peuvent porter sur les niveaux de puissance et/ou de réserves, et ce pour toutes les unités de production. Un délai de neutralisation Δ est nécessaire entre le calcul des re-déclarations et leur mise en œuvre au guichet H_g . Actuellement, le délai de neutralisation est d'1 heure.

Enfin, à chaque guichet, un nombre maximum de re-déclarations est imposé par le gestionnaire du réseau. Ce seuil s'établit aujourd'hui à 30 re-déclarations. La notion de re-déclaration dépend de l'unité de production dont le planning est à modifier. Si la modification du programme d'une unité thermique compte pour une seule re-déclaration, on compte pour une re-déclaration la modification du programme d'un ensemble de groupes hydrauliques interconnectés par un ou plusieurs réservoirs.

3.3.3 Vers une gestion optimisée des re-déclarations

En pratique, *EDF* modifie une dizaine de fois ses programmes dans la journée, et les changements interviennent essentiellement avant les pics de consommation du matin et du soir et lors du "raccordement" à 23h (contrainte du passage d'un processus pour le jour J à celui pour le jour $J + 1$). Ces modifications se font actuellement "à la main". Une étude d'enjeux a montré que le gain d'un outil d'optimisation intégré au processus infra-journalier peut s'élever à *plusieurs millions d'euros par an*. La formulation à l'étude pour un modèle infra-journalier est la formulation du problème résolu par Apogée auquel on ajoute *la contrainte dualisée* des 30 re-déclarations à respecter. Le problème s'écrit :

$$\max_{\mu, \lambda} \min_{P_i} \sum_{i \in I} c_i(P_i) + \sum_{i \in I} \lambda(P_i - D) + \sum_{i \in I} \mu \left(30 - \sum_{i \in I} \left(1 - \delta \left(P_i - P_i^{ref} \right) \right) \right)$$

$$P_i \in \mathcal{X}_i, \quad i \in I$$

avec δ la fonction de Dirac :

$$\delta : \quad \mathbb{R}^n \quad \longrightarrow \{0, 1\} \tag{3.2}$$

$$(x_i)_{1 \leq i \leq n} \longmapsto \begin{cases} 1 & \text{si } x_i = 0 \quad \forall i = 1 \dots n \\ 0 & \text{sinon} \end{cases} \tag{3.3}$$

$$\tag{3.4}$$

La mise en exploitation de l'outil optimisant le placement de la production à horizon infra-journalier sera mis en exploitation dans un horizon proche.

3.4 Aléas physiques et financiers : Limites d'une vision déterministe

Dans les paragraphes précédents, nous avons introduit le processus journalier et infra-journalier de gestion de la production. Avec en entrée une prévision **moyenne** de demande en puissance, Apogée fournit des programmes modifiables le jour même par les exploitants selon la variation de l'écart entre la consommation réelle et les prévisions établies la veille.

Ces variations sont dues à l'observation de différents phénomènes imprévus. De plus ces variations possibles ne sont prises en compte dans les modèles d'optimisation du court-terme. En effet, contrairement au moyen-terme et long-terme, la gestion court-terme a une vision très fine du parc de production, mais s'affranchit d'une modélisation des aléas éventuels.

A horizon court terme, les programmes de production établis sont modifiés pour faire face à des aléas de différentes nature.

L'aléa sur la demande correspond à l'erreur observée entre la demande prévisionnelle et la demande réalisée. La demande prévisionnelle se base sur des paramètres météorologiques tels que la température ou la nébulosité. Par conséquent, une erreur de prévision météorologique impliquera une erreur, parfois très élevée, sur la prévision de demande. Pour avoir un ordre de grandeur de cet impact, il suffit de réaliser que *la baisse de la température de 1°C en hiver correspond à environ 1500MW de consommation en plus vers 19h, soit l'équivalent d'une centrale nucléaire à plein régime.*

L'aléa sur la production correspond à l'incertitude concernant la disponibilité réelle des unités de production, principalement thermiques et nucléaires. La disponibilité d'une unité peut être remise en cause pour trois raisons : une panne fortuite, une limitation de puissance (pour des raisons techniques ou humaines), ou une prolongation de son arrêt pour contrôles techniques. Comme cela a été développé dans la note technique ([54]), il a été développé un générateur de scénarios de disponibilité, qui prend en compte un historique des durées et taux des pannes, durées de retours de panne, et ce pour chaque type d'unité. Ce générateur fournit en sortie l'état de marche de la centrale sur tout l'horizon d'étude.

Enfin, les prévisions des prix d'achat/vente sur le marché de l'électricité (EPEXSPOT [56]) servent aux opérationnels dans leurs choix de placements sur le marché journalier et infra-journalier. Ces prévisions de prix sont évidemment sujettes à de nombreux aléas, non pris en compte à ce stade dans la gestion court-terme de la production.

Conclusions

Dans un marché électrique concurrentiel, EDF adapte ses outils de gestion de production pour permettre une gestion optimale de son portefeuille à long-terme (choix d'investissements, valorisation du mix énergétique), moyen-terme (placement sur les marchés, placement des arrêts des centrales, valorisation de l'hydraulique) et court-terme.

Les horizons journalier et infra-journalier sont les derniers leviers pour une gestion optimisée de la production.

Chaque jour, EDF constitue ainsi un programme de production pour ses unités. Ce programme est réalisé avec des modèles qui se basent sur des techniques d'optimisation classiques (relaxation lagrangienne, programmation linéaire, programmation dynamique ...).

Néanmoins, plus l'horizon d'optimisation s'approche du temps réel, et plus les décisions prises aux instants précédents deviennent structurantes voire limitantes en terme d'actions, ces décisions étant prises sans tenir compte du caractère aléatoire de certaines entrées du modèle.

Pour se prémunir face à ces aléas, des techniques d'optimisation robustes tenant compte du contexte incertain ont été étudiés afin de proposer un modèle de placement optimisé de la production tenant compte des incertitudes sur la demande en puissance. L'introduction de ce modèle et l'interprétation des résultats obtenus feront l'objet des chapitres suivants.

Chapitre 4

Formulation robuste du problème de gestion de production électrique à court terme

4.1 Introduction

Le problème de *placement optimal de la production électrique*, connu sous le nom de "*Unit-Commitment Problem*", a fait l'objet de plusieurs travaux portant sur des modèles d'optimisation déterministes et/ou stochastiques.

Dans le cadre déterministe, des travaux ont étudié de manière spécifique le problème de placement journalier de production à EDF. Ces travaux ont étudié la résolution de problèmes d'optimisation complexes de très grandes taille. On cite notamment les travaux de Renaud [46], et Lemaréchal [35], [25]. Ces travaux ont permis par ailleurs la mise en exploitation de l'outil Apogée introduit au chapitre 3.

Dans un cadre à forte incertitude, des travaux tels que [51], [18], [22] proposent des approches minimisant en espérance le coût de gestion, à travers des méthodes de décomposition de problèmes en deux-étapes et en multi-étapes.

A ce stade, le placement journalier de la production à EDF est étudié dans un cadre déterministe. Nos travaux consistent quant à eux à développer *un modèle d'optimisation qui prenne en compte l'incertitude sur la demande*.

Cette formulation compte parmi les premiers travaux qui répondent au problème de gestion de production sous incertitude en utilisant les techniques de l'optimisation robuste. En effet, Une approche appelée "APOGEE stochastique" consistant à résoudre le problème de gestion de production électrique avec un arbre de scénarios de demande à l'horizon hebdomadaire (horizon de 12 jours à optimiser) a été étudiée sans succès au vu des enjeux

estimés.

D'autres modèles avec incertitude sont utilisés pour répondre à la même problématique mais à horizon moyen terme (entre 1 et 3 ans). Ces modèles sont résolus pour la plupart par programmation dynamique stochastique, ou par programmation sur arbre de scénarios. L'aléa étant modélisé soit par des arbres de scénarios soit par des lois probabilistes.

Nous présenterons dans un premier temps les notations utilisées pour formuler le problème, ainsi que la formulation proposée. Nous étudierons ce modèle en petite dimension, avant de proposer un algorithme de résolution approprié applicable en dimension réelle. Nous concluons par une étude comparative entre notre approche et celle proposée par Zhang et Guan [53].

4.2 Modèle robuste proposé

Le modèle de gestion de production électrique APOGEE optimise le placement de la production journalière avec une prévision moyenne de la demande en puissance. Le coût du programme de production est optimisé en tenant compte des contraintes physiques et économiques du problème.

Néanmoins, le coût total de gestion correspond au coût du programme de production auquel s'ajoute le coût/gain de toutes les décisions prises en cours de journée pour répondre au mieux aux évolutions de la demande en puissance.

D'un point de vue métier, un programme de production dit "robuste" serait un programme de production dont les caractéristiques principales sont une stabilité du coût de gestion associé à une stabilité de la production finale décidée en opérationnel.

Plus explicitement, on cherche un modèle qui optimise le coût total de gestion, c'est à dire le coût du placement de la production ainsi que le coût lié à l'ajustement en temps de réel de ce programme de production. La robustesse du modèle se traduit par l'optimisation des coûts de gestion dans le cas du pire aléa pouvant affecter la demande en puissance.

Dans la suite, nous présenterons les notations nécessaires à la compréhension du modèle robuste proposé.

4.2.1 Notations

Posons les variables et données suivantes :

| | |
|--|---|
| T | : Nombre de pas de temps sur l'horizon du modèle d'optimisation |
| U | : Nombre d'unités de production du parc |
| \mathcal{D} | : Polyèdre d'incertitude sur la demande en puissance |
| \mathcal{X} | : Ensemble des programmes réalisables |
| C^u | : Coût de fonctionnement de l'unité de production u |
| P_t^u | : Production au pas de temps t de l'unité de production u . |
| $P_t = (P_t^u)_{u=1\dots U}$ | : Programme total de production au pas t |
| $P^u = (P_t^u)_{t=1\dots T}$ | : Programme de l'unité u sur tout l'horizon d'optimisation |
| $P = (P_t^u)_{u=1\dots U, t=1\dots T}$ | : Programme de production |
| d_t | : Demande en puissance au pas de temps t |
| $d = (d_1, d_2, \dots, d_T)$ | : Scénario journalier de demande en puissance |

4.2.2 Fonction objectif

On cherche à définir un modèle qui détermine un programme de production optimal robuste face à la pire erreur sur la prévision de consommation. En appliquant comme critère d'optimisation celui du pire cas (ou critère du *Min-Max*), le problème à résoudre s'écrit :

$$\Theta : \min_{P \in \mathcal{X}} \{ \Gamma(P) + \Phi(P) \} \quad (4.1)$$

$\Gamma(P)$ correspond au coût du plan de production P , somme des plans de production P^u des unités u :

$$\Gamma(P) = \sum_{u=1}^U C^u (P^u)$$

La fonction complémentaire $\Phi(P)$ représente le pire surcoût dû à l'aléa, pour un programme P donné :

$$\Phi(P) = \max_{d \in \mathcal{D}} \left\{ \sum_{t=1}^T \phi_t (d_t, P_t) \right\}$$

$\Phi(P)$ est une fonction dont la valeur correspond au **coût d'ajustement pire cas** sur l'ensemble des scénarios de demandes possibles.

4.3 Etude de la fonction de coût d'ajustement $\Phi(P)$

4.3.1 Modélisation d'un contexte industriel

La fonction d'ajustement $\phi_t(d_t, P_t)$ peut être utilisée dans deux contextes applicatifs différents :

- **Première vision** : On pénalise toute défaillance en production par une pénalité π_A qui correspondrait au prix d'achat de la puissance manquante sur le marché de l'énergie. Et lorsqu'on est en surproduction, on vend la puissance en surplus à un prix de vente π_V . Cette vision est celle des agents de trading d'EDF. Ils achètent et revendent de l'électricité sur le marché de l'électricité, avec comme objectif d'effectuer les bons arbitrages entre acheter de l'électricité sur les marchés ou produire sur le parc EDF.
- **Seconde vision** : On pénalise aussi bien la sous-production que la surproduction d'électricité. Une pénalité $\alpha \geq 0$ est payée si le système sous-produit et une pénalité $\beta \geq 0$ si il surproduit. La vision Pénalités force le modèle dans la mesure du possible à fournir des programmes qui s'écartent peu de la demande moyenne.

Les deux visions ont un intérêt industriel tant pour limiter les pénalités d'écart à l'équilibre ou pour prendre des positions optimales sur le marché de l'électricité.

Posons :

| | | |
|-----------------------------|---|--|
| π_t^A | : | Prix d'achat de la puissance manquante ξ_t^- |
| $\pi^A = (\pi_t^A)_{t=1}^T$ | : | Scénario de prix d'achat d'électricité |
| π_t^V | : | Prix de vente du surplus en production ξ_t^+ |
| $\pi^V = (\pi_t^V)_{t=1}^T$ | : | Scénario de prix de vente d'électricité |
| α_t | : | Pénalité pour production inférieure à la demande au pas de temps t |
| α | : | Vecteur de pénalité pour défaillance |
| β_t | : | Pénalité pour production supérieure à la demande au pas de temps t |
| β | : | Vecteur de pénalité pour surproduction |

Selon la vision qui nous intéresse, la fonction d'ajustement s'écrira :

- **Vision 1 ou contexte Marchés** : Prix d'achat-vente sur le marché Spot ou marché d'ajustement :

$$\phi_t(d_t, P_t) = \pi_t^A \max\{0, d_t - P_t\} - \pi_t^V \max\{0, P_t - d_t\}$$

Dans ce contexte, nous cherchons à valoriser au mieux le plan de production par rapport à des prix de marché. Les décideurs sont les traders.

- **Vision 2 ou contexte Pénalités** : Pénalités d'écart à l'équilibre offre-demande :

$$\phi_t(d_t, P_t) = \alpha_t \max\{0, d_t - P_t\} + \beta_t \max\{0, P_t - d_t\}$$

Dans ce contexte, nous cherchons à réduire les coûts de réajustement des programmes. Les décideurs sont les exploitants infra-journalier.

4.3.2 Formulation mathématique

Posons d'abord les notations suivantes :

- ξ_t^- : Différence $d_t - P_t$ si positive
- ξ_t^+ : Différence $P_t - d_t$ si positive

La fonction d'ajustement définie ci-dessus s'écrit alors, par exemple, pour la **Vision Marchés** :

$$\phi_t(d_t, P_t) = \pi_t^A \xi_t^- - \pi_t^V \xi_t^+$$

L'utilisation des variables ξ_t^- et ξ_t^+ impliquant une contrainte de complémentarité ($\xi_t^- \cdot \xi_t^+ = 0$).

Dans la formulation robuste, on cherche à calculer, pour un programme de production P donné, la demande en puissance impliquant le pire coût d'ajustement. Sachant que la demande appartient à un ensemble d'incertitude polyédral, l'objectif revient à maximiser la fonction d'ajustement (qui est linéaire par morceaux) sur l'ensemble auquel appartient la demande d . La pire demande est ainsi un point extrême du polyèdre d'incertitude sur la demande.

Selon la **vision Marchés 1**, $\Phi(P)$ s'écrit ainsi comme solution d'un programme linéaire en nombre entiers sous la forme suivante :

$$\begin{aligned} \Phi(P) &= \max \left\{ \sum_{t=1}^T (\pi_t^A \xi_t^- - \pi_t^V \xi_t^+) \right\} \\ \text{s.c : } & d_t - \xi_t^- + \xi_t^+ = P_t \\ & 0 \leq \xi_t^- \leq v_t \cdot M^- \\ & 0 \leq \xi_t^+ \leq (1 - v_t) \cdot M^+ \\ & v_t \in \{0, 1\}, d \in \mathcal{D} \end{aligned}$$

- M^- : Valeur maximale de la défaillance de production ξ^-
- M^+ : Valeur maximale du surplus de production ξ^+

avec :

$$v_t = \begin{cases} 1 & \text{si } d_t \geq P_t \\ 0 & \text{sinon} \end{cases}$$

Autrement dit, v_t est une variable binaire valant 1 si le système est défaillant, donc $\xi^- \geq 0$ et $\xi^+ = 0$. Cette variable vaut 0 sinon.

La contrainte de complémentarité liant les variables ξ_t^- et ξ_t^+ nécessite l'utilisation d'une

variable binaire pour obtenir une écriture linéaire de l'équation $\xi_t^- \cdot \xi_t^+ = 0$, ainsi que deux contraintes du type "Big M" pour majorer leurs valeurs.

L'utilisation de contraintes du type "Big M" est nécessaire pour modéliser fidèlement la notion de pire coût d'ajustement. Néanmoins, les bornes M^- et M^+ ont des valeurs égales au sup des deux fonctions $\max\{0, d_t - P_t\}$ et $\max\{0, P_t - d_t\}$.

De manière analogue, et selon **la vision 2**, $\Phi(P)$ s'écrit :

$$\begin{aligned} \Phi^2(P) &= \max \left\{ \sum_{t=1}^T (\alpha_t \xi_t^- + \beta_t \xi_t^+) \right\} \\ \text{s.c : } & d_t - \xi_t^- + \xi_t^+ = P_t \\ & 0 \leq \xi_t^- \leq v_t^- \cdot M^- \\ & 0 \leq \xi_t^+ \leq (1 - v_t^-) \cdot M^+ \\ & v_t^- \in \{0, 1\}, d \in \mathcal{D} \end{aligned}$$

4.3.3 Convexité de $\Phi(P)$

La fonction $\Phi(P)$ pourrait ressembler à premier abord à une fonction de perturbation d'un programme linéaire. Or, la contrainte de complémentarité entre les variables ξ^+ et ξ^- n'est pas convexe ($\xi^- \times \xi^+ = 0$). $\Phi(P)$ n'est donc pas une fonction de perturbation d'un programme convexe.

La fonction de pénalisation peut néanmoins s'interpréter comme le sup d'une collection finie de fonctions convexes (il suffit de considérer les points extrêmes du polyèdre \mathcal{D}). Dans le cas de la **Vision Pénalités** (le résultat reste valable pour la **Vision Marchés**), on a :

$$\Phi(P) = \sup_{d \in \overline{\mathcal{D}}} \sum_{t=1}^T \phi_t^d(P)$$

avec $\overline{\mathcal{D}}$ l'ensemble fini de points extrêmes du polyèdre \mathcal{D} . Et :

$$\phi_t^d(P) = \begin{cases} \alpha \times (d_t - P_t) & \text{pour } d_t \geq P_t \\ \beta \times (P_t - d_t) & \text{pour } d_t \leq P_t \end{cases}$$

Les pénalités α et β dans la Vision Pénalités étant positifs, ainsi que les prix d'achat/vente π_A et π_V dans la Vision Marchés, la fonction $\phi_t^d(P)$ est alors convexe. $\Phi(P)$ étant le sup des fonctions $\sum_{t=1}^T \phi_t^d(P)$, pour tout $d \in \overline{\mathcal{D}}$, est donc convexe en P .

4.4 Modèle robuste : Exemples en faible dimension

4.4.1 Exemples en une dimension

Vision Marchés

Dans le cas unidimensionnel, la fonction économique $\Phi(P)$ estimant le pire coût de réajustement dans une vision Marchés s'écrit sous la forme :

$$\Phi(P) = \max_{\mathcal{D}} \{ \pi_A \cdot \xi^- - \pi_V \cdot \xi^+ \}$$

avec \mathcal{D} un intervalle, π_A et π_V des scalaires. En posant le coût de production γ comme linéaire en P , la fonction économique générale s'écrit :

$$\min_{P \in \mathcal{X}} \left\{ \gamma \cdot P + \max_{d, \xi^-, \xi^+} \{ \pi_A \cdot \xi^- - \pi_V \cdot \xi^+ \} \right\}$$

où $P \in \mathcal{X}$ correspond en dimension 1 à $0 \leq P \leq P_{max}$, le terme de maximisation s'écrit explicitement sous la forme suivante :

$$\begin{aligned} \Phi(P) &= \max_{d, \xi^-, \xi^+} \{ (\pi_A \xi^- - \pi_V \xi^+) \} \\ s.c : & d - \xi^- + \xi^+ = P \\ & 0 \leq \xi^- \leq v \cdot M^- \\ & 0 \leq \xi^+ \leq (1 - v) \cdot M^+ \\ & v \in \{0, 1\} \\ & d_{min} \leq d \leq d_{max} \end{aligned}$$

Posons :

$$\pi_V \ll \gamma \ll \pi_A \implies \pi_V = 1, \quad \gamma = 10, \quad \pi_A = 100$$

$$d \in [d_{min}, d_{max}] \implies d \in [1000, 2000]$$

Soit une production P donnée. Le problème admet une solution optimale simple à trouver.

En effet, la valeur maximale de $\Phi(P)$ pour une production P donnée est :

$$\Phi(P) = 100 \cdot (d_{max} - P)$$

Ainsi, pour P fixé, la valeur maximale de $\Phi(P)$ est atteinte en $d = d_{max}$, et ce $\forall P \in [P_{min}, P_{max}]$.

Pour se prémunir contre ce pire coût, et minimiser la valeur de la fonction décrite en 4.4.1, il faut minimiser le terme :

$$10 \cdot P + 100 \cdot (d_{max} - P)$$

ce qui revient à la solution optimale $P^* = P_{max}$. Le graphe ci-dessous illustre ces différentes fonctions :

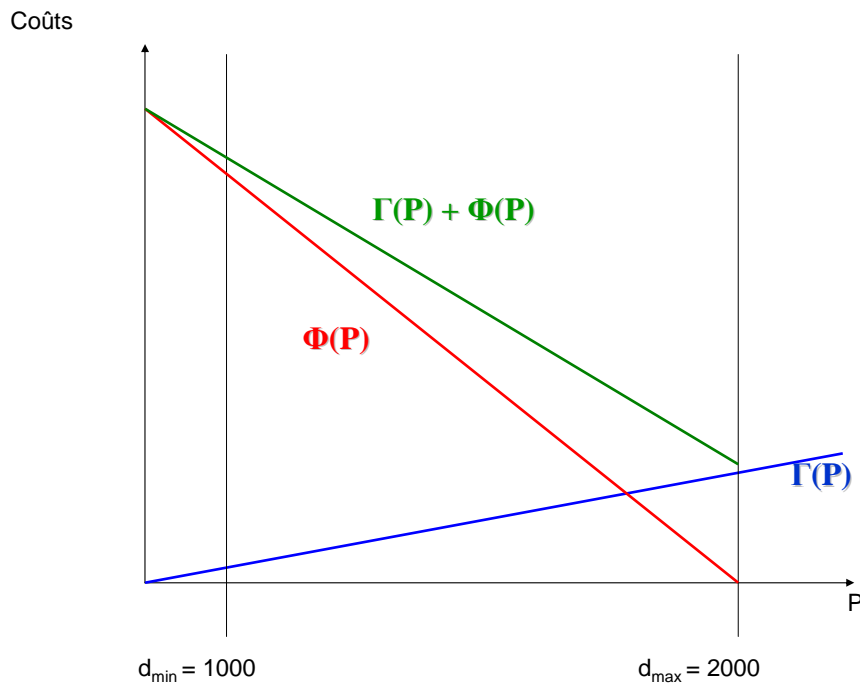


FIGURE 4.1 – Vision Marchés : Fonctions coûts

En faisant l'hypothèse que $P_{max} = d_{max}$, on lit sur ce graphique que la solution optimale du modèle est une production à son maximum. Le modèle n'incite à prendre aucun risque de défaillance, le surplus de production étant profitable à travers sa vente au prix π_V sur les marchés. Or les prix d'achat et de vente de l'électricité ne sont pas, en réalité, des données

exogènes. En effet, si on se place dans une logique de marché, on cherche à estimer ces π_A et π_V pour savoir quels sont les moments propices pour acheter ou vendre.

La vision Marchés nécessiterait donc la modélisation des prix de d'achat et de vente comme variables du modèle, et non comme seuls paramètres d'entrée, afin de mieux modéliser la problématique métier actuelle.

Vision Pénalités : choix des paramètres

De même que pour la vision Marchés, nous cherchons à calculer la solution optimale que proposerait le modèle robuste en dimension 1. Rappelons la fonction de pire réajustement :

$$\Phi(P) = \max_{\mathcal{D}} \{ \alpha \cdot \xi^- + \beta \cdot \xi^+ \}$$

avec \mathcal{D} un intervalle, α et β des scalaires. En posant le coût de production γ comme linéaire en P , la fonction économique générale s'écrit :

$$\min_{P \in \mathcal{X}} \left\{ \gamma \cdot P + \max_{d, \xi^-, \xi^+} \{ \alpha \cdot \xi^- + \beta \cdot \xi^+ \} \right\}$$

où $P \in \mathcal{X}$ correspond en dimension 1 à $0 \leq P \leq P_{max}$, le terme de maximisation s'écrit explicitement sous la forme suivante :

$$\begin{aligned} \Phi(P) &= \max_{d, \xi^-, \xi^+} \{ (\alpha \xi^- + \beta \xi^+) \} \\ s.c : & d - \xi^- + \xi^+ = P \\ & 0 \leq \xi^- \leq v \cdot M^- \\ & 0 \leq \xi^+ \leq (1 - v) \cdot M^+ \\ & v \in \{0, 1\} \\ & d_{min} \leq d \leq d_{max} \end{aligned}$$

Cette modélisation décrit le besoin de la part de l'utilisateur d'obtenir un programme de production satisfaisant au mieux l'équilibre *offre-demande*. tout écart à cet équilibre serait pénalisé. L'implication sur le choix des pénalités est donc :

- Une pénalité α de la puissance défaillante élevée comparée au coût de production de la même quantité de puissance : $\alpha \geq \gamma$
- Une pénalité β de la puissance en surplus qui exprimerait une pénalisation de la surproduction : $\beta \geq \gamma$

Le paramétrage de ces pénalités dépend ainsi du contexte opérationnel dans lequel il s'intégrera. Pour exemple, on peut décider que le coût de 1 MW défaillant est le même que celui de 1 MW mégawatt en surproduction, donc $\alpha = \beta$. Mais cette stratégie ne tient pas compte du coût de production du mégawatt en trop. Ou alors décider qu'un programme avec défaillance de 1 MW et un programme avec une surproduction de 1 MW doivent impliquer une même pénalisation. Ce qui implique : $\alpha = \beta + 2\gamma$.

Preuve :

Soit par exemple une demande $d = 2$ MW. On souhaite pénaliser de manière égale une production défaillante de 1MW ou surproduisant 1MW.

Dans le premier cas de figure, le coût total de la gestion est : $\gamma + \alpha$.

Dans le second cas, le coût total vaut : $3\gamma + \beta$.

Pénaliser de manière égale implique un coût total de gestion identique pour les deux cas de figure, d'où :

$\gamma + \alpha = 3\gamma + \beta$. D'où le résultat : $\alpha = \beta + 2\gamma$.

Vision Pénalités : Etude de la solution

Soit la fonction économique générale, vérifiant $\alpha = \beta + 2\gamma$:

$$\min_{P \in \mathcal{X}} \left\{ \gamma \cdot P + \max_{d, \xi^-, \xi^+} \{ \alpha \cdot \xi^- + \beta \cdot \xi^+ \} \right\}$$

Pour une production P donnée, on cherche à connaître l'expression de la fonction $\Phi(P)$ en étudiant les différents intervalles de valeurs dans lesquels se situe la solution optimale d^* . Ainsi :

- Si $d^* \geq P$, alors $\xi^- (= \max \{0, d^* - P\}) \geq 0$ et $\xi^+ = 0 \implies \Phi_1(P) = \alpha \times (d_{max} - P)$
- Si $d^* \leq P$, alors $\xi^- = 0$ et $\xi^+ \geq 0 \implies \Phi_2(P) = \beta \times (P - d_{min})$

La fonction de pénalisation Φ est convexe linéaire par morceaux, $\Gamma(P)$ est convexe croissante. Le minimum de $\Gamma(P) + \Phi(P)$ est atteint à l'intersection de Φ_1 et Φ_2 .

Ce minimum est atteint avec un programme de production optimisant le placement de la demande :

$$\frac{\alpha \cdot d_{max} + \beta \times d_{min}}{\alpha + \beta}$$

La pire pénalité qu'aurait à payer l'utilisateur est atteinte dans les cas où la demande réalisée vaut d_{max} ou d_{min} .

On a illustré ces résultats à l'aide des graphes ci-dessous, qui représentent la fonction coût de production Γ et celle de pénalisation Φ dans deux cas :

1. Avec $\gamma = 10$, $\beta = 10$, et $\alpha = 40$ pour le graphique 4.2
2. Avec $\alpha = \beta = 20$ et $\gamma = 10$ pour le graphique 4.3

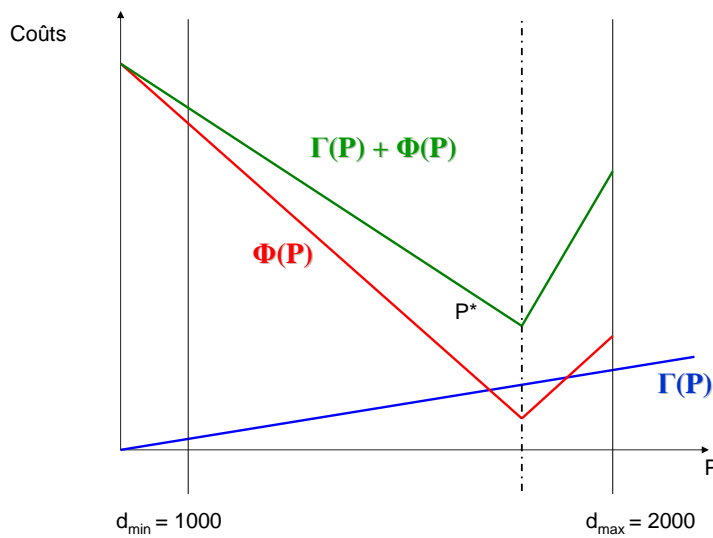


FIGURE 4.2 – Vision Pénalités : Fonctions coûts

On remarque sur ces graphiques que la solution optimale en sortie du modèle dépend fortement de la valeur donnée aux paramètres de pénalisation α et β . La production robuste optimale est géométriquement située au niveau du barycentre de l'intervalle $[d_{min}, d_{max}]$, ce barycentre étant complètement lié au *poids* attribué à chaque point extrême de cet ensemble d'incertitude.

La vision Pénalités semble au premier abord, répondre de manière intéressante au souhait de proposer une solution qui minimise le coût des pénalités. Graphiquement, ces pénalités sont, à un facteur près, égales à la distance entre la solution optimale et les extrémités de l'intervalle $[d_{min}, d_{max}]$.

Nous allons dans la suite explorer la vision Pénalités en dimension 2 puis proposer une résolution du problème général dans le cas de dimensions plus élevées.

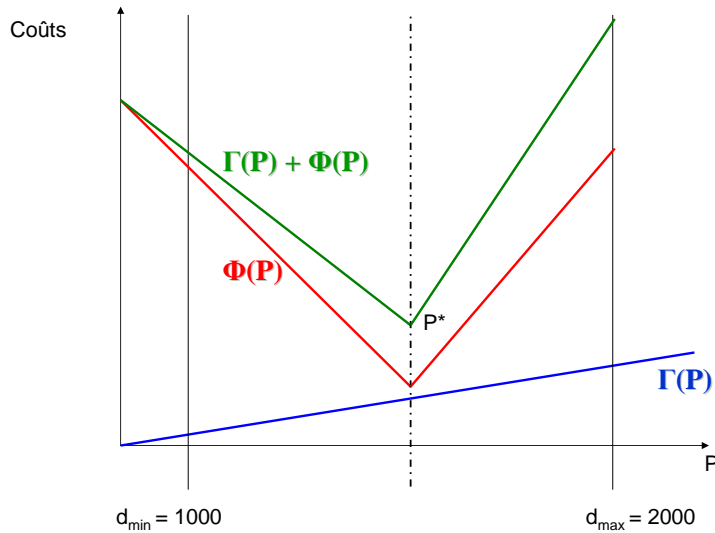


FIGURE 4.3 – Fonctions coûts de production et pénalisation

4.4.2 Exemple en dimension 2 de la vision Pénalités

Ecriture de $\Phi(P)$

La fonction pénalité s'écrit :

$$\begin{aligned} \Phi(P) &= \max_{\mathcal{D}} \sum_{t=1}^2 (\alpha \cdot \xi_t^- + \beta \cdot \xi_t^+) \\ \text{s.c :} & \\ \forall t = 1, 2 & \\ & d_t - \xi_t^- + \xi_t^+ = P_t \\ & \xi_t^- \leq v_t^- \times M_t^- \\ & \xi_t^+ \leq (1 - v_t^-) \times M_t^+ \\ & v_t^- \in \{0, 1\} \end{aligned}$$

Pour définir les grandes constantes M_t^- et M_t^+ , on leur attribue les valeurs suivantes :
 $M_t^- = M_t^+ = d_{max}^t - d_{min}^t$

Remarque :

Même si la résolution du problème est plus longue si le nombre de pas de temps s'accroît (problème en nombres entiers, contraintes du type "Big M"), nous verrons dans le chapitre

suivant que la résolution de ce problème reste raisonnablement efficace jusqu'à 48 pas de temps.

Etude de la solution

A partir d'un P_{init} donné, on cherche à calculer une demande d^* la plus éloignée possible de P_{init} . d^* appartient à un polyèdre défini par :

- quatre contraintes de bornes sur les valeurs d_1 et d_2
- deux contraintes sur le gradient $d_1 - d_2$

Ainsi on compte donc au plus 6 points extrêmes dans ce polyèdre. Quelque soit le programme P choisi, la consommation la plus éloignée de ce programme sera un point extrême (puisque l'on cherche à s'éloigner le plus du programme).

Le graphe suivant illustre notre polyèdre si on considère les 2 premiers pas de temps $t = 0h$ et $t = 0h30$, avec la prévision moyenne initiale du 18 mars 2008.

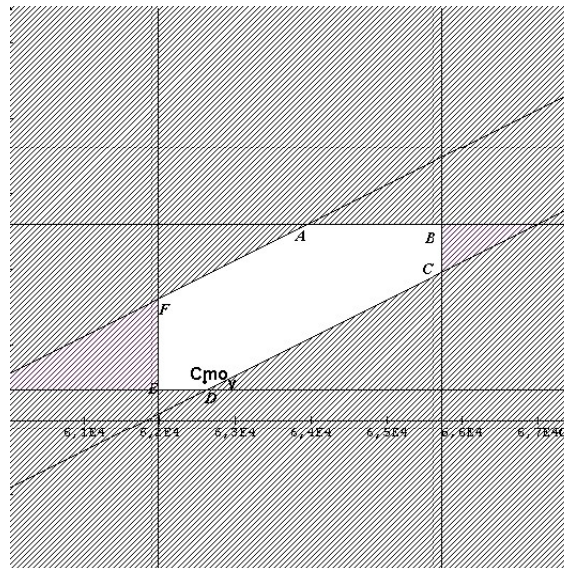


FIGURE 4.4 – Représentation graphique du polyèdre d'incertitude

Le polyèdre présenté ci-dessus a 6 points extrêmes.

On peut visualiser les solutions possibles de ξ^- et ξ^+ sur ce polyèdre : En plaçant $P = d_{moy}$, si d^* se situe en haut à gauche de P , ceci signifie que $\xi_1^+ \geq 0$ et $\xi_2^+ \geq 0$.

La pire demande possible si $P_{init} = d_{moy}$ est $d^* = B$.

Pour minimiser la pénalité maximum à payer, il faudrait que la production P soit située

plus près du point B , donc du centre du polyèdre.

4.5 Résolution du problème général

L'objectif est de résoudre le problème d'optimisation Θ , en dimension plus élevée, typiquement jusqu'à 48 pas de temps. Pour cela, on se place dans le cadre suivant :

- La fonction $\Gamma(P)$ est soit linéaire, soit convexe affine par morceaux.
- $\Phi(P)$ est solution d'un programme linéaire en nombres entiers. L'étude de la fonction d'ajustement a montré que $\Phi(P)$ est une fonction convexe non différentiable.

Nous nous intéressons, dans la suite, au cas général de minimisation de fonction convexe non différentiable à l'aide de l'algorithme proximal. Cette méthode de résolution a été utilisée pour minimiser $(\Gamma(P) + \Phi(P))$ avec $P \in \mathcal{X}$.

4.5.1 Algorithme proximal conceptuel

Le principe de l'algorithme proximal (tel que défini dans [23]) consiste à utiliser la notion de régularisée d'une fonction convexe non partout différentiable. Soit f une fonction convexe propre non partout différentiable, on note \tilde{f}_λ la régularisée de Yosida-Moreau de f (cf Yosida 1965, Moreau 1965), définie pour toute valeur de $\lambda \geq 0$ par :

$$\tilde{f}_\lambda(x) = \min_{y \in \mathbf{R}} \left\{ f(y) + \frac{\lambda}{2} \|y - x\|^2 \right\}$$

Les propriétés de la régularisée de Yosida-Moreau expliquent son utilisation dans l'algorithme proximal et ses variantes (algorithme des faisceaux, algorithme de sous-gradient, etc.). On cite ici les propriétés principales :

- (i) Pour tout $x \in \mathbf{R}$, la fonction en y : $f(y) + \frac{\lambda}{2} \|y - x\|^2$ admet un minimum global unique
- (ii) Pour tout $x \in \mathbf{R}^n$, $\tilde{f}_\lambda(x) \leq f(x)$
- (iii) x^* minimise f sur \mathbf{R}^n si et seulement si

$$x^* = \arg \min_{y \in \mathbf{R}^n} \left\{ f(y) + \frac{\lambda}{2} \|y - x^*\|^2 \right\}$$

c'est à dire, si et seulement si $\tilde{f}_\lambda(x^*) = f(x^*)$

Preuve :

(i) Pour x fixé, la fonction $f(y) + \frac{\lambda}{2} \|y - x\|^2$ est convexe propre en y . Elle admet donc un minimum global unique \bar{y}

(ii) Par définition, pour tout $y \in \mathbb{R}^n$, $\tilde{f}_\lambda(x) \leq f(y) + \frac{\lambda}{2} \|y - x\|^2$. En posant $y = x$, on obtient $\tilde{f}_\lambda(x) \leq f(x)$

(iii) (\Rightarrow)

Soit x^* minimise f sur \mathbb{R}^n . Posons $z^* = \arg \min_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{\lambda}{2} \|y - x^*\|^2 \right\}$, et montrons que $x^* = z^*$.

D'après (ii), on a :

$$\begin{aligned} \forall x \in \mathbb{R}^n \quad \tilde{f}_\lambda(x) &\leq f(x) \\ \text{en posant } x = x^* \quad \tilde{f}_\lambda(x^*) &\leq f(x^*) \\ \Leftrightarrow & \\ \min_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{\lambda}{2} \|y - x^*\|^2 \right\} &\leq f(x^*) \quad (4.2) \\ \Leftrightarrow & \\ \left\{ f(z^*) + \frac{\lambda}{2} \|z^* - x^*\|^2 \right\} &\leq f(x^*) \end{aligned}$$

Or par définition $f(z^*) \leq f(x^*)$. Ce qui implique $x^* = z^*$.

(\Leftarrow)

Soit x^* tel que $f(x^*) = \min_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{\lambda}{2} \|y - x^*\|^2 \right\}$.

Soit $z^* = \arg \min_{x \in \mathbb{R}^n} f(x)$, et montrons que $x^* = z^*$

$$\begin{aligned} \forall y \in \mathbb{R}^n, \quad f(x^*) &\leq f(y) + \frac{\lambda}{2} \|y - x^*\|^2 \\ \text{d'où} & \\ \text{en posant } y = z^*, \quad f(x^*) &\leq f(z^*) + \frac{\lambda}{2} \|z^* - x^*\|^2 \end{aligned} \quad (4.3)$$

Or par définition de z^* , $f(z^*) \leq f(x^*)$. Ce qui implique $x^* = z^*$.

Ainsi, minimiser f revient à minimiser sa régularisée \tilde{f} . Ce qui peut se faire à l'aide de l'algorithme "conceptuel" suivant :

Algorithm 1 Algorithme proximal conceptuel

$x^{(0)}$ solution de départ
 $k \leftarrow 0$
loop
 $x^{(k+1)} = \arg \min_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{\lambda}{2} \|y - x^{(k)}\|^2 \right\}$
 $k \leftarrow k + 1$
end loop

Remarque :

On notera les liens très proches entre l'algorithme proximal et les approches explorées sous le nom de **Méthodes des Faisceaux**, et introduites par Lemaréchal [35] en 1978. La méthode des faisceaux est actuellement employée avec succès dans le modèle en exploitation APOGEE [26].

4.5.2 Algorithme proximal effectif

Minimiser en y la fonction $f(y) + \frac{\lambda}{2} \|y - x^*\|^2$ reste toutefois un problème difficile. Ainsi, pour une résolution efficace, nous remplaçons f par son approximation tangentielle \hat{f} construite de la manière suivante :

$$\hat{f}_J(x) = \max_{j \in J} \{ f(z^j) + (g^j)^T (x - z^j) \}$$

où z^j est un point de \mathbf{R}^n , et $g^j \in \partial f(z^j)$, un sous-gradient de f au point z^j . J est l'ensemble des couples (z^j, g^j) obtenus au cours des itérations précédentes.

Le problème de minimisation revient à résoudre le problème quadratique convexe suivant :

$$PQuad(J, x^k) = \min_{y \in \mathbf{R}^n} \left\{ \hat{f}_J(y) + \frac{\lambda}{2} \|y - x^k\|^2 \right\}$$

sous les contraintes :

$$\forall j \in J, \quad f(z^j) + (g^j)^T (y - z^j) \leq \hat{f}_J(y)$$

L'algorithme *conceptuel* peut nécessiter un nombre infini d'itérations, vu le niveau d'exigence souhaité. En effet, il suppose que la minimisation de $\min_{y \in \mathbf{R}^n} \left\{ f(y) + \frac{\lambda}{2} \|y - x^{(k)}\|^2 \right\}$ soit effectuée de manière exacte.

C'est la raison pour laquelle on décrit ci-dessous un algorithme *effectif* (implémentable) donnant en un temps fini une solution dont la précision dépend de deux paramètres ϵ et ϵ' . L'algorithme proximal effectif tel que décrit dans [38] à une étape courant k s'écrit alors :

Algorithm 2 Algorithme proximal effectif

(A) Soit x^k le point courant
 Initialiser $PQuad(J, x^k)$ avec comme ensemble de couples $S = \{(z^1, g^1)\}$, avec $z^1 = x^k$
 et $g^1 \in \partial f(x^k)$
 $i \leftarrow 1$
loop
 (B) $i \leftarrow i + 1$
 Résoudre $PQuad(J, x^k)$. Soit z^i la solution optimale obtenue
 Calculer $f(z^i)$ et $g^i \in \partial f(z^i)$
 Calculer $\hat{f}_J(z^i) = \max_{j \in i-1} \{f(z^j) + (g^j)^T (z^i - z^j)\}$ et $\epsilon_i = f(z^i) - \hat{f}_J(z^i)$
 $S \leftarrow S \cup \{z^i, g^i\}$ et $J \leftarrow J \cup \{i\}$
 Tester :
if

$$\begin{cases} \epsilon_i \leq \epsilon \\ or \\ \epsilon_i \leq \frac{\lambda}{2} \|z^i - x^k\|^2 \end{cases}$$

then
 poser $x^{k+1} \leftarrow z^i$ et Aller en (C)
else
 aller en (B)
end if
end loop
 (C) Tester :
if $\|x^{k+1} - x^k\| \leq \epsilon'$ **then**
 Fin de l'algorithme . Solution optimale obtenue au point x^{k+1}
else
 Aller en (B)
end if

Pour atteindre, de manière approchée, l'optimum de la fonction régularisée de f , l'algorithme proximal intègre deux itérations complémentaires.

L'itération principale améliore la solution courante, jusqu'à convergence. Pour décider de la solution candidate à chaque itération principale, on améliore l'approximation tangentielle \hat{f} de f et ce à l'aide d'ajout successifs de coupes affinant cette approximation. Lorsque la solution candidate proposée satisfait le niveau d'approximation requis (critère ϵ), ou lorsqu'elle améliore la valeur de la fonction (comparée au candidat de l'itération précédente), alors cette solution est retenue pour l'itération principale.

Remarque :

Une variante à l'algorithme proximal décrit ci-dessus permet de minimiser **la fonction** $f(x)$ avec x **appartenant à un convexe** \mathcal{C} donné. Comme présenté par Dodu, Eve et Minoux dans [23], l'algorithme proximal s'implémente de manière efficace dans le cas où on minimise la fonction sous un ensemble convexe décrit par un ensemble de contraintes linéaires. C'est *cette variante qui sera implémentée* pour la résolution du problème robuste de gestion de production.

4.5.3 Tests d'arrêt

Comme exposé dans l'algorithme ci-dessus, le contrôle de l'algorithme se fait par le choix de deux tolérances ϵ_i et ϵ' qui déterminent l'arrêt ou pas de l'itération en cours. Pour saisir l'intérêt du premier test d'arrêt, nous utilisons la proposition suivante :

Proposition :

$$f(z_i) - \hat{f}_J(z_i) \leq \frac{\lambda}{2} \times \|z_i - x_k\|^2 \rightarrow f(z_i) \leq f(x_k)$$

Preuve :

Pour la clarté des notations, nous poserons $h = \hat{f}_J$, approximation tangentielle de f suite aux J précédentes itérations. z_i est la solution de $PQuad(J, x^k)$, c'est à dire :

$$\tilde{h}_\lambda(x_k) = PQuad(J, x^k) = \min_{y \in \mathbb{R}^n} \left\{ h(y) + \frac{\lambda}{2} \times \|y - x_k\|^2 \right\}$$

Et par construction, $\forall x \in \mathbb{R}^n, h(x) \leq f(x)$. La proposition 4.4 s'écrit :

$$\begin{aligned} f(z_i) - h(z_i) &\leq \frac{\lambda}{2} \times \|z_i - x_k\|^2 \\ &\Leftrightarrow \\ f(z_i) &\leq h(z_i) + \frac{\lambda}{2} \times \|z_i - x_k\|^2 \\ &\Leftrightarrow \\ f(z_i) &\leq \min_{y \in \mathbb{R}^n} \left\{ h(y) + \frac{\lambda}{2} \times \|y - x_k\|^2 \right\} \\ &\Leftrightarrow \\ f(z_i) &\leq \tilde{h}_\lambda(x_k) \leq f(x_k) \end{aligned}$$

Détaillons l'intérêt des deux test d'arrêts de l'algorithme proximal :

- **Test d'arrêt n°1** : Ce test d'arrêt est un double test :
 - il teste la qualité de l'approximation tangentielle \hat{f} . Si \hat{f} approxime parfaitement f avec une tolérance inférieure ou égale à ϵ alors la solution obtenue du sous-problème quadratique est un candidat à conserver.
 - il teste si le nouveau candidat z^i améliore la minimisation de la fonction f , d'après la proposition ci-dessus. Ce test permet de conserver un bon candidat z^i s'il vérifie la seconde inégalité du test.
- **Test d'arrêt n°2** ou test de fin d'algorithme : Il intervient à la dernière étape de l'algorithme. Le paramètre ϵ' permet de finir les itérations si la valeur $\|x^{k+1} - x^k\|$ est assez faible, c'est à dire dès que l'itération effectuée n'améliore pas suffisamment la solution obtenue à l'itération précédente.

4.5.4 Implémentation

En pratique, nous avons développé une maquette informatique développée en langage C/C++ mettant en œuvre l'algorithme de Dodu, Eve et Minoux [23]. L'architecture de la maquette se décompose en 3 parties qu'on retrouve dans l'algorithme proximal :

- Phase d'initialisation : Lecture et import des données. Allocation mémoire et création d'un environnement de travail en interface avec CPLEX 12.0.
- Résolution de sous problèmes : L'algorithme proximal nécessite l'estimation de la fonction à minimiser, ainsi qu'un sous-gradient, et ce à plusieurs reprises. D'où l'implémentation d'un module pour chaque fonction à estimer (fonctions Γ et Φ), ainsi qu'un troisième module pour la résolution du sous-problème quadratique
- Algorithme proximal : Un module principal est dédié au processus itératif de l'algorithme. L'appel aux différents sous-problèmes, la mise à jour de l'oracle (ensemble de couplets dans J) et les tests d'arrêts sont gérés dans ce module.

Nous avons exposé dans ci-dessus une formulation robuste *originale* du problème de gestion de production en contexte incertain, ainsi qu'une méthode de résolution dans le cas de fonctions convexes non partout différentiables. Cette approche a pu être qualifiée d'*originale* au vu des travaux publiés dans la littérature, et en particulier le travail ultérieur de Zhang et Guan [53] sur "*le placement robuste à deux étapes de la production*". Dans la section suivante nous présentons une étude comparative entre leurs travaux et notre approche robuste.

4.6 Comparatif avec les travaux de Zhang et Guan [53]

Zhang et Guan étudient dans leur article le problème de gestion de production électrique journalière dans un contexte d'incertitude sur la demande en puissance. Nos problèmes métier sont donc similaires.

Le modèle robuste proposé est un modèle en "deux étapes" ("Two-Stage"). Il s'écrit de manière simplifiée sous la forme suivante :

$$\begin{aligned}
\min_{x_1} \quad & \left\{ c_1^T x_1 + \max_{d \in \mathcal{D}} \left\{ \min_{x_2} \left\{ \sum_{t \in T} c_2^T x_2(t) \right\} \right\} \right\} \\
s.c : & x_1 \in \mathcal{X}_1 \subset \{0, 1\}^{n_1} \\
& x_2 \in \mathcal{X}_2(x_1) \subset \mathbb{R}^{n_2}
\end{aligned} \tag{4.4}$$

Les variables x_1 de première étape correspondent à des *décisions binaires* d'arrêt/marche des unités de production, \mathcal{X}_1 étant l'ensemble des solutions binaires réalisables.

En seconde étape, les variables de décision x_2 portent sur le niveau de fonctionnement de ces unités et les décisions d'achat/vente sur le marché de l'électricité, et ce pour la pire demande possible. $\mathcal{X}_2(x_1)$ est l'ensemble des décisions continues réalisables, dépendant des décisions de première étape.

La demande d (vecteur sur T pas de temps) appartient à un ensemble d'incertitude \mathcal{D} décrit par des bornes sur la demande, ainsi que d'une limitation du volume de demande :

$$d \in \mathcal{D} = \left\{ \sum_{t=1}^T \pi_t d_t \leq \pi_0, \underline{d}_t \leq d_t \leq \bar{d}_t, t = 1 \dots T \right\}$$

Sous l'hypothèse de séparabilité du sous-problème $\{\min_{x_2} \{\sum_{t \in T} c_2^T x_2(t)\}\}$, le sous-problème est réécrit sous la forme : $\sum_{t \in T} f_t(d_t)$, avec $f_t(d_t)$ une fonction convexe linéaire par morceaux, sa valeur étant le maximum d'une collection finie (mais de grande taille) de fonctions linéaires.

Par dualisation du terme $\max_{d \in \mathcal{D}} \sum_{t \in T} f_t(d_t)$, le problème devient traitable mais possède un nombre exponentiel de contraintes duales (dû à la taille de la collection de fonctions). La résolution se fait de manière itérative à l'aide d'un algorithme de séparation (ajout graduel de contraintes au problème initial).

Les résultats numériques présentés restent préliminaires et en particulier ne donnent pas d'indication sur le comportement de cette approche par rapport à une approche déterministe de référence.

Notre approche et l'approche proposée par Zhang et Guan se distinguent sur plusieurs points.

Le modèle de Zhang ne propose pas, en sortie, *un programme de production* à fournir aux unités de production. Les niveaux de production des unités de production sont en effet des variables de seconde étape.

Le modèle en deux étapes implique que *les décisions obtenues* en sortie du modèle sont **uniquement** les décisions d'arrêt/marche des unités. Notre approche robuste propose quant

à elle un programme de production pour l'ensemble des unités de production.

De plus, le nombre exponentiel de contraintes suite à la dualisation du terme de maximisation présente une limite non négligeable. En effet, le nombre de contraintes résultantes vaut exactement U^T (le nombre d'usines *puissance* le nombre de pas de temps). La résolution du problème, même à l'aide d'un algorithme de séparation, risque d'être difficile dans un contexte industriel réel (l'approche que nous proposons prend en entrée 52 unités de production sur 48 pas de temps, contre 32 unités de production sur 24 périodes pour le modèle de Zhang).

Enfin, notre polyèdre d'incertitude a une vision bien plus fine de l'aléa sur la demande en puissance. Le travail mené pour la construction de notre ensemble d'incertitude permet une description plus aboutie de l'aléa, comme nous le détaillerons dans le prochain chapitre.

Conclusions

Nous avons exposé dans ce chapitre une modélisation robuste du problème de gestion de production en contexte incertain, qui peut être qualifiée d'*originale* au vu des travaux publiés dans la littérature (en particulier le travail ultérieur de Zhang et Guan).

Cette approche permet de se protéger contre la pire incertitude sur la demande. La modélisation peut se décliner selon deux visions, toutes deux intéressantes sur le plan industriel, la vision *Marchés* et la vision *Pénalités*.

Après étude de la régularité des fonctions à optimiser, nous avons étudié le comportement du modèle dans le cas d'exemples simples en faible dimension. Une différence est vite apparue entre les visions proposées, ainsi que le rôle important que peuvent jouer les paramètres en entrée du modèle.

Nous avons ensuite introduit l'algorithme proximal qui a été utilisé pour la résolution du problème. L'algorithme se base sur les propriétés des fonctions à minimiser, et particulièrement sur la notion de régularisée (définie ci-dessus).

Néanmoins cette résolution est valable grâce à la régularité des fonctions à minimiser $\Gamma(P)$ et $\Phi(P)$.

Les différentes classes de fonction impliquent toutefois des approches différentes pour la résolution du modèle robuste.

Le cas le plus favorable est une résolution dans le cas où les fonctions Γ et Φ sont convexes différentiables. Ensuite vient le cas traité dans ce chapitre, à savoir ***la minimisation d'une fonction convexe non-différentiable.***

La propriété de convexité est une propriété importante pour la résolution de l'approche robuste. Toutefois, le contexte et la réalité industrielle sont rarement décrits par des fonctions convexes. La non convexité est même parfois *inévitabile* pour tout modèle réaliste.

Prenons pour exemple le modèle APOGEE. La fonction d'estimation du coût est non convexe, minimisée sur un ensemble non connexe de programmes réalisables.

Une des perspectives suite à ces travaux serait l'implémentation d'une méthode de résolution dans le cas où la fonction Γ ou Φ sont non-convexes. Néanmoins, nous proposons ci-dessous les pistes de réflexion vers lesquelles nous avons convergé.

Dans le cas où *une fonction f est non convexe mais séparable*, le problème reste traitable. On peut procéder à une convexification f^* de la fonction de départ, évaluer un minorant puis minimiser la fonction initiale f à l'aide d'une approche de Branch & Bound. Le cas d'*une fonction non convexe non séparable* est par contre un cas difficile, mais fréquent. Un cas important où des non-convexités apparaissent est lorsque le modèle tient compte des coûts de démarrage.

La non-séparabilité d'une fonction est souvent liée au fait que les solutions réalisables satisfont des contraintes dynamiques structurantes.

Pour exemple, *la contrainte de gradient* dans la gestion de production électrique est une contrainte dynamique. En effet, la production à un pas de temps d'une unité de production dépend fortement des productions aux pas de temps *avant et après*.

Dans ce cas de figure, une approche possible serait d'étudier la nature des contraintes dynamiques du problème, afin de déterminer si possible *un changement de base* permettant de rendre la fonction f séparable.

Chapitre 5

Construction de l'ensemble d'incertitude

5.1 Prédiction de la consommation électrique

5.1.1 Principe

Pour le calcul des prévisions journalières de consommation, plusieurs facteurs sont étudiés. La température est le paramètre le plus important, ensuite s'ajoutent des corrections dues à l'influence du calendrier (jour ouvré, férié, changement d'heure d'été/hiver) et de la nébulosité (taux de couverture nuageuse). Mais ces paramètres ont une influence bien moindre que celle de la température.

Météo France fournit 50 scénarios de température ainsi qu'un scénario moyen de température dont sont déduits 50 scénarios de demande en puissance et un scénario moyen de demande en puissance.

Actuellement, seul le scénario moyen de température est utilisé pour le calcul de prévision journalières. Les 50 scénarios servent uniquement aux prévisions **hebdomadaires** de demande. Il n'existe donc pas d'ensemble discret de scénarios pour la demande en puissance à l'horizon **journalier**.

5.1.2 Outils

Un modèle paramétrique de prévision de consommation est utilisé pour les prévisions à court et moyen terme par EDF. Les estimations obtenues sont déterministes et ne permettent pas de donner une vision probabiliste de la demande en électricité. Cet outil de prévision (appelé "Eventail") fournit pour chaque demi-heure une estimation de la demande en fonction des variables exogènes climatiques (température, nébulosité), et calendaires (date, type de jour particulier).

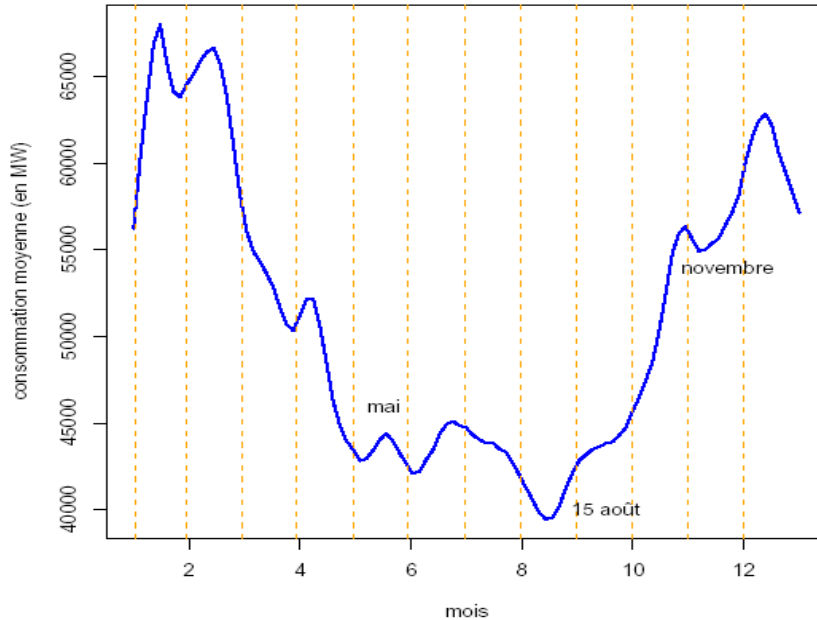


FIGURE 5.1 – Consommation électrique moyenne hebdomadaire sur une année

Comme décrit dans [30], l'estimation de la consommation électrique se décompose en trois termes additifs $\hat{P} = \hat{P}_{hc} + \hat{P}_c + \hat{P}_{eff}$:

- \hat{P}_{hc} modélisant la consommation indépendante du climat. Elle estime la forme de la courbe de charge en fonction du type de jour, de l'heure et de la position dans l'année.
- \hat{P}_c modélisant la consommation qui dépend des conditions climatiques. Ce terme est défini à partir de la température ressentie, et il mesure l'impact du chauffage et/ou de la climatisation sur la consommation
- \hat{P}_{eff} modélisant l'influence des tarifs spéciaux accordés par EDF à certains clients afin de réduire leur consommation en période de pointe.

Pour avoir une vision de l'incertitude autour de la consommation prévue, et ce pour les besoins de notre étude, un outil a été implémenté afin de modéliser le profil de l'erreur issue du modèle de prévision. Ceci permet ainsi de générer *des scénarios d'erreurs de prévision*. Ce générateur sera utile dans la suite pour la construction de notre ensemble d'incertitude.

5.1.3 Hypothèses de travail

Il est donc nécessaire de construire un ensemble d'incertitude pour modéliser les erreurs de prévision sur la demande en puissance. Nous supposons que cet ensemble est continu, incluant une infinité de réalisations possibles, et basé sur l'observation des écarts entre prévision de demande et demande réalisée.

D'autre part on suppose que l'incertitude peut être correctement représentée sous forme d'un polyèdre, c'est à dire un ensemble défini par un système d'inégalités linéaires, que l'on va chercher à construire à partir d'observations réelles (ce polyèdre serait idéalement l'enveloppe convexe des scénarios décrits).

Dans la suite, nous allons expliciter les indicateurs qui nous intéressent pour construire ce polyèdre d'incertitude. Ce polyèdre représentera l'ensemble des écarts possibles entre prévision de demande et réalisation (que nous appellerons erreur de prévision) pour plusieurs journées tests qui serviront à expliciter ces indicateurs.

5.2 Construction du polyèdre

5.2.1 Première construction à partir de scénarios d'erreurs

Soit I l'ensemble des observations des erreurs de prévision. Une observation est une chronique d'erreur de prévision sur une journée. $e_i(t)$ est donc l'erreur de prévision au pas de temps t observée sur la journée test i . Nous avons cherché des indicateurs caractérisant les différentes chroniques d'erreurs. Certains indicateurs sont déductibles par simple observation de l'ensemble des scénarios. Les indicateurs retenus sont les suivants :

- $\bar{e}(t)$ et $\underline{e}(t)$: Les bornes supérieures et inférieures de l'erreur à un instant donné sont les premiers indicateurs à étudier pour borner l'incertitude sur la demande :

$$\bar{e}(t) = \max_{i \in I} e_i(t), \quad \underline{e}(t) = \min_{i \in I} e_i(t)$$

- $\bar{g}(t)$ et $\underline{g}(t)$: la borne supérieure et inférieure du gradient de l'erreur à l'instant t aident à décrire la dynamique de l'erreur sur deux pas de temps consécutifs. En effet, le gradient de puissance des unités de production étant lui-même borné, il est important d'évaluer la propagation de l'erreur sur deux pas de temps consécutifs. D'où :

$$\bar{g}(t) = \max_{i \in I} \{e_i(t+1) - e_i(t)\} \quad \underline{g}(t) = \min_{i \in I} \{e_i(t+1) - e_i(t)\}$$

- Des bornes \bar{e}_a et \underline{e}_a sur la somme des erreurs en valeur absolue indique le niveau minimum et maximum d'écart en énergie sur une journée donnée :

$$\bar{e}_a = \max_{i \in I} \left\{ \sum_{t=1}^T |e_i(t)| \right\}, \quad \underline{e}_a = \min_{i \in I} \left\{ \sum_{t=1}^T |e_i(t)| \right\}$$

- Des bornes \overline{g}_a et \underline{g}_a sur la somme des gradients en valeur absolue indique les niveaux extrêmes de propagation de l'erreur sur une journée donnée :

$$\overline{g}_a = \max_{i \in I} \left\{ \sum_{t=1}^{T-1} |e_i(t+1) - e_i(t)| \right\}, \quad \underline{g}_a = \min_{i \in I} \left\{ \sum_{t=1}^{T-1} |e_i(t+1) - e_i(t)| \right\}$$

un scénario d'erreur de prévision appartient au polyèdre d'incertitude si et seulement s'il respecte les contraintes linéaires associées.

Remarque : Pour les contraintes sur les valeurs absolues de l'erreur (ou du gradient), l'ajout de deux variables supplémentaires (en l'occurrence la partie positive e^+ et la partie négative e^- de la variable e) permet de conserver la structure linéaire des contraintes.

Néanmoins, on peut se douter qu'un polyèdre construit à partir de ces simples contraintes peut manquer de finesse dans la description de l'erreur sur la demande. Afin de mieux saisir la structure contenue dans les chroniques d'erreur étudiées, nous envisageons des indicateurs moins intuitifs qui permettent d'affiner la description polyédrale de l'aléa.

5.2.2 Consolidation de la construction

L'ensemble des observations d'erreur peut contenir des indicateurs non triviaux, non décelables par observation. Face à un nuage d'observations, des méthodes appartenant au domaine de l'analyse de données permettent d'explicitier des liens entre les données afin d'obtenir une description statistique succincte.

Pour mieux exploiter l'information contenue dans ces chroniques d'erreurs, on a fait le choix de la méthode dite d'analyse en composantes principales (ACP) pour mieux décrire notre échantillon d'observations.

Principe de l'ACP

L'Analyse en Composantes Principales est une méthode mathématique d'analyse des données qui consiste à rechercher les directions de l'espace qui représentent le mieux les corrélations entre plusieurs variables aléatoires. Elle permet d'obtenir des axes selon lesquels la dispersion de l'échantillon d'observations est la plus forte.

Soit S le nuage de points qui correspond aux S réalisations de la variable aléatoire $X = (X_1, X_2, \dots, X_T)$, et soit $\tilde{C} \in \mathbb{R}^S \times \mathbb{R}^T$ la matrice centrée de ce nuage de points (chaque ligne de la matrice correspond à une réalisation possible de l'aléa sur T pas de temps). X_i^s correspond à l'observation s de l'erreur de prévision au pas de temps i . L'ACP permet de déterminer les directions dans le nuage de points qui suivent son allongement maximal.

On peut donc obtenir un axe u tel que :

$$\max_{r \in \mathbb{R}^T} \sum_{i,k \in S} d_r^2(X^i, X^k) = \sum_{i,k \in S} d_u^2(X^i, X^k)$$

Autrement dit, La méthode de l'ACP nous permet d'écrire la contrainte suivante :

$$\underline{\nu} \leq \langle u, e \rangle \leq \bar{\nu}$$

avec $\underline{\nu}, \bar{\nu}$ et u obtenus par l'ACP, u étant un vecteur qui **maximise l'écart** $\bar{\nu} - \underline{\nu}$.

Obtention des axes

Soit u l'axe à calculer, de norme 1. La projection de l'échantillon des X s'écrit :

$$\pi_u(\tilde{C}) = \tilde{C} \cdot u = \begin{pmatrix} \sum_{i=1}^T (X_i^1 - X_i^{moy}) \cdot u_i \\ \vdots \\ \sum_{i=1}^T (X_i^S - X_i^{moy}) \cdot u_i \end{pmatrix}$$

La variance empirique de $\pi_u(\tilde{C})$ vaut :

$$\sigma_u^2 = \left(\frac{1}{S}\right) \cdot \pi_u(\tilde{C})^T \cdot \pi_u(\tilde{C}) = \left(\frac{1}{S}\right) \cdot u^T \cdot \tilde{C}^T \cdot \tilde{C} \cdot u$$

Et $V = \left(\frac{1}{S}\right) \cdot \tilde{C}^T \cdot \tilde{C}$ est la matrice de covariance du nuage de points \mathbf{S} . V étant symétrique positive, elle s'écrit dans une base orthonormée sous la forme suivante :

$$V = P^T \cdot D \cdot P \text{ avec } D = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_T)$$

Donc :

$$\begin{aligned} \sigma_u^2 &= (u^T P^T) \cdot D \cdot (Pu) \\ \text{Posons } w &= P \cdot u \\ \Leftrightarrow \sigma_u^2 &= w^T \cdot D \cdot w \end{aligned}$$

D'où σ_u^2 maximale pour $w^T \cdot D \cdot w = \lambda_{max}$

Donc σ_u^2 maximale pour u vecteur propre associé à la valeur propre λ_{max} . La matrice P est la matrice des vecteurs propres orthogonaux associés aux valeurs propres λ_i .

En l'occurrence, la direction qui suit l'allongement maximal du nuage, est l'axe v tel que :

$$\sigma_v^2 = \lambda_{max}, \quad \lambda_{max} = \max_i (\lambda_i \mid \lambda_i \geq 0, 1 \leq i \leq T)$$

Résultats

A partir de 70 000 simulations de l'erreur de prévision sur les journées ouvrées d'hiver, on a construit des contraintes basées sur des indicateurs triviaux et d'autres basées sur les axes de l'ACP.

La journée étant discrétisée en 48 pas de temps demi-horaires, on compte alors 96 contraintes sur les bornes, 94 contraintes sur les gradients, 4 contraintes sur leur somme en valeur absolue et 96 contraintes correspondant aux projections sur les 48 axes de l'ACP.

Pour obtenir les axes de l'ACP, on fait appel à une librairie pré-existante dans le langage *R* [55]. 5 axes principaux se distinguent avec une valeur propre associée supérieure à 1 (cf tableau 5.2.2 et figure 5.2) :

| | Axe1 | Axe 2 | Axe3 | Axe 4 | Axe 5 | Axes 6 à 48 |
|--------------------|------|-------|------|-------|-------|-------------|
| Variance expliquée | 5.2 | 3.04 | 2.14 | 1.48 | 1.10 | < à 1 |

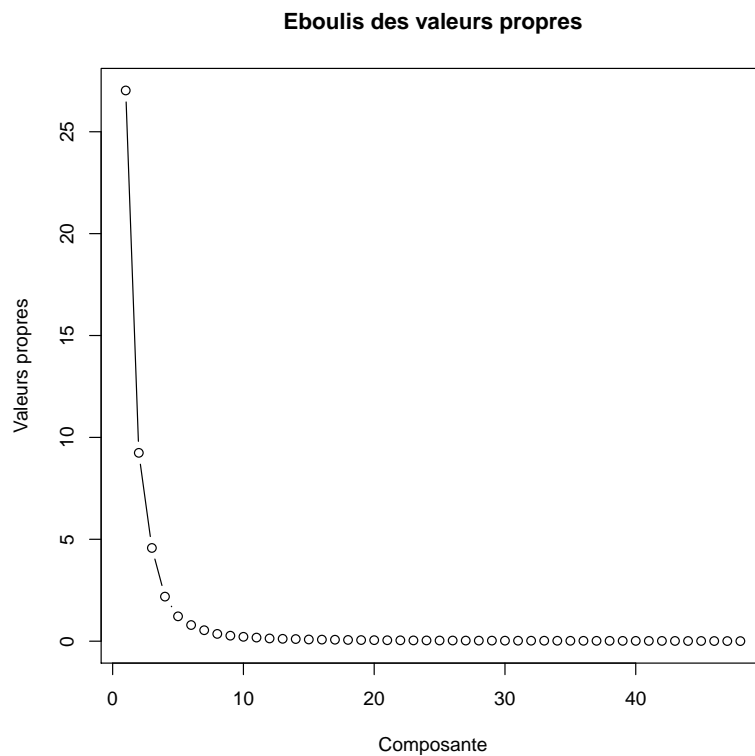


FIGURE 5.2 – Eboulis des valeurs propres

ACP & polyèdre d'incertitude

Les axes obtenus par l'ACP sont utilisés pour affiner la description polyédrale de l'erreur sur la demande. Soit J l'ensemble de ces axes, de cardinal T .

Les axes de l'ACP peuvent s'interpréter comme une base de T vecteurs indépendants qui suivent l'allongement maximal du nuage de points.

Pour tout axe $v_j \in J$, on obtient ainsi les indicateurs de dispersion $\underline{\nu}_j$ et $\bar{\nu}$. Ces indicateurs bornent la projection du nuage de points sur l'axe v_j . Rappelons que I est l'ensemble des scénarios d'erreurs de prévision, e_i étant une chronique d'erreur. Alors les indicateurs ont pour expression :

$$\underline{\nu}_j = \min_{i \in I} \langle v_j \cdot e_i \rangle = \min_{i \in I} \left\{ \sum_{t \in T} v_j(t) e_i(t) \right\}$$

$$\bar{\nu}_j = \max_{i \in I} \langle v_j \cdot e_i \rangle = \max_{i \in I} \left\{ \sum_{t \in T} v_j(t) e_i(t) \right\}$$

5.2.3 Polyèdre final

Partant des premiers indicateurs définis en 5.2.1, et des indicateurs obtenus après l'Analyse en Composantes Principales en 5.2.2, nous avons abouti à la construction du polyèdre sous la forme suivante :

Soit $e = (e_1, \dots, e_T)$ un vecteur d'erreur donné, et soit \mathcal{E} le polyèdre d'incertitude sur l'erreur de prévision. Alors :

$$e \in (\mathcal{E}) \iff \begin{cases} \bar{e}(t) \leq e_t & \leq \underline{e}(t) \quad \forall t = 1 \dots T \\ \underline{g}(t) \leq e_{t+1} - e_t & \leq \bar{g}(t) \quad \forall t = 1 \dots T - 1 \\ \underline{e}_a \leq \sum_{t=1}^T |e_t| & \leq \bar{e}_a \\ \underline{g}_a \leq \sum_{t=1}^{T-1} |e_{t+1} - e_t| & \leq \bar{g}_a \\ \underline{\nu}_j \leq \sum_{t=1}^{T-1} v_j(t) e_t & \leq \bar{\nu}_j \quad \forall j \in J \end{cases}$$

5.3 Validation basée sur un polyèdre

Un certain nombre de méthodes d'optimisation sous incertitude sont valables dans le cas où l'ensemble d'incertitude est un polyèdre. Si la phase d'optimisation a une vision polyédrique de l'aléa, la phase de simulation nécessite des scénarios décrivant les futurs possibles. L'obtention de ces scénarios peut se faire de deux manière différentes.

5.3.1 A partir d'un générateur

Dans le cas où l'on possède un générateur de scénarios avec un nombre suffisant de scénarios générés en un temps raisonnable, les scénarios de simulation peuvent être obtenus directement de ce générateur. Néanmoins, il faut être vigilant et veiller à ce que les scénarios ayant servi à la construction du polyèdre soient distincts des scénarios servant à la simulation. De cette manière, on exclut tout surapprentissage du futur.

5.3.2 Validation avec Tirage uniforme

Il est probable que des scénarios de validation ne soient pas disponibles ou que leur faible nombre nécessite leur utilisation exclusive pour la construction du polyèdre. Dans ce cas, on peut envisager une génération de scénarios à partir du polyèdre déjà construit. Certes, cette solution risque de favoriser la vision de l'incertitude décrite par le polyèdre. On fait ainsi l'hypothèse implicite que l'ensemble des observations possibles se restreint à l'ensemble des réalisations comprises dans le polyèdre. En faisant cette hypothèse, l'emploi d'une méthode de tirage uniforme permet alors de générer des scénarios uniformément répartis dans le polyèdre.

Ensembles simples

On peut effectuer un tirage uniforme multidimensionnel sur certains ensembles tels que l'hypercube, le parallélépipède ou encore l'hypersphère. Dans le cas de l'hypercube, il suffit d'effectuer n tirages uniformes dans l'intervalle $[a_n, b_n]$ (n étant la dimension de l'espace correspondant au côté n de l'hypercube).

Dans le cas de l'hypersphère, on effectue un tirage selon une loi normale centrée réduite sur chaque coordonnée du point.

Dans le cas où l'ensemble ne peut être obtenu à partir d'une sphère ou un rectangle par transformation linéaire, ces techniques de tirage ne s'appliquent pas. D'où l'intérêt porté à d'autres méthodes.

Techniques de rejet

On cherche à effectuer un tirage uniforme dans un ensemble donné, qu'on appellera V . Le tirage uniforme avec rejet consiste à trouver le plus petit ensemble X contenant V et pour lequel une technique de tirage uniforme existe.

Les choix de X sont généralement des hypercubes ou des hypersphères.

Cette approche a pour limite le nombre de points s à générer dans X pour obtenir un point appartenant à V . En effet, lorsque la dimension de l'espace croît, le rapport des volumes $\frac{Vol(V)}{Vol(X)}$ tend vers 0.

Le taux de rejet, qui dépend de ce rapport de volumes, devient ainsi plus élevé en grande dimension. Cette limite a été vite atteinte dans le cas de tirage uniforme dans le polyèdre construit précédemment.

Exemple :

Pour $T = 19$, sur 100.000.000 tirages dans l'hypercube, seuls 2 tirages sont dans le polyèdre (temps de calcul : 30 min).

Pour $T = 25$, aucun tirage dans le polyèdre, après 3h de temps de calcul sur une machine unix standard (les 600.000.000 tirages effectués jusque là n'appartenant pas au polyèdre).

Méthode de *Hit and Run*

Proposé en premier par Robert L. Smith en 1982 ([49]) sous le nom de *Mixing Algorithm*, cet algorithme est considéré comme l'un des plus rapides permettant d'effectuer un tirage uniforme dans un ensemble convexe.

Il est décrit de manière générale de la manière suivante :

Algorithm 3 Algorithme Hit and Run

Soit un point de départ $x_i \in V$, et soit $i = 0$

while $i \leq N$ **do**

 Générer une direction $d \in \mathbb{R}^n$ de manière uniforme à partir du point x_i

 Soit le segment L correspondant à l'intersection $S \cap \{x | x = x_i + \lambda d, \lambda \in \mathbb{R}\}$

 Générer un point aléatoire x_{i+1} tiré uniformément sur L

$i \leftarrow i + 1$

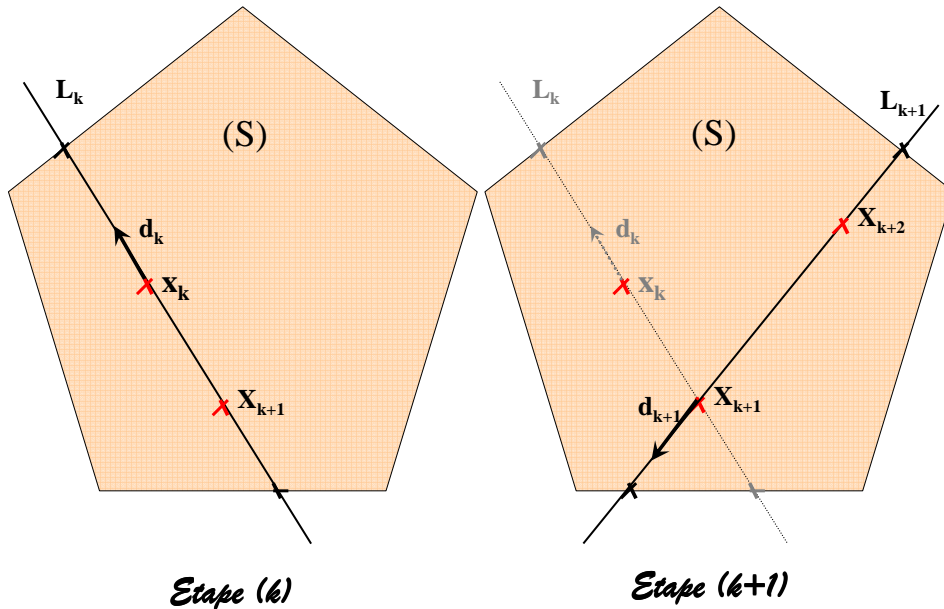
end while

Comme illustré sur la figure 5.3, la méthode *Hit and Run* a pour point de départ un point quelconque x_k dans l'ensemble convexe S . Une direction d_k est générée de manière uniforme (tirage uniforme sur une hypersphère), obtenant ainsi la droite de direction d_k et passant par x_k .

L_k est le segment résultat de l'intersection de la droite ainsi obtenue et l'ensemble convexe S . x_{k+1} est obtenu par tirage uniforme sur le segment L_k .

Les résultats obtenus montrent un bien meilleur taux de tirage que les méthode classiques de rejet, avec un tirage uniforme asymptotique.

La difficulté constatée à l'étude de cette approche réside dans le calcul de l'intersection entre la droite générée et l'ensemble convexe. Dans le cas où l'ensemble convexe S est notre polyèdre d'incertitude, on obtient les deux intersections par résolution de deux programmes linéaires. Ce calcul peut s'avérer coûteux en temps lorsque le nombre de tirages à effectuer devient important.

FIGURE 5.3 – Algorithme *Hit and Run* : Itérations (k) et (k+1)

5.3.3 Validation avec tirage et changement de base

Motivations

Dans le cadre de la démarche de validation, il était nécessaire de pouvoir proposer une technique de tirage uniforme ou pseudo-uniforme de scénarios à partir du polyèdre construit. Il est vite apparu qu'à partir d'un polyèdre de dimension 15, la technique de rejet devenait caduque. La structure d'un polyèdre quelconque fait que le rapport entre le volume du polyèdre et le volume de l'hypercube le contenant tend vers 0 quand la dimension de l'espace devient grande.

En utilisant l'hypercube contenant le polyèdre, on suppose de manière implicite que la forme du polyèdre est plus ou moins proche de la forme de l'hypercube dans la base canonique. Or les différentes coupes formant le polyèdre peuvent lui donner une forme toute différente.

Pour améliorer la technique de rejet, il était opportun de trouver un hypercube de volume plus petit. L'utilisation des axes obtenus par analyse en composantes principales a

présenté un grand intérêt.

En effet, on peut intuitivement penser qu'un hypercube "calé" sur la forme du nuage permettra un taux de rejet plus faible.

Exemple :

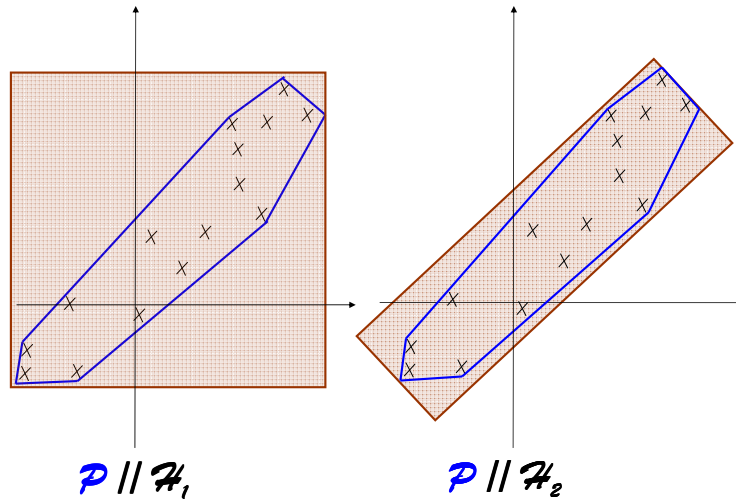


FIGURE 5.4 – Hypercube \mathcal{H}_1 selon les axes de la base canonique, \mathcal{H}_2 l'hypercube "calé" selon la forme du nuage, et \mathcal{P} le polyèdre d'incertitude

Le premier axe suit la dispersion maximale des scénarios, donc a fortiori l'allongement maximal du polyèdre. Le second axe suit l'allongement maximal du polyèdre dans l'hyperplan orthogonal au 1^{er} axe, et ainsi de suite. On peut alors construire un hypercube dans cette base. Sur chaque axe, l'hypercube correspond à l'intervalle compris entre les valeurs min et max de la projection des scénarios sur cet axe.

Soit l'hypercube \mathcal{H}_1 construit sur la base canonique :

$$\mathcal{H}_1 = \{S_i^1\}_{i=1\dots n} \quad S_i^1 = [\underline{e}_i, \bar{e}_i]$$

avec S_i^1 l'arête i de l'hypercube, et \underline{e}_i et \bar{e}_i respectivement les bornes min et max de l'erreur au pas de temps i , qui correspond dans la base canonique à la i -ème coordonnée.

Et soit \mathcal{H}_2 l'hypercube construit sur la base des axes de l'ACP :

$$\mathcal{H}_2 = \{S_i^2\}_{i=1\dots n} \quad S_i^2 = [\underline{\gamma}_i, \bar{\gamma}_i]$$

avec S_i^2 l'arête i de l'hypercube, et $\underline{\gamma}_i$ et $\bar{\gamma}_i$ respectivement les bornes min et max des projections sur l'axe i de l'ACP, qui correspond dans la nouvelle base à la i -ème coordonnée.

On constate à première vue que le volume de H_2 est plus faible que celui de H_1 . Les résultats montrent un écart de facteur 100 entre les deux volumes en dimension 4, et augmentant très fortement en dimension 10 ou encore 48 :

| | \mathbb{R}^4 | \mathbb{R}^{10} | \mathbb{R}^{48} |
|----------------------|------------------|-----------------------|------------------------|
| Volume H_1 | $4.18e + 13$ | $2.49e + 34$ | $1.01e + 164$ |
| Volume H_2 | $5.55e + 11$ | $3.53e + 29$ | $2.3e + 136$ |
| Facteur de réduction | $\sim \times 75$ | $\sim \times 7e + 04$ | $\sim \times 44e + 27$ |

Méthode de tirage améliorée

La technique de rejet améliorée s'inspire d'un changement de la base canonique à la base obtenue avec les axes de l'ACP.

Soit \mathcal{E} la base canonique, \mathcal{A} la base formée par les axes de l'ACP. Le polyèdre \mathcal{P} est défini dans la base canonique \mathcal{E} .

Posons $P_{\mathcal{A}}^{\mathcal{E}}$ la matrice de changement de base de \mathcal{A} vers \mathcal{E} .

Pour le tirage de N points dans le polyèdre de manière uniforme, nous avons utilisé l'algorithme suivant :

Algorithm 4 Tirage avec changement de base

```

while  $i \leq N$  do
  Générer  $X^{\mathcal{A}}$  uniformément dans  $H_2^{\mathcal{A}}$ 
  Calculer  $X^{\mathcal{E}} = P_{\mathcal{A}}^{\mathcal{E}} X^{\mathcal{A}}$  i.e. les coordonnées du point généré dans la base canonique  $\mathcal{E}$ 
  if  $X^{\mathcal{E}} \in \mathcal{P}$  then
     $i \leftarrow i + 1$ 
  end if
end while

```

Avec cette méthode de tirage uniforme améliorée, le taux de rejet a été nettement diminué, et ce pour des dimensions variables. Par exemple, en dimension 4, la méthode de tirage avec rejet dans la base canonique avait un taux de réussite d'environ 1%, contre 58% pour la méthode de tirage avec changement de base. Cet écart de taux de réussite reste important dans des espace à dimension plus élevée (exemple en dimension 10). En dimension 48, la méthode de tirage classique n'a généré aucun scénario valide (sur 1 500 000 tirages effectués). L'heuristique a permis de générer 2 scénarios valides.

Le tableau ci-dessous récapitule les taux de réussite obtenus avec les deux méthodes de

tirage, et ce avec **1 500 000 tirages initiaux** :

| | \mathbb{R}^4 | \mathbb{R}^{10} | \mathbb{R}^{48} |
|-------------------|----------------|-------------------|-------------------|
| Méthode classique | 1.0% | 0.0004% | ND |
| Temps de calcul | 1s | 1s | 28s |
| Heuristique | 58.69% | 13.018% | 0.0001% |
| Temps de calcul | 4s | 4s | 57s |

Cette méthode nécessite un temps de calcul sensiblement supérieur à une approche classique. Ceci étant dû aux opérations de changement de base. L'heuristique reste néanmoins rapide et efficace, même en dimension élevée.

Conclusion

Nous avons abordé dans ce chapitre l'approche suivie pour la construction de notre ensemble d'incertitude. Le polyèdre est construit autour d'un grand nombre d'observations, et les contraintes linéaires qui le délimitent sont basées sur des indicateurs variés (gradients, axes d'Acp, etc.).

Enfin, nous proposons une alternative de validation des modèles dans le cas où il n'y aurait pas suffisamment d'échantillons pour la phase d'optimisation (construction de l'ensemble d'incertitude) et la phase de simulation (validation des résultats obtenus).

La méthode de tirage uniforme décrite en dernière partie permet d'améliorer considérablement le taux de réussite des méthodes classiques de tirage avec rejet, et ce pour des dimensions élevées.

Nous présenterons dans les chapitres suivants les résultats obtenus avec l'approche robuste, basée sur l'ensemble d'incertitude décrit ci-dessus.

Chapitre 6

Le modèle robuste avec pénalités : Analyse des résultats

6.1 Hypothèses de travail

L'étude a pour objectif de comparer les solutions proposées par deux modèles d'optimisation : le premier modèle d'optimisation, qu'on appellera modèle déterministe, et le modèle robuste avec vision Pénalités.

Le modèle déterministe optimise le placement de la production avec *une vision moyenne* de la demande.

Le modèle robuste optimise le placement de la production en tenant compte du pire coût d'ajustement dû à l'écart entre la production et la demande réalisée. Il fournit un plan de production robuste en réponse au polyèdre d'incertitude autour de la prévision moyenne de la demande.

6.1.1 Modèle déterministe

Notations

Pour une bonne compréhension du problème que nous venons d'énoncer, détaillons les variables de décision, les données et les contraintes générales puis celles pour le parc thermique (nucléaire et thermique à flamme) et enfin, celles du parc hydraulique :

Générales

- Données
 - T : Nombre de pas de temps optimisés
 - $D(t)$: Demande en puissance à satisfaire pour tout $t = 1, \dots, T$ (en MW)
 - C_{def} : Coût de la défaillance en demande (en €/MWh)

- Variables
 - $P_{def}(t)$: Puissance de défaillance de la demande pour tout $t = 1, \dots, T$ (en MW)
 - $P_{surp}(t)$: Puissance en surplus de la demande pour tout $t = 1, \dots, T$ (en MW)
- Contraintes
 - Contrainte d'équilibre offre-demande (1)

Parc Thermique (nucléaire et thermique à flamme)

- Données
 - C : Nombre de centrales thermiques
 - $\overline{P}_T(c)$: Puissance maximale pour toute centrale thermique $c = 1 \dots C$
 - $G_T^+(c)$: Gradient de montée (en MW/pas de temps) pour toute centrale thermique $c = 1 \dots C$
 - $G_T^-(c)$: Gradient de descente (en MW/pas de temps) pour toute centrale thermique $c = 1 \dots C$
 - $C_T(c)$: Coût de fonctionnement proportionnel de la centrale thermique c (en €/MW)
- Variables
 - $P_{Th}(c, t)$: Production en puissance de la centrale thermique c au pas de temps t
- Contraintes
 - Contrainte sur la production de puissance (2)
 - Contrainte de gradient de puissance des centrales (3)

Parc Hydraulique

Par souci de clarté du modèle de production, le stock d'eau est vu comme la puissance qu'elle pourrait produire. L'unité utilisée est donc le MW.

- Données
 - U : Nombre d'usines
 - R : Nombre de réservoirs
 - $\overline{P}_H(u)$: Puissance maximale pour toute usine $u = 1 \dots U$
 - $App(r)$: Apport en eau pour le réservoir $r = 1 \dots R$ (en MW)
 - $\omega(r)$: Valeur de l'eau pour le réservoir r (en €/MW)
 - $\underline{S}(r)$: Stock minimum du réservoir r (en MW)
 - $\overline{S}(r)$: Stock maximum du réservoir r (en MW)
 - $S_0(r)$: Stock initial du réservoir r (en MW)

- $ReUs(r, u) = \begin{cases} 1 & \text{si le réservoir } r \text{ est lié à l'usine } u \\ 0 & \text{sinon} \end{cases}$
- $Kh(u)$: Rendement de l'usine selon son mode de fonctionnement
avec $Kh(u) = \begin{cases} 1 & \text{si } u \text{ est une turbine} \\ -0,74 & \text{si } u \text{ est une pompe} \end{cases}$

- Variables

- $P_H(u, t)$: Production en puissance de l'usine u au pas de temps t
- $S(r, t)$: Stock dans le réservoir r au pas de temps t

- Contraintes

- Contrainte sur la production de puissance (4)
- Contraintes sur les stocks (5) et (6)

Remarques :

- Pour conserver la convexité du problème d'optimisation, nous n'avons volontairement pas pris en compte certaines contraintes comme la durée minimale de marche (respectivement d'arrêt) d'une centrale ou les coûts de démarrage qui apportent de la non convexité et des discontinuités.
- Dans la formulation du problème que nous venons d'énoncer, nous avons fait le choix de ne parler que de la puissance active et de laisser de côté la réserve primaire et de la réserve secondaire.

Formulation explicite

Le modèle déterministe s'écrit sous la forme suivante :

$$\begin{aligned}
\min \quad & \sum_{t=1}^T \sum_{c=1}^C C_{Th}(c) \cdot P_{Th}(c, t) - \sum_{r=1}^R S(r, T_{fin}) \cdot \omega(r) + \sum_{t=1}^T C_{def} \cdot (P_{def}(t) + P_{surp}(t)) \\
s.c. : \quad & \sum_{c=1}^C P_{Th}(c, t) + \sum_{u=1}^U sgn(Kh(u)) \cdot P_H(u, t) + P_{def}(t) - P_{surp}(t) = D(t) \\
& 0 \leq P_{Th}(c, t) \leq \bar{P}_{Th}(c) \\
& G_T^-(c) \leq P_{Th}(c, t+1) - P_{Th}(c, t) \leq G_T^+(c) \\
& 0 \leq P_H(u, t) \leq \bar{P}_H(u) \\
& \underline{S}(r) \leq Stockh(r, t) \leq \bar{S}(r) \\
& S(r, t) = S(r, t-1) + App(r) - \sum_{u=1}^U Kh(u) \cdot P_H(u, t) \cdot ReUs(r, u)
\end{aligned} \tag{6.1}$$

La formulation déterministe peut s'écrire de manière simplifiée. En posant :

- u : Unité de production (nucléaire, thermique, hydraulique ou unité fictive de défaillance/surproduction)
- P^u : Vecteur de décision pour l'unité u (puissance électrique), avec $P^u = (P_1^u, \dots, P_T^u)$, T étant le nombre de pas de temps
- \mathcal{X}^u : Ensemble de contraintes de l'unité u

Le problème **déterministe**, s'écrit alors :

$$\begin{aligned}
\min \quad & \sum_{u=1}^U C_u(P_u) \\
s.c. : \quad & \sum_{u=1}^U g_u(P_u) = 0 \in \mathbb{R}^m \\
& P_u \in \mathcal{X}_u \subset \mathbb{R}^n
\end{aligned}$$

L'objectif est de minimiser les "coûts" de production, qui sont de différentes natures. Ils correspondent aux coûts de fonctionnement des centrales thermiques et nucléaires, aux pénalités de défaillance/surproduction des centrales fictives de défaillance/surproduction, et enfin aux valeurs de l'eau (donc coûts négatifs) pour les différents réservoirs.

Ce problème considère deux types de contraintes. Les premières sont appelées *contraintes liantes*, couplant les productions des différentes unités de production (exemple : contrainte sur la demande). Les seconds types de contraintes sont spécifiques à chaque unité de production, appelées *contraintes de fonctionnement* (exemple : contrainte sur le gradient).

6.1.2 Modèle robuste

Le modèle robuste minimise les coûts de production ainsi que le pire surcoût lié à un écart entre la demande prévue et la demande réalisée. La fonction d'ajustement décrit dans ce chapitre la Vision Pénalités.

En reprenant l'écriture simplifiée du problème déterministe (\mathcal{P}_{det}), le problème robuste s'écrit :

$$(\Theta) : \min_{P \in \mathcal{X}} \{ \Gamma(P) + \Phi(P) \} \quad (6.2)$$

avec $P = (P_1, \dots, P_u, \dots, P_U)$. L'ensemble des programmes réalisables \mathcal{X} répond aux contraintes décrites pour le modèle déterministe :

$$P \in \mathcal{X} \iff \begin{cases} \sum_{u=1}^U g_u(P_u) & = 0 \in \mathbb{R}^m \\ P_u \in \mathcal{X}_u & \subset \mathbb{R}^n \end{cases} \quad (6.3)$$

Ainsi, $\Gamma(P)$ évalue le coût de production :

$$\Gamma(P) = \sum_{u=1}^U C_u(P_u) \quad (6.4)$$

Et la fonction $\Phi(P)$ évalue le pire surcoût dû à l'aléa dans la vision Pénalités (cf Chapitre 4) :

$$\Phi(P) = \max_{d \in \mathcal{D}} \sum_{t=1}^T (\alpha_t \cdot \xi_t^- + \beta_t \cdot \xi_t^+)$$

s.c :

$$\forall t = 1 \dots T$$

$$\begin{aligned} d_t - \xi_t^- + \xi_t^+ &= \sum_{u=1}^U P_u^t \\ \xi_t^- &\leq v_t^- \times M_t^- \\ \xi_t^+ &\leq (1 - v_t^-) \times M_t^+ \\ v_t^- &\in \{0, 1\} \end{aligned}$$

6.1.3 Développement

Un logiciel informatique a été développé pour résoudre les approches déterministe et robuste. Le problème linéaire du modèle déterministe est traité par appel au solveur Cplex 12. Le modèle robuste (comme décrit au chapitre 4) combine l'appel au solveur et l'algorithme proximal pour la résolution du problème. Le logiciel propose une modularité qui permet de résoudre soit le problème déterministe, soit le problème robuste (et ce pour toute écriture de la fonction d'ajustement).

6.1.4 Données générales

On travaille sur l'horizon journalier avec un pas demi-horaire, le nombre de période considéré est donc $T = 48$.

La prévision moyenne de la demande est celle estimée pour la journée du mardi 13 janvier 2009. Les données sur l'état du parc de production, les coûts de fonctionnement et les valeurs d'usage de l'eau ont été recueillis dans les différents fichiers paramètres en entrée du modèle déterministe APOGEE aujourd'hui utilisé à EDF.

6.1.5 Parc de production

On considère pour ce problème l'intégralité du parc de production thermique, c.à.d :

- 58 centrales nucléaires
- 31 centrales thermiques classiques (fioul, charbon, cycle combiné)

Le parc hydraulique est modélisé sous la forme agrégée de 3 vallées hydrauliques. Cette modélisation est utilisée à EDF en opérationnel pour la gestion de production à horizon Moyen-Terme (entre 1 et 5 années).

Pour des raisons de confidentialité, nous n'explicitons pas les valeurs numériques des données en entrée du modèle déterministe et robuste (ex : coûts de fonctionnement, valeurs d'usage de l'eau, pénalités d'écart calées sur les coûts de fonctionnement).

6.2 Simulations : Mode opératoire

Nous allons présenter dans cette section les hypothèses de travail et la méthode de validation choisie.

Le corps d'hypothèses pour les fonctions Γ et Φ est identique à celui introduit dans le chapitre 4. L'algorithme proximal est utilisé pour la résolution du problème général (cf chapitre 4).

6.2.1 Objectif

Nous cherchons à comparer le plan de production proposé par le modèle robuste P_{opt} , à celui satisfaisant la demande moyenne d_{moy} , qu'on appellera dans la suite P_{moy} . Cette comparaison s'effectuera sur un ensemble de réalisations futures de la demande obtenues à l'aide d'un générateur d'erreur de prévision.

13 000 scénarios d'erreur de prévision ont été générés pour servir à cette phase de simulation. Ces vecteurs de demande en puissance, notés d_r (demande en puissance réalisée) simuleront 13 000 *futurs possibles* sur lesquels nous testerons l'approche robuste en comparaison à l'approche déterministe classique.

Remarque :

Les 13 000 scénarios utilisés dans la phase de validation n'ont pas servi à la construction du polyèdre d'incertitude. Il n'y a donc aucune dépendance directe entre l'ensemble d'incertitude et l'ensemble de réalisations.

6.2.2 Valorisation des écarts en puissance

On calcule le coût du programme et le coût des pénalités à payer pour non respect de l'équilibre offre-demande ($P(t) = d_r(t)$) pour les programmes P_{opt} et P_{moy} . Il est à noter qu'on s'intéresse aux coûts sur tout l'horizon d'étude (48 pas de temps dans ce qui suit). Pour le tirage $n \in N$, cela revient à calculer :

$$\Theta^n(P_{moy}) = \Gamma(P_{moy}) + \phi(d_r^n, P_{moy})$$

$$\Theta^n(P_{opt}) = \Gamma(P_{opt}) + \phi(d_r^n, P_{opt})$$

Dans la suite, la comparaison des deux modèles sera menée sous deux angles différents.

Après une comparaison du coût de la solution robuste et de la solution moyenne, nous présenterons pour le premier angle de vue les courbes associées aux deux distributions $\Theta^n(P_{moy})$ et $\Theta^n(P_{opt})$. La **courbe bleue** représentera la distribution des coûts avec le

programme moyen P_{moy} et la **courbe rouge** la distribution des coûts avec le **programme robuste** P_{opt} .

Un second graphique illustrera la distribution des écarts, c'est à dire :

$$\Upsilon^n = \Theta^n(P_{opt}) - \Theta^n(P_{moy}) \quad (6.5)$$

Ces graphiques seront complétés par un tableau avec des indicateurs statistiques (à savoir la moyenne, la dispersion et le quantile à 95%) pour chacune des deux distributions $(\Theta^n(P_{opt}))_{n \in N}$ $(\Theta^n(P_{moy}))_{n \in N}$, ainsi que l'écart en pourcentage entre les deux solutions. Prenons l'exemple de la dispersion, nous indiquons **l'écart en pourcentage** entre la dispersion de la distribution $(\Theta^n(P_{opt}))_{n \in N}$ et la dispersion de la distribution $(\Theta^n(P_{moy}))_{n \in N}$.

Remarque :

Rappelons que le modèle à valider n'utilise pas de distribution de probabilités et ne minimise pas le coût en espérance. Néanmoins, pour valider les résultats obtenus, on compare les moyennes empiriques de chaque distribution de coûts.

La seconde approche consiste à s'intéresser aux différences entre le programme robuste et le programme obtenu sous une vision déterministe. Un graphique traçant l'écart en puissance entre les deux programmes $P_{opt} - P_{moy}$ sur l'horizon d'étude illustrera cet écart.

6.3 Distribution des coûts : impact des pénalités

Dans le cas d'une vision Pénalités, les valeurs attribuées aux facteurs α et β jouent un rôle significatif dans les caractéristiques du programme robuste obtenu.

Dans la suite, nous varions à la hausse et à la baisse l'écart entre les pénalités de sous-production et de sur-production dans le modèle robuste. Nous commenterons succinctement les résultats obtenus, et une analyse globale du comportement du modèle pour ces trois études de cas conclura ces expérimentations.

6.3.1 Pénalisation plus marquée pour la sous-production

Une attente de l'utilisateur du modèle robuste peut être de se prémunir de manière plus accentuée contre une défaillance de la production électrique, que pour un excès de production. Nous avons alors testé la qualité de la solution robuste dans le cas où les pénalités α de sous-production étaient plus élevées que les pénalités β de sur-production.

D'après l'étude sur les pénalités dans le chapitre 4, section 4.1, nous avons posé $\alpha \geq \beta + 2\gamma$. Le coût de production n'étant pas constant, γ avait pour valeur le **coût marginal du système** dans le cas d'une résolution déterministe.

Le coût marginal au pas de temps t est le coût induit par la production d'une unité supplémentaire de puissance (1 MW dans notre cas) au pas de temps t . On obtient le coût

marginal du système par résolution de problème dual dans le cas déterministe.

On pose la pénalité de sur-production comme égale au coût de fonctionnement d'une des unités les plus chères du parc de production.

Le programme robuste et le programme moyen génèrent les coûts suivants :

| Coûts | $\Gamma(\cdot)$ | $\Phi(\cdot)$ | Total |
|-----------|-----------------|---------------|------------|
| P_{moy} | 39 868 000 | 9 694 000 | 49 562 000 |
| P_{rob} | 40 931 000 | 5 815 000 | 46 746 000 |
| Ecart | + 2.66 % | - 40.01 % | - 5.62 % |

Le programme robuste admet *par construction* un coût total de gestion inférieur au coût de gestion du programme moyen, puisque P_{rob} est la solution optimale du problème minimisant le coût $\Gamma + \Phi$.

Néanmoins, il est intéressant pour une première lecture, de connaître l'écart entre la solution "actuelle" qu'est le programme P_{moy} , et celle que propose notre modèle robuste.

On constate dans ce cas de figure une réduction du coût de gestion de plus de 5%. Même si les coûts de production sont plus élevés (+2.66%) dans le cas robuste (surcoût lié à la démarche prudente du modèle), cet écart est **totalement résorbé** puisque le coût des pénalités est très réduit (réduction de 40%).

Cette première analyse montre la pertinence de la solution robuste comparée à l'approche déterministe, et ce pour répondre à une problématique de minimisation des coûts de production et des coûts induits par l'observation de l'aléa.

Néanmoins, il est important d'étudier le comportement de la solution robuste dans un contexte où les observations de l'aléa ont une forte variance. La distribution des coûts est illustrée par la figure 6.1 suivante :

L'écart $\Theta^n(P_{opt}) - \Theta^n(P_{moy})$ a pour distribution la courbe 6.2.

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------------------|---------------------|------------|-------------|
| $(\Theta^n(P_{moy}))_{n \in N}$ | 41 066 221 | 7 951 000 | 42 804 725 |
| $(\Theta^n(P_{opt}))_{n \in N}$ | 43 028 501 | 5 063 400 | 44 051 140 |
| Ecart | +4.778% | -36.317% | +2.912% |

Une première lecture de ces indicateurs permet de constater le surcoût du modèle robuste comparé au modèle initial. Le surcoût de la solution robuste comparée à la solution moyenne est logique, puisque le modèle robuste se prémunit contre les pires aléas. De plus, le premier indicateur du tableau signifie que le coût total de la gestion de production reste plus cher avec un modèle robuste. Enfin, le quantile à 95% et le graphique 6.2 appuient cette première conclusion.

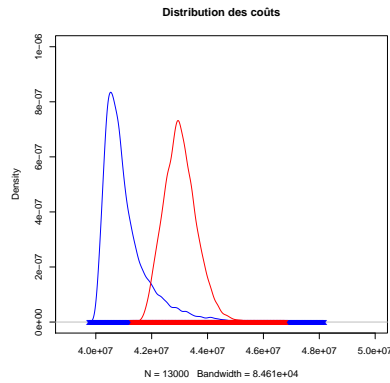


FIGURE 6.1 – Distribution des coûts pour le cas 6.3.1

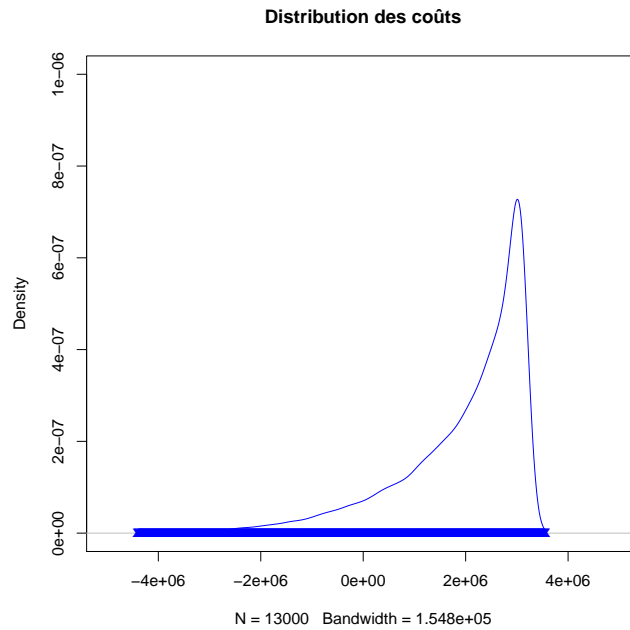


FIGURE 6.2 – Distribution des écarts de coûts pour le cas 6.3.1

Néanmoins, l'écart en dispersion indique une réduction très intéressante de la dispersion et ce grâce au modèle robuste. Cette amélioration indique une stabilité des coûts de gestion robustes comparé aux coûts moyens de gestion, et ce, face à 13000 réalisations possibles de l'aléa.

Pour compléter notre première analyse, nous observons le graphique illustrant l'écart de production :

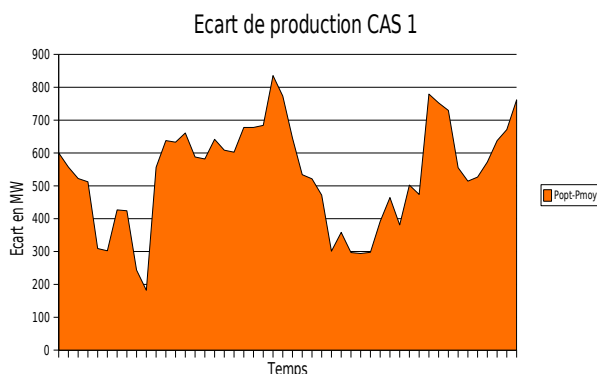


FIGURE 6.3 – Ecart de production pour le cas 6.3.1

La solution robuste implique un niveau de production sensiblement plus élevé que la solution moyenne, comme on le voit sur le graphique 6.3.

6.3.2 Pénalisation équivalente de la sous-production et sur-production

Une pénalisation équivalente de la défaillance et de la sur-production décrit le contexte industriel réel. En effet, le contexte réglementaire incite fortement le producteur électrique à s'équilibrer autant que possible, sans préférence aucune pour une situation de défaillance ou de sur-production. Le cas étudié ci-dessous est ainsi le plus proche du contexte réel. Une pénalisation équivalente de la défaillance et de la sur-production implique l'égalité suivante : $\alpha = \beta + 2\gamma$, γ étant le coût marginal du système.

Le programme robuste et le programme moyen génèrent les coûts suivants :

| Coûts | $\Gamma(\cdot)$ | $\Phi(\cdot)$ | Total |
|-----------|-----------------|---------------|------------|
| P_{moy} | 39 868 000 | 7 122 000 | 46 990 000 |
| P_{rob} | 40 456 000 | 5 288 000 | 46 219 000 |
| Ecart | + 1.45 % | - 25.75 % | - 1.64 % |

Dans ce cas de figure, l'écart des coûts de programmes est réduit de moitié (comparé à la section précédente). Le gain en coût de gestion est réduit de plus de 1.5%. La réduction importante des coûts liés à la pénalisation combinés à un surcoût faible des coûts de

production présentent un compromis intéressant entre conservatisme et stabilité.

Les courbes de distributions sont représentées en 6.4

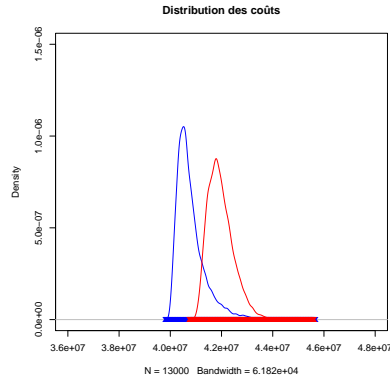


FIGURE 6.4 – Distribution des coûts pour le cas 6.3.2

L'écart entre la valeur de $\Theta^n(P_{moy})$ et $\Theta^n(P_{opt})$ a pour distribution la courbe suivante :

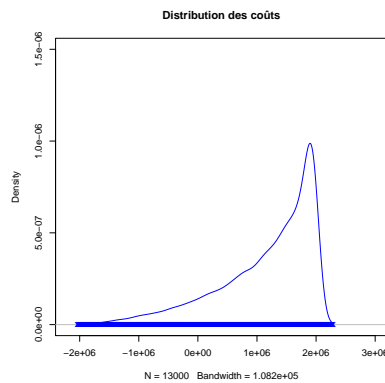


FIGURE 6.5 – Distribution des écarts de coûts pour le cas 6.3.2

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------------------|---------------------|------------|-------------|
| $(\Theta^n(P_{moy}))_{n \in N}$ | 40805041 | 5541200 | 41966610 |
| $(\Theta^n(P_{opt}))_{n \in N}$ | 41970330 | 4508600 | 42908540 |
| Ecart | +2.85 | -18.63495% | +2.24% |

Une pénalisation moins forte de la sous-production diminue la stabilité du programme robuste (comparée au cas précédent). Le surcoût de gestion moyen du modèle robuste com-

paré au modèle initial est quant à lui réduit.

Le graphique de l'écart de production entre les deux programmes explique en partie cet écart :

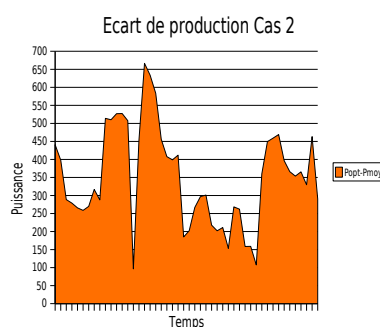


FIGURE 6.6 – Ecart de production pour le cas 6.3.2

En effet, la solution robuste propose dans ce cas un niveau de production moins élevé que dans l'exemple précédent. Ceci réduit en partie le surcoût de gestion.

6.3.3 Pénalisation plus marquée pour la sur-production

Dans ce dernier cas, La sur-production est plus pénalisée que la sous-production. On a posé comme égalité $\alpha = \beta$.

Le programme robuste et le programme moyen génèrent les coûts suivants :

| Coûts | $\Gamma(.)$ | $\Phi(.)$ | Total |
|-----------|-------------|-----------|------------|
| P_{moy} | 39 868 000 | 4 168 000 | 44 036 000 |
| P_{rob} | 39 757 000 | 4 152 000 | 43 909 000 |
| Ecart | - 0.28 % | - 0.38 % | - 0.29 % |

La solution robuste obtenue manque dans ce cas de pertinence. En effet, elle n'améliore que très sensiblement les coûts de gestion obtenus avec une approche déterministe. Ces résultats nous confirment l'intérêt discutable d'un modèle robuste *favorisant* une sous-production, et ce en pénalisant la sur-production plus fortement que la défaillance.

En effet, en réponse à cette politique, le modèle robuste propose, à première vue, un programme très proche du programme moyen. Les résultats qui suivent confirment cette première interprétation.

Les distributions de coûts sont les suivants :

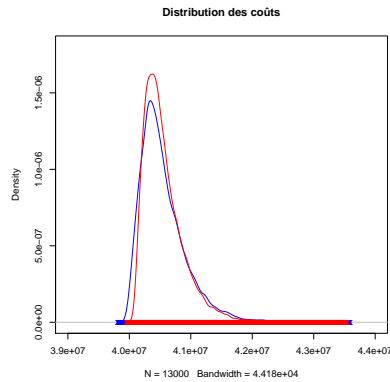


FIGURE 6.7 – Distribution des coûts pour le cas 6.3.3

L'écart entre la valeur de $\Theta^n(P_{moy})$ et $\Theta^n(P_{opt})$ a pour distribution la courbe suivante :

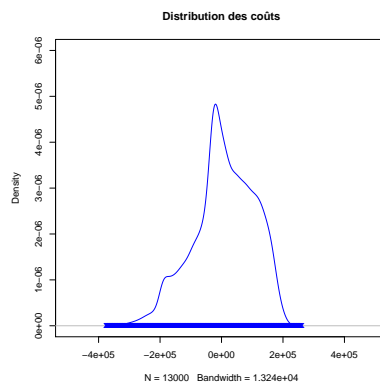


FIGURE 6.8 – Distribution des écarts de coûts pour le cas 6.3.3

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------|---------------------|------------|-------------|
| $\Theta^N(P_{moy})$ | 40 548 035 | 3525200 | 41260905 |
| $\Theta^N(P_{opt})$ | 40 556 189 | 3370400 | 41192615 |
| Ecart | +0.020 % | -4.391% | -0.165 % |

La dispersion du modèle déterministe initial est réduite lors du passage à une optimisation robuste. Néanmoins, parmi les trois cas étudiés jusque là, *le niveau d'amélioration* (environ 4.4%) obtenu est **le plus faible des trois cas d'étude**. Par contre, le surcoût moyen de gestion est réduit dans ce cas de figure.

Si on regarde de plus près l'écart en production, on obtient le graphique suivant :

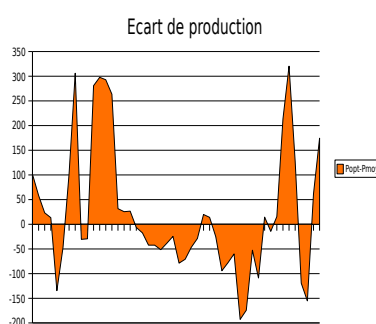


FIGURE 6.9 – Ecart de production pour le cas 6.3.3

Le programme robuste répond globalement au même volume de puissance (seulement 1000 MWh d'écart). Comme indiqué avec les premiers résultats, le modèle robuste paramétré de manière à pénaliser plus fortement tout sur-production tend à proposer *une solution moyenne*, avec une faible réduction de la stabilité, et des coûts moyens de gestion comparables à la solution déterministe.

6.3.4 Interprétation et analyse

Définissons certaines notations qui nous aiderons à résumer les résultats obtenus avec ces trois études de cas :

- $\mathcal{E}\S 1$: Cas où $\alpha \gg \beta + 2\gamma$ (sous-production beaucoup plus pénalisée que la sur-production)
- $\mathcal{E}\S 2$: Cas où $\alpha \simeq \beta + 2\gamma$ (pénalisation équivalente de la sous-production et sur-production)
- $\mathcal{E}\S 3$: Cas où $\alpha = \beta$ (sous-production moins pénalisée que la sur-production)

Voici un tableau récapitulant les écarts entre le modèle robuste et le modèle déterministe concernant les 3 principaux indicateurs. Chaque ligne rappellera l'écart entre la solution robuste et la solution déterministe :

| | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|-------------------|---------------------|------------|-------------|
| $\mathcal{E}\S 1$ | +4.778% | -36.317% | +2.912% |
| $\mathcal{E}\S 2$ | +2.85% | -18.63% | +2.24% |
| $\mathcal{E}\S 3$ | +0.020% | -4.391% | -0.165 % |

Posons Δ_i l'écart total de production entre le programme robuste et le programme moyen, et ce pour l'étude du cas i . $|\Delta_i|$ est la somme en valeur absolue de l'écart de production.

Rappelons que pour une solution robuste P_{rob} associée à un cas donné, $\Gamma(P_{rob})$ représente son coût du programme de production.

Les indicateurs relatifs aux écarts de production sont résumés ci-dessous :

| | Δ_i | $ \Delta_i $ | $\Gamma(\cdot)$ |
|-------------------|------------|--------------|-----------------|
| $\mathcal{E}\S 1$ | 25 674 MW | 25 674 MW | 4.093 e+07 |
| en % | +0.76 % | +0.76 % | +2.66 % |
| $\mathcal{E}\S 2$ | 16 794 MW | 16 794 MW | 4.049 e+07 |
| en % | +0.5 % | +0.5 % | +1.55 % |
| $\mathcal{E}\S 3$ | 1 022 MW | 4 549 MW | 3.97573e+07 |
| en % | +0.03 % | +0.13 % | -0.26 % |

Les principales conclusions que l'on peut tirer des résultats présentés sont les suivantes :

D'une part, l'utilisation d'une approche robuste répond aux besoins opérationnels **d'un programme plus stable** face aux différents aléas pouvant impacter la demande en puissance. En effet, nous avons observé la réduction (importante pour certains exemples) de la dispersion des coûts de gestion. Cette qualité est *intrinsèque* au modèle, même si elle peut être amplifiée ou réduite en fonction des termes de pénalisation.

D'autre part, il n'est pas surprenant de constater qu'une approche du type robuste admet **un coût sensiblement plus élevé** qu'une approche déterministe classique. *La pertinence du modèle robuste est visible dans l'étude du cas 2*, où on cherche à se prémunir de *manière équivalente* contre les pires scénarios de défaillance et de sur-production. *Le rapport* entre le surcoût de la solution et le gain en stabilité est alors *intéressant*. Les cas 1 et 3 décrivent des contextes mineurs, et l'intérêt d'une approche robuste pour ces cas est alors discutable.

Pour conclure, on ressent à travers cette étude comparative le rôle que jouent les pénalités de défaillance et de sur-production sur les résultats de l'approche robuste. L'utilisateur

de ce modèle, en fonction de ses besoins et de ses priorités pourra à travers ses paramètres, obtenir une solution robuste satisfaisant au mieux ses critères sur le coût moyen de gestion et sur la stabilité de la solution.

Chapitre 7

Le modèle robuste avec coûts de réorganisation

7.1 Introduction

Dans le processus infra-journalier, les programmes de marche des unités de production ne restent pas figés au cours de la journée. Si le programme de marche initial exécuté à 22h est issu d'une optimisation des placements de production, le programme est modifié à plusieurs reprises au cours des 24h qui suivent. Le programme initial étant calculé avec une prévision moyenne de la demande en puissance, toute déviation de la demande en puissance réalisée par rapport à celle prévue entraîne un changement des programmes, afin de satisfaire autant que possible l'équilibre *offre = demande*.

Néanmoins, ces modifications ont un coût, dû à la structure même du parc de production :

Les centrales nucléaires ont une production, certes peu coûteuse, mais assez rigide, qui ne permet pas de modifications à moins de 2h d'intervalle (l'arrêt/démarrage d'une centrale nécessite plusieurs heures).

Parmi les moyens de production les plus flexibles, on dispose des usines hydrauliques dont le programme de marche est modifiable rapidement. De ce fait, les usines hydrauliques sont très sollicitées lors de modifications dans les programmes de marche. Le système de turbinage et les niveaux d'eau suffisants dans les réservoirs permettent de modifier instantanément leur production si nécessaire. Actuellement, plus de 50% des modifications sont effectuées sur les programmes de marche des vallées hydrauliques.

Nous aborderons dans une première partie les écritures possibles d'une *fonction évaluant le pire coût de réorganisation* du programme de production. La complexité du problème nous amènera à étudier des approximations de la fonction initiale, et plus particulièrement une approximation convexe du pire coût de réorganisation.

Nous présenterons dans un second temps les résultats du modèle robuste avec cette vision approximative du coût de réorganisation. Enfin, nous étudions l'impact du polyèdre

d'incertitude sur la qualité de la solution robuste.

7.2 Coût de réorganisation : définitions & propriétés

Le modèle d'optimisation robuste défini dans le chapitre 4 propose comme solution un programme robuste face aux pires changements - en termes de coûts - de la demande.

En se plaçant dans le contexte infra-journalinier, *Le pire ajustement* correspond à une demande qui implique un nouveau programme de production avec *un coût de changement* de programme le plus élevé possible.

En effet, soit un programme de production P donné. Le pire ajustement possible correspond à une certaine demande d^* qui va impliquer de **changer le programme de production P en P'** , P' satisfaisant ainsi la demande d^* .

De plus, le programme P' *dépend* de "l'ancien programme" P . En effet, P' est calculé en fonction notamment de l'état *initial* du parc de production qui dépend de P .

Pour plus de clarté, nous noterons cette dépendance sous la forme : $P' \in \mathcal{I}_n(P)$, $\mathcal{I}_n(P)$ étant l'ensemble des programmes dont l'état initial est déterminé par le programme P .

La fonction d'ajustement pire cas a donc pour expression :

$$\Phi(P) = \max_{d \in \mathcal{D}} \phi(d, P)$$

On cherche la demande d^* qui implique *un coût de réorganisation le plus élevé possible*. $\phi(d, P)$ a donc pour valeur :

$$\begin{aligned} \phi(d, P) = \min & \quad \sum_{u=1}^U C_u(P'_u) \\ \text{s.c. :} & \quad \sum_{u=1}^U P'_u = d \in \mathbb{R}^m \\ & \quad P'_u \in \mathcal{X}_u \quad \subset \mathbb{R}^n \\ & \quad P' \in \mathcal{I}_n(P) \end{aligned}$$

où $C_u(P'_u)$ est le coût d'une production P'_u de l'unité u , d est une demande appartenant au polyèdre \mathcal{D} .

Respecter le nombre maximum de re-déclarations

Le calcul de la fonction d'ajustement pire cas $\Phi(P)$ telle que présentée ci-dessus est un problème difficile. A cette écriture s'ajoute une contrainte réglementaire très structurante appelée *la contrainte du nombre de re-déclarations*.

En effet, et comme abordé au chapitre 3, au cours du processus infra-journalier, le nouveau programme de production P' ne doit pas compter plus de N_{max} unités de production dont le nouveau programme est différent de celui décidé "initialement" $P_u \in P$. Actuellement, N_{max} est fixé à 30.

Pour exprimer de manière claire cette nouvelle contrainte, on fait appel à la fonction de **Dirac** de la manière suivante :

$$\delta : \quad \mathbb{R}^n \quad \longrightarrow \{0, 1\} \quad (7.1)$$

$$(x_i)_{1 \leq i \leq n} \longmapsto \begin{cases} 1 & \text{si } x_i = 0 \ \forall i = 1 \dots n \\ 0 & \text{sinon} \end{cases} \quad (7.2)$$

$$(7.3)$$

La fonction de Dirac δ a pour entrée *l'écart* entre les deux programmes de production d'une même unité. Elle permet de savoir si un écart existe entre ces deux programmes (sa valeur vaut alors 0), ou s'ils sont *exactement identiques* (c'est à dire pour toutes les unités et sur tous les pas de temps), elle vaut alors 1.

La fonction d'ajustement décrivant exactement la problématique de re-déclaration des programmes de production s'écrit alors :

$$\phi(d, P) = \min \sum_{u=1}^U C_u(P'_u)$$

$$s.c. : \sum_{u=1}^U P'_u = d \in \mathbb{R}^m \quad (7.4)$$

$$P'_u \in \mathcal{X}_u \subset \mathbb{R}^n$$

$$\sum_{u=1}^U (1 - \delta(P'_u - P_u)) \leq N_{max}$$

Ce problème peut également s'écrire sous la forme de fonctions d'ensemble de la manière suivante :

- Soit S un sous-ensemble de U . S correspond aux unités de production dont le programme sera modifié pour satisfaire la nouvelle demande. Ainsi, $|S| \leq N_{max}$
- la fonction $\Phi(P)$ peut donc s'écrire sous la forme suivante :

$$\max_{d \in \mathcal{D}} \min_{|S| \leq 30} \Lambda(S, d, P)$$

Λ étant le coût de réorganisation modifiant les unités appartenant à S , avec un programme de production initial P et une nouvelle demande à satisfaire d .

On remarque par ailleurs que si la cardinalité de S augmente, il est plus facile d'obtenir un nouveau plan de production à moindre coût, donc la valeur de Λ diminue. Ainsi, Λ est décroissante quand $|S|$ augmente.

Néanmoins, et quelque soit l'écriture de la fonction évaluant le coût de réorganisation du parc de production, nous sommes confrontés à un problème du type *Min-Max-Min complexe* dans sa résolution. Cette complexité est détectable dans l'écriture ensembliste du problème. En minimisant une collection finie de fonctions convexes polyédrales, nous obtenons *une fonction ni convexe ni concave* à maximiser, puis à minimiser.

La difficulté rencontrée pour la résolution de l'approche robuste avec une formulation exacte de la fonction d'ajustement (énoncée ci-dessus) nous a alors amené à étudier des approximations de la fonction initiale, pour obtenir un problème traitable.

7.3 Approximations de la fonction initiale

Deux pistes de travail étaient susceptibles d'aider à approcher la fonction calculant le coût de réorganisation.

La première approche calcule à l'aide d'un algorithme type glouton le coût de réorganisation pour une demande donnée et un programme de référence. Cette approche fut initiée dans le cadre des travaux de Ben-Abbes et Diemer dans [4] sur les algorithmes d'apprentissage supervisé pour approcher la fonction de coût de réorganisation. Les méthodes d'apprentissage supervisé furent étudiés avec intérêt, et proposent une alternative intéressante pour de futurs travaux.

La seconde piste de travail permettait d'écrire la fonction $\Phi(P)$ d'ajustement pire cas comme solution d'un programme linéaire en nombre entiers. Cette approximation fut source de résultats prometteurs, qui seront développés dans la section suivante. Nous introduisons dans la suite l'algorithme type glouton.

7.3.1 Algorithme glouton

Principe

Lors de la résolution d'un problème d'optimisation, la construction d'une solution se fait souvent de manière séquentielle, l'algorithme faisant à chaque étape un certain nombre de choix. Le principe glouton consiste à faire le choix qui semble le meilleur sur le moment (choix local), sans se préoccuper des conséquences dans l'avenir, et sans revenir en arrière.

Ainsi, un algorithme glouton consiste à trier les éléments d'un ensemble par valuation croissante, puis à choisir, à chaque pas, l'élément de plus grande valeur qui satisfait les contraintes.

Les algorithmes gloutons constituent une technique de résolution générale dans nombre de problèmes d'optimisation. Cette technique n'est pas assurée d'atteindre la solution optimale, sauf sous certaines conditions bien précises.

Soit E un ensemble fini d'éléments, v une valuation des éléments de E ($v : E \rightarrow \mathbf{R}^+$), et \mathcal{F} la famille des parties de E . L'algorithme glouton consiste à trouver un sous-ensemble $F \in \mathcal{F}$ tel que $\sum_{f \in F} v(f)$ soit maximal. Soit la formalisation suivante de l'algorithme glouton :

Algorithm 5 Algorithme glouton conceptuel

Classer les éléments de E par valuations décroissantes :

$$v(e_1) \geq v(e_2) \dots \geq v(e_n)$$

```

F = ∅
while i ≤ n do
  if F ∪ {ei} ∈ ℱ then
    F := F ∪ {ei}
  end if
end while
    
```

Pour les algorithmes décrits sous cette forme, il existe un théorème majeur d'optimalité (Edmonds 1971 [27]) :

L'algorithme glouton donne une solution optimale quelle que soit la fonction de valuation si et seulement si le système (E, \mathcal{F}) est un matroïde.

Matroïde :

Introduits par Whitney en 1935, les matroïdes vérifient les axiomes suivants :

- Trivial : $\emptyset \in \mathcal{F}$
- Héritéité : $\forall X \in \mathcal{F}, \forall Y \subset E, Y \subset X \Rightarrow Y \in \mathcal{F}$

– Augmentation : $\forall X, Y \in \mathcal{F}, |X| = |Y| + 1 \Rightarrow \exists x \in X, Y \cup x \in \mathcal{F}$

Application pour le calcul de $\Phi(P)$

Un algorithme glouton ne garantit donc pas nécessairement l'obtention d'une solution optimale, mais il fournit néanmoins un résultat rapidement. Même si la solution n'est pas optimale, il n'est pas rare que l'on s'en contente (par exemple pour un problème NP-difficile). Pour que la méthode gloutonne ait une chance de trouver l'optimum exact, il faut que le choix local aboutisse à un problème similaire plus petit.

Pour notre problème, on cherche un algorithme qui prend en entrée la production initialement prévue et le nouvel aléa sur la demande et rend un nouveau programme à appliquer à chaque centrale.

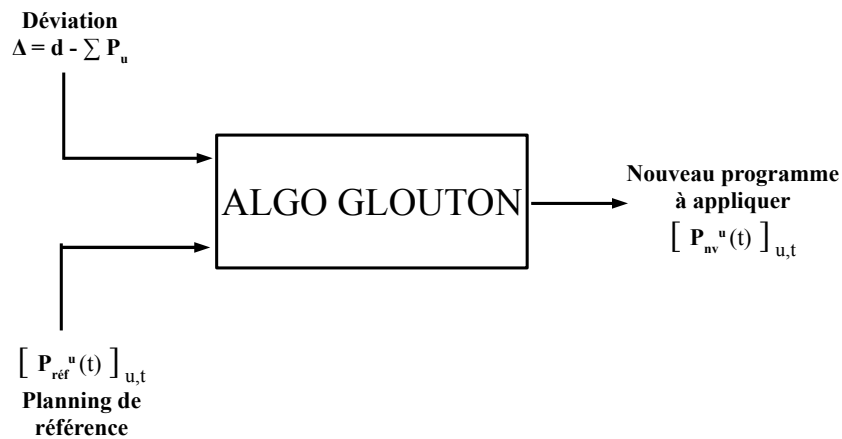


FIGURE 7.1 – Algorithme glouton

Notre approche gloutonne consiste à considérer, pour une demande d et un programme P , le potentiel de chaque unité de production.

Nous définissons le **potentiel** d'une unité de production comme *sa capacité à réduire l'écart* Δ entre la demande d et la production décidée par le programme P , et ce en *modifiant uniquement* le programme de l'unité u . Ce potentiel, correspondant à la fonction de valuation introduite dans le cas général, est défini de la façon suivante :

$$Q(u) = \min_{P'_u \in \mathcal{X}_u} C_u(P'_u) + C_{def}(P_{def})$$

$$s.c. P'_u + P_{def} = \Delta$$

où P_{def} est la production d'une unité de défaillance fictive au coût très élevé. L'unité de production dont le potentiel est le plus élevé voit son programme de production modifié (plus exactement remplacé par P'_u) et fixé.

A l'itération suivante, on s'intéresse au sous-problème formé par le reste des unités de production et de l'écart entre la demande d et le programme P dont on déduit l'écart qui vient d'être rattrapé.

Algorithme

L'algorithme s'énonce de manière plus claire sous la forme suivante :

Algorithm 6 Algorithme glouton

Initialisation :

$$\begin{aligned} \text{Nbredec} &\leftarrow 0; \\ \Delta &\leftarrow d - \sum_u P_u; \\ \mathcal{U} &\leftarrow \{Usines\}; \end{aligned}$$

while $\Delta \neq 0$ & $\text{Nbredec} \leq \text{Nbredec}_{max}$ **do**

$$\forall u \in \mathcal{U}, Q(u) = \min_{P_u} [C_u(P_{nv}^u) + C_{def}(P_{def})];$$

$$s.c. P_{nv}^u + P_{def} = d$$

Récupérer $u^* = \text{argmin}_{u \in \mathcal{U}} Q(u)$

$$\mathcal{U} \leftarrow \mathcal{U} \setminus \{u^*\}$$

$$\Delta \leftarrow \Delta - (P_{nv}(u^*) - P_{ref}(u^*))$$

$$\text{Nbredec} \leftarrow \text{Nbredec} + 1$$

end while

Retourner le nouveau programme

L'algorithme glouton décrit pour minimiser le coût de réorganisation du parc ne peut s'écrire sous la forme nécessaire à l'application de ce théorème.

En effet, réactualiser à chaque itération certains paramètres (ex : demande à satisfaire, coûts de production des centrales restantes) font que **les valuations** associées aux unités de production **ne sont pas fixes** pour toutes les itérations.

Nous n'avons donc pas de certitude quant à l'optimalité de cette approche.

Toutefois, il peut être intéressant de formuler une fonction de valuation pour les unités de production, tel que le choix des unités de production selon un ordre croissant/décroissant de leur valuations permet d'obtenir un plan de réorganisation optimal. Les résultats préliminaires obtenus avec cette approche restent finalement à consolider.

7.4 Approximation convexe

7.4.1 Evolution sur la modélisation du parc électrique

En utilisant la modélisation présentée dans le chapitre précédent, nous avons vite observé une limite due au parc hydraulique agrégé en uniquement trois vallées. En effet, quelque soit le scénario d'erreur qui se réalisait, cet aléa était comblé par la re-declaration d'une seule usine hydraulique, les moyens hydrauliques ayant un fort niveau de production d'après notre modélisation.

Pour la suite de l'analyse, nous avons supprimé les moyens hydrauliques de la modélisation actuelle, et ce jusqu'à validation des résultats. Il sera par la suite aisé de réintroduire ces moyens de production, mais de manière désagrégée, pour correspondre réellement à l'état du parc électrique d'EDF.

Nous avons donc un parc de production avec 89 unités de production (thermiques et nucléaires). La demande en puissance a été réduite à 90% de la demande en puissance de la journée de référence (pour rappel, journée du mardi 13 janvier 2009).

7.4.2 Etude expérimentale de la fonction φ

Dans un premier temps, nous avons essayé de représenter graphiquement le comportement du coût de réorganisation dans le cas où l'aléa qui est observé est nul sur tous les pas de temps sauf un seul.

De manière plus explicite, nous avons fait le choix d'un échantillon de 400 scénarios d'aléas de la forme $(0, \dots, x_t, \dots, 0)$ avec $x_t \succeq$, puis nous avons calculé le coût de réorganisation associé (le choix de l'indice t a été établi en fonction des pics de consommation durant la journée).

Cette expérimentation nous a permis d'avoir les résultats graphiques suivants :

Sous l'hypothèse que la fonction qui modélise le coût de réorganisation est séparable dans le temps, on peut observer sur ces graphiques qu'elle est très proche d'une fonction linéaire par morceaux, et d'allure convexe.

Le graphique 7.3 représente la pente de la courbe 7.2 en fonction de l'aléa pour le pas de temps $T = 39$.

Ce travail expérimental a permis d'étudier le comportement de la fonction de réorganisation du programme de production.

Dans la suite, le **pire coût de réorganisation** pour un ensemble d'aléas possible sera décrit comme solution d'un **programme linéaire en nombre entiers**.

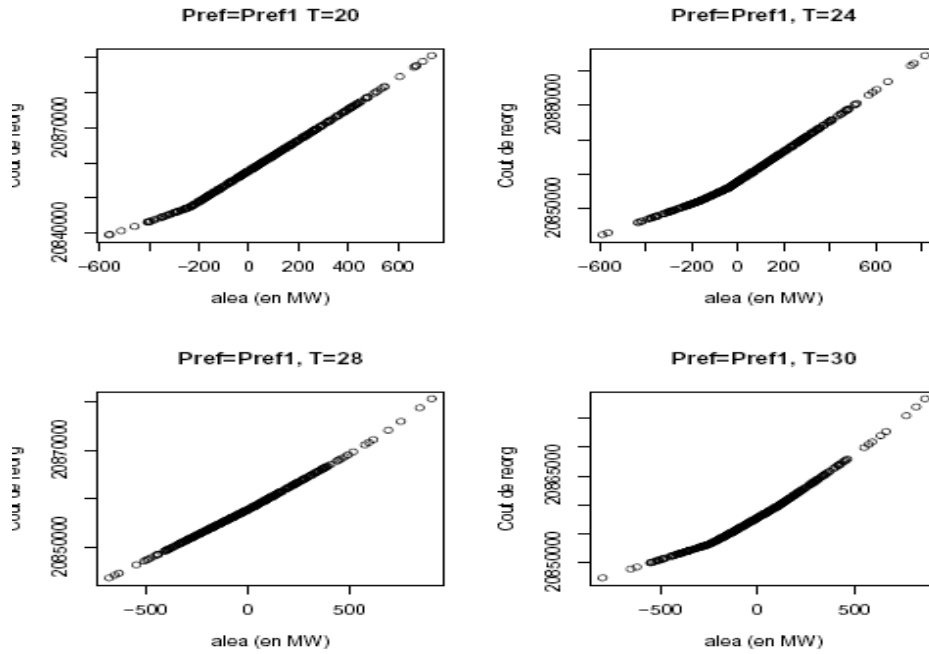


FIGURE 7.2 – Coût de réorganisation en fonction de l'aléa observé

Ecriture en PLNE de la fonction de réorganisation

Soient les notations suivantes :

Données

- K : Ensemble des segments sur lesquels ϕ est linéaire
- α_k^t : Pente de la fonction ϕ sur le segment k au pas de temps t
- β_k^t : Valeur à l'origine de ϕ sur le segment k au pas de temps t
- E_k^t : Borne gauche du k -ième intervalle

Variables

- d^t : Demande en puissance
- v_k^t : Variable binaire valant 1 si la solution d^* appartient au segment k
- u_k^t : Variable continue valant l'écart optimal entre la production P et la pire demande

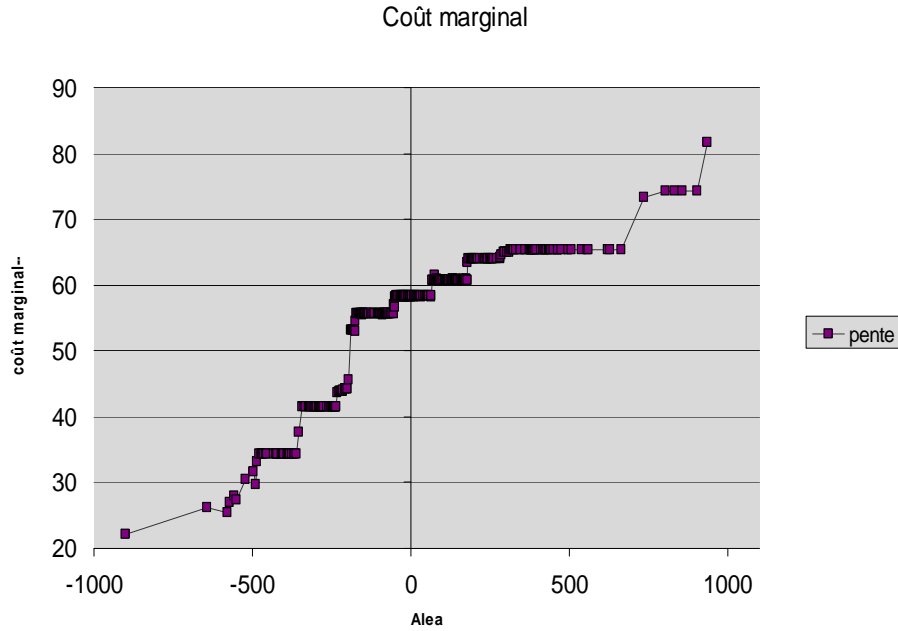


FIGURE 7.3 – Pente de la fonction coût pour T=39

$$\begin{aligned}\Phi(P) &= \max_{d \in \mathcal{D}} \phi(P, d) \\ &= \max_{d \in \mathcal{D}} \sum_{t=1}^T \sum_{k \in K} \alpha_k^t u_k^t + \beta_k^t\end{aligned}$$

s.c :

$$\begin{aligned}\forall t \in T, \forall k \in K & \quad v_k^t E_k^t + (1 - v_k^t) E_0^t \leq P^t - d^t \leq v_k^t E_{k+1}^t + (1 - v_k^t) E_k^t \\ \forall t = 1 \dots T, \forall k \in K & \quad v_k^t E_k^t \leq u_k^t \leq v_k^t E_{k+1}^t \\ \forall t = 1 \dots T, & \quad \sum_{k \in K} u_k^t = P^t - d^t \\ \forall t = 1 \dots T, & \quad \sum_{k \in K} v_k^t = 1\end{aligned}$$

Cette modélisation nous permet d'approcher le pire coût de réorganisation pour une production P donnée. On est face à un programme linéaire en nombre entiers, traitable de manière exacte avec un solveur commercial (tel que CPLEX), pour une taille de problème raisonnable (typiquement jusqu'à 48 pas de temps et K inférieur à 5).

Cependant, les contraintes de types "Big M" peuvent dégrader la qualité de la relaxation linéaire et rendre la résolution difficile dans le cas de problèmes en dimension plus élevée.

Implémentation

La phase d'optimisation et de simulation sont identiques au cas de la Vision Pénalités. La phase d'optimisation consiste à minimiser le coût de production et le pire coût de réorganisation de la production.

La phase de simulation consiste à évaluer le coût de la solution robuste et la solution déterministe face à 13 000 réalisations possibles de l'aléa.

La fonction de réorganisation comporte 3 paliers. La pente sur chaque intervalle a été fixée à l'aide des expérimentations décrites dans les sections précédentes.

7.4.3 Distributions des coûts

De manière similaire au chapitre 6, nous allons présenter les résultats du modèles robuste avec cette vision originale et concrète du pire surcoût. L'analyse se fera sous deux angles différents : le premier s'intéresse aux distributions de coûts associées à chaque solution (robuste et moyenne). Le second point de vue aide à comparer les programmes de production, et voir ce qui les distingue.

Les distributions de coûts $\Theta^n(P_{moy})$ et $\Theta^n(P_{opt})$ sont illustrés dans le graphique 7.4 suivant :

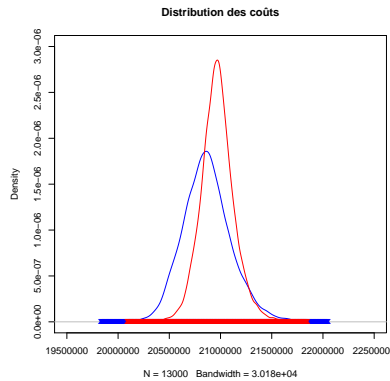


FIGURE 7.4 – Distribution des coûts

L'écart $\Theta^n(P_{opt}) - \Theta^n(P_{moy})$ a pour distribution la courbe suivante :

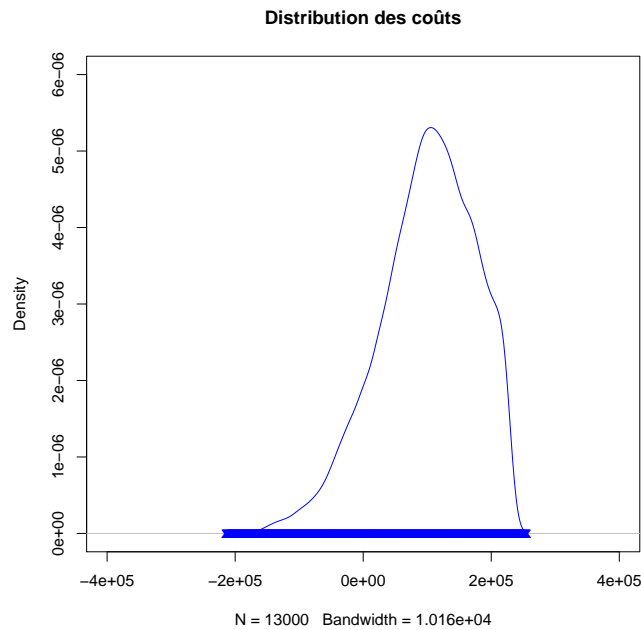


FIGURE 7.5 – Distribution des écarts de coûts

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------|---------------------|------------|-------------|
| $\Theta^N(P_{moy})$ | 20 870 414 | 2 044 900 | 21 282 005 |
| $\Theta^N(P_{opt})$ | 20 970 221 | 1 647 200 | 21 245 705 |
| Ecart | +0.478% | -19.448% | -0.170% |

La solution robuste proposée par le modèle apparaît ainsi intéressante. Pour un surcoût limité (moins de 0.5 % en plus), le programme robuste réduit la dispersion du coût de près de 20%. Une grande stabilité de la solution à laquelle s'ajoute une amélioration du coût dans les "pires cas", avec un quantile à 95% favorable à l'approche robuste.

Les programmes de production robustes et moyens sont comme suit :

L'écart total de production entre le programme robuste et le programme est de 31 GWH, soit en moyenne 1% d'écart par rapport au programme moyen. La solution robuste se prémunit ainsi contre les pires réalisations possibles avec un niveau de production proche du niveau de production moyen. De plus, l'écart de près de 1% en production engendre un surcoût moyen de seulement 0.48%.

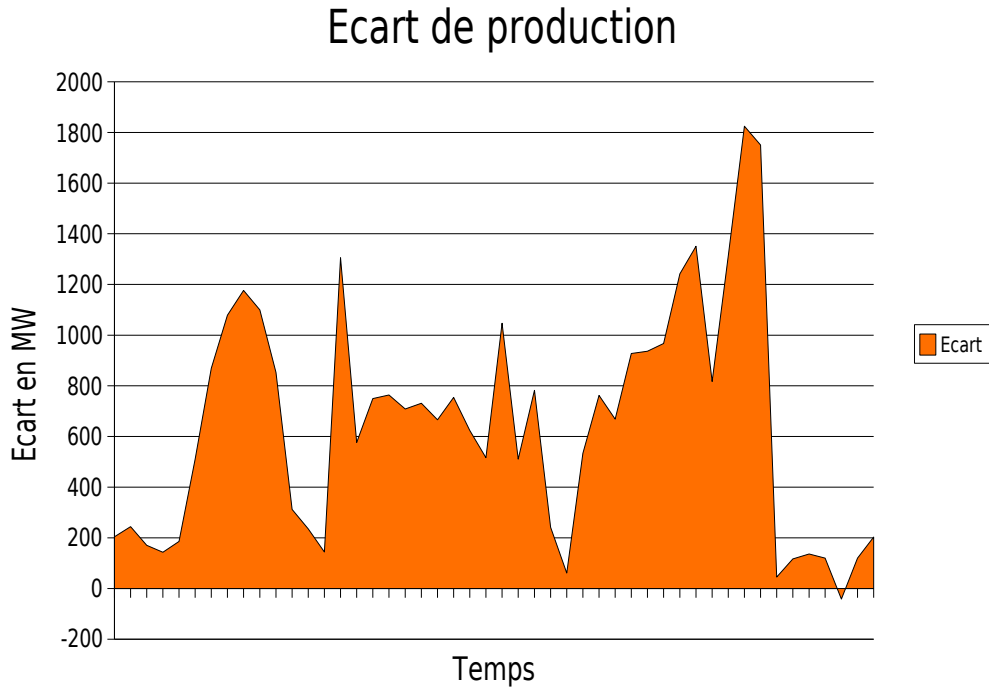


FIGURE 7.6 – Ecart de production $P_{opt} - P_{moy}$

7.5 Impact du polyèdre sur la distribution des coûts

7.5.1 Principe

Nous avons étudié l'impact du polyèdre d'incertitude sur les solutions robustes obtenues. Pour cela, nous avons comparé la solution robuste initiale présentée ci-dessus, à d'autres solutions robustes obtenues, et ce, après modification du polyèdre d'incertitude. En réduisant sensiblement cette enveloppe, la solution robuste est modifiée.

Pour cela, on a procédé à l'ajout de coupes particulières à notre polyèdre. Soit le polyèdre initial \mathcal{D} :

$$\mathcal{D} = \left\{ x \in \mathbb{R}^{48} \mid \sum_j a_{ij} x_{ij} \leq b_i, \forall i \in I \right\}$$

avec I l'ensemble des inéquations décrivant cet ensemble d'incertitude. Le polyèdre "réduit" $\mathcal{D}_k^{I_0}$ est alors défini de la manière suivante :

$$\mathcal{D}_k^{I_0} = \left\{ x \in \mathbb{R}^{48} \mid \sum_j a_{ij}x_j \leq \alpha_k b_i \forall i \in I_0 \text{ et } \sum_j a_{ij}x_j \leq b_i, \forall i \in I \setminus I_0 \right\}$$

où I_0 est un sous-ensemble de I . Les inéquations appartenant à I_0 ont leur second membre réduit par un facteur $\alpha_k \lesssim 1$.

7.5.2 Implémentation

On a effectué l'étude de sensibilité sur les coupes sur les bornes. Ces coupes concernent le sous-ensemble I_0 suivant :

$$I_0 = \{ \underline{e}_t \leq e_t \leq \bar{e}_t \}$$

Pour le polyèdre \mathcal{D}_k , cet ensemble d'inéquations s'écrit :

$$I_0^k = \{ \underline{e}_t \alpha_k \leq e_t \leq \bar{e}_t \alpha_k \}$$

Notre étude comparative s'est basée sur l'étude des solutions robustes obtenues avec trois polyèdres distincts $\mathcal{D}_1, \mathcal{D}_2$ et \mathcal{D}_3 , obtenus à partir du polyèdre initial \mathcal{D} et du facteur réducteur : $\alpha_k = \{0.95, 0.99, 0.995\}$ pour $k = \{1, 2, 3\}$.

Ainsi, ces ensembles d'incertitude vérifient les inclusions suivantes :

$$\mathcal{D}_1 \subset \mathcal{D}_2 \subset \mathcal{D}_3 \subset \mathcal{D}$$

Nous allons dans un premier temps présenter succinctement les courbes et indicateurs associés aux distributions de coûts obtenues avec chaque polyèdre d'incertitude. Ces résultats seront synthétisés et commentés dans un second temps, avec une étude des niveaux de production associé à chaque solution robuste.

7.5.3 Résultats

Résultats avec le polyèdre \mathcal{D}_1

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------|---------------------|------------|-------------|
| $\Theta^N(P_{moy})$ | 20 870 414 | 2 044 900 | 21 282 005 |
| $\Theta^N(P_{opt})$ | 20 835 941 | 1 931 200 | 21 213 205 |
| Ecart | -0.165 % | -5.56 % | -0.323% |

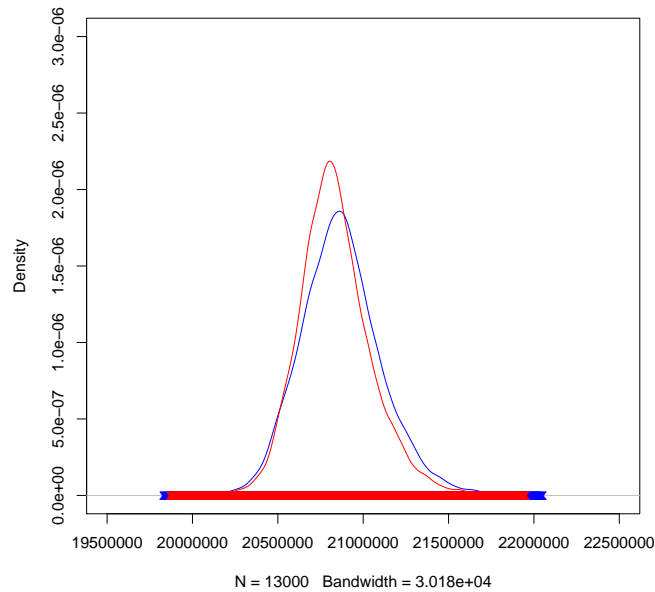


FIGURE 7.7 – Distribution des coûts

Résultats avec le polyèdre \mathcal{D}_2

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------|---------------------|------------|-------------|
| $\Theta^N(P_{moy})$ | 20 870 414 | 2 044 900 | 21 282 005 |
| $\Theta^N(P_{opt})$ | 20 843 144 | 1 860 500 | 21 192 210 |
| Ecart | -0.130 % | -9.017 % | -0.422 % |

Résultats avec le polyèdre \mathcal{D}_3

| Distribution | $\mathbb{E}(\cdot)$ | Dispersion | Quant à 95% |
|---------------------|---------------------|------------|-------------|
| $\Theta^N(P_{moy})$ | 20 870 414 | 2 044 900 | 21 282 005 |
| $\Theta^N(P_{opt})$ | 20 850 747 | 1 828 200 | 21 189 300 |
| Ecart | 0.0942 % | 10.597 % | 0.435 % |

Comparaison des indicateurs & programmes

Pour chaque polyèdre \mathcal{D}_k , on présente l'écart (positif ou négatif) de la solution robuste comparée à la solution moyenne. Les flèches décrivent la tendance de cet écart lors du passage de la solution initiale (avec polyèdre \mathcal{D}) à une solution obtenue avec le polyèdre \mathcal{D}_k .

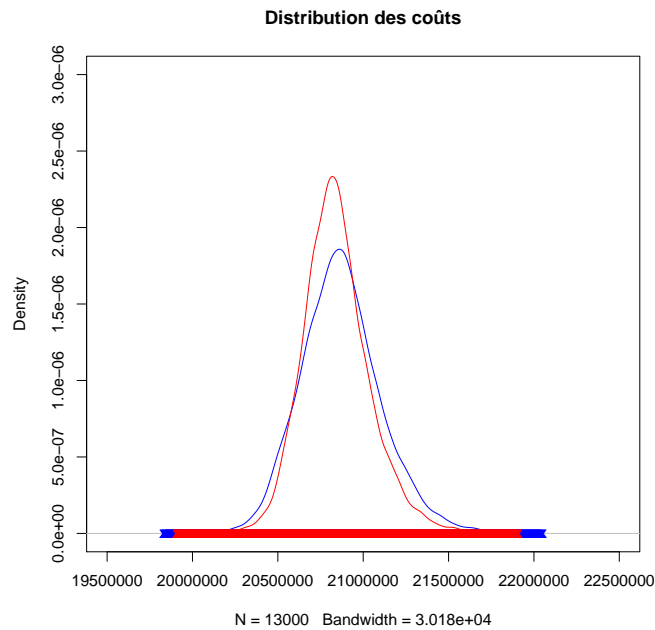


FIGURE 7.8 – Distribution des coûts

| Set | \mathcal{D}_1 | \subset | \mathcal{D}_2 | \subset | \mathcal{D}_3 | \subset | \mathcal{D} |
|---------------------|-----------------|-----------|-----------------|-----------|-----------------|-----------|---------------|
| Dispersion | +5.56% | | +9.01% | | +10.6% | | +19.45% |
| $\mathbb{E}(\cdot)$ | +0.16% | | +0.13% | | +0.09% | | -0.48% |

TABLE 7.1 – Comparatif des solutions robustes avec polyèdre variable

Ce que l'on remarque à l'aide du tableau ci-dessus est que d'une part la stabilité de la solution robuste augmente avec l'augmentation de l'ensemble d'incertitude en entrée du modèle.

D'autre part, le coût moyen du programme est amélioré avec l'approche robuste dans le cas de polyèdres réduits. La solution robuste devient en moyenne plus chère lorsque le polyèdre devient plus grand, mais ce surcoût reste toujours raisonnable.

Ces résultats valident une intuition que l'on peut avoir : Plus l'ensemble d'incertitude est grand, et plus le modèle robuste va proposer une solution optimale limitant le surcoût des pires réalisations.

Cette prévention est d'autant plus coûteuse que la région à couvrir est grande. En

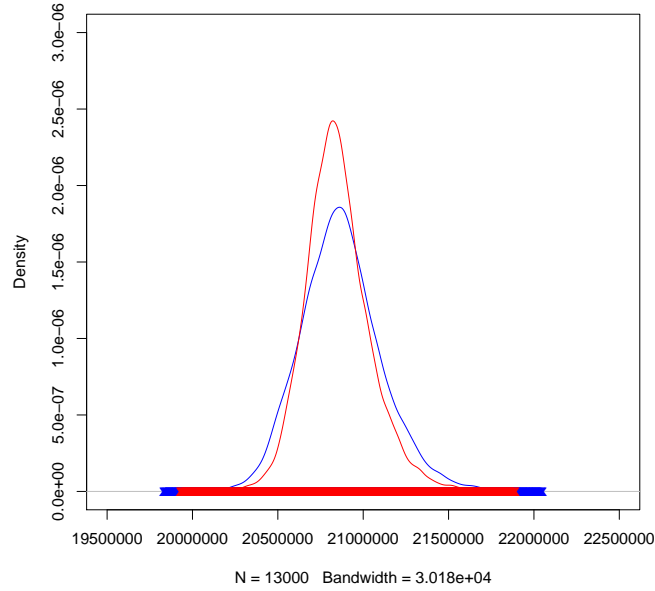


FIGURE 7.9 – Distribution des coûts

| Set | \mathcal{D}_1 | \subset | \mathcal{D}_2 | \subset | \mathcal{D}_3 | \subset | \mathcal{D} |
|---------------------------|-----------------|-----------|-----------------|-----------|-----------------|-----------|---------------|
| Ecart $P_{opt} - P_{moy}$ | 7 911,6 | | 13 334,9 | | 15 596,4 | | 31 056,8 |
| | ↗ | | ↗ | | ↗ | | |

TABLE 7.2 – Comparaison des écarts de production

contrepartie, la solution robuste améliore sa stabilité comparée à la solution moyenne. Le modèle "capte" les réalisations extrêmes qui peuvent se produire, ce qui affine sa solution avec un coût total de gestion moins dispersé que dans une vision moyenne.

Pour étayer ces propos, nous nous sommes intéressés aux programmes de production obtenus avec ces différents ensembles d'incertitude.

Le graphique 7.10 illustre l'écart de production entre les différents programmes robustes de production et le programme de production moyen. L'écart sur tout l'horizon d'étude entre les solutions robustes et la solution moyenne sont présentés dans le tableau ci-dessus.

De ces résultats, on observe des niveaux de production sensiblement plus importants lorsque le polyèdre est plus grand. Ce qui est néanmoins intéressant, c'est l'écart entre le polyèdre \mathcal{D}_3 et le polyèdre de référence. Le passage de près 31 GW à près de 15 GW d'écart

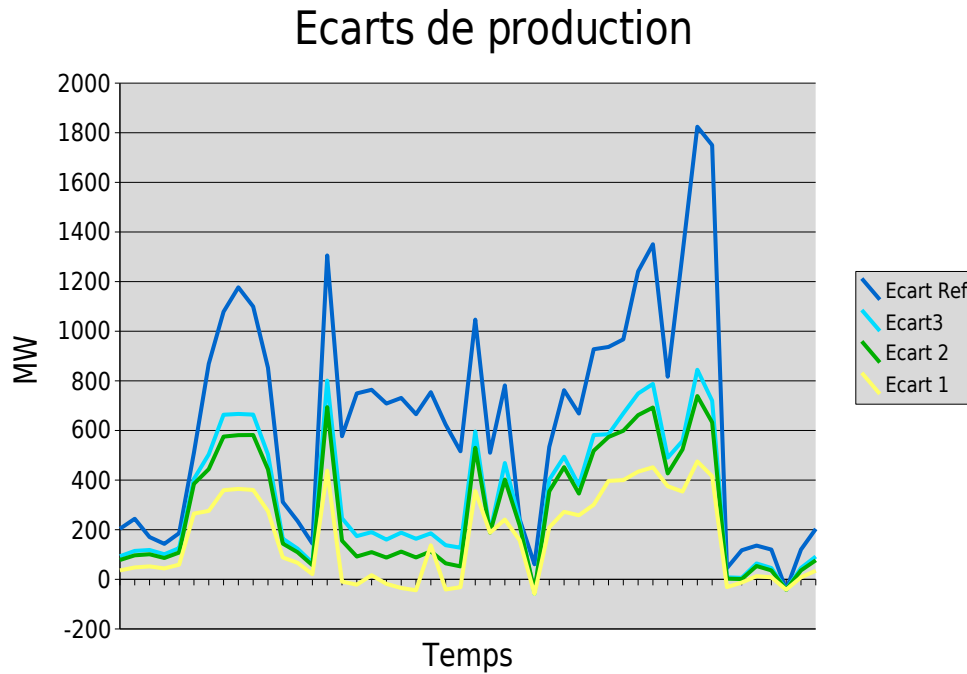


FIGURE 7.10 – Ecart de production entre solution robuste et solution moyenne

s'explique par une réduction de certains points extrêmes du polyèdre qui nécessitaient un niveau de production assez élevé (comparé à la solution moyenne).

Ce qui explique d'ailleurs que cet écart décroît beaucoup moins vite entre les polyèdres \mathcal{D}_3 et \mathcal{D}_2 et les polyèdres \mathcal{D}_2 et \mathcal{D}_1 .

En résumé, le modèle robuste associé à un paramétrage adéquat du polyèdre d'incertitude est capable de proposer des solutions robustes satisfaisant le degré de conservatisme souhaité, et le coût moyen que l'on est prêt à payer.

Conclusions & Perspectives

L'évolution des problèmes d'optimisation déterministes vers la prise en compte de facteurs aléatoires a donné lieu à de nombreux travaux.

Nous avons discuté dans le chapitre 1 de deux familles de problèmes sous incertitude, largement étudiés en contexte industriel. L'intérêt des modèles avec recours réside dans leur formulation, qui introduit la notion de *mesures correctives* que l'on peut apporter une fois l'aléa observé. De l'autre côté, on retrouve les modèles avec contraintes probabilistes, où la probabilité représente un *indicateur de la fiabilité* du système, ce qui rend ces modèles présents dans les problèmes de contrôle et de sûreté.

Nous avons abordé dans le chapitre 2 la programmation robuste. Les différents résultats ont été introduits avec une distinction faite entre les problèmes robustes linéaires et les problèmes robustes discrets. On s'est également intéressé aux modèles robustes dynamiques, et plus précisément ceux avec recours.

Les récentes avancées dans le domaine de l'optimisation robuste ont permis de donner un aperçu des possibilités ouvertes par une telle approche, en fonction de la modélisation de l'incertitude et de la structure du problème nominal. Ainsi, un choix pertinent du modèle d'incertitude combiné à un problème nominal traitable permet d'obtenir une version robuste traitable avec des résultats intéressants.

Le chapitre 3 a permis d'introduire le problème de gestion de production électrique. Dans un marché électrique concurrentiel, EDF a adapté ses outils de gestion de production pour permettre une gestion optimale de son portefeuille, et particulièrement sur les horizons journalier et infra-journalier, derniers leviers pour une gestion optimisée de la production. Et plus l'horizon d'optimisation s'approche du temps réel, plus les décisions prises aux instants précédents deviennent structurantes voire limitantes en terme d'actions, ces décisions étant prises *sans tenir compte du caractère aléatoire de certaines entrées du modèle*.

Pour se prémunir face à ces aléas, *des techniques d'optimisation robuste* tenant compte du contexte incertain ont fait l'objet de ces travaux de thèse.

Dans le chapitre 4, nous avons ainsi proposé *un modèle robuste de placement de la production tenant compte des incertitudes sur la demande en puissance*.

Cette approche permet de se protéger contre le pire scénario de demande dans le domaine d'incertitude considéré. La modélisation peut se décliner selon deux visions, toutes deux intéressantes sur le plan industriel, la vision *Marchés* et la vision *Pénalités*. Après *étude de la régularité* des fonctions à optimiser, nous avons étudié le comportement du modèle dans le cas d'exemples en faible dimension, avant de proposer une approche de résolution dans le cas général, et ce en employant l'algorithme proximal.

Nous avons abordé dans le chapitre 5 *la construction de notre ensemble d'incertitude*. Le polyèdre est construit autour d'un grand nombre d'observations, et les contraintes linéaires qui le délimitent sont basées sur des indicateurs variés (gradients, axes d'Acp, etc.).

La troisième partie de cette thèse est consacrée aux différents résultats obtenus avec le modèle robuste.

D'une part, l'utilisation d'une approche robuste répond aux besoins opérationnels d'*un programme plus stable face aux différents aléas* pouvant impacter la demande en puissance. En effet, nous avons observé la *réduction importante de la dispersion* des coûts de gestion. Cette qualité est *intrinsèque* au modèle, même si elle peut être amplifiée ou réduite en fonction des valeurs attribuées aux termes de pénalisation.

D'autre part, il n'est pas surprenant de constater que les solutions optimales du modèle robuste conduisent à des coûts plus élevés qu'une approche déterministe classique : il y a toujours un coût à payer pour se prémunir de l'incertitude. Nos résultats de calcul montrent cependant que ce surcoût reste faible.

L'étude spécifique de la notion de coût de réorganisation des programmes de production a permis de présenter une application concrète du modèle robuste, avec des solutions stables et performantes.

L'analyse de sensibilité de notre modèle par rapport aux paramètres définissant le polyèdre d'incertitude a d'ailleurs permis de valider l'idée suivante :

Le modèle robuste, associé à un paramétrage adéquat du polyèdre d'incertitude, est capable de proposer des solutions robustes satisfaisant le degré de conservatisme souhaité, dans la limite d'un coût moyen que l'on est prêt à payer.

Le travail mené dans le cadre de cette thèse a permis d'ouvrir la voie sur de nombreuses pistes de réflexion qui, faute de temps, n'ont pas été complètement poursuivies. Néanmoins, nous les énonçons ci-dessous comme perspectives intéressantes faisant suite à ce travail de recherche.

Concernant la formulation robuste, les travaux qui nous semblent les plus intéressants sont l'extension de cette approche robuste au *cadre non-convexe*. En effet, le contexte opérationnel à court-terme nécessite une description fine du problème, impliquant ainsi des

coûts de production non convexes, des variables binaires pour décrire les arrêts/démarrages des centrales, tout cela sous un ensemble de solutions réalisables non connexe.

Il serait d'ailleurs pertinent d'étudier la possibilité d'"*adapter*" le modèle robuste pour une classe de problèmes traitables par décomposition, comme c'est le cas actuellement avec l'outil APOGEE.

Par ailleurs, les résultats encourageants du modèle avec coût de réorganisation plaident pour l'étude d'autres classes de fonctions de coût. Ces dernières pouvant être traitées par l'approche proposée, tout en présentant un caractère plus général que celui développé dans le chapitre 7 (par exemple en remettant en cause la séparabilité).

Nous avons également exploré la voie d'un modèle robuste avec *une vision Marchés*, mais où *les prix de marchés étaient également sujets à de l'incertitude*. Cette voie intéressante et très proche du contexte industriel semble difficile à résoudre à premier abord. Néanmoins, les enjeux économiques et industriels derrière un tel modèle sont très importants.

Enfin, nous portons un intérêt certain aux travaux sur les modèles robustes dynamiques, et plus particulièrement aux règles de décisions linéaires appliquées aux problèmes robustes (cf. [8],[12],[36]). L'insertion du modèle robuste dans un processus décisionnel dynamique est une approche très intéressante, pouvant proposer dans l'avenir *une étude comparative rigoureuse entre les modèles robustes et les modèles stochastiques avec recours*.

Bibliographie

- [1] L. Andrieu, R. Henrion *Optimisation sous contraintes en probabilité : modèles, propriétés, stabilité et algorithmes* EDF R&D, H-R36-2010-00059-FR, mars 2010
- [2] R. Apparigliato *Application de la simulation avec recours pour la gestion de risque aux horizons hebdomadaire et mensuel*. Thèse (2008), Ecole Polytechnique.
- [3] R. E. Bellman *Dynamic Programming*. Princeton University Press, 2010
- [4] A. Ben Abbes , S. Diemer *Utilisation des techniques d'apprentissage automatique pour la gestion de production court-terme* EDF R&D, H-R36-2009-01914-FR, juillet 2009
- [5] A. Ben-Tal and A. Nemirovski *Robust solutions to uncertain programs*. Mathematical Operations Research (1998), 23 :769-805
- [6] A. Ben-Tal and A. Nemirovski *Robust solutions of uncertain linear programs*. Operations Research Letters (1999), 25 :1-13
- [7] A. Ben-Tal and A. Nemirovski *Robust solutions of linear programming problems contaminated with uncertain data*. Mathematical Programming (2000), 88 :411-424
- [8] A. Ben-Tal, A. Goryashki, E. Guslitzer, A. Nemirovski *Adjustable robust solutions of uncertain linear programs*. Mathematical Programming (2004), 99 :351-376
- [9] J.F. Benders *Partitioning procedures for solving mixed-variables programming problems*. Numerische Mathematik,(1962), 4 :238-252
- [10] D. Bertsimas and M. Sim *The Price of Robustness*. Operations Research (2004), 52 :35-53
- [11] D. Bertsimas and M. Sim *Robust Discrete Optimization and Network Flows*. Mathematical Programming (2004), 98 :49-71
- [12] D. Bertsimas and C. Caramanis *Finite adaptability in linear optimization*. The university of Texas at Austin (2007) available in : <http://users.ece.utexas.edu/cmcararam/pubs/>
- [13] BIBLOS *Processus et méthodes pour la gestion d'un portefeuille d'actifs dans les marchés de l'électricité* Manuel du département OSIRIS, Décembre 2005
- [14] J. R. Birge and F. Louveaux *Introduction to stochastic programming*. Springer Series in Operations Research, 1997
- [15] J. R. Birge and F. Louveaux. *A multicut algorithm for two-stage stochastic linear programs* European Journal of Operational Research, (1988), 34(3) :384-392

- [16] J.F. Bonnans, J.C. Gilbert, C. Lemaréchal, C. Sagastizabal *Optimisation numérique. aspects théoriques et pratiques* Springer-Verlag, 1997
- [17] C. Caroe and J. Tind *L-shaped decomposition of two-stage stochastic programs with integer recourse*. Mathematical Programming, (1998), 83 :451-464
- [18] P. Carpentier, G. Cohen, J.C. Culioli and A. Renaud *Stochastic optimization of unit commitment : a new decomposition framework*. IEEE Transactions on Power Systems, (1996), 11 :1067-1073
- [19] A. Charnes, W.W. Cooper, and G.H. Symonds. *Cost horizons and certainty equivalents : an approach to stochastic programming of heating oil*. Management Science (1958), 4 :235-263
- [20] G.B. Dantzig *Linear programming under uncertainty*. Management Science, (1955), 1 : 197-206
- [21] D. Dentcheva, A. Prékopa and A. Ruszczyński *On convex probabilistic programming with discrete distributions*. Nonlinear Analysis 47 (2001) 1997-2009
- [22] D. Dentcheva and W. Römisch *Optimal power generation under uncertainty via stochastic programming*. Economics and Mathematical Systems (1997) 22-56
- [23] J.C. Dodu, T. Eve, M. Minoux *Implementation of a proximal algorithm for linearly constrained nonsmooth optimization problems and computational results*. Numerical Algorithms (1994), 6 :245-273
- [24] G. Doukopoulos, J. Quenu *APOGENE : modélisation de l'hydraulique en discret* EDF R&D, H-R36-2009-01392-FR, juillet 2009
- [25] L. Dubost, R. Gonzalez and C. Lemaréchal *A primal-proximal heuristic applied to the French Unit-commitment problem* Mathematical Programming Ser A, (2005), 104 : 129-151
- [26] L. Dubost, R. Gonzalez, C. Lemaréchal, C. Sagastizabal *A primal-proximal heuristic applied to the French Unit-commitment problem* Mathematical Programming, Series A (2005), (104) :129-151
- [27] J. Edmonds *Matroids and the greedy algorithm* Mathematical Programming 1 (1971) : 127 - 136
- [28] L. El-Gahoui and H. Lebret *Robust solutions to least square problems to uncertain data matrices*. SIAM J. Matrix Analysis Applications (1997), 23 :1035-1064
- [29] L. El-Gahoui, F. Oustry and H. Lebret *Robust solutions to uncertain semidefinite programs*. SIAM J. Optimization (1998), 9 :33-52
- [30] P. Grente *Construction d'un intervalle de confiance autour de la prévision de consommation électrique* EDF R&D, Rapport de projet de fin d'études, septembre 2009
- [31] G. Iyengar *Robust Dynamic Programming* Mathematics of Operations Research (2005), 30(2) : 257-280
- [32] P. Kall and S.W. Wallace. *Stochastic programming*. New York, (1997), NY : Wiley

- [33] P. Kouvelis and G. Yu *Robust discrete optimization and its applications*. Kluwer Academic Publishers (1997), Boston
- [34] G. Laporte and F.V. Louveaux. *The integer L-shaped method for stochastic integer programs*. Operations Research Letters, (1993), 13 :48-62
- [35] C. Lemaréchal, F. Pellegrino, A. Renaud and C. Sagastizabal *Bundle Methods Applied to the Unit-Commitment Problem* System Modelling and Optimization, Chapman and Hall, (1996), 395-402
- [36] C. Lemaréchal, A. Ouorou, G. Petrou *Robust network design in telecommunications under polytope demand uncertainty* European Journal of Operational Research, (2010), 10.1016
- [37] J. Louyrette, V. Lavier *Synthèse sur le Mécanisme d'Ajustement Français - Fonctionnement et Perspectives* EDF R&D, H-R19-2007-03261-FR, mars 2009
- [38] M. Minoux *Programmation mathématique. Théorie et algorithmes*. Lavoisier (2008)
- [39] M. Minoux *Robust linear programming with right-hand-side uncertainty, duality and applications*. Encyclopedia of Optimization (2009), 3317-3327
- [40] M. Minoux *On robust maximum flow with polyhedral uncertainty sets*. Optimization Letters (2009), 3 :367-376
- [41] M. Minoux *Robust network optimization under polyhedral demand uncertainty is NP-hard*. Discrete Applied Mathematics (2010), 158 :597-603
- [42] M. Minoux *Solving some Multistage Robust Decision Problems with huge implicitly-defined Scenario Trees*. Algorithmic Operations Research (2009) 4, 1 :1-18
- [43] A. Prékopa *On probabilistic constrained programming*. Princeton University Press (1970), 113-138
- [44] A. Prékopa *Stochastic programming*. Kluwer, Dordrecht, Boston (1995)
- [45] A. Renaud *Algorithmes de régularisation et de décomposition pour les problèmes variationnels monotones*. Thèse (1993), Ecole des Mines de Paris
- [46] A. Renaud *Daily generation management at Electricité de France : from planning towards real time*. IEEE Transactions on Automatic Control (1993) 38, 7 :1080-1093
- [47] A. Ruszczyński and A. Shapiro *Probabilistic programming*. Computers and chemicals Engineering (2004), 28 :971-983
- [48] N. V. Sahinidis. *Optimization under uncertainty : state-of-the-art and opportunities*. Handbooks in Operations Research and Management Science, 10. North-Holland, 2003.
- [49] R. L. Smith *Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions* University of Michigan, (1982), Technical report : 81-1
- [50] A. L. Soyster *Convex programming with set-inclusive constraints and applications to inexact linear programming*. Operations Research (1973), 21 :1154-1157
- [51] S. Takriti, J.R. Birge and E. Long *A stochastic model for the unit commitment problem*. IEEE Transactions on Power Systems (1996), 11 :1497-1508

- [52] R. Van Slyke and R. J. B. Wets. *L-Shaped linear programs with applications to optimal control and stochastic programming*. SIAM Journal on Applied Mathematics, (1969), 17 :638663
- [53] M. Zhang, Y. Guan *Two-stage robust Unit Commitment problem* Optimization Online, Octobre 2009
- [54] R. Zorgati, P. Thomas, W. Van Ackooij, J.P. Argaud *Projet Cible Court-terme à l'horizon 2008. Lot 8.1* EDF R&D, H-R37-2005-013-A, novembre 2005
- [55] <http://www.r-project.org/> *Langage R*
- [56] <http://www.epexspot.com> *European Power EXchange*