**Résumé**

***Optimisation sous incertitudes par métamodèle pour la planification à court terme de la production conjointe d’une chaîne hydroélectrique et d’actifs de production variable.***

La planification à court terme de la production conjointe d'une chaîne hydroélectrique et d'actifs de production variable est un problème d'optimisation. Le programme de production hydroélectrique est construit à un horizon de quelques jours pour respecter les contraintes d'exploitation et pour maximiser le chiffre d'affaires obtenu par la vente sur le marché *day-ahead* et par la pénalisation des écarts, en considérant la production totale de tous les actifs.

Le travail de la thèse consiste à prendre en compte les incertitudes qui entachent les prévisions des apports en eau, des prix de l'électricité et des productions variables dans le problème d'optimisation. À partir d'une modélisation des incertitudes par un ensemble fini de scénarios multivariés, le problème d'optimisation en univers probabiliste proposé consiste en un problème de programmation dynamique stochastique à deux niveaux linéaires mixtes.

La problématique des temps de calcul étant cruciale pour un usage opérationnel, le problème d'optimisation probabiliste est résolu numériquement en remplaçant la valeur optimale du second niveau par un métamodèle calibré par apprentissage supervisé durant une phase de pré-traitement. Pour calibrer ce métamodèle, les données d'entrée sont simplifiées par des approches spécifiques. Un échantillonnage par analogie permet d'abord d'approcher le domaine de la donnée d'entrée associée aux variables de décision du premier niveau. Les données d'entrée fonctionnelles sont ensuite réduites par analyse en composantes principales. Un plan d'expérience peut ainsi être construit par échantillonnage par hypercube latin pour former le jeu de données nécessaire à l'apprentissage du métamodèle. Plusieurs métamodèles linéaires sont ensuite proposés.

La méthodologie proposée est testée sur un cas d'étude réel simplifié. Le métamodèle obtenu par régression linéaire sur les données d'entrée réduites donne des performances acceptables et il permet d'obtenir une solution du problème d'optimisation en univers probabiliste. Néanmoins, en l'absence de productions variables dans le cas d'étude, le problème d'optimisation en univers probabiliste ne permet pas d'apporter des gains significatifs par rapport à celui en univers déterministe. En outre, les temps de calcul ne permettent pas une utilisation en opérationnel sans calcul distribué. Plusieurs pistes de recherche sont toutefois proposées pour améliorer la méthodologie. Une validation sur plusieurs cas d'étude, voire sur un horizon roulant, permettrait d'estimer les performances de la méthodologie de manière plus robuste.

**Mots-clés :** programmation linéaire mixte, programmation stochastique à deux niveaux, programmation dynamique stochastique, prévision d’ensemble, prévision par analogie, réduction de données fonctionnelles, réduction de scénarios, plan d’expérience, régression linéaire.

**Abstract**

***Optimization under uncertainty with a surrogate model for the short-term production planning of a hydropower cascade combined with variable energy sources.***

Short-term production planning of a hydropower cascade combined with variable energy sources is an optimization problem. The hydropower production plan is constructed within a time horizon of a few days to comply with operational requirements and to maximize revenues obtained by selling the total production from all energy sources on the day-ahead market and by penalizing imbalances.

This PhD thesis focuses on the optimization problem considering data uncertainty on water inflows, electricity prices and variable generation, modeled by a finite set of multivariate scenarios. The proposed optimization problem in the stochastic framework is then written with a two-stage stochastic dynamic mixed-integer linear programming formulation.

Since computation time is a crucial issue for an operational use, the optimization problem is solved by replacing the value function of the second stage with a surrogate model fitted by supervised learning during a pre-processing step. The surrogate model is fitted on input data that are simplified by specific approaches. The domain of the input data related to the decision variables of the first stage is first estimated by analog-based sampling, and the functional inputs are reduced by principal components analysis. The learning data set can then be obtained from a design of experiments using latin hypercube sampling. Several linear models are finally proposed for the surrogate model.

The proposed methodology is tested on a real but simplified case study. The surrogate model obtained by linear regression on the reduced inputs provides acceptable performance, and it can be used to get a solution to the optimization problem in the stochastic framework. Since variable generation is not considered in the case study, the stochastic framework provides, however, minor profit compared to the deterministic framework. Besides, computation time is not compatible with an operational use without distributed computing. Several research tracks are then proposed to improve the methodology. A validation procedure on several case studies or on a rolling horizon could be used to assess the performance of the methodology in a more robust way.

**Keywords:** mixed-integer linear programming, two-stage stochastic programming, stochastic dynamic programming, ensemble forecasting, analog forecasting, functional reduction, scenario reduction, design of experiments, linear regression.