

REMPLACEMENT DU RECALAGE DE SIMULATIONS PAR DE L'AGRÉGATION POUR LA PRÉVISION D'HYDROCARBURES

Raphaël Deswarte¹

¹ *raphael.deswarte@cmap.polytechnique.fr*

Résumé. Pour les entreprises pétrolières, la capacité de prévoir les quantités et qualité des hydrocarbures extraits des puits d'un champ, ainsi que certaines grandeurs physiques, est essentielle à la gestion efficace des puits. Les incertitudes liées à l'impossibilité de disposer de toutes les mesures amènent à considérer différentes simulations, aboutissant à des scénarios différents, sans savoir à l'avance laquelle se rapprochera le plus de la réalité. Les méthodes utilisées à l'heure actuelle nécessitent un "recalage" des paramètres à de multiples reprises. Nous proposons dans cette présentation une application de méthodes d'agrégation des prévisions des simulations, utilisant efficacement l'information fournie au fil du temps par les observations, et permettant d'être compétitif sur le long terme avec la simulation la plus performante, sans avoir à effectuer ce recalage. Le cadre mathématique choisi est celui, très robuste, des "suites individuelles". Nous mettons tout d'abord en évidence les principes et le formalisme de l'agrégation séquentielle. Puis nous présentons les résultats obtenus avec trois algorithmes d'agrégation. Enfin, nous présentons une problématique légèrement différente : l'obtention de faisceaux de prévision dans le cadre de suites individuelles, et nous proposons une méthodologie innovante.

Mots-clés. Prévision de séries temporelles, apprentissage séquentiel, suites individuelle.

Abstract. For oil companies, the ability to forecast the quantity and the quality of the oil extracted from fields, and some other physical quantities, is crucial to manage oil fields efficiently. Operational conditions make it impossible to have all the required measurements, so some uncertainties remain. This implies the necessity of taking into account several simulations, which lead to various scenarios, and one cannot know beforehand which one will be closer to the reality. The current methods use several times a "history matching" for parameters. We show in this presentation how to apply aggregation methods on the simulation predictions, in order to use efficiently the growing information at our disposal, and to compete on the long run with the best simulation, without having to do this history matching. The mathematical framework is a very robust one : the prediction of "individual sequences". We first introduce the principles and formalism of sequential aggregation. Then we show the results we obtained with three different aggregation algorithms. At last, we tackle a slightly different issue : being able to provide "forecasting tubes" in the individual sequences framework ; and we present a new methodology about it.

Keywords. Time series forecasting, sequential learning, individual sequences.

1 Introduction

Cette présentation montre des résultats issus d’une collaboration avec l’IFP Energies Nouvelles. Elle s’appuie sur un jeu de données classique dans le domaine de la gestion de puits, le benchmark Brugge, dont une présentation est faite par Da Veiga et al. (2012).

Pour les entreprises pétrolières, il est crucial de pouvoir prévoir mois après mois la production des puits, tant en quantité qu’en qualité (quantité d’eau et de gaz mélangés au pétrole), ainsi que certaines grandeurs physiques (pression dans les puits), afin de pouvoir réaliser une gestion et une exploitation efficaces.

Ces prévisions s’appuient sur la résolution, sur une grille, d’équations aux dérivées partielles issues de modèles géophysiques, utilisant notamment des paramètres de perméabilité et de porosité. La difficulté majeure provient de contraintes techniques limitant les mesures sur le terrain, ce qui fait qu’un certain nombre de paramètres sont inconnus. Pour pallier ce manque, plusieurs “répartitions géologiques” (une centaine dans notre étude de cas) sont proposées, et les simulations correspondantes sont alors calculées. Il n’est pas envisageable de relancer les simulations à chaque pas de temps, d’une part pour des raisons de temps de calculs, d’autre part car les données obtenues ne permettent pas de remonter directement à la “répartition géologique” et donc d’améliorer la résolution des équations du problème. Toutefois, il est souhaitable d’utiliser au mieux les informations supplémentaires fournies à chaque pas de temps par les observations afin d’améliorer les prévisions.

Nous proposons dans notre étude l’utilisation (nouvelle dans ce domaine) d’une théorie efficace dans ce cadre : celle de l’agrégation d’experts dans le cadre de suites individuelles.

2 Agrégation séquentielle d’experts

La théorie de l’agrégation d’experts dans le cadre de suites individuelles fait partie du domaine de l’apprentissage séquentiel, dont l’objectif est de prévoir à chaque pas de temps t , une certaine quantité (“observation”), en disposant des variables explicatives jusqu’au temps t , et des observations jusqu’au temps $t - 1$. On cherche donc à utiliser au mieux l’information croissante dont on dispose au fil du temps.

On parle d’agrégation d’experts, lorsque la prévision est réalisée en combinant (“agrégant”) des prévisions fournies aux statisticiens par divers “experts”. Ces derniers peuvent être des méthodes statistiques diverses, ou un même modèle pris avec des paramètres différents, voire plusieurs experts humains. Dans cette théorie, on ne s’intéresse pas à la façon dont les experts fonctionnent, on cherche uniquement à utiliser au mieux les prévisions qu’ils nous fournissent, en tenant compte notamment de leurs performances passées. C’est une situation très commune dans la réalité, où l’on dispose souvent de diverses méthodes de prévision, sans savoir a priori laquelle est la meilleure et donc sans savoir sur laquelle s’appuyer.

Dans le cadre des suites individuelles, les observations ne sont pas considérées comme la réalisation d’une variable aléatoire sous-jacente, mais comme une suite déterministe, susceptible de prendre à chaque pas de temps n’importe quelle valeur dans un ensemble connu. On cherche ainsi une borne de performance déterministe uniforme sur toutes les suites d’observations possibles. Ce type d’approche a été proposée par Littlestone et Warmuth (1994). Il est illusoire d’espérer dans ce cadre des résultats de performance absolues, les observations pouvant toujours prendre des valeurs éloignées des prévisions. On cherche plutôt à obtenir de bonnes performances relatives : on se compare “sur le long terme” (i.e. en perte cumulée) aux performances des experts utilisés pour notre prévision. Divers points de comparaison (“oracles”) sont envisageables : le meilleur expert constant, la meilleure combinaison constante d’experts, la meilleure combinaison linéaire... Ainsi, avoir une perte cumulée plus faible que le meilleur expert constant signifie que l’on a “fait de meilleures prévisions que toutes les méthodes disponibles précédemment” !

Une présentation détaillée de l’agrégation séquentielle d’experts est effectuée dans le livre de Cesa-Bianchi et Lugosi (2006).

3 Application à la gestion de réservoirs

3.1 Cadre formel du problème

On s’intéresse pour chacun des 20 puits du jeu de données, à deux quantités : la pression en fonds de puits (“BHP : Bottom Hole Pressure”) et le débit volumique de pétrole (“QO”). Chacune des propriétés est traitée indépendamment.

Pour une propriété donnée, le cadre formel du problème est le suivant. Avant chaque instant t , on cherche à prévoir la valeur y_t que prendra la propriété (“observation”). On dispose pour ce faire des prévisions de K experts $x_{1,s}, x_{2,s}, \dots, x_{K,s}$ pour tous les instants s jusqu’à t inclus ; dans notre cas, il s’agit des prévisions des simulations pour les $K = 104$ répartitions géologiques proposées. On dispose aussi des observations y_s pour tous les instants passés (autrement dit jusqu’à $t - 1$). En agrégeant les K prévisions d’experts $x_{1,t}, x_{2,t}, \dots, x_{K,t}$, on obtient notre prévision \hat{y}_t . Dans notre étude, on se limite à l’agrégation linéaire : \hat{y}_t est de la forme $\sum_j p_{j,t} x_{j,t}$ (les $p_{j,t}$ sont les “poids” attribués aux différents experts). La vraie valeur de l’observation y_t est alors révélée. On peut dès lors via la perte quadratique faire le bilan des performances de chaque expert $((x_{j,t} - y_t)^2$ pour le j -ième expert) et de notre prévision : $(\hat{y}_t - y_t)^2$. On peut effectuer la décomposition suivante de la perte cumulée de notre algorithme :

$$\sum_{t=1}^T (\hat{y}_t - y_t)^2 = \inf_{f \in F} \sum_{s=1}^T \left(f(x_{1,s}, x_{2,s}, \dots, x_{K,s}) - y_t \right)^2 + R_T,$$

F pouvant être l’ensemble des combinaisons convexes, linéaires, chacun des K experts,... Le premier terme (“erreur d’approximation”) correspond à la performance de l’oracle

(“meilleure combinaison convexe constante”, “meilleur expert constant”, etc). Le deuxième terme, appelé “regret” (ou “erreur d’estimation séquentielle”) compare la performance de nos prévisions à celle de l’oracle. Notre but est d’avoir un regret le plus faible possible (négatif si l’on peut), et a minima sous-linéaire : $o(T)$.

3.2 Un algorithme performant : “ML-Poly”

Nous présentons ici les résultats obtenus par un algorithme d’agrégation récent : Polynomially Weighted Averages with Multiple Learning rates (“ML-Poly”). Il a été proposé par Gaillard, Stoltz et Van Erven (2014). Notre point de référence pour chaque propriété est le meilleur expert (celui ayant la perte cumulée la plus faible) pour la propriété en question. On note $L(j, t) = \sum_{u=1}^t (x_{j,u} - y_u)^2$ et $\hat{L}_t = \sum_{u=1}^t (\hat{y}_u - y_u)^2$ les pertes cumulées respectivement du j -ième expert et de l’algorithme à l’instant t . On voit figure 1 un exemple de prévision d’une propriété par ML-Poly, dans une version dite “gradient” s’appuyant sur des pertes modifiées (en vert les prévisions des experts, en rouge les observations, en violet le meilleur expert, en bleu les prévisions de ML-Poly).

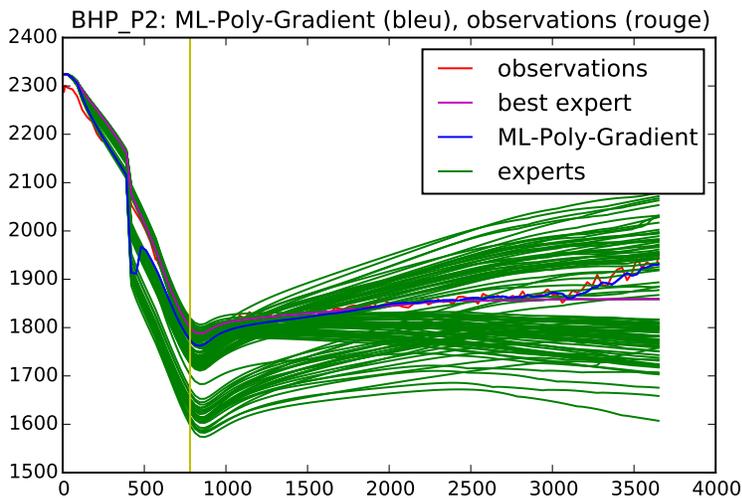


Figure 1: Un exemple d’observations et de prévisions des experts et de ML-Poly

Il s’agit d’agrégation convexe, les poids sont définis de la manière suivante :

$$\eta_{j,t} = \frac{1}{1 + \sum_{u=1}^t \left((\hat{y}_u - y_u)^2 - (x_{j,u} - y_u)^2 \right)^2}$$

$$R_{j,t} = \hat{L}_t - L_{j,t}$$

et

$$p_{j,t} = \frac{\eta_{j,t-1}(R_{j,t-1})_+}{\sum_{i=1}^K \eta_{i,t-1}(R_{i,t-1})_+}$$

Cet algorithme présente le grand avantage de ne pas avoir de paramètre à calibrer, et d’être donc parfaitement opérationnel. Il obtient de bonnes performances à la fois en théorie et en pratique. La figure 2 résume les performances de l’algorithme ML-Poly. Pour chaque puits et chaque propriété du jeu de données étudié, les rectangles verts correspondent à la RMSE du meilleur expert (les puits sont rangés par ordre croissant pour cette RMSE) et les barres de couleur indiquent la RMSE de l’algorithme (bleu quand il fait mieux que le meilleur expert, rouge lorsqu’il fait moins bien).

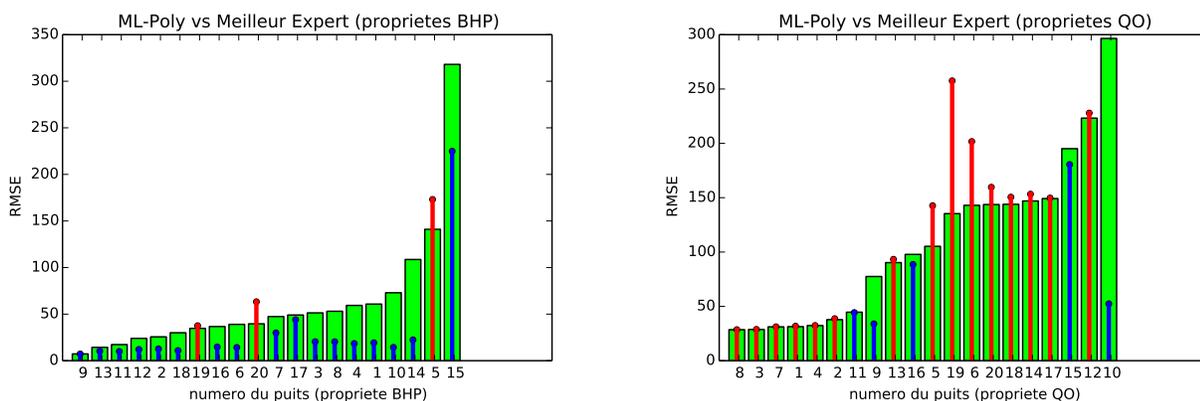


Figure 2: Comparaison des performances de ML-Poly et du meilleur expert

4 Faisceaux de prévision

Une autre demande forte des pétroliers est de disposer de prévisions “moyen-terme” (de quelques mois à quelques années), pour lesquelles, plus qu’une unique prévision, ils souhaitent disposer de régions de confiance, de “faisceaux de prévisions”.

Si cela ne change rien aux simulations numériques aboutissant aux prévisions d’experts, satisfaire ce nouveau cadre nécessite en revanche des adaptations de l’approche. En effet, si les modélisations statistiques traditionnelles disposent d’outils bien connus pour ce type de problèmes (méthodes de Monte-Carlo par exemple), ce cadre est beaucoup plus inhabituel pour l’agrégation en suites individuelles. Sans modèle statistique, il est en effet délicat de dériver des régions de confiance. De plus, la mise à jour séquentielle des poids n’est plus possible dans une prévision “moyen-terme”.

L'approche la plus fréquente consiste à utiliser les algorithmes classiques appliqués à des pertes quantiles (voir Koenker and Bassett, 1978), mais elle ne convient pas aux prévisions moyen-terme. Nous proposons ici une approche novatrice pour répondre à ce problème. Il s'agit de s'appuyer sur un ensemble de scénarios envisageables afin d'en déduire un faisceau de prévisions pour les algorithmes utilisés.

Bibliographie

- [1] Le Ravalec, M., Tillier, E., Da Veiga, S., Enchery, G. and Gervais, V. (2012) , *Advanced Integrated Workflows for Incorporating Both Production and 4D Seismic-Related Data into Reservoir Models*, Oil & Gas Science and Technology
- [2] Littlestone, N. et Warmuth, M. (1994), *The Weighted Majority Algorithm*, Information and Computation
- [3] Cesa-Bianchi, N. et Lugosi, G. (2006), *Prediction, Learning and Games* , Cambridge University Press
- [4] Gaillard, P., Stoltz, G. et Van Erven, T. (2014), *A Second-order Bound with Excess Losses*
- [5] Koenker, R. et Bassett, G. (1998), *Regression quantiles*, Econometrica