

# Estimation de composante non stationnaire et modélisation par modèle additif généralisé (GAM), application aux barrages voûtes

Camille ROQUENCOURT, Alexandre GIRARD<sup>1</sup>, Pauline LAVIRON PETIT<sup>1</sup>, Guy D'URSO<sup>1</sup>

<sup>1</sup>EDF R&D

6 Quai Watier, Chatou 1

camille.roquencourt@hotmail.fr, alexandre.girard@edf.fr, pauline.laviron-petit@edf.fr, guy.durso@edf.fr

**Résumé** – Pouvoir estimer une composante non stationnaire faible et sans forme paramétrique définie dans un signal est un objectif global de traitement qui concerne notamment beaucoup de problèmes de surveillance. Cette notion de composante non-stationnaire exprime en particulier le fait que le temps est en soi une variable explicative du comportement du signal, par exemple pour suivre les conséquences du vieillissement des ouvrages en suivi des structures. Il est proposé dans cet article d’adopter une démarche de Machine Learning en déterminant dans un premier temps des modèles, tels que les Modèles Additifs Généralisés (GAM), les Réseaux de Neurones (RN) (ou tout autre modèle dérivable), prenant comme entrées les variables explicatives connues et le temps. Puis, pour estimer la contribution propre au temps, nous utilisons la dérivée partielle par rapport à celui-ci, par analogie avec la notion d’instationnarité en mécanique des fluides. Une application aux barrages voûtes est proposée dans cet article, où le phénomène étudié est le déplacement du barrage au cours du temps.

**Abstract** – Being able to estimate a weak non-stationary component without a parametric form defined in a signal is an overall processing objective that concerns many problems. This notion of non-stationary component expresses the fact that time is itself an explanatory variable of signal behaviour, for example to monitor the consequences of ageing structures in structural monitoring. This paper proposes to adopt a Machine Learning approach by first using models, such as Generalized Additive Models (GAM), Neural Networks (RN) (or any other derivable model), taking as inputs the known explanatory variables and time. Then, to estimate the specific contribution of time, we use the partial derivative with respect to it, by analogy with the notion of instationarity in fluid mechanics. An application to arch dams is proposed in this article, where the phenomenon studied is the displacement of the dam over time.

## 1 Introduction

EDF (Electricité de France), premier producteur et fournisseur d’électricité en France et en Europe, exploite un parc hydroélectrique qui assure 15% de la puissance produite. Ce parc comprend principalement des barrages, dont il faut assurer la surveillance. En effet, un barrage subit tout au long de sa vie plusieurs déplacements, qui sont induits par différents types de sollicitations (voir Figure 1). Ces déplacements sont de deux types : les **déplacements réversibles**, induits par la pression de l’eau dans le réservoir et les variations de température, et les **déplacements irréversibles**, regroupant l’ensemble des phénomènes qui évoluent dans le temps : vieillissement du béton, pathologies du barrage, etc. (composante non-stationnaire). Cet article s’intéresse plus spécifiquement à l’influence du temps.

Ces déplacements sont surveillés à l’aide de piézomètres<sup>1</sup> placés au niveau de l’interface roche-béton : la mesure obtenue est un niveau piézométrique (en mètres) proportionnel à la pression de l’eau circulant à l’interface. Plus le niveau piézométrique est grand, plus l’ouverture est supposée importante, et inversement.

Le but de notre étude est donc d’estimer le mouvement ir-

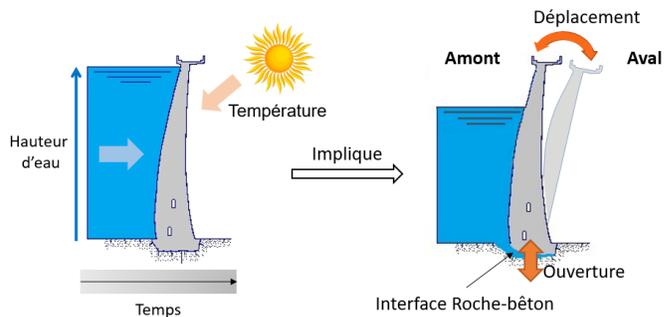


FIGURE 1 – Représentation d’un barrage voûte vu de profil

réversible du barrage au cours du temps, à partir d’un historique  $[(Z_t, S_t, t), Y_t]$ , où  $Y_t$  représente le niveau piézométrique au temps  $t$ ,  $Z_t$  la hauteur d’eau normalisée dans le bassin,  $S_t = 2\pi (t/365.25 - \lfloor t/365.25 \rfloor)$ , une variable de saisonnalité comprise entre 0 (1er Janvier) et  $2\pi$  (31 décembre), et enfin  $t$  le temps continu, variable d’entrée du mouvement irréversible. Nous cherchons donc à modéliser un phénomène physique, prenant en compte trois variables observées, dont **les effets ne sont pas indépendants**.

1. Dispositif pour mesurer la pression des liquides

## 2 État de l'art

Nous cherchons à estimer le déplacement du barrage uniquement induit par la variable irréversible, i.e. le temps. Pour cela, nous allons construire dans un premier temps une fonction de prédiction  $\hat{Y}$  à partir des données que nous possédons (historique sur les 5 dernières années, à échantillonnage irrégulier) :

On pose  $Y_t = f_\theta(Z_t, S_t, t) + \epsilon_t$  puis on estime les paramètres  $\theta$  :

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \| \mathbf{Y} - f_\theta(\mathbf{Z}, \mathbf{S}, \mathbf{t}) \|^2, \text{ et enfin : } \hat{Y}_t = f_{\hat{\theta}}(Z, S, t)$$

où  $\mathbf{Y}$ ,  $\mathbf{Z}$ ,  $\mathbf{S}$  et  $\mathbf{t}$  sont les vecteurs d'observation,  $f_\theta$  le modèle, et  $\epsilon_t$  l'erreur de modélisation.

Plusieurs modèles existent déjà : le plus connu est le modèle additif HST [1] qui ne prend pas en compte l'interaction des variables, et estime le mouvement irréversible par une fonction linéaire :  $Y_t = f_1^\gamma(S_t) + f_2^\theta(Z_t) + \beta t$ , avec  $f_1^\gamma$  une fonction saisonnière et  $f_2^\theta$  un polynôme de degré 4.

Deux autres modèles ont récemment été développés par M. de Granrut (dans le cadre de sa thèse à EDF/DTG, Grenoble) : un modèle physico-statistique non linéaire (NL-HST)[2], qui contraint le mouvement irréversible à être monotone, et un réseau de neurones (RN) [3], très performant en termes de prédiction, mais difficilement interprétable et modulable. De plus, aucune méthode n'est encore satisfaisante d'un point de vue mathématique et physique pour estimer le mouvement irréversible avec ces derniers modèles.

Nous proposons donc dans le cadre de cet article une nouvelle méthode de calcul du mouvement irréversible, ainsi qu'une nouvelle modélisation s'appuyant sur les modèles additifs généralisés (GAM) [4], qui ont pour avantages d'être flexibles, interprétables, et performants.

## 3 Modèle Additif Généralisé (GAM)

Les modèles GAM (Generalized Additive Model) sont une généralisation des modèles linéaires. Pour expliquer une variable  $Y$  à partir des variables  $(X_1, \dots, X_K)$ , on pose le modèle non paramétrique suivant :  $Y = \sum_{i=1}^K f_i(X_i)$ . On paramètre ensuite le modèle en décomposant chaque  $f_i$  dans une base de fonction  $(b_1, \dots, b_{K_i})$  :  $f_i(x) = \sum_{j=1}^{K_i} \beta_j b_j(x)$ . Étant donné que les effets de nos variables ne sont pas indépendants, utiliser un tel modèle additif ne serait pas adapté. Nous allons donc utiliser un modèle couplé, décrit dans le livre de Simon Wood [4].

### 3.1 Formulation du modèle

Le modèle utilisé se présente sous la forme générale suivante :

$$Y_t = f(Z_t, S_t, t)$$

Pour estimer la fonction  $f$ , on associe à chaque variable une décomposition générale d'une fonction donnée de cette variable

dans une base de fonctions choisie (fonctions cubiques, B-spline, base gaussienne, ... [4]) :

$$f_z(z) = \sum_{i=1}^{K_z} \delta_i d_i(z), \quad f_s(s) = \sum_{j=1}^{K_s} \alpha_j a_j(s), \quad f_t(t) = \sum_{k=1}^{K_t} \beta_k b_k(t)$$

où les coefficients  $(\delta_j, \alpha_j, \beta_j)$  sont des paramètres inconnus à estimer. Les hyperparamètres à prédéfinir sont les bases de fonctions  $(d_j, a_j, b_j)$  et les dimensions des bases  $(K_z, K_s, K_t)$  (voir 3.2). On écrit ensuite la fonction  $f_z$  comme une fonction de  $s$  et  $z$ . Pour cela, on décompose les coefficients  $\delta_i$  dans la base de  $f_s$  (Figure 2) :

$$\delta_i(s) = \sum_{j=1}^{K_s} \alpha_{ij} a_j(s) \rightarrow f(z, s) = \sum_{i=1}^{K_z} \sum_{j=1}^{K_s} \alpha_{ij} a_j(s) d_i(z)$$

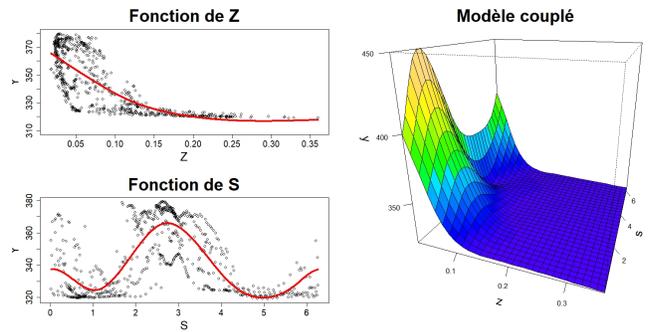


FIGURE 2 – A gauche, les fonctions  $f_z$  et  $f_s$  estimées séparément, et à droite la fonction couplée  $f(z, s)$

Puis en décomposant de la même manière les coefficients  $\alpha_{ij}$  dans la base de  $f_t$ , on obtient :

$$f_\beta(s, z, t) = \sum_{i=1}^{K_z} \sum_{j=1}^{K_s} \sum_{k=1}^{K_t} \beta_{ijk} b_k(t) a_j(s) d_i(z) = \mathbf{X} \beta$$

Avec :  $X_i = X_{zi} \otimes X_{si} \otimes X_{ti}$ , où  $\otimes$  est le produit de Kronecker  $X_i$  la  $i^{\text{ème}}$  ligne de la matrice  $\mathbf{X}$ ,  $X_s = (a_1(\mathbf{S}), \dots, a_{K_s}(\mathbf{S}))$ ,  $X_z = (d_1(\mathbf{Z}), \dots, d_{K_z}(\mathbf{Z}))$ ,  $X_t = (b_1(\mathbf{t}), \dots, b_{K_t}(\mathbf{t}))$ , et  $\beta \in \mathbb{R}^{K_z \times K_s \times K_t}$

On résout ensuite le même problème d'optimisation :

$$\underset{\beta}{\operatorname{argmin}} \| \mathbf{Y} - \mathbf{X} \beta \|^2 + \sum_i^3 \lambda_i \int [f_i''(x)]^2 dx$$

où les  $\lambda_i$  sont des paramètres de lissage relatifs à chaque variable. Plus  $\lambda_i$  est grand, moins la fonction est libre de varier sur l'axe de la variable  $i$ . Il est donc possible grâce à ce modèle de choisir une contrainte de lissage différente en fonction des variables.

La pénalisation  $\sum_i \lambda_i \int [f_i''(x)]^2 dx$  peut, en développant les fonctions  $f_i$ , se réécrire :  $\lambda \beta^T S \beta$  où  $S$  est une matrice aux coefficients connus. On obtient alors une formule explicite de la solution :

$$\hat{\beta} = (X^T X + \lambda S)^{-1} X^T Y$$

Ce modèle a donc pour avantage d'être simple, d'avoir des coefficients interprétables selon les bases choisies, et d'être flexible :

le choix des bases et des paramètres de lissage permet à l'utilisateur d'affiner et de contraindre son modèle très facilement en fonction du problème.

### 3.2 Ajout de contrainte physique

Le problème d'optimisation du modèle GAM présenté étant quadratique, il est aisé d'y rajouter des contraintes. En effet, pour notre problème, le niveau piézométrique est strictement croissant en fonction de la hauteur d'eau. Il suffit donc de rajouter la contrainte que la dérivée partielle en  $z$  est strictement positive et de résoudre un problème d'optimisation quadratique sous contrainte d'inégalité suivant :

$$\min_{\beta} \|\mathbf{Y} - \mathbf{X}\beta\| + \lambda\beta^T S\beta \quad t.q \quad X'_z\beta < 0 \quad \forall (z, s, t)$$

où  $X'_z$  correspond à la dérivée partielle du modèle (1) par rapport à la variable  $\mathbf{Z}$ , avec  $X'_{z_i} = X'_{z_i} \otimes X_{s_i} \otimes X_{t_i}$ , et  $X'_z = (d'_1(\mathbf{Z}), \dots, d'_{K_z}(\mathbf{Z}))$ .

### 3.3 Choix des hyperparamètres et critères de validation

Les hyperparamètres à déterminer pour chaque variable avant d'optimiser le modèle GAM sont : le choix des bases de fonctions, les dimensions des bases et les paramètres de lissage. Leur choix peut fortement influencer le modèle. Pour déterminer les hyperparamètres optimaux, une validation croisée est effectuée pour plusieurs valeurs du paramètre d'intérêt, et la valeur qui donne la plus petite erreur quadratique moyenne est alors retenue.

Étant donné que nos observations ne sont pas indépendantes (série temporelle), nous n'utilisons pas une validation croisée K-fold classique, où la séparation en sous échantillons se fait aléatoirement, mais nous effectuons une séparation par année, puisque le phénomène observé est saisonnier sur une année.

### 3.4 Résultats

Après avoir testé plusieurs hyperparamètres pour le modèle GAM, nous obtenons finalement les résultats détaillés dans les Tables 1 et 2. Le réseau de neurones calculé ici est celui utilisé par M. De Granrut [3]. La première colonne correspond à la racine de l'erreur quadratique moyenne (RMSE) calculée sur l'ensemble des données d'entraînement, la deuxième à l'erreur en validation croisée, la troisième au  $R_2$  ajusté, qui quantifie la variance des erreurs, et la complexité du modèle (plus il est proche de 1, plus le modèle est performant). La dernière colonne correspond au nombre de paramètres à estimer (p).

Le modèle GAM que nous avons construit est donc tout aussi performant que le réseau de neurones et a pour avantage d'être interprétable, et facile à contraindre. De plus, l'optimisation de ce modèle est simple et robuste (problème convexe), contrairement au réseau de neurones.

	RMSE Train	RMSE CV	$R_2$	p
HST	7.495	10.717	0.868	10
GAM	2.581	6.534	0.985	120
RN	2.731	6.610	0.981	54
NL HST	4.152	8.545	0.959	13

TABLE 1 – Résultats pour chaque modèle

	Z	S	t
Base	cubic spline	cubic cyclic spline	B-spline d= 2
Dim	6	5	4
$\lambda$	0.05	0.01	0.1

TABLE 2 – Hyperparamètres du modèle GAM utilisé

Maintenant que nous avons construit et entraîné chaque modèle, nous nous intéressons au calcul du mouvement irréversible.

## 4 Estimation du mouvement irréversible

La méthode utilisée actuellement pour retracer le mouvement irréversible se base sur le modèle HST, où la composante irréversible est en fait représentée par le signe du coefficient obtenu par régression linéaire. Nous proposons ici une méthode qui permet d'estimer la composante irréversible à partir de modèles qui prennent en compte l'interaction des effets entre chaque variable. Cette méthode peut s'appliquer à tout modèle dérivable.

### 4.1 Méthode de calcul

Pour éviter toute confusion, nous notons  $\tau$  le temps courant et  $t$  la variable pilotant l'irréversible. On exprime donc les niveaux piézométriques  $Y$  par une fonction  $f$  dérivable de la cote  $Z$ , de la saison  $S$  et du temps  $t$  :

$$Y_{\tau} = f(Z(\tau), S(\tau), t(\tau)) + \varepsilon \quad (1)$$

$$= f(\tau) + \varepsilon$$

où  $f$  correspond au modèle,  $\varepsilon$  à l'erreur de modélisation, et  $(Z(\tau), S(\tau), t(\tau))$  représente une observation de nos variables explicatives au temps  $\tau$ , avec  $t(\tau) = \tau$ ,  $S(\tau)$  fonction saisonnière dépendant de  $\tau$  et  $Z(\tau)$  fonction d'interpolation linéaire entre deux observation de  $Z$ .

La dérivée de  $f$ , qui correspond à estimer la dérivée temporelle des mesures de niveaux piézométriques, est alors donnée par :

$$\begin{aligned} \frac{df}{d\tau}(\tau) &= \frac{\partial f}{\partial Z} \frac{\partial Z}{\partial \tau}(Z(\tau), S(\tau), t(\tau)) + \frac{\partial f}{\partial S} \frac{\partial S}{\partial \tau}(Z(\tau), S(\tau), t(\tau)) \\ &+ \frac{\partial f}{\partial t} \frac{\partial t}{\partial \tau}(Z(\tau), S(\tau), t(\tau)) \\ &= \frac{\partial f}{\partial Z} \frac{\partial Z}{\partial \tau}(Z(\tau), S(\tau), t(\tau)) + \frac{\partial f}{\partial S} \frac{\partial S}{\partial \tau}(Z(\tau), S(\tau), t(\tau)) \\ &+ \frac{\partial f}{\partial t} \frac{\partial t}{\partial \tau}(Z(\tau), S(\tau), t(\tau)) \text{ car } t(\tau) = \tau \text{ donc } \frac{\partial t}{\partial \tau} = 1 \end{aligned}$$

Si l'on veut se placer dans une situation où la cote et la saison n'influent plus sur l'évolution de  $f$ , et donc ne varient pas au premier ordre du temps, on pose :  $\frac{\partial Z}{\partial \tau} = 0$  et  $\frac{\partial S}{\partial \tau} = 0$ .

$$\text{On a alors : } \frac{df}{d\tau}(\tau) = \frac{\partial f}{\partial t}(Z(\tau), S(\tau), t(\tau))$$

Sur un intervalle  $\Delta\tau$ , cela revient à dire que la variation de  $f$  à hauteur d'eau et saison fixées, qui représente donc le déplacement irréversible local, est approximée par :

$$\frac{\partial f}{\partial t}(Z(\tau), S(\tau), t(\tau))\Delta\tau$$

L'estimation du déplacement irréversible global au temps  $\bar{\tau}$  est alors fournie par intégration de ces déplacements irréversibles locaux :

$$D_{irrev}(\bar{\tau}) = \int_{\tau_1}^{\bar{\tau}} \frac{\partial f}{\partial t}(Z(\tau), S(\tau), t(\tau))d\tau$$

où  $\tau_1$  correspond à la première date d'observation.  $D_{irrev}$  représente un déplacement relatif. Pour le positionner en absolu, il faut alors définir un point de référence qui peut être assez intuitivement  $f(\tau_1)$ . La courbe ainsi définie :

$$D_{irrev}(\tau) + f(Z(\tau_1), S(\tau_1), t(\tau_1))$$

permet de prendre en compte les déplacements irréversibles locaux fournis par le modèle tout en éliminant les contributions de la hauteur d'eau et de la saison.

On peut aussi définir de même les déplacements à hauteur d'eau et temps fixés et à saison et temps fixés :

$$\begin{aligned} \frac{\partial S}{\partial \tau} = 0 \quad \frac{\partial t}{\partial \tau} = 0 \quad D_Z(\bar{\tau}) &= \int_{\tau_1}^{\bar{\tau}} \frac{\partial f}{\partial Z} \frac{\partial Z}{\partial \tau}(Z(\tau), S(\tau), t(\tau))d\tau \\ \frac{\partial S}{\partial \tau} = 0 \quad \frac{\partial Z}{\partial \tau} = 0 \quad D_S(\bar{\tau}) &= \int_{\tau_1}^{\bar{\tau}} \frac{\partial f}{\partial S} \frac{\partial S}{\partial \tau}(Z(\tau), S(\tau), t(\tau))d\tau \end{aligned}$$

On se ramène ainsi à un modèle équivalent à HST :

$$f(Z(\tau), S(\tau), t(\tau)) = \underbrace{D_S(\tau) + D_Z(\tau)}_{D_{reversible}(\tau)} + D_{irrev}(\tau) + f(\tau_1)$$

## 4.2 Résultats

Nous avons appliqué la méthode décrite précédemment pour chaque modèle : HST, NLHST, GAM et RN, en langage R (R Development Core Team). La figure 3 représente la décomposition des déplacements réversibles et irréversibles, ainsi que l'erreur de modélisation. Pour chaque modèle, la somme des trois courbes est donc égale aux niveaux piézométriques observés.

Grâce à cette méthode, il est donc possible d'estimer une composante non stationnaire à partir de n'importe quel modèle prédictif dérivable. Les exploitants peuvent alors observer le mouvement irréversible du barrage en temps réel, et peuvent ainsi détecter des anomalies plus rapidement et de manière plus fiable.

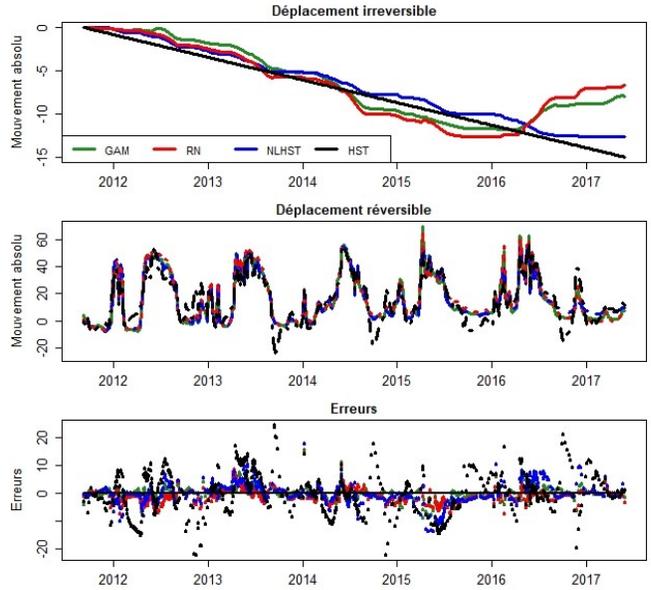


FIGURE 3 – Décomposition des déplacements réversible/irréversible pour chaque modèle

## 5 Conclusion

Le déplacement des barrages est un exemple de phénomène physique, influencé par plusieurs variables explicatives dont les effets ne sont pas indépendants, et le temps. Cet article a présenté une approche pour estimer les déplacements dits irréversibles, c'est-à-dire non induits par les variations de température ou la pression d'eau dans le bassin.

Une variable temporelle  $t$  a été introduite dans la modélisation, afin de représenter la composante irréversible. Un modèle GAM, aux coefficients interprétables, performant et peu coûteux en termes d'optimisation, a alors été construit. La composante non stationnaire (ici le déplacement irréversible du barrage) a alors été estimée en intégrant la dérivée partielle du modèle par rapport au temps.

## Références

- [1] F Salazar, R Morán, M Toledo, and E Oñate. Data-Based Models for the Prediction of Dam Behaviour : A Review and Some Methodological Considerations. *Archives of Computational Methods in Engineering*, 2015.
- [2] M. de Granrut J-P. Fabre D.Dias D F.Letué A.Simon. Prediction of piezometric levels at the rock concrete interface considering the non linearity of permeability in the foundations. *Third Dam World Conf., Foz Do Iguacu : 2018*.
- [3] M de Granrut, D Dias, and A Simon. Artificial neural networks for the interpretation of piezometric levels at the rock-concrete interface of arch dams. *Engineering Structures*, 178, 10 2018.
- [4] S Wood. *Generalized Additive Models : an introduction with R*. Boca Raton, FL : Chapman Hall/CRC, 2006.