

---

# Identification de paramètres dans les équations de Saint-Venant

Céline Cheviet

EDF – R&D

Laboratoire National d'Hydraulique et Environnement

6, quai Watier

F-78401 Chatou

celine.cheviet@edf.fr

---

*RÉSUMÉ. Les codes numériques fondés sur les éléments finis sont à l'heure actuelle couramment utilisés pour la simulation d'écoulements à surface libre, tels le calcul des marées et des crues, la dispersion de panaches thermiques, la prédiction des ondes de submersion liées à la rupture de barrage, l'évaluation de l'impact de travaux publics sur l'environnement, etc. La plupart des étapes effectuées par des pré ou post-processeurs dans une étude hydraulique sont maintenant automatisées et l'estimation de paramètres inconnus (par exemple les coefficients de frottement) s'affiche désormais comme la tâche la plus longue lors du calage d'un modèle numérique. Ici nous traiterons du problème de l'estimation de paramètres dans les équations de Saint-Venant, dans le cadre du système Telemac.*

*ABSTRACT. Finite element software systems are nowadays current practice for simulating free surface flow problems, such as the computation of tides and floods, dispersion of thermal plumes, prediction of dambreak floodwaves, assessment of the impact of public works on environment, etc. Most pre and post-processing steps in a hydraulic study are now automated and the estimation of unknown parameters, such as bottom friction coefficients, becomes more and more the main time-consuming task for calibrating a numerical model. Here the problem of parameter estimation in Shallow Water Equations is dealt with, in the framework of the Telemac system.*

*MOTS-CLÉS : identification, estimation, paramètres, calage, optimisation.*

*KEYWORDS: identification, estimation, parameters, calibration, optimisation.*

---

## 1. Introduction

Le but de l'identification de paramètres est d'estimer des paramètres inconnus, dont la valeur est habituellement déterminée par calage manuel. Ce calage est effectué par expertise de l'hydraulicien (en particulier en ce qui concerne la valeur d'un tir initial pour le paramètre à caler), associée à une méthode d'essais-erreurs qui permet d'approcher empiriquement la valeur optimale du paramètre en question. Ces essais représentent autant de simulations numériques qui peuvent s'avérer coûteuses et longues.

La méthode que nous allons présenter ici traite de la mise en œuvre d'une méthode automatique d'estimation de paramètres dans le code numérique de simulation des écoulements à surface libre Telemac.

## 2. Technique de l'estimation de paramètres

Les paramètres inconnus, notés  $S_i$ , peuvent être les coefficients de frottement en chaque nœud d'un domaine bidimensionnel, les conditions aux limites ou les coefficients de diffusion. La démarche de résolution choisie est basée sur la théorie de la modélisation inverse et du contrôle optimal (Chavent, 1971 et Barros, 1996).

Une fonction coût  $J$  (dépendant des paramètres recherchés) est utilisée, en général égale à l'écart type entre les résultats numériques et les mesures :

$$J = \sum_{it=1}^{nit} \sum_{ip=1}^{np} W_{ip} (X_{ip}^{it} - M_{ip}^{it})^2$$

Ce coût est minimisé en faisant appel à une méthode de gradient conjugué. Dans la formule ci-dessus  $nit$  est le nombre d'itérations en temps et  $np$  est le nombre de nœuds de calcul.  $W_{ip}$  est un facteur de pondération tenant compte de la confiance dans les données ou de leur importance (pour une donnée indisponible le poids sera fixé à zéro).  $X$  représente les variables d'états (dans Telemac-2D les vitesses moyennées sur la verticale et la hauteur d'eau) et  $M$  les mesures. Le principal problème technique rencontré lors de cette minimisation est l'évaluation du gradient

de  $J$  par rapport aux paramètres inconnus :  $\frac{\partial J}{\partial S_i}$ . Cette évaluation fait appel à la

théorie des multiplicateurs de Lagrange. On forme le Lagrangien  $L$  :

$$L(X, S, P) = J(S) + (P, AX - B)$$

où  $P$  est le vecteur des multiplicateurs de Lagrange (variables arbitrairement introduites que l'on va ensuite préciser). On cherche à minimiser  $J(S)$  sous

contrainte, la contrainte étant que l'on ait toujours  $AX = B$ . Le minimum recherché correspond au triplet  $(X, S, P)$  qui satisfait les conditions d'optimalité suivantes :

- $\frac{\partial L(X, S, P)}{\partial P} = 0$  qui donne  $AX - B = 0$  c'est-à-dire le système « direct », résolu par le code pour les équations de Saint-Venant. En le résolvant nous obtenons  $X$ .
- $\frac{\partial L(X, S, P)}{\partial X} = 0$  qui donne le système « adjoint ». En le résolvant nous obtenons  $P$  appelé dans ce cas le vecteur des « variables adjointes » et que nous noterons  $X_{adj}$  dans la suite de cet article.
- $\frac{\partial L(X, S, P)}{\partial S} = 0$  qui donne la formulation du gradient de  $J$ .

On voit donc que le calcul de  $\frac{\partial J}{\partial S_i}$  nécessite au préalable la résolution d'un

système linéaire adjoint, dont la matrice se trouve être la transposée de celle présente dans le système linéaire résolu par le code pour les équations de Saint-Venant. Les inconnues de ce système adjoint sont appelées variables adjointes. Si  $A^i X^i = B^i$  est le système direct, alors le système adjoint sera :

$${}^t A^i X_{adj}^i = 2W(X^i - M^i) - \left( \frac{\partial A^{i+1}}{\partial X^i} X^{i+1} - \frac{\partial B^{i+1}}{\partial X^i}, X_{adj}^{i+1} \right)$$

Une caractéristique intéressante à noter est que le système adjoint ne dépend pas du paramètre à estimer. Seul le gradient de  $J$  en dépend. La mise en œuvre de cette technique peut donc être rapidement adaptée à un nouveau type de paramètre à estimer.

Pour les problèmes non stationnaires, le système direct est une série de systèmes linéaires dans l'ordre chronologique, chaque nouveau système dépendant des résultats du précédent,  $X^{i+1}$  étant obtenu après  $X^i$ . Au contraire, dans le cas des systèmes adjoints, les variables adjointes au temps  $i+1$  apparaissent dans le système correspondant au temps  $i$ . Selon la théorie, les séries correspondantes de systèmes adjoints seront donc résolues en sens inverse du temps, à partir de la dernière itération et en posant  $X_{adj}^{nit+1} = 0$  comme condition « initiale ». Pour un problème non stationnaire, chaque itération de l'algorithme du gradient conjugué comprendra donc :

- La résolution d'une série de systèmes directs (c'est-à-dire une résolution classique des équations de Saint-Venant avec un ensemble donné de paramètres) pour chaque itération en temps  $it$  de 1 à  $nit$ .

– Le calcul de la fonction coût  $J$ , et la comparaison de sa valeur avec un critère d'arrêt (l'algorithme a atteint un minimum local si  $J < \varepsilon$  ou si  $J$  augmente par rapport à sa valeur précédente).

– La solution d'une série de systèmes adjoints, en sens inverse du temps de  $it = nit$  à 1.

Le calcul du gradient de la fonction coût par la formule :

$$\frac{\partial J}{\partial S_i} = - \sum_{it=1}^{nit} (X_{adj}^{it}, \frac{\partial A^{it}}{\partial S_i} X^{it})$$
 où  $X_{adj}^{it}$  sont les variables adjointes au temps  $it$  et  $A^{it}$  la matrice du système direct au même temps.

– Le calcul d'un nouveau jeu de paramètres par une méthode de descente utilisant la valeur du gradient de la fonction coût :

$$S_i^{m+1} = S_i^m - \rho_m \frac{\partial J}{\partial S_i}.$$

En pratique, le nombre d'itérations avant d'atteindre un minimum de  $J$  varie entre 10 et 100. Le nombre de paramètres inconnus doit être cohérent avec le nombre de données disponibles, et la théorie exposée ici peut être couplée à une stratégie de zonage où le domaine d'étude est divisé en un nombre réduit de zones à frottement constant.

### 3. Mise en œuvre pour les équations de Saint-Venant

La principale difficulté pour la mise en œuvre de la méthode appliquée aux équations de Saint-Venant est la construction du système adjoint. Si sa matrice est obtenue simplement par transposition de celle du système direct, il en va tout autrement pour le second membre. Il s'écrit en effet :

$$2W(X^{it} - M^{it}) - \left( \frac{\partial A^{it+1}}{\partial X^{it}} X^{it+1} - \frac{\partial B^{it+1}}{\partial X^{it}}, X_{adj}^{it+1} \right) - \left( \frac{\partial A^{it}}{\partial X^{it}} X^{it} - \frac{\partial B^{it}}{\partial X^{it}}, X_{adj}^{it} \right)$$

ce qui suppose la construction de nombreux termes.

Le dernier terme, non cité dans le paragraphe précédent, provient des conditions aux limites de type Dirichlet (cas particulier où  $A^{it}$  et  $B^{it}$  sont des fonctions de  $X^{it}$ ).

Avec l'algorithme de Telemac-2D, qui couple dans sa résolution trois variables : la hauteur d'eau et les deux composantes de la vitesse horizontale, il faut examiner la sensibilité de sept matrices et trois seconds membres par rapport à ces trois variables. Ce sont finalement plus de 30 termes qui doivent être construits. On comprend ainsi que le système adjoint est intimement lié au système direct, et qu'il doit être maintenu et évoluer conjointement.

## 4. Validation et applications

### 4.1. Test préliminaire : validation du calcul du gradient de la fonction coût

Toute la procédure est basée sur le calcul du gradient de la fonction coût. Cependant un minimum de  $J$  peut être atteint même si ce gradient est légèrement faux. Un cas test réussi n'est donc pas une preuve que le gradient de  $J$  est correct. Pour vérifier la valeur de ce gradient, nous en avons calculé une valeur approchée en utilisant la définition d'une dérivée : l'un des paramètres  $S$  est changé en  $S + dS$ , et  $grad(J)$  est évalué par la formule : 
$$\frac{J(S + dS) - J(S)}{dS}.$$

Les tests effectués montrent que le calcul du gradient s'effectue avec une erreur généralement inférieure à 3 %.

### 4.2. Cas tests

Une seconde approche pour la validation de la procédure a été mise en œuvre. Un calcul est fait avec un jeu donné de coefficients, qui est ensuite « oublié ». Seul le résultat du calcul est donné en entrée de la procédure d'estimation. Afin de se mettre dans des conditions plus proches de celles rencontrées dans des cas réels, on peut ne retenir qu'une partie des données et ne fournir au code que celles-là comme référence de sa procédure d'estimation.

Prenons l'un des exemples testés : le cas du canal rectangulaire à fond plat.

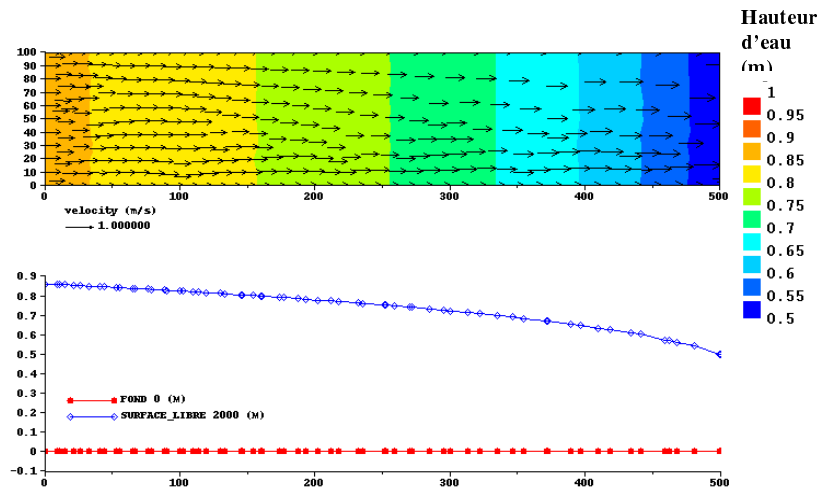


Figure 1. Cas-test pour l'estimation du frottement : canal rectangulaire

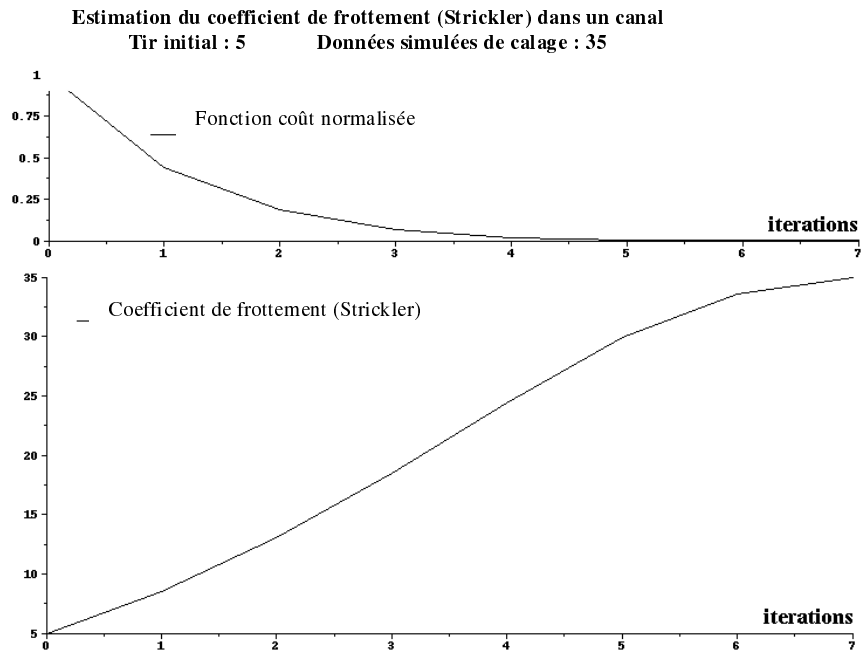
On peut voir sur la figure 1 la géométrie du canal ainsi qu'un exemple d'écoulement (hauteur d'eau, vecteurs vitesse).

Une simulation « de référence » a été menée avec un coefficient de frottement (Strickler) de 35. Puis ce coefficient a été « oublié » et les données issues de la simulation de référence ont été fournies au code afin qu'il estime le coefficient de frottement correspondant. Le tir initial du coefficient de Strickler pour ce test a été fixé à 5.

Les résultats de l'estimation de paramètres sont tracés sur la figure 2. On peut voir dans la partie supérieure de l'image un graphe montrant l'évolution de la fonction coût normalisée. Une valeur suffisamment petite de cette fonction coût (et donc un jeu de coefficients de frottement satisfaisant au regard des données de référence) est atteinte au bout de 7 itérations.

Dans la partie inférieure de la figure 2, on voit l'évolution du coefficient de Strickler avec les itérations de l'algorithme. A partir de la valeur initiale de 5, le coefficient de Strickler augmente jusqu'à atteindre 35, ce qui était la valeur recherchée.

Ce cas-test montre donc une bonne précision de l'algorithme mis en œuvre, pour un cas où les données de référence couvrent l'ensemble du domaine.



**Figure 2.** Cas-test pour l'estimation du frottement : canal rectangulaire

### 4.3. Cas réels d'application

#### 4.3.1. Onde M2 en Manche – Mer du Nord

La procédure d'estimation a été appliquée à la simulation de la marée sur un domaine de 520 km x 950 km couvrant la côte Atlantique française, la Manche et la mer du Nord. Le bord ouest de ce domaine se trouve au-delà du plateau continental, à une longitude d'environ 12,5° W. Le maillage comporte 9 414 éléments et 5 007 nœuds et est raffiné près des côtes avec une taille de maille variant entre 10 km et 50 km. Le système de coordonnées est une projection de Mercator (option présente dans Telemac-2D). Quatre marées (c'est-à-dire 181 500 s) sont simulées, en partant d'une surface libre au repos. Sur les frontières liquides, les cotes de la surface libre correspondant à l'onde M2 sont imposées. Ces cotes sont déduites d'une analyse de Fourier des mesures (qui sont des marnages). Le pas de temps est 150 s, avec donc un nombre de courant bien supérieur à 1 si le calcul utilise la célérité des ondes. Le coefficient de diffusion est fixé à zéro et la convection des vitesses est négligée. Une loi de Chézy est utilisée pour le frottement. Les données sont le marnage et la phase de l'onde M2 dans 7 ports et à une bouée mouillée dans l'Atlantique. Le domaine est divisé en 5 zones pour lesquelles le coefficient de frottement est constant. D'après une étude précédente effectuée sans procédure automatique d'estimation et dont le calage précis nécessita plusieurs mois (appelons-la estimation « manuelle »), nous possédions un jeu initial de coefficients. Ces coefficients ont donc été utilisés comme point de départ pour la procédure d'estimation. Le tableau 1 donne les jeux de coefficients de frottements obtenus par calage « manuel » et « automatique ».

Numéro de la zone	Estimation manuelle	Estimation automatique
1	60	54.3
2	45	47.2
3	90	95.3
4	73	62.2
5	60	76.5

**Tableau 1.** Jeux de coefficients de frottement (Chézy) manuels et automatiques

Les tableaux 2 et 3 montrent le marnage et la phase de l'onde M2 et comparent les mesures, le calage manuel et les résultats de la procédure automatique d'estimation.

Avec la procédure automatique, l'erreur est réduite de 16 cm (Saint-Malo) à 12 cm (Calais). L'erreur sur la phase est réduite de 7° (Zeebrugge) à 5° (Dieppe). Il faut une centaine d'itérations, c'est-à-dire une centaine de calculs directs plus une centaine de calculs du système adjoint pour atteindre un minimum. Ceci est rendu

possible par le fait que le temps de calcul pour 1 itération (calcul de 4 marées) prend environ 10 mn sur un PC cadencé à 1,5 GHz.

Lieu	Mesure	Estimation manuelle	Estimation automatique
Zeebrugge	1.59 m	1.50 m	1.57 m
Calais	2.46 m	2.35 m	2.34 m
Boulogne	2.93 m	3.04 m	2.89 m
Dieppe	3.08 m	3.17 m	3.00 m
Cherbourg	1.87 m	1.96 m	1.95 m
Saint-Malo	3.74 m	3.58 m	3.64 m
Balise C1	1.29 m	1.29 m	1.29 m
Ouessant	2.06 m	1.96 m	1.97 m

**Tableau 2.** *Marnage de l'onde M2*

Lieu	Mesure (degrés)	Estimation manuelle	Estimation automatique
Zeebrugge	14	21	16
Calais	345	347	345
Boulogne	331	333	334
Dieppe	311	312	316
Cherbourg	228	227	228
Saint-Malo	178	179	181
Balise C1	103	105	105
Ouessant	111	114	115

**Tableau 3.** *Phase de l'onde M2*

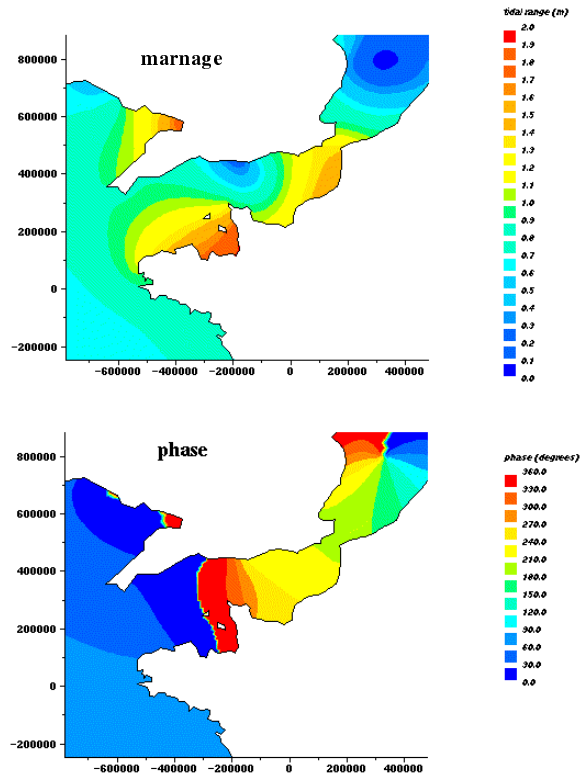
Nous voyons donc qu'en partant d'un calage manuel du frottement, il est possible de l'améliorer par une procédure automatique. Cependant, d'autres tests furent effectués en partant de coefficients de frottement éloignés de la solution. Dans ce cas, un minimum local différent est parfois trouvé, d'une qualité moindre.

La figure 3 montre les résultats en termes de marnage et de phase dans le domaine en son entier, obtenus pour le meilleur calage.

La figure 4 représente le maillage et la bathymétrie (figure du haut) et les zones choisies avec leur valeur correspondante pour le coefficient de Chézy.



**Onde de marée M2 dans la Manche : marnage et phase**



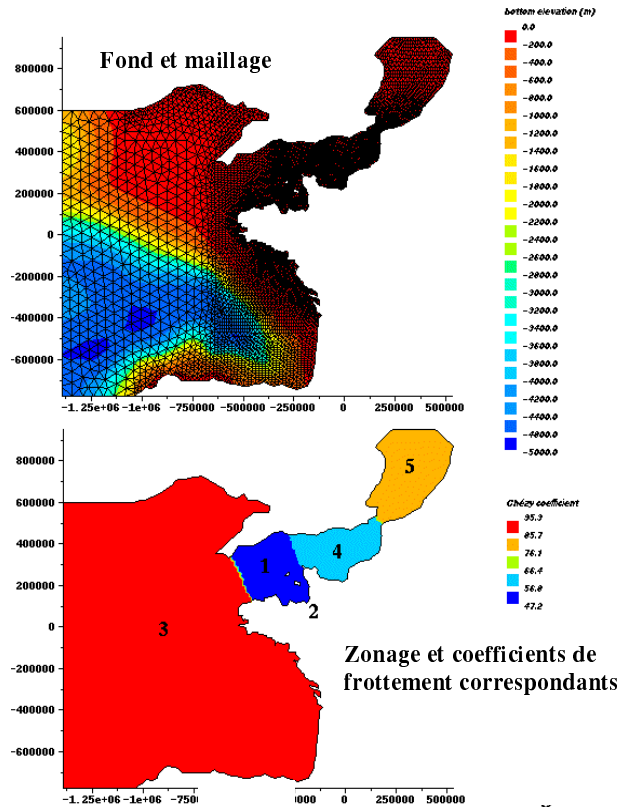
**Figure 3.** *Marnage et phase obtenus avec le meilleur calage*

**4.3.2. Crue fluviale en Loire moyenne**

La méthode d'estimation a été testée sur un autre cas réel : le calage du frottement pour un modèle local d'évaluation de risque d'inondation vis-à-vis d'une centrale nucléaire située sur la Loire moyenne. Ce modèle s'étend sur 40 km de Loire environ, comporte 16154 nœuds et 31799 éléments triangulaires.

La figure 5 montre la bathymétrie du modèle (fait avec Telemac-2D).

**Onde de marée M2 dans la Manche : fond, maillage et coefficients de frottement**



**Figure 4.** Bathymétrie, maillage, zonage et valeurs du coefficient de frottement (Chézy)

La figure 6 montre le maillage réalisé, ainsi que des zooms sur le raffinement du maillage autour des levées entourant le lit mineur (en bas à gauche), et autour des digues de protection de la centrale nucléaire (en haut à droite).

La procédure d'estimation a été lancée en utilisant une seule donnée de surface libre à l'amont. Le zonage, minimal, ne comportait qu'une seule zone de coefficients de frottement sur tout le lit mineur du modèle (le calage a été fait sur une crue non débordante).

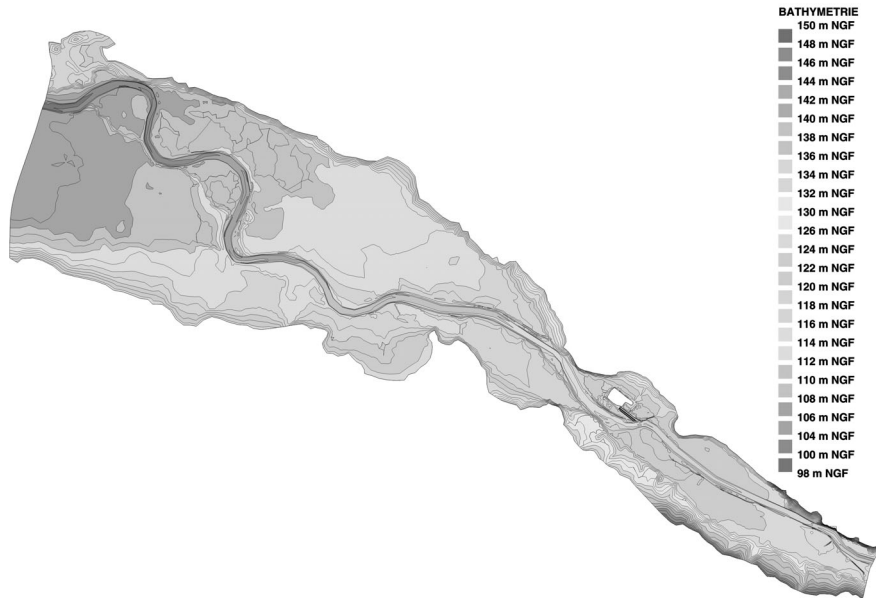


Figure 5. Bathymétrie pour le modèle local de la Loire

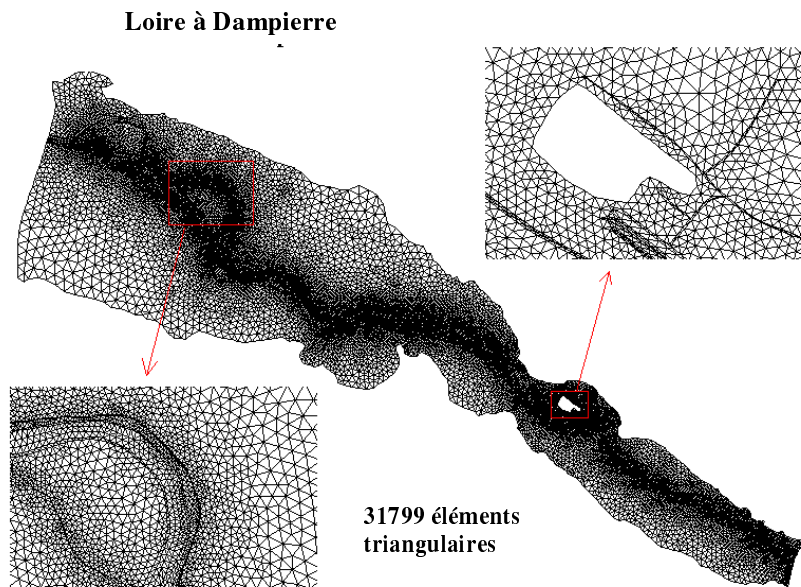


Figure 6. Maillage et détails du maillage pour le modèle local de la Loire

La figure 7 montre les résultats de la procédure d'estimation. On peut voir l'évolution du coût normalisé (traits pointillés larges) en fonction du nombre d'itérations nécessaires à la procédure d'identification. On peut constater que, sur le cas présenté, la fonction coût relative atteint en 7 itérations le critère de convergence fixé.

L'erreur sur la hauteur d'eau (pointillés fins), égale à 1,88 m initialement, décroît jusqu'à atteindre une valeur tout à fait satisfaisante de 1 cm à la fin de la procédure.

En trait plein est tracée l'évolution de la valeur du coefficient de frottement pour l'unique zone définie. Le tir initial était à un coefficient de Strickler de 15, pour une valeur finale de 32,3.

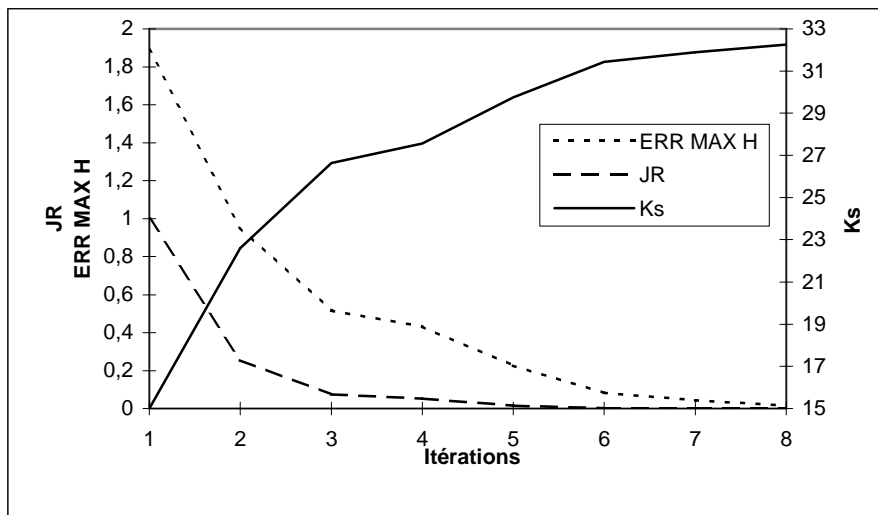


Figure 7. Résultats du calage automatique pour le modèle local de la Loire

## 5. Perspectives : le projet CLACIC

### 5.1. Contexte, partenaires

Le projet CLACIC (CaLAge de Codes par Ingénierie des Connaissances), subventionné par le Réseau National des Technologies Logicielles (émanation du Ministère de la Recherche et de l'Industrie), a débuté le 1<sup>er</sup> décembre 2002 pour une durée de 2 ans. Les partenaires du projet sont :

- Les industriels CEA/DEN, CETMEF et EDF/ LNHE : expérience du calage de codes de calcul dédiés respectivement à la simulation de la sûreté des installations

nucléaires, la simulation de manœuvres de navires et la simulation numérique des crues ;

– Le LARIA, laboratoire de recherche de l'Université de Picardie Jules Verne, et l'association VERNE-ADER : travail dans le domaine de la modélisation des connaissances ;

– Les sociétés DeltaCAD et DIVERGENT (UTC) : réalisation de logiciels scientifiques et techniques dans le domaine de la CAO et la simulation.

## **5.2. Objectifs et innovation**

Les objectifs de ce projet sont les suivants :

– Définir une méthodologie de calage des codes de calcul associant méthodes mathématiques et connaissances ;

– Développer une architecture générique (atelier) d'assistance au calage ;

– Développer des applications spécifiques pour chaque industriel partenaire du projet à partir de l'outil générique ;

– Spécifier et développer des bases de données de résultats de calage (Bases de Cas) dans le but d'archiver des données mais aussi de faciliter les calages futurs.

Le principe est que l'assistant de calage s'appuie sur une bibliothèque de méthodes de résolution. Ces méthodes de résolution seront de plusieurs types : des méthodes numériques, des règles d'expertise, des méthodes de raisonnement à Base de Cas.

On voit donc que l'originalité principale du projet est l'approche combinée symbolique – numérique pour la résolution d'un problème complexe, celui du calage d'un code de calcul, généralement traité par une approche exclusivement « numérique ». Un apport important sera notamment l'appui sur une Base de Cas de systèmes calés pour guider la réalisation de la tâche de calage.

Les retombées scientifiques attendues sont l'expérience d'une méthode hybride associant méthodes numériques et ingénierie des connaissances pour la résolution d'une tâche.

Les retombées industrielles attendues sont de trois types : la diffusion des trois applications spécifiques de calage, la diffusion de l'application générique de calage, et enfin la diffusion des bases de données associant le système physique et le jeu de coefficients de calage.

## **6. Conclusion**

L'identification de paramètres dans les équations de Saint-Venant, grâce à la technique des systèmes adjoints, offre la possibilité de réduire l'une des tâches les

plus longues et coûteuses dans les études : le calage. La méthode mise en œuvre dans Telemac-2D et testée sur plusieurs cas schématiques et réels a montré son intérêt et a fourni des résultats satisfaisants.

De plus, grâce au caractère générique du système adjoint, la théorie peut être étendue facilement à l'estimation simultanée de plusieurs paramètres différents, tels que le frottement et les conditions aux limites.

Cependant, la principale limitation actuelle de la méthode est que les techniques de minimisation employées (comme le gradient conjugué) ne permettent de trouver qu'un minimum local. C'est pourquoi un avis d'expert sur le résultat obtenu est toujours nécessaire, d'où la réflexion en cours sur des approches qui associeraient des procédures de minimisation classiques et des systèmes experts, à travers le projet CLACIC.

De plus, la définition des zones de frottement uniforme peut générer des cas mal posés où l'algorithme ne converge pas.

Enfin, un problème plus global est soulevé par les incertitudes dans les données. L'étape ultérieure de développement sera le recours à des techniques de filtrage de Kalman pour traiter ces incertitudes. Cette technique a déjà été utilisée avec succès au LNHE pour des solveurs unidimensionnels (N'Kaoua, 1985).

## 7. Bibliographie

- Barros E., Estimation de paramètres dans les équations de Saint-Venant. Thèse de l'Université Pierre et Marie Curie (Paris VI). Spécialité Mathématiques Appliquées, 1996.
- Chavent G., Analyse fonctionnelle et identification de coefficients répartis dans les équations aux dérivées partielles. Thèse de la Faculté des Sciences de Paris, spécialité Mathématiques, 1971.
- N'Kaoua T., Identification du coefficient de Strickler dans les équations de Saint-Venant. Thèse de l'Université Pierre et Marie Curie (Paris VI). Spécialité Mathématiques Appliquées, 1985.