



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *Institut National Polytechnique de Toulouse*
Discipline ou spécialité : *Hydrologie*

Présentée et soutenue par *BESSIERE Hélène*
Le *18 décembre 2008*

Titre :
*Assimilation de données variationnelle
pour la modélisation hydrologique distribuée
des crues à cinétique rapide*

JURY

Isabelle BRAUD : Directrice de Recherche au CEMAGREF LYON
Denis DARTUS : Professeur de l'INP TOULOUSE
Jacques GEORGE : Professeur de l'INP TOULOUSE
David LABAT : Maître de conférences de l'UNIVERSITÉ TOULOUSE 3
Roger MOUSSA : Directeur de recherche à l'INRA MONTPELLIER
Catherine OTTLE : Directrice de recherche CNRS
Hélène ROUX : Maître de conférences de l'INP TOULOUSE
Jeffrey P. WALKER : Professeur associé de l'UNIVERSITÉ DE MELBOURNE

Ecole doctorale : Sciences de l'Univers, de l'Environnement et de l'Espace
Unité de recherche : Institut de Mécanique des Fluides de Toulouse
Directeur(s) de Thèse : Denis DARTUS / David LABAT
Rapporteurs : *Isabelle BRAUD / Catherine OTTLE / Roger MOUSSA*

Résumé

Dans la région Méditerranéenne les crues « éclair » représentent un risque des plus destructeurs ayant coûté la vie à de nombreuses personnes et entraînant des dommages matériels importants ces dernières années (Nîmes en 1988, Vaison-la-Romaine en 1992, Aude en 1999, Gard en 2002). Le groupe HydroEco de l'IMFT (Institut de Mécanique des Fluides de Toulouse) a développé le code MARINE (Modélisation de l'Anticipation du Ruissellement et des Inondations pour des événements Extrêmes) qui est un modèle hydrologique distribué à base physique pour la prévision de ce type de crues.

Une méthode d'assimilation variationnelle de données, la méthode de l'état adjoint, est implémentée dans le modèle MARINE. Cette méthode considère les paramètres d'entrée comme des variables de contrôle du modèle et les optimise en minimisant les écarts entre les observations et les simulations par l'intermédiaire d'une fonction coût. Le bassin d'étude est le bassin versant du Gardon d'Anduze sur lequel nous disposons d'une base de données d'évènements extrêmes.

Une analyse de sensibilité préliminaire permet d'identifier les paramètres les plus sensibles du modèle afin de réduire l'équifinalité. Une première application de la méthode de l'état adjoint est alors utilisée pour la calibration du modèle. Elle permet de tester la validité des hypothèses physiques en évaluant les performances du modèle puis de suggérer des pistes d'amélioration. Un jeu de paramètres consistant, cohérent avec les résultats de simulations utilisant une méthode de Monte Carlo, est alors identifié. Les hydrogrammes de crues obtenus sont compris dans une plage d'incertitude réduite.

La seconde application de cette méthode s'intéresse à la prévision des crues par l'emploi de deux stratégies différentes. La première consiste à utiliser les informations à l'exutoire pendant la phase de montée de crue pour prédire ce pic de crue. La seconde consiste à utiliser les observations d'une station intermédiaire à l'amont du bassin pour prédire le débit à l'exutoire. Ces deux stratégies permettent une anticipation du pic de crue de quelques heures, résultat essentiel d'un point de vue opérationnel.

Mots clefs : Crues éclair, modèle hydrologique distribué à base physique, méthode de l'état adjoint, bassin du Gardon d'Anduze, estimation de paramètres, prévision en temps réel.

Abstract

Intense flash floods, produced by severe thunderstorms, represent the most destructive hazard in the Mediterranean region: Nîmes in 1988, Vaison-la-Romaine in 1992, Aude in 1999, Gard in 2002. Human lives and properties are increasingly exposed to flood risks. Therefore, there is a clear need to improve real-time monitoring and flood forecasting. The HydroEco research group at IMFT (Institute of Fluid Mechanics of Toulouse) developed the MARINE model (Modélisation de l'Anticipation du Ruissellement et des Inondations pour des événements Extrêmes), a distributed physically-based model, in order to forecast this kind of flood.

A data assimilation process called the adjoint state method is implemented into the MARINE model. The principle of the adjoint method is to consider input parameters as control variables and optimize them by minimizing a cost function measuring the discrepancy between observations and model simulations. The procedure is tested on the Mediterranean Gardon d'Anduze catchment.

A previous sensitivity analysis allows identifying the most sensitive parameters in order to avoid the "non-uniqueness problem" and overparameterization. Then a first application of the adjoint method aims at calibrating the model. The method enables to evaluate the model performance and contributes to either the validation of physical hypothesis or perspectives for the model structure improvements. A consistent set of parameters, in accordance with results achieved with a Monte-Carlo procedure, is then identified and simulated hydrographs are limited into a reduced range of uncertainties.

The second application of the method is dedicated to flash flood forecasting. The first strategy employed is to use outlet flow information during the rising flood phase to estimate the peak. The second strategy employed is to use information of intermediate upstream gauging stations to predict the outlet flow. Both strategies allow anticipating, a few hours in advance, the peak which is crucial from an operational point of view.

Keywords: Flash flood, distributed physically-based model, adjoint model, Gardon d'Anduze catchment, parameter estimation, real-time forecasting

Table des matières

Table des figures	ix
Liste des tableaux	xxi
Introduction générale	2
1 La modélisation hydrologique et la problématique des crues éclair	8
1.1 Introduction	8
1.2 La modélisation hydrologique	9
1.2.1 Typologie des modèles	9
1.2.2 Problème d'échelle des modèles distribués	11
1.2.3 Procédure de modélisation	12
1.3 Les crues « éclair »	17
1.3.1 Processus de genèse des crues éclair	17
1.3.2 La prévision de crues à l'échelle nationale	21
1.3.3 Apport de la modélisation hydrologique pour une prévision opérationnelle	25
1.3.4 Apport de la distribution spatiale de l'information	27
1.4 Synthèse	30
2 Assimilation de données variationnelle	34
2.1 Théorie générale de l'assimilation de données	34
2.2 Assimilation de données variationnelle appliquée à la modélisation hydrologique	34
2.2.1 Introduction	34
2.2.2 Application des méthodes d'assimilation de données	35
2.2.3 La méthode de l'état adjoint : une méthode variationnelle	36
2.2.4 Synthèse	45

3	Procédures d'analyse des modèles	48
3.1	Analyse de sensibilité	48
3.1.1	Introduction	48
3.1.2	Méthode globale	49
3.1.3	Méthode locale	52
3.1.4	Conclusion	59
3.2	Optimisation	61
3.2.1	Introduction	61
3.2.2	Critère d'évaluation des performances d'un modèle	62
3.2.3	Equifinalité	64
3.2.4	Différentes stratégies d'optimisation des modèles hydrologiques	64
3.2.5	Algorithme d'optimisation utilisant la méthode de l'état adjoint	68
3.2.6	Méthodologie des expériences jumelles	69
3.2.7	Synthèse	73
4	Présentation du modèle, du bassin versant, des observations et des critères de performance	76
4.1	MARINE : un modèle distribuée à base physique pour la modélisation des crues à cinétiques rapide	76
4.1.1	Introduction	76
4.1.2	Données nécessaires au modèle MARINE	77
4.1.3	Equations représentatives du comportement des processus physiques	78
4.1.4	Prédétermination des paramètres du modèle	89
4.1.5	Synthèse	92
4.2	Bassin d'étude : le Gardon d'Anduze	93
4.2.1	Introduction	93
4.2.2	Description	93
4.2.3	Caractéristiques	94
4.2.4	Base de données disponibles	100
4.3	Les événements sélectionnés	108

4.3.1	Critères de sélection	108
4.3.2	Caractéristiques des événements sélectionnés	108
4.3.3	Critiques des données	109
4.3.4	Bilan des événements sélectionnés	113
4.4	Validité de l'approximation de l'onde cinématique sur le bassin versant	114
4.5	Formulation des critères de performance	115
4.5.1	Un critère classique	115
4.5.2	Un critère généralisé	115
4.6	Synthèse	117
5	Calibration du modèle et prise en compte des incertitudes sur le débit	120
5.1	Introduction	120
5.2	Calibration du modèle	121
5.2.1	Détermination des paramètres de calibration : analyse de sensibilité et expériences jumelles	121
5.2.2	Résultats des simulations de Monte-Carlo	129
5.2.3	Calibration avec la méthode d'optimisation	138
5.2.4	Synthèse	161
5.3	Prise en compte des incertitudes sur le débit	162
5.3.1	Résultats des simulations de Monte-Carlo	162
5.3.2	Utilisation de la méthode d'optimisation	168
5.3.3	Synthèse	172
5.4	Synthèse de la calibration et validation	173
5.4.1	Agglomération des résultats	173
5.4.2	Validation	174
5.5	Conclusion	178
6	Estimation de paramètres et assimilation de données	182
6.1	Introduction	182
6.2	Assimilation des observations partielles à Anduze	183
6.2.1	Positionnement du problème	183

6.2.2	Résultats	183
6.2.3	Bilan	189
6.3	Assimilation des observations partielles à Saumane	190
6.3.1	Objectif	190
6.3.2	Résultats de l'estimation des paramètres	192
6.3.3	Bilan	216
6.4	Conclusion	217
7	Contributions de l'étude et perspectives	220
7.1	Contributions de l'étude	220
7.1.1	Modélisation des écoulements de subsurface	220
7.1.2	Loi logarithmique du coefficient de Strickler	221
7.2	Perspectives d'étude	223
7.2.1	L'analyse de sensibilité – ASF	223
7.2.2	Synthèse	228
	Conclusion générale	232
	Bibliographie	238
	Annexes	255
A	Chapitre 10 du guide OMM – Assimilation de données	257
B	Compléments au chapitre 4	307
C	Compléments au chapitre 5	311

Table des figures

Fig. 1.1 :	Définition des trois caractéristiques d'échelle selon Blöschl et Sivapalan (1995) (extent, spacing et support)	12
Fig. 1.2 :	Procédure (standard) simplifiée pour la modélisation pluie-débit d'après Wagener et al. (2004)	13
Fig. 1.3 :	Relation entre la complexité du modèle, la disponibilité des données et la performance du modèle selon Grayson et Blösch (2000)	14
Fig. 1.4 :	Stades de dégradation de la surface du sol d'après Boiffin (1984)	18
Fig. 1.5 :	Habitations proches du lit mineur à Anduze (photo du haut) et maisons dans le lit majeur (photo du bas) en amont d'Anduze (département du Gard).	21
Fig. 1.6 :	La gestion du risque d'inondation d'après Estupina (2004)	22
Fig. 1.7 :	Le réseau ARAMIS de Météo-France	24
Fig. 1.8 :	Comparaison entre les cumuls de pluie radar de Bollène : (a) ancienne lame d'eau HYDRAM, (b) nouvelle lame d'eau PANTHERE. (c) est l'image des poids pour chaque pixel, d'après Parent du Châtelet et al., 2005.	24
Fig. 1.9 :	Test de spatialisation de la hauteur d'infiltration	28
Fig. 1.10 :	Données de pluie spatialisée pour la crue du 08/09/2002 sur le Gardon d'Anduze.	29
Fig. 1.11 :	Test de spatialisation de la pluie.	29
Fig. 2.1 :	Comparaison du modèle direct (schéma du haut) avec le modèle couplé à une procédure de mise à jour (schéma du bas) d'après Drécourt (2004).	35
Fig. 2.2 :	Tests du produit scalaire et test du gradient pour la validation du code adjoint du modèle MARINE.	41

Fig. 2.3 :	Approches possibles pour l'implémentation du code adjoint	41
Fig. 2.4 :	Arborescence d'un code de calcul nécessaire à la différentiation automatique.	42
Fig. 3.1 :	Principe de l'analyse de sensibilité généralisée, d'après Hornberger et Spear (1981)	51
Fig. 3.2 :	Tests du produit scalaire et test du gradient pour la validation du code adjoint de MARINE.	71
Fig. 3.3 :	Schéma représentant le processus d'estimation des paramètres du modèle MARINE avec la méthode adjoint.	72
Fig. 4.1 :	Schématisation du modèle MARINE	77
Fig. 4.2 :	Schématisation de l'infiltration, d'après Green et Ampt (1911)	79
Fig. 4.3 :	Géométrie du profil en travers du drain	85
Fig. 4.4 :	Flux à travers les mailles du réseau de drainage	86
Fig. 4.5 :	Domaine de validité de l'application de l'onde cinématique pour la modélisation du ruissellement de surface, d'après Moore et Foster (1990)	88
Fig. 4.6 :	Critère de validité de l'approximation de l'onde cinématique (zone ombrée) (l'intersection des régions provident de Lighthill and Whitham (1955); Morris and Woolhiser (1980); Woolhiser and Liggett (1967)); et (gris clair) de Daluz Vieira (1983), d'après Singh (1996).	88
Fig. 4.7 :	Pourcentage d'argile, de limon et de sable dans les textures classiques de sol.	91
Fig. 4.8 :	Situation géographique des Cévennes (source IGN)	93
Fig. 4.9 :	Réseau de drainage du Gardon d'Anduze	94
Fig. 4.10 :	L'amont du Gardon d'Anduze est caractérisé par des hautes montagnes.	96
Fig. 4.11 :	Les vallées encaissées du Gardon d'Anduze.	96
Fig. 4.12 :	A l'aval le lit mineur s'élargie à l'intérieur d'une plaine alluviale	96
Fig. 4.13 :	Topographie du Gardon d'Anduze	97
Fig. 4.14 :	Pentes du Gardon d'Anduze	97
Fig. 4.15 :	Epaisseur de sol du Gardon d'Anduze	97

Fig. 4.16 :	Occupation du sol du Gardon d'Anduze	97
Fig. 4.17 :	Chronologie des principales crues sur le Gardon à Anduze	99
Fig. 4.18 :	Echelles limnimétriques à Anduze (à gauche) et Mialet (à droite).	100
Fig. 4.19 :	Composition du sol du Gardon d'Anduze	106
Fig. 4.20 :	Distribution spatiale de la porosité sur Gardon d'Anduze	106
Fig. 4.21 :	Distribution spatiale de la conductivité hydraulique	106
Fig. 4.22 :	Distribution spatiale de la force de succion	106
Fig. 4.23 :	Evolution de l'humidité initiale réduite moyenne sur le Gardon d'Anduze de 1981 à 2006	107
Fig. 4.24 :	Evolution annuelle de la moyenne de l'humidité initiale réduite sur le Gardon d'Anduze de 1994 à 2002	107
Fig. 4.25 :	Principales sources d'erreurs pour les mesures des précipitations par radar : échos de sol, bande brillante et profil vertical de réflectivité, d'après Parent du Châtelet (2005)	111
Fig. 4.26 :	Interpolation des débits sur certaines plages horaires pour la crue du 17/05/1999 (événement n°11) à Anduze.	112
Fig. 4.27 :	Pour l'événement n°2, l'hydrogramme observé ne présente aucune réaction au second pic de pluie, pourtant assez intense. Il y a probablement un problème au niveau de la mesure des pluies ou des débits.	113
Fig. 4.28 :	Validité de l'approximation de l'onde cinématique pour la crue de Septembre 2002, à 26h, d'après Singh (1996) : la zone en noir correspond aux pixels avec une valeur de $F_0^2 k_k$ inférieur à 5.	114
Fig. 4.29 :	Représentation de J1 en fonction de Q_{sim}/Q_{obs}	115
Fig. 4.30 :	Représentation de J2 en fonction de Q_{sim}/Q_{obs}	116
Fig. 4.31 :	Les cercles correspondent aux débits observés et les traits discontinus représentent la plage des débits acceptables avec une erreur de $\pm 20\%$ sur les valeurs observées.	116
Fig. 5.1 :	Schéma représentant le processus d'estimation des paramètres du modèle MARINE avec la méthode adjoint.	121

Fig. 5.2 :	Convergence des paramètres C_K (a), C_{Sf} (b), C_n (c), C_Z (d), C_θ (e), n_d (f) et de la fonction coût (g) pour les deux initialisations. Hydrogrammes initiales (trait vert continu et discontinu) et assimilé (trait noir) comparés à l'hydrogramme de référence (points rouges) (h) – test I.	123
Fig. 5.3 :	Hydrogrammes issus de trois simulations avec un produit $C_{\theta s} \times C_Z$ constant pour la crue de Septembre 1994.	125
Fig. 5.4 :	Influence directe de C_{Sf} pour la crue de Septembre 1994.	125
Fig. 5.5 :	Convergence des paramètres C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) pour les deux initialisations. Hydrogrammes initiales (trait vert continu et discontinu) et assimilé (trait noir) comparés à l'hydrogramme de référence (points rouges) (e) – test II.	125
Fig. 5.6 :	Valeurs du coefficient de Nash obtenues avec les simulations de Monte-Carlo pour les paramètres C_K , C_Z et n_d pour les événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g) et n°25 (h).	134
Fig. 5.7 :	Tracé des cent meilleures simulations (traits noirs) comparés aux observations (points rouges) pour les événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g) et n°25 (h).	136
Fig. 5.8 :	Bilan des intervalles de valeurs acceptables pour les paramètres C_K et C_Z avec le critère de Nash.	137
Fig. 5.9 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°1.	139
Fig. 5.10 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu – événement n°1.	141
Fig. 5.11 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) – événement n°1.	141
Fig. 5.12 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°3.	142

Fig. 5.13 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°3.	142
Fig. 5.14 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°3.	143
Fig. 5.15 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°4.	144
Fig. 5.16 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°4.	145
Fig. 5.17 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°4.	145
Fig. 5.18 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°11.	146
Fig. 5.19 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°11.	147
Fig. 5.20 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°11.	147
Fig. 5.21 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°13.	149
Fig. 5.22 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°11.	150
Fig. 5.23 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°11.	150

Fig. 5.24 :	Hydrogrammes avant optimisation (trait vert), après 2 itérations (trait mauve), après 9 itérations (trait bleu) et après 14 itérations (trait noir) comparés aux observations (points rouges) pour le test 1- événement n°13.	151
Fig. 5.25 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°15.	152
Fig. 5.26 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°15.	153
Fig. 5.27 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°15.	153
Fig. 5.28 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°19.	154
Fig. 5.29 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°19.	155
Fig. 5.30 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°19.	155
Fig. 5.31 :	Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°25.	156
Fig. 5.32 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°25.	157
Fig. 5.33 :	Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°25.	158
Fig. 5.34 :	Loi du coefficient de Strickler en fonction du débit de pointe.	160

Fig. 5.35 :	Tracé des couples (C_K, C_Z) , (C_K, n_d) et (C_Z, n_d) correspondant aux 200 meilleures simulations de Monte-Carlo pour le critère de Nash (points noirs) et le Nash généralisé (points verts) des événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g) et n°25 (h).	164
Fig. 5.36 :	Comparaison des simulations de Monte-Carlo pour le critère de Nash et le Nash généralisé pour les paramètres C_K , C_Z et n_d pour les événements n°3 (a) et n°19 (b).	166
Fig. 5.37 :	Bilan des intervalles de valeurs acceptables pour les paramètres C_K et C_Z avec le critère de Nash généralisé.	167
Fig. 5.38 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu.	168
Fig. 5.39 :	Hydrogrammes avant (trait vert discontinu) et après optimisation (trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°3.	169
Fig. 5.40 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu.	170
Fig. 5.41 :	Hydrogrammes avant (trait vert discontinu) et après optimisation (trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°19	170
Fig. 5.42 :	Choix du couple (C_K, C_Z) par rapport aux simulations de Monte-Carlo avec le critère de Nash (a) et le Nash généralisé (b).	173
Fig. 5.43 :	Validation du couple de paramètres $(C_K, C_Z) = (6.5, 3.5)$ et comparaison des coefficients de Manning estimés (trait vert) et extrapolés (trait noir) pour les événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g), n°25 (h), n°2 (i) et n°8 (j).	177

Fig. 6.1 :	Résultat de l'assimilation (trait noir) et prévision de l'hydrogramme (trait bleu discontinu) comparées aux observations (points rouges) pour : (a) l'événement n°1, (b) l'événement n°4 – t est 1, (c) l'événement n°4 – test 2, (d) l'événement n°11, (e) l'événement n°13, (f) l'événement n°15, (g) l'événement n°19 – test 1, (h) l'événement n°19 – test 2 et (i) l'événement n°25.	188
Fig. 6.2 :	Hydrogrammes simulés à Mialet (a) et Saumane (b) à l'issu d'un premier calage manuel comparés aux observations – crue de Septembre 1994.	190
Fig. 6.3 :	Localisation de la station hydrométrique de Saumane	191
Fig. 6.4 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) (a) et total des pluies cumulées – événement n°1.	192
Fig. 6.5 :	Convergences de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'assimilation des observations partielles à Saumane pour le test 1 (noir) et le test 2 (bleu) – événement n°1.	193
Fig. 6.6 :	Hydrogrammes avant assimilation (trait noir discontinu) et après assimilation (trait noir continu) comparés aux observations (points rouges) pour le test 1 (a) et simulation après 40 h (trait bleu) à Saumane (b) – événement n°1.	194
Fig. 6.7 :	Prévisions après 40 h (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé à Saumane (a) et avec correction du coefficient de Manning (b) comparées aux observations (points rouges) – événement n°1.	194
Fig. 6.8 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) (a) et cumul des pluies (b) – événement n°3.	196
Fig. 6.9 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'assimilation pour les deux différents tests d'initialisation : test 1 en noir, test 2 en bleu – événement n°3 à Saumane.	197

Fig. 6.10 :	Hydrogrammes avant assimilation (trait noir discontinu) et après assimilation (trait noir continu) comparés aux observations (points rouges) pour le test 1 (a) et simulation après 40 h (trait bleu) à Saumane (b) – événement n°3.	198
Fig. 6.11 :	Prévisions après 40 h (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé à Saumane comparées aux observations (points rouges) – événement n°3.	198
Fig. 6.12 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) (a) et cumul des pluies (b) – événement n°11.	199
Fig. 6.13 :	Convergences de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'assimilation des observations partielles à Saumane pour le test 1 (noir) et le test 2 (bleu) – événement n°1.	200
Fig. 6.14 :	(a) Hydrogrammes avant assimilation (trait noir discontinu) et après assimilation (trait noir continu) comparés aux observations (points rouges) pour le test 1 et (b) simulation après 18 h (trait bleu) à Saumane – événement n°11.	201
Fig. 6.15 :	Prévisions après 18 h (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé à Saumane (a) et correction du coefficient de Manning (b) comparées aux observations (points rouges) – événement n°11.	201
Fig. 6.16 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) (a) et cumul des pluies (b) – événement n°13.	203
Fig. 6.17 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'assimilation pour les deux différents tests d'initialisation : test 1 en noir, test 2 en bleu – événement n°13 à Saumane.	204
Fig. 6.18 :	(a) Hydrogrammes avant assimilation (trait noir discontinu) et après assimilation (trait noir continu) comparés aux observations (points rouges) pour le test 1 et (b) simulation après 30 h (trait bleu) à Saumane – événement n°13.	204

Fig. 6.19 :	Prévisions après 30 h (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé à Saumane (a) et correction du coefficient de Manning (b) comparées aux observations (points rouges) – événement n°13.	205
Fig. 6.20 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) – événement n°19.	206
Fig. 6.21 :	Cumul des pluies de la phase 1 (a), de la phase 2 (b), de la phase 3 (c) et cumul total (d) – événement n°19.	207
Fig. 6.22 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'assimilation pour les deux différents tests d'initialisation : test 1 en noir, test 2 en bleu – événement n°19 à Saumane.	208
Fig. 6.23 :	(a) Hydrogrammes avant assimilation (trait noir discontinu) et après assimilation (trait noir continu) comparés aux observations (points rouges) pour le test 1 et (b) simulation après 28h40 (trait bleu) à Saumane – événement n°19.	209
Fig. 6.24 :	Prévisions après 28h40 (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé à Saumane (a) et correction du coefficient de Manning (b) comparées aux observations (points rouges) – événement n°19.	209
Fig. 6.25 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) (a) et cumul des pluies (b) – événement n°25.	211
Fig. 6.26 :	Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'assimilation pour les deux différents tests d'initialisation : test 1 en noir, test 2 en bleu – événement n°25 à Saumane.	211
Fig. 6.27 :	(a) Hydrogrammes avant assimilation (trait noir discontinu) et après assimilation pour le test 1 (trait noir continu) et le test 2 (trait jaune) comparés aux observations (points rouges) et (b) simulation après 33h10 (trait bleu) à Saumane – événement n°25.	212

Fig. 6.28 :	Prévisions après 23h10 (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé au test 1 à Saumane (a) et correction du coefficient de Manning (b) comparées aux observations (points rouges) – événement n°25.	212
Fig. 6.29 :	Prévisions après 23h10 (trait bleu discontinu) de l'hydrogramme de crue à Anduze avec le jeu de paramètres estimé au test 2 à Saumane (a) et correction du coefficient de Manning (b) comparées aux observations (points rouges) – événement n°25.	213
Fig. 6.30 :	Comparaison entre les hydrogrammes observés à Saumane (trait vert) et à Anduze (trait rouge) (a) et cumul des pluies (b) – événement n°2.	214
Fig. 6.31 :	Comparaison des hydrogrammes avant (trait vert) et après (trait noir) assimilation des observations jusqu'à 9h12 à Saumane (a) et prévision après 9h12 (trait noir) de l'hydrogramme de crue à Anduze comparées aux observations (points rouges) – événement n°3.	215
Fig. 7.1 :	Apport des écoulements de subsurface – crue de Septembre 1994.	221
Fig. 7.2 :	Influence de la variation du coefficient de Manning en fonction de l'épaisseur de la lame d'eau – crue de Septembre 2002.	222
Fig. 7.3 :	Sensibilité de g à la conductivité hydraulique K pour les événements n°1 (a), n°13 (b) et n°19 (c).	225
Fig. 7.4 :	Sensibilité de g à la profondeur de sol Z pour les événements n°1 (a), n°13 (b) et n°19 (c).	226
Fig. 7.5 :	Sensibilité de g au coefficient de Manning du lit majeur n_d et cumul des pluies pour les événements n°1 (a), n°13 (b) et n°19 (c).	227

Liste des tableaux

Tableau 3.1.	Exemples de méthodes locales	66
Tableau 3.2.	Exemples de méthodes globales	67
Tableau 4.1.	Paramètres d'entrée du modèle MARINE	89
Tableau 4.2.	Paramètres de Green et Ampt, d'après Maidment (1993)	90
Tableau 4.3.	Valeurs du coefficient de Manning en fonction de l'occupation du sol, d'après Chow (1959)	91
Tableau 4.4.	Temps de réaction des différents bassins (source : fiche BVNE : Gardons, Hansard et Louis, 2005)	98
Tableau 4.5.	Données hydrologiques des principales crues	99
Tableau 4.6.	Débits caractéristiques en m ³ /s à différentes stations. Qi2, Qi10 et Qi100 sont les débits instantanés de période retour 2, 10 et 100 ans. Les intervalles de confiance sont notés entre crochets.	99
Tableau 4.7.	Données radar météorologiques disponibles	101
Tableau 4.8.	Hydrogrammes disponibles aux trois stations : Anduze, Mialet et Saumane	103
Tableau 4.9.	Caractéristiques des paramètres d'infiltration, de l'épaisseur du sol et du coefficient de frottement de Manning sur le Gardon d'Anduze.	105
Tableau 4.10.	Caractéristiques des crues sélectionnées	109
Tableau 5.1.	Valeurs de référence, initiales et assimilées des six paramètres, de la fonction coût et statistiques d'erreur du test I.	122

Tableau 5.2.	Valeurs de référence, initiales et assimilées des trois paramètres, de la fonction coût et statistiques d'erreur pour le test II.	126
Tableau 5.3.	Intervalles de variation des paramètres utilisés dans les simulations des Monte-Carlo (* la valeur maximale dépend des crues)	129
Tableau 5.4.	Intervalles de variation des paramètres lors de l'optimisation	138
Tableau 5.5.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux – événement n°1.	140
Tableau 5.6.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux – événement n°3.	143
Tableau 5.7.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux– événement n°4.	144
Tableau 5.8.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°11.	148
Tableau 5.9.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°11.	149
Tableau 5.10.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°15.	154
Tableau 5.11.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux- événement n°19.	156
Tableau 5.12.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°25.	157
Tableau 5.13.	Tableau récapitulatif des paramètres sélectionnés	159
Tableau 5.14.	Comparaison des coefficients de Strickler ($1/n_d$) estimé et obtenu par extrapolation logarithmique	160
Tableau 5.15.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux	169
Tableau 5.16.	Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux	171
Tableau 5.17.	Tableau récapitulatif des coefficients sélectionnés	172

Tableau 5.18.	Résultats de la validation du modèle avec le jeu de paramètres $(C_K, C_Z) = (6.5, 3.5)$ et le coefficient de Manning estimé ou extrapolé.	175
Tableau 6.1.	Dates et intensités des pics de crue et des derniers débits assimilés à Anduze pour chaque événement.	185
Tableau 6.2.	Résultats de l'assimilation des observations partielles à Anduze pour chaque événement. Les valeurs de la fonction coût sur la fenêtre d'assimilation et sur l'événement entier sont listées ainsi que les erreurs de prévision des pics de crue pour chaque événement testés.	186
Tableau 6.3.	Comparaison des date et intensité du premier pic de crue respectivement à Saumane et Anduze pour l'événement n°1.	192
Tableau 6.4.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°1.	193
Tableau 6.5.	Valeurs des fonctions coût, des erreurs aux pics de crue et des coefficients de ruissellement obtenus avec le jeu de paramètres estimé à partir des observations partielles à Saumane avec et sans correction du coefficient de Manning n_d – événement n°1.	195
Tableau 6.6.	Comparaison de la date et de l'intensité des pics de crue pour l'événement n°3 respectivement à Saumane et Anduze.	195
Tableau 6.7.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°3.	197
Tableau 6.8.	Comparaison des date et intensité du pic de crue respectivement à Saumane et Anduze pour l'événement n°11.	199
Tableau 6.9.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°11.	200

Tableau 6.10.	Valeurs des fonctions coût, des erreurs aux pics de crue et des coefficients de ruissellement obtenus avec le jeu de paramètres estimé à partir des observations partielles à Saumane avec et sans correction du coefficient de Manning n_d – événement n°1.	202
Tableau 6.11.	Comparaison des date et intensité du premier pic de crue respectivement à Saumane et Anduze pour l'événement n°13.	202
Tableau 6.12.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°13.	204
Tableau 6.13.	Valeurs des fonctions coût, des erreurs aux pics de crue et des coefficients de ruissellement avec le jeu de paramètres estimé à partir des observations partielles à Saumane avec et sans correction du coefficient de Manning n_d – événement n°13.	205
Tableau 6.14.	Comparaison de la date et de l'intensité des pics de crue pour l'événement n°19 respectivement à Saumane et Anduze.	206
Tableau 6.15.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°19.	209
Tableau 6.16.	Valeurs des fonctions coût, des erreurs aux pics de crue et des coefficients de ruissellement avec le jeu de paramètres estimé à partir des observations partielles à Saumane avec et sans correction du coefficient de Manning n_d – événement n°19.	210
Tableau 6.17.	Comparaison de la date et de l'intensité du premier pic de crue pour l'événement n°3 respectivement à Saumane et Anduze.	210
Tableau 6.18.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°25.	212

Tableau 6.19.	Valeurs des fonctions coût, des erreurs aux pics de crue et des coefficients de ruissellement obtenus avec le jeu de paramètres estimé à partir des observations partielles à Saumane avec et sans correction du coefficient de Manning n_d – événement n°25.	213
Tableau 6.20.	Comparaison de la date et de l'intensité des pics de crue pour l'événement n°2 respectivement à Saumane et Anduze.	214
Tableau 6.21.	Résultats de l'assimilation des observations partielles à Saumane pour les deux tests – événement n°2.	215

Introduction générale

Introduction générale

Contexte de l'étude

Les crues éclair sont définies comme des crues soudaines, de fortes intensités et de courte durée dues à de violentes pluies localisées dont les impacts humains et socio-économiques sont lourds de conséquence (IAHS-UNESCO-WMO, 1974), elles sont parfois aussi appelées crues à cinétique rapide. Dans la région Méditerranéenne les crues « éclair » représentent un risque des plus destructeurs ayant coûté la vie à de nombreuses personnes et entraînant des dommages matériels importants ces dernières années (Nîmes en 1988, Vaison-la-Romaine en 1992, Aude en 1999, Gard en 2002). La dernière crue de Septembre 2002 témoigne encore de la gravité de ces phénomènes.

Face à l'augmentation de la vulnérabilité des personnes et des biens due au risque d'inondation, la prévision hydrologique est de plus en plus nécessaire. Cependant, les processus hydrologiques responsables de la genèse de ces crues sont très variables et difficiles à prévoir car d'une part, les pluies responsables sont très variables en temps et en espace et d'autre part, les bassins versants concernés sont généralement de petite taille avec des temps de réaction extrêmement rapides.

Plusieurs raisons ont justifié le choix de l'utilisation du modèle MARINE (Modélisation de l'Anticipation du Ruissellement et des Inondations pour des Evénements Extrêmes) pour la modélisation des crues éclair (Estupina, 2004). Le caractère distribué du modèle permet de prendre en compte la forte variabilité spatio-temporelle des caractéristiques physiques du bassin versant et du forçage par les pluies par l'intégration des données provenant de l'imagerie satellitale et des champs de pluie distribués (données radar CALAMAR ou PANTHERE). MARINE est un modèle perceptuel qui fait du ruissellement par dépassement de la capacité d'infiltration le processus responsable de la genèse des crues, l'unique terme de perte du bilan hydrologique est l'infiltration dans la zone non saturée. C'est un modèle événementiel adapté au mode opérationnel car les temps de calcul sont relativement courts. Enfin, le choix d'un modèle à base physique est adapté au cadre d'application des méthodes variationnelles d'assimilation de données.

Le groupe HydroEco de l'IMFT (Institut de Mécanique des Fluides Toulouse) participe au programme BVNE (Bassins Versants Numériques Expérimentaux) mis en place par le SCHAPI (Service Central Hydrométéorologique d'Appui à la Prévision des Inondations) depuis 2005 (Tanguy et al., 2005). Dans ce cadre, le bassin versant du Gardon d'Anduze a été retenu comme base de test des modèles hydrologiques de prévision des crues éclair. Le travail de thèse utilise donc ce bassin comme site d'étude et les données récoltées lors de divers épisodes pluvieux.

Le bassin versant du Gardon d'Anduze est soumis à un aléa pluviométrique intense en particulier au cours des mois d'automne. Les intensités de pluie varient entre quelques dizaines et plusieurs centaines de millimètres par heure sur plusieurs heures consécutives. D'autre part, la bordure des Cévennes est caractérisée par des reliefs accidentés. Ainsi, la combinaison de ces différents facteurs cause des crues violentes, voire meurtrières, et ce risque de crue est d'autant plus problématique que l'urbanisation à proximité des cours d'eau augmente.

La compréhension et la connaissance de la genèse de ces crues et de leurs paramètres hydrologiques sont nécessaires pour prévenir les risques associés à ces crues. D'un point de vue modélisation hydrologique, il s'agit de d'obtenir une représentation la plus fidèle possible du phénomène physique.

Objectifs de la thèse

Cependant, de nombreux « défis » sont à relever dans la modélisation hydrologique distribuée (Braud, 2008) car son intérêt a été remis en question par de nombreux auteurs (Beven et Binley, 1992 ; Grayson et al., 1992 ; Beven, 2001) à cause des problèmes de surparamétrisation, d'estimation de paramètres ou encore des limitations de la validation.

En effet, de multiples sources d'incertitudes et d'erreurs dans les modèles et les données d'observation (Kitanisis et Bras, 1980) et des problèmes d'échelle (Dehotin et Braud, 2007), dus à la présence de seuil de fonctionnement et à la non-linéarité des processus hydrologiques (Indarto, 2002 ; Sivapalan, 2003), interviennent dans la modélisation distribuée. Par conséquent, un des problèmes majeurs de la modélisation hydrologique distribuée est l'identification des paramètres.

Un des objectifs de cette thèse est de réduire les incertitudes paramétriques essentiellement grâce à une méthode d'assimilation de données variationnelle, la méthode de l'état adjoint (Le Dimet et Talagrand, 1986). Ces techniques d'assimilation de données, initialement utilisées pour la prévision climatique, ont été intégrées aux modèles hydrologiques. En effet, l'amélioration de la modélisation des processus hydrologiques et des systèmes d'observation des bassins versants a encouragé les hydrologues à implémenter ces techniques dans leurs modèles. Différentes méthodes d'assimilation de données ont déjà été exploitées pour divers types de problèmes hydrologiques (Young, 2002 ; Thiemann et al., 2001 ; Kavetski et al., 2002 ; Krystofowicz, 2002 ; Reichle et al., 2001 ; Moradkhani et al., 2005).

Dans cette étude, la méthode variationnelle de l'état adjoint est utilisée comme méthode d'estimation de paramètres. L'application de cette méthode à la modélisation pluie-débit est assez récente : les premières tentatives d'estimation de paramètres et de conditions initiales datent de 2003 (White et al., 2003 ; Seo et al., 2003 ; Castaings et al., 2007).

Dans ce travail de thèse, la calibration du modèle MARINE est réalisée par l'emploi de deux méthodes : le processus d'estimation des paramètres basé sur la méthode de l'état adjoint et l'exploitation des simulations de Monte-Carlo. Ces deux méthodes sont utilisées pour déterminer les paramètres du modèle les plus sensibles déterminés par une analyse de sensibilité globale préliminaire (Le, 2008).

Deux types de fonctions objectifs sont utilisées, l'une classique similaire au critère de Nash ; l'autre plus généralisée permettant d'intégrer des incertitudes sur les débits observés. Il s'agit d'utiliser toutes les observations disponibles pour améliorer la compréhension des processus de surface et cerner la réponse dominante du bassin en contraignant les incertitudes liées la caractérisation des systèmes hydrologiques, notamment les incertitudes paramétriques. La méthodologie employée dans cette étude consiste à utiliser des facteurs multiplicatifs (ou coefficients correctifs) pour ajuster de manière relative une distribution spatiale fixée a priori des paramètres du modèle (Refsgaard, 1997 ; Senarath et al., 2000 ; Madsen, 2003). Il en résulte une paramétrisation d'ordre réduit souvent de faible dimension.

Un autre objectif de la thèse est la recherche d'une stratégie de prévision et d'anticipation des crues toujours grâce à la méthode de l'état adjoint. La première consiste à utiliser les informations à l'exutoire pendant la phase de montée de crue pour prédire ce pic de crue. La seconde consiste à utiliser les observations d'une station intermédiaire à l'amont du bassin pour prédire le débit à l'exutoire. Le but de la prévision hydrologique est d'anticiper de manière assez fidèle la réponse du bassin versant afin de pouvoir alerter la population, appliquer des actions préventives et agir là où il y a des risques.

Plan du mémoire

Le mémoire est organisé en 7 chapitres principaux.

Le chapitre 1 est composé de deux parties. La première fait un état de l'art de la modélisation hydrologique à travers une présentation des différents types de modèles rencontrés dans la littérature et d'un exemple de procédure de modélisation. La seconde partie de ce chapitre est consacrée à la problématique des crues éclair, à leurs processus générateurs et à leurs modélisations.

Le chapitre 2 expose la théorie générale de l'assimilation de données puis traite de l'application de ces méthodes à la modélisation hydrologique. L'accent est mis sur la méthode de l'état adjoint et l'utilisation des outils de différentiation automatique.

Le chapitre 3 propose une synthèse des procédures d'analyse des modèles. Dans un premier temps les méthodes d'analyse de sensibilité sont présentées puis celles relatives à l'optimisation.

Le chapitre 4 présente le cadre d'étude de la thèse en présentant successivement le modèle testé, le bassin versant étudié, la base de données utilisée et enfin les critères d'évaluation du modèle.

Le chapitre 5 est dédié à la calibration du modèle par l'utilisation de deux critères de performances, l'une classique et l'une plus générale permettant de prendre en compte les incertitudes sur la mesure du débit. Pour chacun de ces critères, la confrontation des résultats des simulations de Monte-Carlo et du processus d'estimation des paramètres est réalisée.

Le chapitre 6 présente les deux démarches adoptées pour la prévision de crue.

Le chapitre 7 propose des voix d'amélioration du modèle et les perspectives d'étude.

Chapitre 1

La modélisation hydrologique et la problématique des crues éclair

1.1	Introduction	8
1.2	La modélisation hydrologique	9
	1.2.1 Typologie des modèles	9
	1.2.2 Problème d'échelle des modèles distribués	11
	1.2.3 Procédure de modélisation	12
1.3	Les crues « éclair »	17
	1.3.1 Processus de genèse des crues éclair	17
	1.3.2 La prévision de crues à l'échelle nationale	21
	1.3.3 Apport de la modélisation hydrologique pour une prévision opérationnelle	25
	1.3.4 Apport de la distribution spatiale de l'information	27
1.4	Synthèse	30

CHAPITRE 1 La modélisation hydrologique et la problématique des crues éclair

1.1 Introduction

La première partie de ce chapitre est consacrée à la modélisation hydrologique à travers une présentation des différents types de modèles rencontrés dans la littérature et d'un exemple de procédure de modélisation. La modélisation hydrologique a pour objectif de représenter mathématiquement les processus hydrologiques. La modélisation du comportement hydrologique des bassins versants touche plusieurs domaines d'application (gestion des ressources en eau, aménagement du territoire ...). Dans un contexte opérationnel, les modèles hydrologiques sont utilisés pour répondre à deux problématiques principales : l'évaluation des risques hydrologiques liés aux phénomènes de crues et la gestion de la ressource en eau liée à l'apparition des étiages. Ce travail de thèse se concentre sur la problématique des crues éclair. La seconde partie de ce chapitre est alors consacrée à cette thématique, à leurs processus générateurs et à leurs modélisations.

1.2 La modélisation hydrologique

Un modèle n'est qu'une représentation imparfaite d'une réalité perçue par le modélisateur. L'action de modéliser sous-entend un certain nombre d'hypothèses implicites. A chaque étape de la modélisation, de la perception du phénomène à sa formalisation dans un cadre conceptuel et sa traduction en langage de programmation, des approximations sont réalisées (Ambroise, 1999). Un modèle est un compromis entre généralité, réalisme et précision.

Les modèles pluie-débit sont des outils couramment utilisés aujourd'hui pour des études hydrologiques dans l'ingénierie et dans la recherche. Ils sont utilisés pour différents objectifs :

- étendre ou combler les séries chronologiques de débits (Sefton et Howarth, 1998),
- reconstituer des chroniques de débits à partir de chroniques des pluies,
- évaluer la réponse d'un bassin versant à la variabilité climatique ou à la variabilité de la végétation (Binley et al., 1991),
- simuler l'impact d'aménagements anthropiques sur l'hydrologie du bassin versant (par exemple la construction d'un barrage),
- gérer la qualité des eaux (Mroczkowski et al., 1997),
- gérer des ressources en eau en période d'étiage,
- faire de la prévision de crue temps réel (Moore, 1999),
- apporter les conditions limites des modèles de circulation atmosphérique (Dümenil et Todini, 1992).

La modélisation du comportement hydrologique des bassins versants doit pouvoir décrire les différentes étapes de la transformation pluie-débit. Un grand nombre de modèles de structures différentes, combinant généralement des fonctions linéaires et non-linéaires, ont été développés depuis le début des années 1960. Todini (1988) fait une revue historique de cette modélisation pluie-débit.

1.2.1 Typologie des modèles

Plusieurs classifications ont été proposées par différents auteurs (Wheater et al., 1993, Singh, 1995, Ambroise, 1999). Une première classification composée de trois classes distinctes est usuellement employée (Wheater et al., 1993) : les modèles empiriques, les modèles conceptuels et les modèles à base physique. Cette classification dépend également de la description spatiale des processus au niveau du bassin versant en tant que modèles globaux ou distribués (Refsgaard, 1997).

1.2.1.1 Les modèles empiriques, conceptuels et à base physique

Les modèles **empiriques** (empirical, lumped model) utilisent les séries de données chronologiques disponibles pour en déduire la structure du modèle ainsi que les valeurs des paramètres correspondants. Ils sont alors basés sur les informations rapportées des données et n'incluent aucune connaissance a priori sur le comportement du bassin versant et des processus physiques de l'écoulement. Ces modèles portent aussi le nom de boîtes noires (black box). Des exemples de ce type de modèles sont les réseaux de neurones artificiels (Artificial Neural Networks) (Hsu et al., 1995), les modèles GR (Edijando et Michel, 1989, Loumagne, 1988), le modèle SCS. Ces modèles considèrent généralement le bassin versant comme une entité unique. Les réseaux de neurones artificiels nécessitent une phase d'apprentissage et ne peuvent être utilisés sur les bassins non jaugés. Les modèles GR et SCS ont une structure simple et le rôle de leurs paramètres est bien identifié. A partir d'un travail de régionalisation des paramètres, il peut être envisageable de les appliquer à des bassins non jaugés.

Les modèles **conceptuels** (conceptual, grey box) utilisent des éléments de stockage ou réservoirs comme composants principaux. Ces réservoirs se remplissent par des flux : précipitation, infiltration ou percolation et se vident par évapotranspiration, ruissellement, drainage etc. Contrairement aux modèles empiriques, la structure des modèles conceptuels est déterminée avant leur utilisation. Elle est définie par la compréhension du système hydrologique du modélisateur. Cependant, ces modèles reposent également sur des séries temporelles de données de la sortie du système pour la détermination des valeurs des paramètres dans un processus de calibration. Les paramètres décrivent la taille des réservoirs ou la répartition des écoulements entre eux. Un seul paramètre peut décrire plusieurs processus et ne peut alors pas être directement déduit des mesures de terrain. La plupart de ces modèles considèrent le bassin comme une unité homogène (modèles globaux). Cependant, l'approche semi distribuée consiste à diviser le bassin versant en plusieurs sous-bassins (Boyle et al., 2001) et tient compte de la variabilité spatiale à travers des classes ayant des comportements hydrologiques similaires. La plupart des modèles utilisés en pratique sont des modèles conceptuels. Le modèle TOPMODEL, modèle conceptuel distribué (Beven et Kirkby, 1979 ; Obed, 2003) est une illustration de ce type de modèles. Ils ne sont pas directement applicables à des bassins non jaugés du fait de leur dépendance aux mesures de l'écoulement. Cependant, des méthodes de régionalisation utilisant des relations statistiques entre les paramètres du modèle et les caractéristiques du bassin versant ont été développées pour leur utilisation sur des bassins non jaugés (Burn and Boorman, 1992 ; Sefton and Howarth, 1998).

Les modèles à **base physique** (physically-based, white box) (Freeze et Harlan, 1969 ; Beven, 1989 ; Beven, 2002) sont fondés sur les principes mécanistes de conservation de la masse, du moment et de l'énergie. Concrètement, les équations qui génèrent ces modèles sont issues des lois de la mécanique des fluides ou de la physique des écoulements d'eau dans les sols (loi de Richards en milieu non saturé, loi de Darcy en milieu saturé et loi de Barré Saint-Venant pour le ruissellement de surface). Leur utilisation pratique date des années 1980 grâce aux progrès de l'informatique.

Ils ont été développés dans l'espoir que la représentation physique de leurs processus était suffisamment réelle pour que leurs paramètres soient directement reliés aux caractéristiques physiques du bassin versant sans processus de calibration (Todini, 1988). Par exemple, ces paramètres sont représentatifs de l'humidité du sol, de la conductivité hydraulique dans la zone non saturée, des coefficients de frottement. Les modèles à base physique utilisent une distribution spatiale basée sur des unités élémentaires : une maille carrée régulière ou irrégulière, des triangles irréguliers (Triangular Irregular Networks – TINS, Ivanov et al., 2004), des tubes de courant (Grayson et al., 1992), des versants ou d'autres types d'unités hydrologiques. Cette distribution spatiale permet de représenter la variabilité spatiale des facteurs géographiques et météorologiques qui déterminent la réponse hydrologique. Elle est facilitée aujourd'hui par l'acquisition des données issues de l'imagerie radar ou satellitale. Le modèle SHE (Abott et al., 1986) et le modèle TOPKAPI (Liu et Todini, 2005) en sont les exemples les plus connus. Cependant, les modèles à base physique souffrent d'une extrême demande de données, des problèmes d'échelles (les échelles de mesures diffèrent de celles des processus et des paramètres) (1.2.2) et de surparamétrisation (overparameterization). Une conséquence directe est alors que les paramètres ne peuvent être directement déduits des mesures et que les modèles à base physique nécessitent également une étape de calibration, généralement des paramètres les plus sensibles.

1.2.1.2 Les modèles continus et événementiels

Les modèles **continus** tournent tout au long de l'année. Ils utilisent une représentation complète des différents composants du cycle de l'eau. Ces modèles ne nécessitent pas d'initialisation, ce qui est avantageux d'un point de vue opérationnel.

Les modèles **événementiels** ne fonctionnent que lorsqu'un événement pluvieux se produit. Certains phénomènes sont considérés comme prépondérants dans la représentation de ces modèles tandis que d'autres phénomènes peuvent être négligés. Par exemple pour la prévision des crues éclair, seule la génération du ruissellement de surface et la décrue peuvent être représentées ; les processus plus lents comme l'évapotranspiration et la redistribution des eaux dans les sols sont généralement négligés. Contrairement aux modèles continus, les modèles événementiels doivent être initialisés (humidité initiale des sols, indice des pluies antérieures, débit de base, etc.).

1.2.2 Problème d'échelle des modèles distribués

La variabilité spatio-temporelle des processus hydrologiques est influencée par différents facteurs (paragraphe 1.3.1). Afin de capturer cette variabilité les observations, souvent issues de mesures ponctuelles, doivent être acquises à des échelles spatio-temporelles adéquates. En pratique seule une partie de cette variabilité est capturée à cause des contraintes techniques, logistiques et financières engendrées. Selon Blöschl et Sivapalan (1995), l'échantillonnage des données conditionne la représentativité de la variabilité spatio-temporelle : ils décrivent les caractéristiques d'un échantillonnage par le triplet (Fig. 1.1) : extent, spacing et support.

L'étendue (extent) correspond à la couverture globale, l'espacement (spacing) exprime la distance entre deux mesures et l'échelle de support (support) caractérise le volume (ou l'aire) sur lequel la mesure est moyennée.

Cette conceptualisation peut également être appliquée à la dimension temporelle : la durée de la mesure, l'intervalle et la constante de temps d'échantillonnage de la mesure.

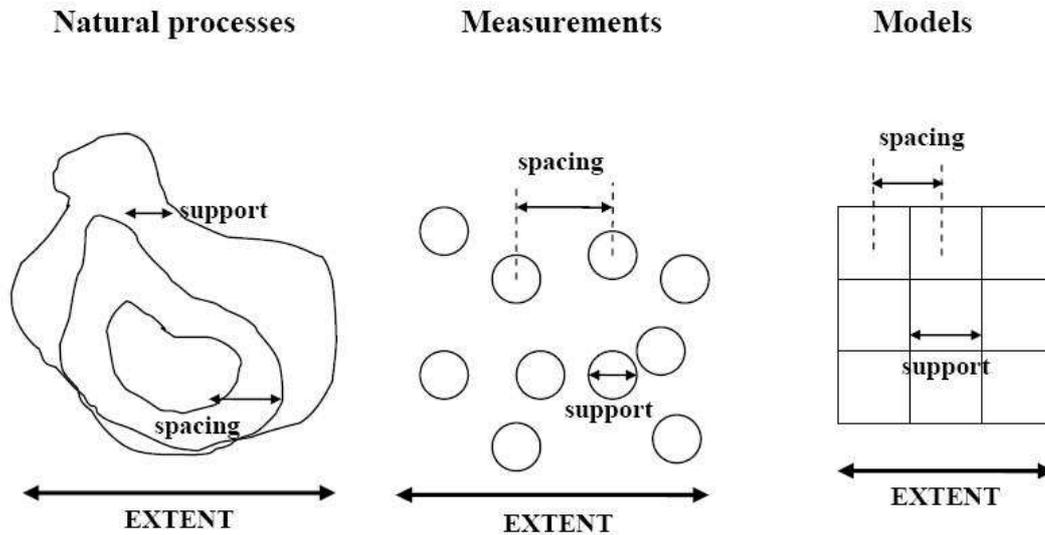


Fig. 1.1 : Définition des trois caractéristiques d'échelle selon Blöschl et Sivapalan (1995) (extent, spacing et support)

L'échantillonnage des données pose le problème du passage d'une unité spatiale à une autre. Des méthodes d'interpolation permettent ce changement d'échelle pour la paramétrisation du forçage, des conditions initiales et des propriétés physiques du milieu. Il existe des méthodes d'interpolation classiques (méthode de Thiessen, moyenne pondérée...), des méthodes d'agrégation et de désagrégation et des méthodes géostatistiques (krigeage).

Les échelles spatiales et temporelles de la modélisation sont caractéristiques de la description des processus hydrologiques. La modélisation des processus à un pas de temps court (comme les crues éclair) requiert une description des processus hydrologiques assez complexe.

1.2.3 Procédure de modélisation

Différentes procédures de modélisation des modèles pluie-débit ont été proposées dans la littérature. Il est cependant possible de trouver une structure commune à l'ensemble des procédures proposées. Elle est composée des étapes suivantes : la sélection de la structure du modèle, l'analyse de sensibilité, la calibration, la validation et la prévision incluant les incertitudes (Fig. 1.2).

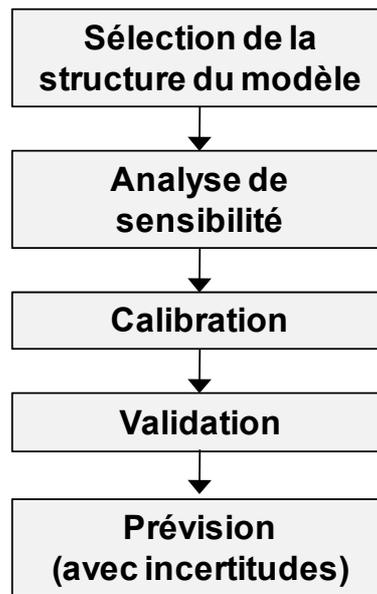


Fig. 1.2 : Procédure (standard) simplifiée pour la modélisation pluie-débit d'après Wagener et al. (2004)

1.2.3.1 Sélection de la structure du modèle

En fonction des processus hydrologiques étudiés et du pas de temps de modélisation requis, la structure du modèle utilisée est différente. Il n'existe pas de méthode objective pour choisir la structure la mieux adaptée à l'objectif de modélisation et il reste une part de subjectivité dans le choix de cette structure (par exemple dans le choix des équations physiques utilisées).

Par ailleurs, la performance d'un modèle est liée à son niveau de complexité ainsi qu'au nombre de données disponibles. La relation entre ces trois composantes est illustrée par la figure 1.3 proposée par Grayson et Blöschl (2000). Pour une quantité de données disponibles (trait plein sur la figure), la performance du modèle passe par un maximum pour une complexité du modèle donnée, c'est le calage optimal. Si le modèle est trop complexe, les paramètres du modèle à déterminer sont trop nombreux par rapport aux données disponibles pour obtenir un calage satisfaisant et la performance du modèle est alors réduite. Inversement, si le modèle est trop simple, il ne peut pas exploiter toutes les informations disponibles et sa performance est moindre. Parallèlement, pour un modèle donné (trait discontinu sur la figure), la performance du modèle est maximale pour une certaine quantité de données disponibles. Si le nombre de données augmente, le modèle doit être plus complexe pour pouvoir exploiter toute la richesse des données disponibles.

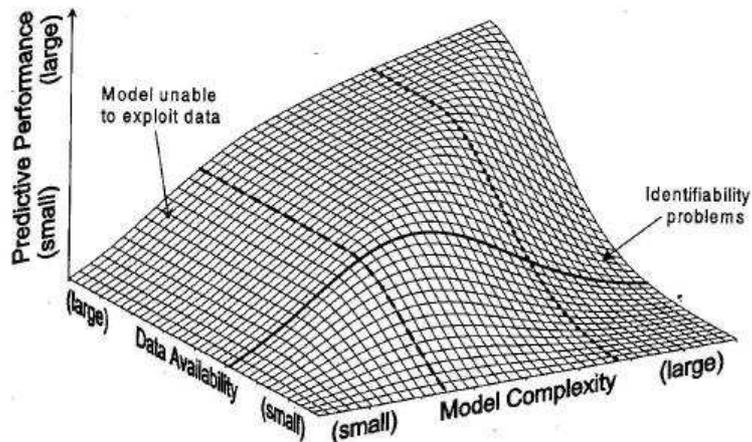


Fig. 1.3 : Relation entre la complexité du modèle, la disponibilité des données et la performance du modèle selon Grayson et Blöschl (2000)

1.2.3.2 Analyse de sensibilité

L'analyse de sensibilité consiste à étudier comment les perturbations sur les entrées du modèle engendrent des perturbations sur la réponse. Les entrées du modèle sont les paramètres, les conditions initiales et aux limites. Cette analyse peut avoir lieu à deux stades différents du processus de modélisation. Elle peut soit être placée avant l'étape de calibration afin d'identifier les paramètres les « plus importants » par rapport à la performance du modèle. Les paramètres les moins sensibles sont eux généralement fixés à des valeurs adaptées pour réduire la dimensionnalité du processus de calage (Bastidas et al., 1999). L'analyse de sensibilité peut aussi avoir lieu après le processus de calage pour juger si les paramètres ont bien été identifiés ou non.

Le lecteur peut se reporter à la première partie du chapitre 3 consacrée à une description des méthodes d'analyse de sensibilité.

1.2.3.3 Calage

Le calage d'un modèle est l'opération qui consiste à trouver, à partir d'un échantillon d'événements de référence, les valeurs d'un ou de plusieurs jeux de paramètres du modèle qui minimisent l'erreur de modélisation. Concrètement, les simulations du modèle sont comparées aux observations afin d'évaluer la capacité du modèle à représenter la réalité.

Après avoir défini une base de données de référence (hydrogrammes), des méthodes d'optimisation sont utilisées pour le calage du modèle. L'objectif est de déterminer un jeu de paramètres optimal pour lequel l'écart entre les simulations du modèle et les observations est minimal. Des fonctions objectif ou critères de performance sont définies afin de mesurer cet écart.

Pour plus de détails sur les différentes méthodes d'optimisation, le lecteur peut se reporter à la seconde partie du chapitre 3.

Les crues éclair sont des phénomènes violents et rares, la base de données de référence est de ce fait réduite ce qui rend délicat le calage des modèles hydrologiques représentant ces processus.

1.2.3.4 Validation

La validation d'un modèle consiste à en évaluer les performances sur un jeu de données non utilisé dans la phase de calage. Une première approche consiste à choisir aléatoirement une partie des données pour le calage du modèle et une seconde partie pour la validation. Le jeu de paramètres obtenu après calage est réutilisé pour la validation. Pour les modèles distribués, une autre approche est basée sur la validation multi-critères et multi-échelles. Elle consiste à confronter les résultats du modèle à d'autres types de données, par exemple des hydrogrammes des stations intermédiaires. Une validation du modèle sur d'autres bassins versants peut aussi être pratiquée lorsque les données sont disponibles.

Pour évaluer la capacité d'un modèle à réaliser certaines tâches spécifiques et évaluer sa performance, le schéma de validation de Klemes (1986) comprend quatre étapes : (1) test du modèle avec des données non utilisées pour le calage, (2) test de la capacité du modèle à prédire le débit d'un autre bassin aux caractéristiques hydrologiques similaires, (3) test avec des données de caractéristiques hydrologiques différentes (par exemple, une période sèche et une période humide), (4) combinaison des tests (2) et (3).

1.2.3.5 Incertitudes

Plusieurs sources d'incertitude règnent dans la modélisation hydrologique (Kitanidis et Bras, 1980 ; Melching, 1995 ; Bastidas et al., 1999) :

- Incertitude dans les processus naturels (Melchin et al., 1991). Elle résulte de la variabilité spatio-temporelle des processus naturels, comme les précipitations, responsable de la genèse du ruissellement. Cette incertitude ne peut être réduite.
- Incertitude sur les données, les observations de débit par exemple. Elle est due aux erreurs de mesures ou du prétraitement, aux erreurs au niveau des techniques d'acquisition des données et des méthodes d'interpolation.
- Incertitude sur la spécification du modèle (ou encore le problème d'identifiabilité du problème) qui se traduit par l'incapacité du modèle à converger vers un unique jeu de paramètre « optimal » en utilisant les données disponibles. Cette incertitude provient essentiellement de l'incertitude sur les données et sur la structure du modèle. En effet, le processus de calage propage les incertitudes sur les données sur les paramètres du modèle. D'autre part, certaines caractéristiques du modèle comme les processus de seuil ou des paramètres interdépendants conduisent à de multiples zones d'attraction et d'optimum locaux ; ce qui rend la localisation de l'optimum global du modèle difficile.
- Incertitude sur la structure du modèle introduite lors des simplifications et des approximations faites sur la description des processus réels.
- Incertitude due à la détermination des conditions initiales (par exemple l'humidité initiale du sol) car ces états sont généralement inconnus au début de la période de calage ou de simulation.

Plusieurs approches ont été proposées pour propager l'incertitude présente dans les différentes composantes du processus de modélisation dans les prédictions du modèle (Melching et al., 1990, Beven et Binley, 1992, Gupta et al., 1998).

Des approches classiques d'estimation de l'incertitude des prévisions consistent à estimer la moyenne et la variance des prévisions en propageant les incertitudes sur les entrées en utilisant des dérivées au premier ordre. Les écarts types sont des mesures directes pour caractériser l'incertitude sur les paramètres d'entrée. Cependant ces méthodes sont limitées par la non-linéarité du modèle et à une multi-modalité de la surface de réponse.

D'autres méthodes basées sur des simulations de Monte-Carlo ont été introduites : la méthode GLUE (Generalised Likelihood Uncertainty Estimation) de Beven et Binley (1992) (3.2.4.3), la méthode MCSM (Monte Carlo Set Membership) de Keesman (1989), la PU (Prediction Uncertainty) de Klepper et al. (1991).

1.3 Les crues « éclair »

Les crues lentes évoluent suivant une échelle temporelle de l'ordre de la journée ou plus et concernent en général des bassins versants moyens à grands, de plus de 1000 km². Les crues éclair ou rapides sont provoquées par des pluies localisées d'intensité extrême qui s'abattent sur des petites à moyennes surfaces de quelques centaines de km². Leur évolution est soudaine et de courte durée, le temps de réaction du bassin est de l'ordre de quelques heures (Montz et Grunfest, 2002). Le débit de pointe est relativement élevé et il est difficile de prévoir ces événements (Carbonnel, 1995).

Georgakakos (1986) définit une crue éclair comme : « a local hydrometeorological phenomenon, which allows short lead times for warnings » (i.e. quelques heures) après un événement pluvieux intense.

Dans un premier temps, les processus responsables de la genèse des crues éclair sont décrits. L'accent est mis sur le cas des bassins méditerranéens du Sud de la France car ceux-ci sont particulièrement touchés par ce phénomène des crues éclair. Par exemple la crue de Novembre 1999 qui a touché le département de l'Aude, les départements voisins de l'Hérault, des Pyrénées Orientales et du Tarn est responsable du décès de 35 personnes (Gaume, 2002). La crue de Septembre 2002 dans le département du Gard a été également dévastatrice (24 victimes humaines et 1.2 milliards d'euros de dégâts matériels). Le bassin versant du Gardon d'Anduze, dans le département du Gard, sera le bassin d'étude de cette thèse (4.2).

La prévision de crue à l'échelle nationale sera abordée dans un second temps. Finalement cette partie se terminera par l'apport de la modélisation hydrologique pour la prévision d'inondation.

1.3.1 Processus de genèse des crues éclair

Plusieurs processus physiques de nature différente peuvent expliquer la formation du ruissellement contribuant à l'écoulement rapide d'une crue. Les concepts majeurs de la genèse des écoulements de crue, incluant les processus superficiels et souterrains, sont présentés dans cette partie. Pour plus de détails, le lecteur pourra se référer aux différentes références mentionnées. L'accent sera mis sur le cas particulier des processus observés sur les bassins méditerranéens. Les crues éclair touchant ces bassins, de petite et moyenne taille (jusqu'à 800 km²) se déclenchent majoritairement en automne. Ces crues sont caractérisées par un temps de réponse court du bassin versant (quelques heures) et des coefficients de ruissellement élevés.

1.3.1.1 Les écoulements superficiels

La battance

Lorsque le sol est peu couvert par la végétation, la pluie peut provoquer une déstructuration de la surface du sol, sous forme d' « organisations pelliculaires superficielles ».

La pluie provoque un tassement de la surface du sol, un éclatement des agrégats (effet « splash ») et une redistribution des particules les plus fines à la surface. Ce processus aboutit à la fermeture progressive des pores et à l'imperméabilisation du sol. Le sol passe d'un état meuble et poreux à un état plus compact. La couche superficielle s'individualise par rapport au reste du profil et forme une croûte de battance qui diminue considérablement la perméabilité du sol. Ce n'est plus la conductivité hydraulique du sol qui constitue la limite à l'infiltration mais celle de la couche superficielle qui peut se modifier très rapidement. Ceci conduit à une intensification du ruissellement. Boiffin (1982) a étudié les stades de dégradation d'un sol ; il estime que la capacité d'infiltration d'un sol initialement de 30 à 60 mm/h peut atteindre 2 à 6 mm/h une fois que l'effet splash a formé la croûte de battance (Fig. 1.4)

Ce processus est spécifique des sols nus ou quasi-nus (Hillel, 1984) (sols agricoles travaillés, sols des régions arides et semi-arides) car l'interception de la couverture végétale diminue l'effet destructurant de l'impact des gouttes de pluie et les racines contribuent à stabiliser le sol.



Fig. 1.4 : Stades de dégradation de la surface du sol d'après Boiffin (1982)

Ruissellement hortonien

D'après la théorie de Horton (Horton, 1933) le ruissellement de surface est généré par le refus d'infiltration d'un sol initialement non saturé ou encore par dépassement de la capacité d'infiltration des sols (infiltration excess surface runoff). Tant que la capacité d'infiltration du sol n'est pas dépassée, la totalité de la pluie s'infiltré. Le ruissellement apparaît lorsque la capacité d'infiltration devient inférieure à l'intensité des précipitations. L'eau qui ne peut s'infiltrer forme l'écoulement rapide de crue. Cependant, l'eau infiltrée ne contribue que très lentement à la recharge des nappes et au maintien du débit de base. Les transferts latéraux par la nappe profonde sont donc suffisamment retardés pour ne pas intervenir dans le volume des crues. Ce type d'écoulement a été observé dans des bassins en climat semi-aride avec des sols peu filtrants et dépourvus de couvert végétal (Albergel et al., 2003 ; Esteves et al., 2000).

Le ruissellement hortonien et la battance peuvent jouer un rôle majeur dans la genèse des crues éclair compte tenu des intensités de pluie auxquelles les bassins versants sont soumis.

Ruissellement par saturation (sur surfaces saturées)

Le ruissellement peut également se produire lorsque tout le profil de sol est saturé. La saturation du sol peut être due à la remontée progressive de la nappe d'eau souterraine ou à un cumul de pluie suffisant pour saturer un sol localement peu profond ou une nappe suffisamment proche. L'extension et la localisation de ces surfaces saturées (Betson, 1964) varient au cours d'un événement pluvieux ; ce sont des zones contributives variables. Par exemple, ces zones saturées peuvent apparaître sur les versants des secteurs où le sol est peu profond et la roche mère imperméable (Cosandey, 1994). La théorie des surfaces contributives variables date des années 60 (Freeze, 1974) ; elle est aujourd'hui fréquemment employée pour expliquer la genèse des crues dans les bassins versants sous climat tempéré. Cependant, les mécanismes de saturation des zones contributives restent discutés : saturation par les pluies incidentes (Dunne et Black, 1970) ou remontée progressive de la nappe de versant au cours de la crue (Gascuel-Oudoux et al., 1999, Obled, 2003).

Ce type de ruissellement peut être observé sur les hauts bassins cévenols aux abords des thalwegs drainants (Marchandise, 2007). Selon Cosandey (1994), les crues cévenoles, soudaines et brutales sur le mont Lozère s'expliquent en deux temps. Dans un premier temps, les zones contributives saturées à l'amont du bassin (où les sols sont minces), s'étendent et génèrent des écoulements qui se réinfiltrent ensuite au cours de leur transfert vers l'exutoire. Dans un second temps, si la pluie persiste, l'aval du bassin versant se transforme en zones contributives saturées et l'écoulement est généralisé sur l'ensemble du bassin versant.

1.3.1.2 Les écoulements souterrains

Les écoulements hypodermiques

Hewlett et Hibbert (1963) ont suggéré que l'horizon superficiel du sol (l'humus et la zone explorée par les racines) était plus perméable et pouvait constituer une zone d'écoulement préférentielle. Une nappe perchée se développe dans l'horizon superficiel et génère un écoulement latéral « hypodermique » ou « subsuperficiel » guidé par la topographie. Cependant, cette théorie est mise en doute pour diverses raisons : les conductivités hydrauliques observées restent en général insuffisantes (inférieure au cm/s) pour expliquer une contribution rapide des écoulements souterrains ; certains auteurs (Jackson, 1992 ; Philip, 1991) pensent que les écoulements sont essentiellement verticaux tant que le milieu est non saturé. Cependant, un ruissellement hypodermique non négligeable a pu être observé lors d'expérimentations sur des sols schisteux en amont du bassin versant d'Anduze (Ayrat, 2004). La conjonction d'une couche d'humus bien développée (30 cm par endroits), de sols grossiers et de pentes importantes pourraient être à l'origine de cet écoulement.

Les écoulements par macropores

Les macropores présents dans les horizons superficiels du sol sont généralement créés par des animaux, des fissures naturelles du sol ou résultent de la décomposition des racines.

Ils représentent, selon certains auteurs (Beven et Germann, 1982, Kirkby, 1988), des drains naturels ou des chemins préférentiels à travers lesquels l'eau peut transiter rapidement vers la rivière. Ces hétérogénéités contribuent à un drainage plus rapide que celui qui est estimé avec la loi de Darcy.

L'intumescence de nappe

L'intumescence de nappe correspond à une augmentation temporaire du gradient de charge de la nappe située à proximité des thalwegs. Son niveau étant proche de la surface, la nappe réagit rapidement aux pluies ; ce qui a pour effet d'accélérer la contribution des écoulements souterrains aux crues (Sklash et Farvolden, 1979). Ce processus peut se déclencher même si le profil de sol n'est pas saturé. Le débit de drainage de la nappe est proportionnel à son gradient de charge selon la loi de Darcy. Ce processus joue un rôle essentiellement lors de l'amorce des écoulements ou lors des crues provoquées pas des pluies de courte durée aux coefficients d'écoulements très faibles (Gaume, cours hydrologie)

L'écoulement dans les zones contributives saturées

La réactivité de la nappe peut être influencée par la présence d'une frange capillaire, zone quasi-saturée au dessus de la surface libre de la nappe. Dans ce cas, un faible cumul des précipitations suffit à saturer le profil de sol et les écoulements souterrains se mettent en place pour participer à l'écoulement de crue. Un écoulement de type piston se met en place dans le versant ; l'eau nouvellement infiltrée dans la partie haute du versant saturé chasse l'eau préexistante dans le sol à l'aval qui est ensuite exfiltrée dans la partie basse du versant.

Des expériences de traçage montrent que par ce type d'écoulement, l'eau du sol contribue à l'écoulement de crue.

1.3.1.3 Synthèse

Les processus prépondérants lors des crues éclair sont difficilement identifiables. Ambroise (1998) a déclaré que « les débits de crue sont générés le plus souvent par plusieurs processus simultanément ou successivement, dans des combinaisons variables dans le temps et dans l'espace ».

Ainsi, l'identification d'un processus prépondérant à l'échelle de la parcelle ne signifie pas que ce dernier soit prépondérant à l'échelle du bassin versant. Cette succession de processus a pour conséquence des effets de seuil qui expliquent en partie la rapidité de montée des crues éclair. Outre cette variabilité des processus responsables de la genèse des crues, la variabilité spatiale et temporelle des profondeurs et des niveaux de saturation des sols provoque des réponses hydrologiques variées des bassins versants. En particulier, la forte variabilité des terrains montagneux où se produisent les crues éclair (épaisseurs de sols, pentes, altitude, etc.) entraîne une importante variabilité spatiale des processus responsables de la génération et des transferts des écoulements.

La réaction hydrologique d'un bassin versant à un même apport d'énergie n'est pas forcément identique.

Cette non-linéarité provient de différentes combinaisons et intensités des processus en jeu mais également de conditions initiales différentes (Ambroise, 1998). Le comportement hydrologique d'un bassin versant se modifie également en fonction des saisons ou des changements climatiques.

Par ailleurs, en cas de crues majeures, ces processus ne sont pas toujours observables car l'accessibilité aux bassins est dangereuse et délicate. En outre, l'endommagement ou la destruction possible des appareils de mesure lors d'événements extrêmes ne permet pas un travail d'analyse a posteriori.

Il n'existe pas de lois a priori sur le fonctionnement de la genèse des crues éclair. La genèse de ces crues peut résulter du fonctionnement courant de certains bassins ou du fonctionnement extrême d'autres (Obled, 1999).

1.3.2 La prévision de crues à l'échelle nationale

1.3.2.1 Nécessité de la prévention

Le risque d'inondation constitue un enjeu majeur, c'est le premier risque naturel en France. Il touche environ 8000 communes, soit près d'une commune sur quatre et l'étendue des zones inondables s'élève à plus de 20 000 km² (source www.ecologie.gouv.fr, Ministère de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du territoire). Depuis 50 ans, l'urbanisation s'est largement développée dans les zones inondables (4.5 millions de français vivent dans ces zones) (Fig. 1.5) ; ce qui a eu pour conséquence une augmentation de la vulnérabilité face au risque d'inondation. Les dommages occasionnés se traduisent par des coûts considérables (plusieurs dizaines de milliards d'euros) qui affectent durablement l'économie locale voire même nationale.



Fig. 1.5 : Habitations proches du lit mineur à Anduze (photo du haut) et maisons dans le lit majeur (photo du bas) en amont d'Anduze (département du Gard).

1.3.2.2 La gestion du risque d'inondation

La gestion du risque d'inondation se décompose en trois phases (Fig. 1.6):

- La phase de prévention, c'est la phase préliminaire de la gestion du risque qui consiste à déterminer l'aléa. Il s'agit de déterminer une hauteur ou un débit de retour donné en chaque endroit du territoire (éventuellement à partir d'outils statistiques basés sur des chroniques de données de pluie ou de débit). Au cours de cette phase, une cartographie des zones à risques est réalisée pour décider de l'occupation du territoire et de l'autorisation de construire. C'est également au cours de cette phase que se fait le dimensionnement des ouvrages hydrauliques (barrages, ponts ...) et que des études géographiques et socio-économiques permettent de localiser et de quantifier la vulnérabilité des territoires.
- La phase de gestion de la crise au cours de laquelle les organismes spécialisés (SPC : Service de Prévion de Crue) observent l'événement en cours et émettent des prévisions de la crue potentielle afin de minimiser les risques et organiser un service de secours (Sécurité Civile). Au cours de cette phase, les météorologues et les hydrologues utilisent des modèles de prévision opérationnelle.
- L'analyse post crise consiste à analyser un événement passé pour en dresser le bilan des conséquences et éviter certaines des erreurs passées. Cette analyse constitue une aide aux deux premières phases car elle permet d'accroître la qualité de ces deux phases grâce au retour d'expérience (Gaume, 2003).

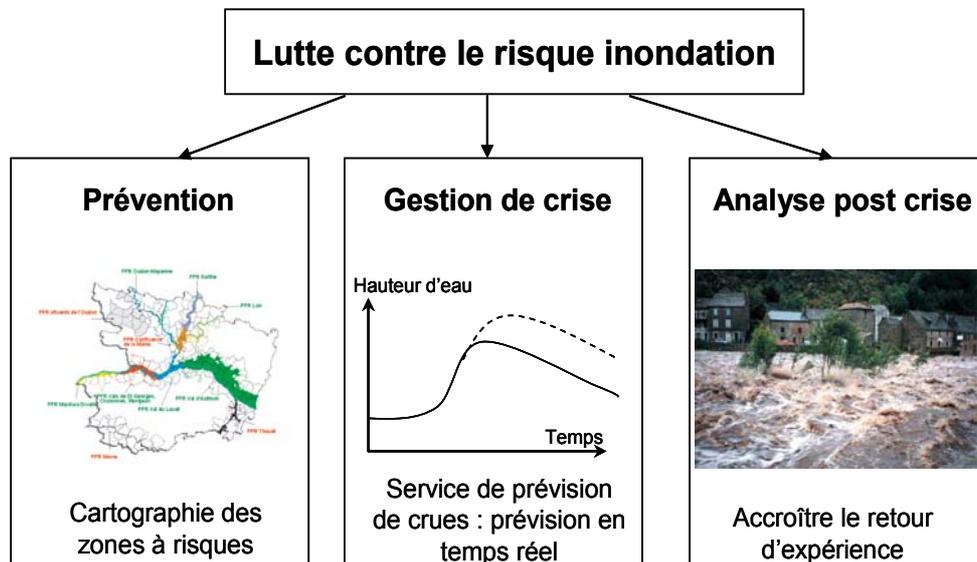


Fig. 1.6 : La gestion du risque d'inondation d'après Estupina (2004)

Contrairement aux crues relativement lentes sur des zones qui bénéficient de moyens d'instrumentation, de télésurveillance et de prévision généralement performants permettant une gestion de crise efficace, la prévision des crues éclair est un exercice difficile.

En 2002, le Ministère de l'Ecologie et du Développement Durable a mis en place la réforme de l'annonce des crues pour améliorer le service rendu au public et aux maires pour plus de 6000 communes inondables.

Cette réforme a abouti à la création du SCHAPI (service central d'hydrométéorologie et d'appui à la prévision d'inondation) à Toulouse (Tanguy et al., 2005) et au remplacement des 52 services d'annonce des crues (SAC) préexistants par 22 services de prévision des crues (SPC) afin de concentrer les compétences.

La réforme de la politique de prévention s'articule autour de la mise en place des Programmes d'Action de Prévention des Inondations (PAPI) qui regroupent les propositions des collectivités et des services de l'Etat.

1.3.2.3 La prévision des pluies

Selon Krajewski et al. (1991), Corradini et Singh (1985), les petits bassins versants naturels présentent une forte sensibilité aux hétérogénéités de la pluie. Les pluies diluviennes à l'origine des crues éclair possèdent une forte variabilité spatio-temporelle. Cosandey et Robinson (2000) estiment que les intensités, durées et extensions spatiales des précipitations sont des facteurs déterminants de la formation de ces crues. Par ailleurs, une anticipation des pluies est nécessaire à une bonne prévision des crues.

L'imagerie radar est un outil exploitable pour la prévision des crues éclair. En effet, l'image radar rend compte de la distribution spatiale de la pluie et fournit des observations pour des intervalles de temps relativement courts et constants, de l'ordre de 6 minutes. De plus, le radar ne subit pas les conséquences des intempéries contrairement aux pluviomètres qui se retrouvent parfois détériorés au cours des orages. Toutefois, ces données nécessitent une première étape de calage à partir des mesures pluviométriques.

Creutin et al. (2003) pensent que : « the combined use of distributed hydrological models and radar detection rainfall is profitable to spatially monitor flash flood hazards and to understand physical processes involved in runoff production mechanisms ».

L'ensemble du territoire français est couvert par un réseau de radars météorologiques, le réseau ARAMIS (Fig. 1.7) qui permet l'observation et la détection en temps réel des phénomènes pluvieux pour la prévision météorologique immédiate. Des mesures quantitatives de la pluie au sol : « lame d'eau HYDRAM », avec une résolution de 1 km toutes les 15 min, ont été apportées grâce à ce réseau. Des travaux ont récemment été menés sur l'amélioration de l'estimation de la lame d'eau (Parent du Châtelet, 2005).

Un nouveau calcul de la lame d'eau sur le réseau ARAMIS, la Lame d'Eau PANTHERE, corrige les données de réflectivité des effets de masques, de bandes brillantes, de sous estimation à grande distance et de profil vertical de réflectivité, principales sources d'incertitudes des mesures radar (4.3.3.2). Aussi, le poids attribué aux données dépend de leur qualité, il diminue avec la distance au radar et dans des zones d'échos fixes (zones masquées par le relief) (Fig. 1.8). Les données sont disponibles au pas de temps 5 min.

Depuis 2006, de nouvelles lames d'eau sont produites par la méthode ANTILOPE. Elle s'appuie sur la séparation des échelles et représente la somme de deux contributions :

- la partie "petite échelle" est extraite des mosaïques de lame d'eau « PANTHERE 5 min » par détection des cellules convectives,
- et la partie "grande échelle" est issue d'une spatialisation par krigeage des données pluviométriques dont la partie convective a été retirée.

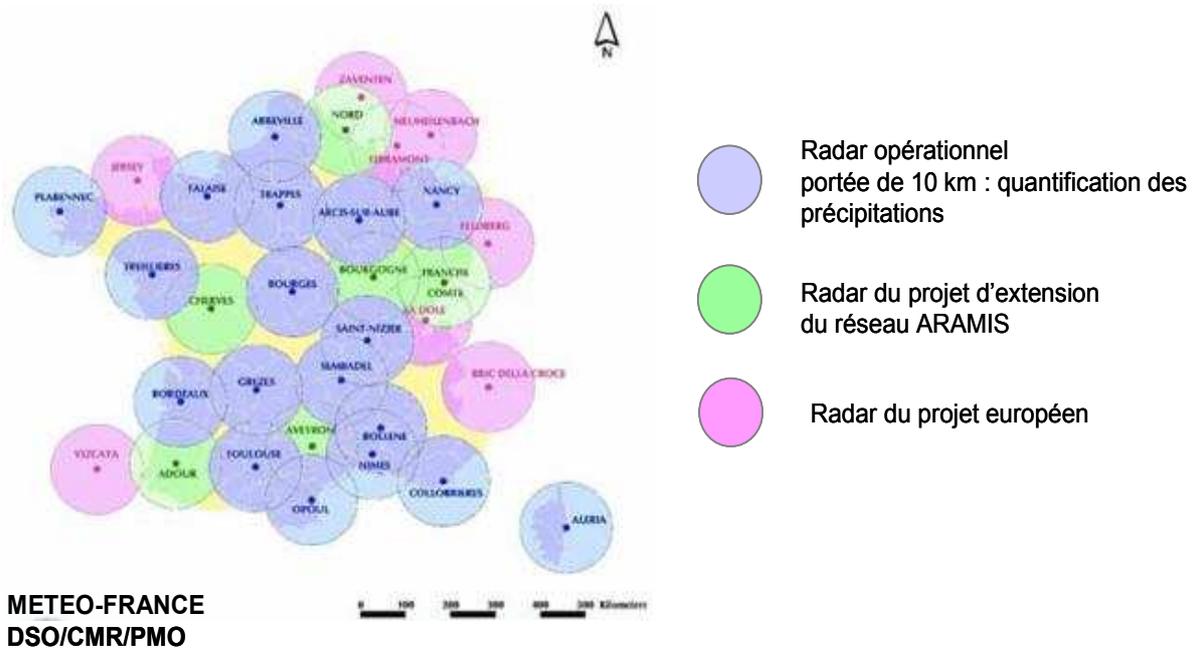


Fig. 1.7 : Le réseau ARAMIS de Météo-France

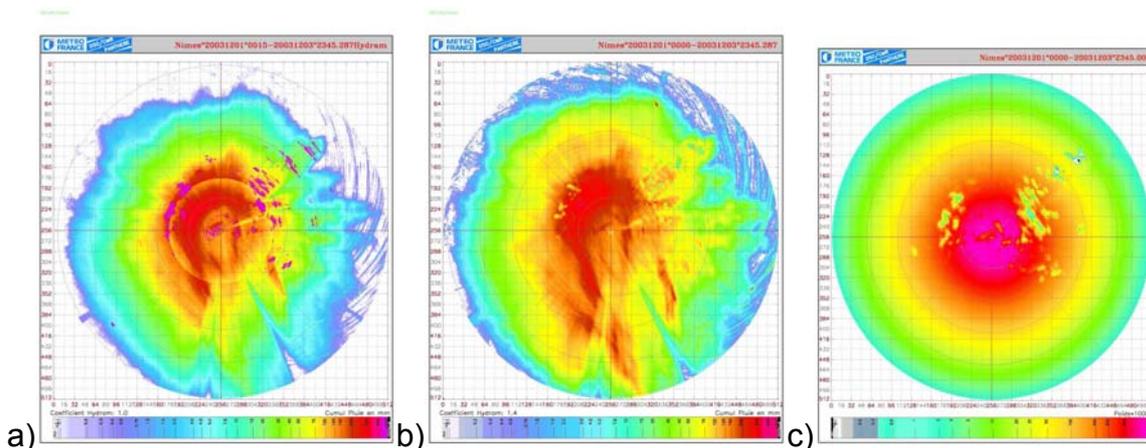


Fig. 1.8 : Comparaison entre les cumuls de pluie radar de Bollène : (a) ancienne lame d'eau HYDRAM, (b) nouvelle lame d'eau PANTHERE. (c) est l'image des poids pour chaque pixel, d'après Parent du Châtelet et al. (2005).

Concernant la prévision des pluies à courte échéance plusieurs méthodes existent :

- la méthode des analogues (Guilbaud et Obled, 1998) : cette méthode permet de prévoir des cumuls journaliers de précipitations sous forme de probabilités à l'échelle locale à partir de prévisions météorologiques établies à l'échelle synoptique¹. Les situations synoptiques à venir sont supposées déjà avoir été observées dans le passé, il s'agit alors de reconnaître des situations « analogues » pour en déduire des effets statistiquement similaires,
- la méthode d'advection qui consiste à suivre le déplacement des cellules orageuses (Boudevillain et al., 2004),
- les méthodes déterministes (ARPEGE de Météo-France),
- les générateurs stochastique de pluie (Eagleson, 1972 ; Datin, 1998) qui s'appuient sur des historiques de pluies,
- l'exploration volumique radar consiste à utiliser plusieurs tours d'antenne, à des angles d'élévation différents, pour obtenir une image volumique du système pluvieux. L'exploitation de ces données a permis de mettre au point une méthode d'estimation en temps du profil vertical de réflectivité du radar (qui est ensuite convertit en intensité de pluie) (Guéguen et Tabary, 2004).

Un nouveau modèle opérationnel AROME doit être mis en exploitation opérationnelle à partir de fin 2008 sur le nouveau calculateur de Météo-France. C'est un modèle à mailles très fines (2.5 km au lieu de 10 km pour ces prédécesseurs) de prévision numérique du temps à courte échéance développé pour améliorer la prévision des phénomènes dangereux tels que les pluies torrentielles méditerranéennes ou les orages violents. Ce modèle sera ainsi mieux adapté pour l'anticipation des crues et des risques d'inondation.

1.3.3 Apport de la modélisation hydrologique pour une prévision opérationnelle

Le prévisionniste doit, dans un délai très court, prendre des décisions dont les enjeux peuvent être importants. Son objectif est d'alerter à temps les organismes chargés de la sécurité civile. Des outils destinés à répondre aux besoins de prévention et de gestion du risque hydrologique en temps réel sont utilisés. Le modèle hydrologique calcule un débit ou une hauteur d'eau prévue en un point d'un cours d'eau à un instant donné. En d'autres termes, il détermine à partir de la connaissance de la pluie observée ou prévue et de l'état initial du bassin versant, la quantité d'eau qui participe à l'écoulement et le délai dans lequel celle-ci va rejoindre le cours d'eau et l'exutoire du bassin versant.

En France, les Service de Prévision des Crues, les services techniques et municipaux et les services de la sécurité civile ont développé leurs propres outils d'alerte répondant à un besoin précis sur un milieu donné.

¹ Adjectif employé pour qualifier les phénomènes atmosphériques dont l'ordre de grandeur est de quelques milliers de kilomètres pour les dimensions horizontales, de quelques kilomètres pour la dimension verticale et de quelques jours pour la durée. L'échelle synoptique constitue par excellence le cadre de la prévision opérationnelle sur une échéance de un à trois jours dans les zones tempérées.

Les SPC gèrent des bassins versants de tailles relativement importantes (supérieures à quelques dizaines de km²). Ils utilisent des modèles de propagation hydraulique, des modèles pluie-débit et des modèles mixtes. D'autres modèles ont été développés par les universitaires. Les principaux modèles de prévision de crue « éclair » fonctionnant en opérationnel ou en phase de test sont : ALTHAÏR (pour les bassins versants du Gard), SOPHIE, MERCEDES et MARINE.

Le modèle SOPHIE (Système Ouvert de Prévisions Hydrologiques Informatisées avec Expertise) est utilisé dans plusieurs SPC. Il est constitué d'un ensemble de modèles hydrologiques : des modèles débit-débit (avec des relations linéaires, puissances ou linéaires multiples), des modèles pluie-débit (modèle GR3H par exemple ; Edijatno et Michel, 1989) et un modèle multi-agent auto-adaptatif. Cette constitution permet un fonctionnement en mode multi-modèles. Ce modèle est le plus utilisé au sein des SPC et tourne en temps réel sur un grand nombre de bassins versants. La structure du modèle ne permet cependant aucun déterminisme ni aucune spatialisation des variables.

Le modèle de prévision pluie-débit distribué ALTHAÏR (Alarme Hydrologique Territoriale Automatisée par Indicateur de Risque, Bressand 2002) est développé et utilisé en mode opérationnel par le Service d'Annonce des Crues du Gard (SAC-30). Il a pour objectif la production d'un hydrogramme de crue en temps réel sur des bassins de superficie inférieure à 500 km². Il intègre un module de production inspiré de la loi de Horton (Horton, 1933) et un module de propagation basé sur les trajectoires d'écoulement et d'isochrones à partir d'un Modèle Numérique de Terrain (MNT). Il utilise des images radars de précipitation, les pluies CALAMAR (CALcul de LAMes d'eau à l'Aide du Radar) (4.2.2.2). Deux versions de ce modèle existent : ALTHAÏR en mode « bassin versant » qui nécessite un calage des paramètres du modèle par bassin versant étudié et ALTHAÏR en mode « régionalisé » qui permet de réaliser un calage en fonction de certaines caractéristiques comme la capacité d'infiltration ou l'occupation des sols.

Le modèle MERCEDES (Maillage Élémentaire Régulier Carré pour l'Etude Des Ecoulements Superficiels ; Bouvier et al., 1994) est un modèle distribué utilisant des mailles carrées régulières.

Ce modèle contient différentes fonctions de production basées sur les lois de Horton, de Green et Ampt, sur le principe des nappes contributives (TOPMODEL – Beven et Kirkby, 1979 – ; TOPKAPI – Ciarapica et Todini, 2002), sur des lois empiriques de type SCS (Soil Conservation Service). Il possède également diverses fonctions de transfert fondées sur des schémas géomorphologiques, sur l'équation de l'onde cinématique, etc. MERCEDES utilise des pluies radar ou distribuées.

Le modèle MARINE (Modélisation de l'Anticipation du Ruissellement et des Inondations pour des Evénements Extrêmes, Estupina et Dartus, 2003) développé à l'IMFT (Institut de Mécanique des Fluides de Toulouse) est un modèle distribué à base physique conçu pour la prévision des crues éclair.

Le processus responsable de la genèse des crues est le ruissellement par dépassement de la capacité d'infiltration. Le modèle utilise une loi de Green et Ampt pour le calcul de la répartition des eaux infiltrées et ruisselées et une équation d'onde cinématique pour le transfert des eaux sur le bassin versant et dans le réseau de drainage. Au moment de l'étude, les écoulements de subsurface ne sont pas

représentés. En effet, compte tenu des temps de réponse des conductivités hydrauliques des sols usuellement mesurées (vitesse horizontale de 10^{-2} à 10^{-3} cm/s pour des sols non argileux), les écoulements souterrains ne peuvent pas, sauf cas exceptionnel, contribuer à l'onde de crue. Une réponse rapide des nappes de versants ne peut être envisagée que dans le cas où la nappe est proche de la surface et la conductivité hydraulique à saturation élevée (Beven, 1982). Ces écoulements n'expliquent donc pas la genèse des débits de crues, en revanche ils peuvent contribuer à la décrue ou au tarissement.

Les informations nécessaires au fonctionnement du modèle sont de deux natures : statique et dynamique. Les caractéristiques du bassin versant considérées comme statiques aux échelles de temps considérées sont la topographie, la géologie, la géographie, la géomorphologie, l'occupation du sol, la végétation et les propriétés dynamiques des processus actifs sur le bassin sont les pluies, les transferts des eaux dans le sol, l'humidité des sols etc.

La structure de ce modèle prenant en compte la variabilité spatiale des pluies et des paramètres géomorphologiques semble être adaptée à la dynamique des crues éclair (Marchandise, 2007), en outre son temps de calcul limité lui permet d'être utilisable en opérationnel.

Le modèle MARINE est étudié en détails dans ce travail de thèse et une description de sa structure est présentée au chapitre 4. Ce modèle fait partie du projet BVNE (Bassins Versants Numériques Expérimentaux) sur le bassin versant du Gardon d'Anduze.

Dans le cadre de son travail de thèse, Marchandise (2007) a proposé une méthodologie de comparaison de certains de ces modèles hydrologiques (SCS, TOPMODEL et ALTHAÏR), basée sur la construction d'un hydrogramme médian.

La structure spatialisée des trois derniers modèles permet de travailler à partir des Modèles Numériques de Terrain, d'utiliser des données issues de l'imagerie satellitale (occupation des sols, épaisseur de sol ...) et enfin d'intégrer l'information de radars météorologiques. Cette distribution spatiale des données est intéressante pour la gestion des crues éclair dont la dynamique est caractérisée par une grande sensibilité à la variabilité spatiale des facteurs climatiques et géomorphologiques.

1.3.4 Apport de la distribution spatiale de l'information

L'apport de la distribution spatiale de l'information a fait l'objet de nombreuses discussions dans la littérature (Fortin et al., 2006). La synthèse sur les processus générateur des crues « éclair » a montré l'importance de la variabilité spatio-temporelle des caractéristiques du bassin versant et des précipitations.

A titre d'exemple, deux résultats significatifs restreints à notre cas d'étude avec modèle MARINE (chapitre 4) comparant des hydrogrammes simulés avec ou sans distribution spatiale des informations aux hydrogrammes observés, sont illustrés ci-après pour démontrer l'importance de la distribution spatiale de l'information. On ne discutera pas dans ce paragraphe de la paramétrisation du modèle (qui sera faite dans le chapitre 4), le but de d'illustrer de manière simple l'influence d'une spatialisation de l'information.

1.3.4.1 Distribution spatiale des caractéristiques physiques du bassin versant

Un premier test mené sur la crue de Septembre 1994 sur le Gardon d'Anduze a montré l'intérêt de la spatialisation de la profondeur de sol.

En utilisant une méthode de calage manuel, les résultats de deux simulations ont été effectués : la première utilisant une profondeur de sol constante sur le bassin versant et l'autre avec des profondeurs de sol spatialisées; la moyenne de ces profondeurs étant la même pour les deux simulations (les autres paramètres du modèle sont eux identiques pour les deux simulations).

Comme on peut le constater sur la figure 1.9, lors de la crue de 1994, l'absence de la distribution spatiale de la profondeur de sol ne permet pas de rendre compte de la première montée du débit à l'exutoire.

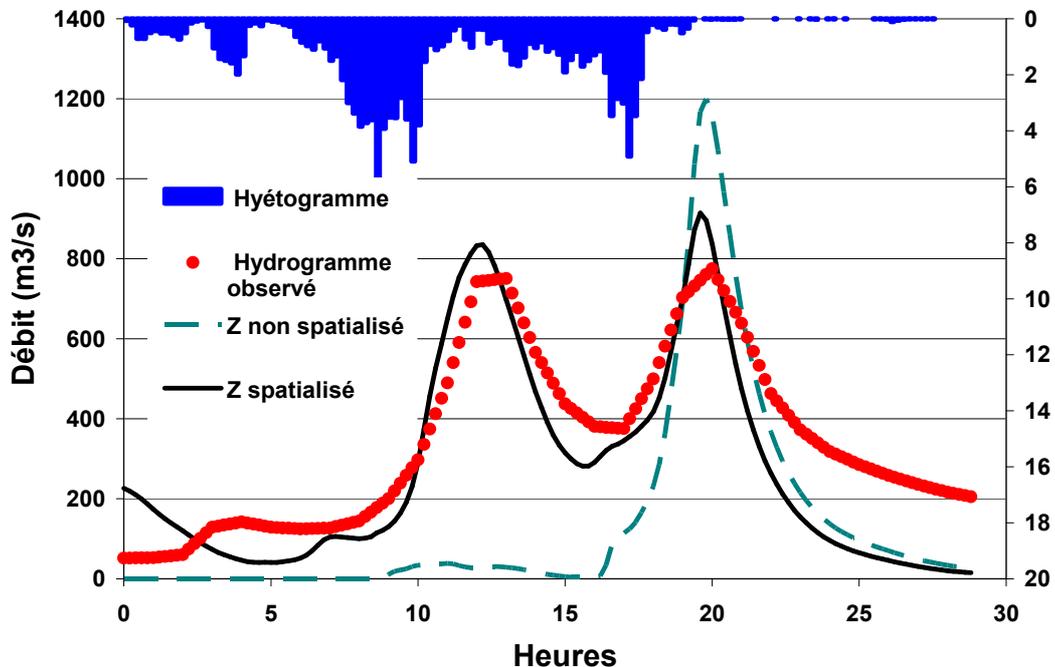


Fig. 1.9 : Test de spatialisation de la hauteur d'infiltration

1.3.4.2 Distribution spatiale des données de pluie

Un deuxième test, effectué celui-ci sur la crue de 2002 (toujours sur le Gardon d'Anduze), a montré qu'une mauvaise prise en compte de la répartition de la pluie sur le bassin versant pouvait entraîner des erreurs importantes : pour ce faire la pluie distribuée (Fig. 1.10) a été remplacée par une pluie uniforme donnant le même hyétogramme moyen sur le bassin versant.

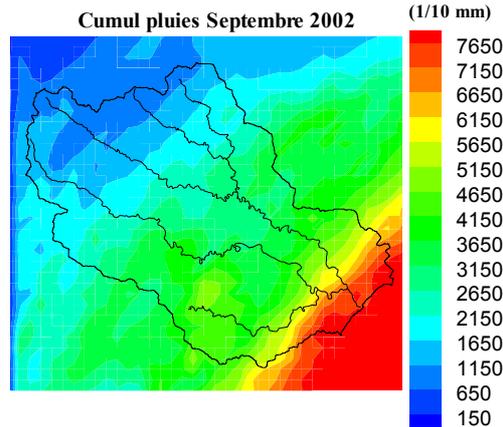


Fig. 1.10 : Données de pluie spatialisée pour la crue du 08/09/2002 sur le Gardon d'Anduze.

Comme le montrent les résultats de la figure 1.11, la crue est lissée de façon exagérée et la décrue intermédiaire constatée à l'exutoire est gommée du fait d'un temps de concentration moyen trop long sur l'ensemble du bassin. On peut noter ici que l'importance de la spatialisation des données de pluie est préoccupante car, dans les simulations en temps réel, il faudra non seulement anticiper des scénarios d'évolution de l'intensité de la pluie, mais il faudra aussi bâtir des scénarios de vent de afin d'anticiper le déplacement de la cellule précipitante.

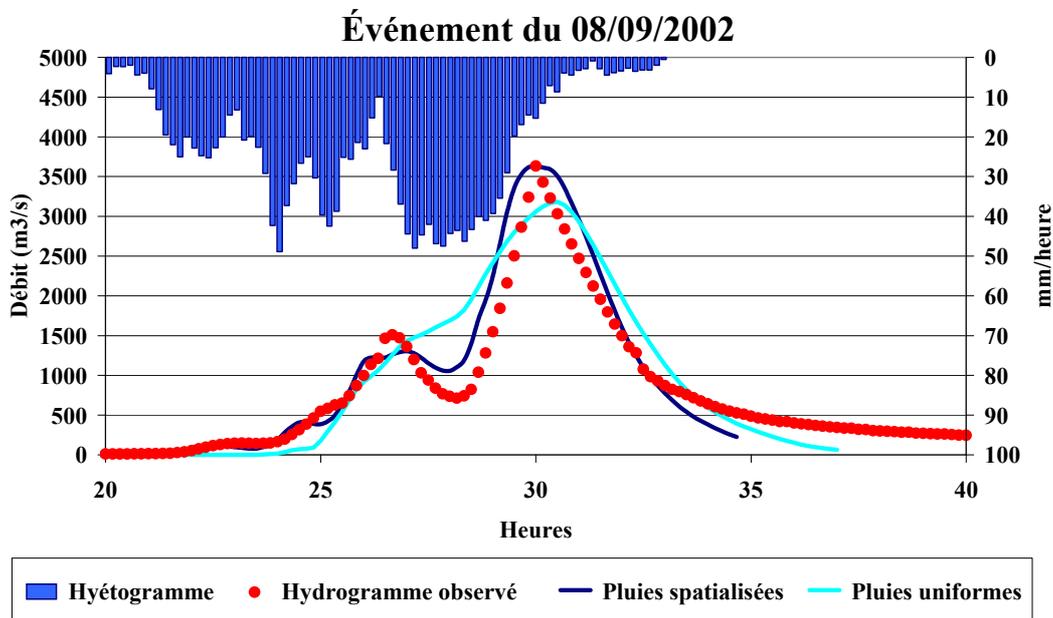


Fig. 1.11 : Test de spatialisation de la pluie

Ces deux tests ont montré l'intérêt de la spatialisation des données pour notre étude. En sus d'une amélioration des hydrogrammes à l'exutoire, la spatialisation des informations sera exploitée lors de l'utilisation de stations hydrométriques en différents points du bassin versant (chapitre 6).

1.4 Synthèse

Après cette brève présentation des processus et facteurs susceptibles d'être responsables de la genèse des crues éclair, il paraît important de retenir la part importante de la variabilité spatio-temporelle des processus eux-mêmes aussi bien que celle des caractéristiques du bassin versant et des précipitations. La représentation de l'hétérogénéité des caractéristiques de sol, de la végétation des bassins versants ; la connaissance des intensité, durée et extension spatiale des pluies ainsi que leur prévision à court terme sont autant d'éléments qui doivent être pris en compte dans la modélisation des crues éclair.

Dans cette thèse, plusieurs raisons ont fait que le modèle MARINE a été choisi pour la modélisation des crues éclair :

- le choix d'une modélisation distribuée, elle se justifie à cause de la forte variabilité spatio-temporelle des caractéristiques physiques du bassin versant et du forçage par les pluies responsables. Le modèle intègre donc des données provenant de l'imagerie satellitale et des champs de pluie distribués (données radar CALAMAR ou PANTHERE),
- le choix d'une structure originale, intégrant une fonction de production et une fonction de transfert pour les versants et le réseau de drainage,
- le choix d'un modèle adapté au mode opérationnel qui nécessite des temps de calcul relativement courts,
- le choix d'un modèle à base physique, adapté au cadre d'application des méthodes variationnelles d'assimilation de données (chapitre 2),
- le choix d'un modèle perceptuel car malgré la multitude des processus entrant en jeu dans la genèse d'une crue sur différentes parties du bassin versant, les modèles hydrologiques se limitent généralement à modéliser le ou les processus dominants responsables des écoulements. Par exemple, il est admis que le ruissellement (de surface et de subsurface) est le processus dominant dans le cas des crues éclair. Le modèle fait donc du ruissellement par dépassement de la capacité d'infiltration le processus responsable de la genèse des crues. La prise en compte des écoulements de subsurface n'est pas forcément nécessaire lors de la modélisation des crues éclair ou à cinétiques rapides (Beven, 1982). Dans cette étude, ils ne seront alors pas considérés,
- le choix d'un modèle événementiel qui permet en termes de modélisation de se concentrer sur la phase de crue.

Le modèle sera exposé plus en détails au chapitre 4.

Cependant, des problèmes d'incertitude (1.2.3.5) et des problèmes d'échelle (1.2.2), dus à la non-linéarité des processus hydrologiques (Indarto, 2002) interviennent dans la modélisation distribuée. Un des objectifs de cette thèse est de réduire les incertitudes paramétriques essentiellement grâce à une méthode d'assimilation de données variationnelle, la méthode de l'état adjoint (chapitre 2). Cette méthode sera utilisée pour estimer les paramètres du modèle les plus sensibles.

Dans un premier temps, l'incertitude sur les mesures de débit ne sera pas prise en compte dans l'estimation des paramètres, puis dans un second temps cette incertitude sera intégrée. Un autre objectif de la thèse est la recherche d'une stratégie de prévision et d'anticipation des crues toujours grâce à la méthode de l'état adjoint.

Chapitre 2

Assimilation de données variationnelle

2.1	Théorie générale de l'assimilation de données	34
2.2	Assimilation de données variationnelle appliquée à la modélisation hydrologique	34
2.2.1	Introduction	34
2.2.2	Application des méthodes d'assimilation de données	35
2.2.3	La méthode de l'état adjoint : une méthode variationnelle	36
2.2.4	Synthèse	45

CHAPITRE 2 Assimilation de données variationnelle

2.1 Théorie générale de l'assimilation de données

La théorie générale de l'assimilation de donnée est décrite dans l'annexe A, un chapitre qui paraîtra dans le guide OMM. Le lecteur peut y trouver, entre autre, une application détaillée de la méthode de l'état adjoint avec l'équation de Burgers.

2.2 Assimilation de données variationnelle appliquée à la modélisation hydrologique

2.2.1 Introduction

Face à l'augmentation de la vulnérabilité des personnes et des biens due au risque d'inondation, la prévision hydrologique est de plus en plus nécessaire (chapitre1). Cependant à cause des multiples sources d'incertitudes et d'erreurs dans les modèles et les données d'observation, un des problèmes majeurs de la modélisation hydrologique est l'identification des paramètres. En outre, les modèles hydrologiques distribués souffrent d'une extrême demande de données, des problèmes d'échelle et de surparamétrisation. Pour faire face à ces problèmes, les techniques d'assimilation de données, initialement utilisées pour la prévision climatique, ont été intégrées aux modèles hydrologiques.

Robinson et al. (1998) définissent l'assimilation de données comme : « the insertion of the reliable data into the dynamical model [...] to improve the quality and accuracy of the estimate »

Selon, Beven (2004) : « Data assimilation / model updating is a highly recommended strategy for real-time or nowcasting applications ».

En effet, l'amélioration de la modélisation des processus hydrologiques et des systèmes d'observation des bassins versants a encouragé les hydrologues à implémenter ces techniques d'assimilation de données dans leurs modèles. Différentes techniques d'assimilation de données ont déjà été exploitées pour divers types de problèmes hydrologiques.

Young (2002) a utilisé le Filtre de Kalman Etendu (EKF) pour la prévision des crues en temps réel.

Des méthodes Bayésiennes ont été testées par Thiemann et al. (2001) et Kavetski et al. (2002) pour des modèles pluie-débit et par Krystofowicz (2002) pour la prévision des crues en temps réel. Reichle et al. (2001) ont développé une méthode d'assimilation variationnelle 4D (4D-Var) prenant en compte les erreurs de la structure du modèle et de la dynamique des processus pour l'assimilation de l'humidité du sol. Moradkhani et al. (2005) appliquent une approche duale d'estimation des variables d'état et des paramètres du modèle basée sur le Filtre de Kalman d'Ensemble (EnKF) à un modèle conceptuel.

Dans cette étude, la méthode variationnelle de l'adjoint, est utilisée (Le Dimet et Talagrand, 1986). L'application de cette méthode à la modélisation pluie-débit est assez récente : les premières tentatives d'estimation de paramètres et de conditions initiales datent de 2003 (White et al., 2003 ; Seo et al., 2003 ; Castaings et al., 2007).

2.2.2 Application des méthodes d'assimilation de données

Il existe quatre types d'applications en assimilation de données (Fig. 2.1) :

- (1) mise à jour du forçage (précipitations, température de l'air...). Il s'agit de l'application la plus classique car l'incertitude sur ces données d'entrée est souvent l'erreur dominante dans le cadre de la prévision opérationnelle,
- (2) mise à jour des variables d'état (hauteur d'eau par exemple),
- (3) mise à jour des paramètres du modèle. C'est l'application la moins fréquente car les paramètres sont généralement supposés constants dans le temps,
- (4) mise à jour des variables de sortie (correction d'erreur)

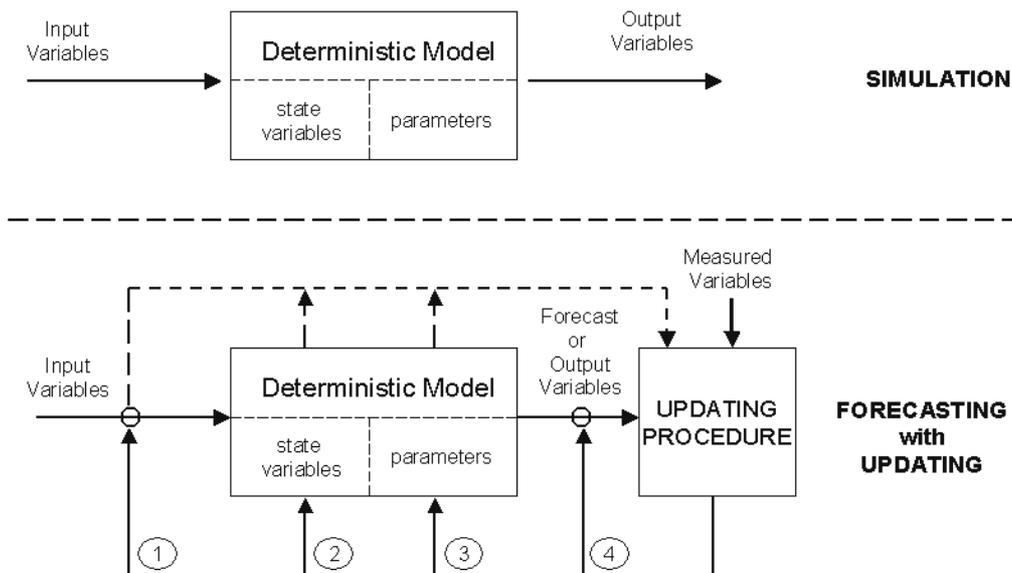


Fig. 2.1 : Comparaison du modèle direct (schéma du haut) avec le modèle couplé à une procédure de mise à jour (schéma du bas) d'après Drécourt (2004).

2.2.3 La méthode de l'état adjoint : une méthode variationnelle

Les méthodes variationnelles sont adaptées aux problèmes représentés par des équations aux dérivées partielles issues de la physique. Dans l'annexe A, le principe général de la méthode de l'état adjoint a été présenté pour l'estimation des conditions initiales. Compte tenu du fait que dans cette thèse ce sont essentiellement les paramètres du modèle qui sont ajustés, une exposition de la méthode dans le cas du contrôle de paramètres est présentée ci après.

2.2.3.1 Contrôle / estimation de paramètres

Le modèle hydrologique peut être décrit par un système d'équations différentielles non-linéaires :

$$\begin{cases} \frac{d\mathbf{X}}{dt} = F(\mathbf{X}, \boldsymbol{\alpha}) \\ \mathbf{X}(0) = 0 \end{cases} \quad (2.1)$$

où \mathbf{X} représente la variable d'état du système (hauteur d'eau dans la maille), elle dépend du temps et est de grande dimension, typiquement de l'ordre de 10^5 dans cette étude. F est un opérateur non-linéaire qui gouverne l'évolution de la hauteur d'eau et le vecteur $\boldsymbol{\alpha}$ représente les paramètres du modèle. Le système est résolu sur l'intervalle de temps $[0, T]$.

Soit \mathbf{X}_{obs} un jeu d'observations supposé continu en temps. Le principe de base est de considérer $\boldsymbol{\alpha}$ comme variable de contrôle et de l'optimiser dans le but de minimiser l'écart entre les observations \mathbf{X}_{obs} et les prévisions du modèle.

Le modèle est supposé avoir une solution unique pour un jeu de paramètres données.

Les observations, distribuées en temps et en espace, et la variable d'état n'appartiennent pas forcément au même espace. Par exemple, les observations et les variables d'état peuvent être localisées à des points différents de l'espace physique ou ces grandeurs ne représentent tout simplement pas la même quantité. Par conséquent, un opérateur \mathbf{H} projette une variable de l'espace des prévisions sur celui des observations.

Une fonction objectif, appelée également fonction coût, est définie. Il s'agit d'une norme mesurant l'écart entre les variables simulées et les observations :

$$J(\boldsymbol{\alpha}) = \frac{1}{2} \int_0^T \|\mathbf{H}\mathbf{X}(\boldsymbol{\alpha}) - \mathbf{X}_{obs}\|_{obs}^2 dt \quad (2.2)$$

Le problème est alors de trouver $\boldsymbol{\alpha}_{opt}$ minimisant J . Si la fonction J est différentiable, une condition nécessaire pour que $\boldsymbol{\alpha}_{opt}$ soit solution du problème d'optimisation est donnée par l'équation d'Euler-Lagrange :

$$\nabla J(\boldsymbol{\alpha}_{opt}) = 0 \quad (2.3)$$

En pratique $\boldsymbol{\alpha}_{opt}$ est calculé par un algorithme d'optimisation de type gradient.

Pour le calcul du gradient, le formalisme de dérivée de Gâteaux, une généralisation du concept de dérivée directionnelle (Cacuci, 1981), est employée.

La dérivée de Gâteaux du modèle (2.1) suivant la direction $\hat{\alpha}$ de α s'écrit :

$$\begin{cases} \frac{d\hat{\mathbf{X}}}{dt} = \left[\frac{\partial F(\mathbf{X}, \alpha)}{\partial \mathbf{X}} \right] \cdot \hat{\mathbf{X}} + \left[\frac{\partial F(\mathbf{X}, \alpha)}{\partial \alpha} \right] \cdot \hat{\alpha} \\ \hat{\mathbf{X}}(0) = 0 \end{cases} \quad (2.4)$$

où $\hat{\mathbf{X}}$ est la dérivée de Gâteaux de \mathbf{X} et $\frac{\partial F}{\partial \mathbf{X}}$ le Jacobien du modèle par rapport à la variable d'état. L'équation (2.4) est connue sous le nom de modèle linéaire tangent. Parallèlement, la dérivée de Gâteaux de la fonction coût est :

$$\hat{J}(\alpha) = \int_0^T (\mathbf{H}\mathbf{X} - \mathbf{X}_{obs}, \mathbf{H}\hat{\mathbf{X}}) dt = \langle \nabla_{\alpha} J, \hat{\alpha} \rangle \quad (2.5)$$

Une variable auxiliaire appelée variable adjointe \mathbf{P} est introduite dans le but de faire apparaître la dépendance linéaire de \hat{J} par rapport à $\hat{\alpha}$. Cette stratégie permet d'éviter la répétition du calcul de dérivation pour toutes les directions de l'espace des paramètres. Le système adjoint est obtenu en multipliant le système de Gâteaux (2.4) par \mathbf{P} et en intégrant de 0 à T .

$$\begin{aligned} \int_0^T \left(\frac{d\hat{\mathbf{X}}}{dt}, \mathbf{P} \right) dt &= \int_0^T \left(\left[\frac{\partial F}{\partial \mathbf{X}} \right] \cdot \hat{\mathbf{X}}, \mathbf{P} \right) dt + \int_0^T \left(\left[\frac{\partial F}{\partial \alpha} \right] \cdot \hat{\alpha}, \mathbf{P} \right) dt \\ \left\langle \frac{d\hat{\mathbf{X}}}{dt}, \mathbf{P} \right\rangle &= \left\langle \left[\frac{\partial F}{\partial \mathbf{X}} \right] \cdot \hat{\mathbf{X}}, \mathbf{P} \right\rangle + \left\langle \left[\frac{\partial F}{\partial \alpha} \right] \cdot \hat{\alpha}, \mathbf{P} \right\rangle \end{aligned} \quad (2.6)$$

Après intégration par parties sur le premier terme de l'équation (2.6), il vient que :

$$\left\langle \frac{d\hat{\mathbf{X}}}{dt}, \mathbf{P} \right\rangle = (\hat{\mathbf{X}}(T), \mathbf{P}(T)) - (\hat{\mathbf{X}}(0), \mathbf{P}(0)) - \left\langle \hat{\mathbf{X}}, \frac{\partial \mathbf{P}}{\partial t} \right\rangle \quad (2.7)$$

En utilisant les propriétés du produit scalaire le second terme de l'équation (2.6) devient :

$$\left\langle \left[\frac{\partial F}{\partial \mathbf{X}} \right] \cdot \hat{\mathbf{X}}, \mathbf{P} \right\rangle + \left\langle \left[\frac{\partial F}{\partial \alpha} \right] \cdot \hat{\alpha}, \mathbf{P} \right\rangle = \left\langle \hat{\mathbf{X}}, \left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \mathbf{P} \right\rangle + \left\langle \hat{\alpha}, \left[\frac{\partial F}{\partial \alpha} \right]^T \cdot \mathbf{P} \right\rangle \quad (2.8)$$

Les expressions (2.7) et (2.8) sont égales, donc :

$$\langle \hat{\mathbf{X}}(T), \mathbf{P}(T) \rangle - \langle \hat{\mathbf{X}}(0), \mathbf{P}(0) \rangle - \left\langle \hat{\mathbf{X}}, \frac{\partial \mathbf{P}}{\partial t} \right\rangle = \left\langle \hat{\mathbf{X}}, \left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \mathbf{P} \right\rangle + \left\langle \hat{\boldsymbol{\alpha}}, \left[\frac{\partial F}{\partial \boldsymbol{\alpha}} \right]^T \cdot \mathbf{P} \right\rangle$$

soit,

$$\langle \hat{\mathbf{X}}(T), \mathbf{P}(T) \rangle = \left\langle \hat{\mathbf{X}}, \frac{\partial \mathbf{P}}{\partial t} + \left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \mathbf{P} \right\rangle + \left\langle \hat{\boldsymbol{\alpha}}, \left[\frac{\partial F}{\partial \boldsymbol{\alpha}} \right]^T \cdot \mathbf{P} \right\rangle \quad (2.9)$$

Il est possible de réécrire l'équation (2.5) de la manière suivante :

$$\hat{J}(\boldsymbol{\alpha}) = \langle \nabla_{\boldsymbol{\alpha}} J, \hat{\boldsymbol{\alpha}} \rangle = \langle \mathbf{H}^T (\mathbf{H}\mathbf{X} - \mathbf{X}_{obs}), \hat{\mathbf{X}} \rangle \quad (2.10)$$

Par conséquent si \mathbf{P} est solution du système suivant :

$$\begin{cases} \frac{d\mathbf{P}}{dt} + \left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \cdot \mathbf{P} = \mathbf{H}^T (\mathbf{H}\mathbf{X} - \mathbf{X}_{obs}) \\ \mathbf{P}(T) = 0 \end{cases} \quad (2.11)$$

alors en comparant (2.9) et (2.10), le gradient de J est donné par :

$$\nabla_{\boldsymbol{\alpha}} J = - \left[\frac{\partial F}{\partial \boldsymbol{\alpha}} \right]^T \cdot \mathbf{P} \quad (2.12)$$

où $\left[\frac{\partial F}{\partial \mathbf{X}} \right]^T$ est la transposée du Jacobien du modèle par rapport aux variables d'état.

L'équation donne une caractérisation du minimum de J en fonction de la variable adjointe \mathbf{P} .

Par conséquent, la minimisation de J passe par la résolution du système d'optimalité¹ (S.O) suivant :

¹ On appelle système d'optimalité (S.O.) le système constitué du modèle (dit direct), du modèle adjoint et de la condition d'optimalité.

$$\left\{ \begin{array}{l} \frac{d\mathbf{X}}{dt} = F(\mathbf{X}, \boldsymbol{\alpha}) \\ \mathbf{X}(0) = 0 \\ \frac{d\mathbf{P}}{dt} + \left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \cdot \mathbf{P} = \mathbf{H}^T (\mathbf{H}\mathbf{X} - \mathbf{X}_{obs}) \\ \mathbf{P}(T) = 0 \\ \nabla_{\boldsymbol{\alpha}} J = - \left[\frac{\partial F}{\partial \boldsymbol{\alpha}} \right]^T \cdot \mathbf{P} \end{array} \right. \quad (2.13)$$

2.2.3.2 Cas général : estimation des paramètres et des conditions initiales

Dans le cas général de l'estimation des paramètres et des conditions initiales, le système d'optimalité s'écrit de la façon suivante :

$$\left\{ \begin{array}{l} \frac{d\mathbf{X}}{dt} = F(\mathbf{X}, \boldsymbol{\alpha}) \\ \mathbf{X}(0) = v \\ \frac{d\mathbf{P}}{dt} + \left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \cdot \mathbf{P} = \mathbf{H}^T (\mathbf{H}\mathbf{X} - \mathbf{X}_{obs}) \\ \mathbf{P}(T) = 0 \\ \nabla_{\boldsymbol{\alpha}} J = - \left[\frac{\partial F}{\partial \boldsymbol{\alpha}} \right]^T \cdot \mathbf{P} \\ \nabla_v J = - \mathbf{P}(0) \end{array} \right. \quad (2.14)$$

2.2.3.3 Bilan de la méthode l'état adjoint

La méthode de l'état adjoint permet un calcul efficace du gradient de la fonction coût, par rapport à toutes les variables de contrôle, avec un coût de calcul indépendant de la dimension de ces variables.

Il est important de noter que le système d'optimalité contient toute l'information disponible : modèles, données, statistiques. A ce titre on pourra considérer le S.O. comme un modèle généralisé car il est le seul capable de rendre compte de l'unicité d'une situation géophysique.

2.2.3.4 Algorithme de calcul de la condition initiale optimale et des paramètres du modèle

Un algorithme d'optimisation est ensuite utilisé pour estimer les paramètres de contrôle (condition initiale et/ou paramètres du modèle) optimaux. Une méthode numérique de type descente est définie par la suite suivante :

$$\begin{pmatrix} \boldsymbol{\alpha}_n \\ v_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_{n-1} \\ v_{n-1} \end{pmatrix} + \rho_n D_n \quad (2.15)$$

Où ρ_n est le pas de descente et la direction D_n de descente. Cette direction peut être calculée avec des méthodes de gradient conjugué ou de type Newton.

Pour des problèmes non-linéaires, la détermination du pas de descente peut exiger un grand nombre d'itérations du modèle et le coût de calcul résultant est élevé.

La mise en œuvre algorithmique de l'estimation des paramètres optimaux contient les étapes suivantes de l'itération k à l'itération $k+1$:

- intégration du modèle direct, de 0 à T pour calculer J
- intégration du modèle adjoint, de T à 0 pour calculer \mathbf{P}
- calcul du gradient ∇J
- calcul d'une direction de descente D_k
- calcul d'un pas de descente ρ_k minimisant J suivant la direction D_k
- nouvelle évaluation de la condition initiale et des paramètres,
- si le critère de convergence (par ex : $\nabla J(\alpha_k) \geq \varepsilon$) n'est pas vérifié, on recommence une itération.

En présence de modèles non-linéaires, la surface de réponse peut présenter des optima locaux mineurs risquant de piéger l'algorithme de minimisation. La détermination d'une solution acceptable est alors conditionnée par le point de départ de l'algorithme : il faut que la valeur initiale du jeu de paramètres soit relativement proche du minimum global (Le Dimet et Blum, 2002). Si le point de départ de l'algorithme n'est pas bien choisi, l'algorithme peut converger vers une solution mathématiquement correcte mais pas nécessairement acceptable d'un point de vue physique. Finalement, en plus de la détermination du point de départ de l'algorithme de minimisation, il est nécessaire d'introduire des intervalles de variations des paramètres physiquement cohérents.

2.2.3.5 Obtention du modèle adjoint

Différentes approches possibles

La construction du modèle adjoint nécessite l'obtention du terme $\left[\frac{\partial F}{\partial \mathbf{X}} \right]^T \cdot \mathbf{P}$ de l'équation (2.13) ou (2.14). Il y a trois approches possibles pour l'implémentation du code adjoint (Fig. 2.2) :

- l'adjoint du modèle continu : cette méthode est très fastidieuse et ne donne pas un calcul exact du gradient de la fonction coût. En effet, la discrétisation du modèle adjoint continu peut introduire des erreurs d'approximation numérique sur la précision du gradient,
- l'adjoint du modèle discret, cette méthode fournit des dérivées très précises car elle assure une parfaite compatibilité entre le modèle direct discret et le modèle adjoint discret,

- l'adjoint du code source à l'aide de la différentiation automatique (annexe A). A partir d'un programme calculant une fonction F , les outils de différentiation automatique calculent directement la dérivée de cette fonction F .

Comme les opérations de dérivation (« adjointisation ») et de discrétisation ne sont pas commutatives, les résultats obtenus avec les diverses approches possibles sont différents.

Le code adjoint doit ensuite être validé de façon rigoureuse (test du produit scalaire et test du gradient) pour garantir la pertinence, la précision et l'efficacité du modèle adjoint (annexe A et Fig. 2.2).

Généralement les méthodes variationnelles s'appliquent aux systèmes gouvernés par des équations aux dérivées ordinaires ou partielles. Cependant, les outils de différentiation automatique permettent l'application de ces techniques à d'autres types de modèles basés sur des formalismes mathématiques moins complexes. Par exemple Castaing (2007) utilise la méthode adjoint pour l'analyse de sensibilité du modèle TOPMODEL (Beven et Kirkby, 1979) basé sur la notion d'indice de similarité hydrologique.

(X_d, X_b)	$= 1.771 \cdot 10^{-2}$
(Y_d, Y_b)	$= 1.771 \cdot 10^{-2}$
diff	$= 1.016 \cdot 10^{-11}$

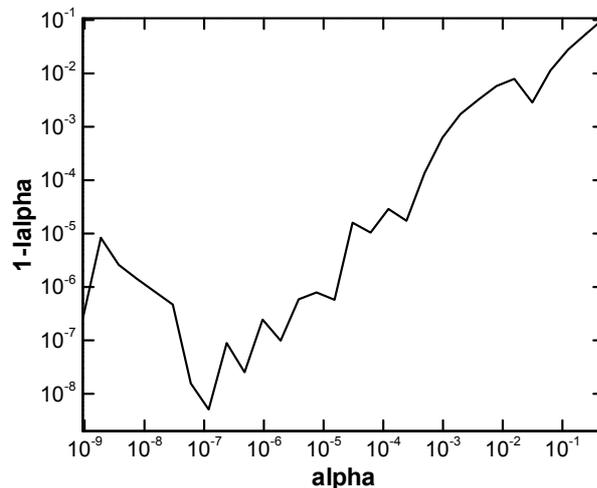


Fig. 2.2 : Tests du produit scalaire et test du gradient pour la validation du code adjoint du modèle MARINE.

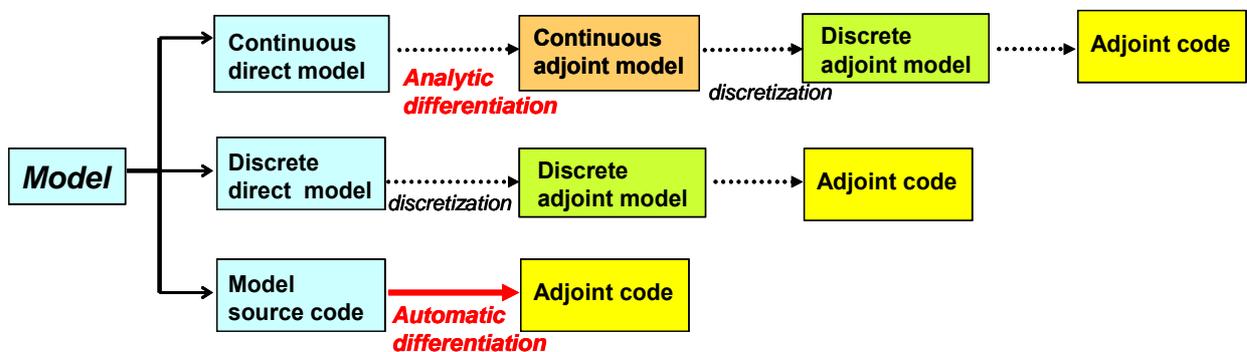


Fig. 2.3 : Approches possibles pour l'implémentation du code adjoint

La différentiation automatique

Dans cette thèse, l'outil de différentiation automatique TAPENADE (Tangent and Adjoint PENultimate Automatic Differentiation Engine – Hacoët et Pascual, 2004), adapté au code source écrit en langage FORTRAN, est utilisé. Le code adjoint (mathématiquement la transposée de la jacobienne) est construit en deux parties : d'abord, le code direct est écrit en sauvegardant toutes les trajectoires puis le code direct est réécrit de la fin vers le début en remplaçant les sauvegardes par des lectures et en dérivant chaque expression du code avec une méthode de transposition (voir ci-dessous).

Dans la seconde partie, le code adjoint relit le code direct de la fin vers le début. L'arbre d'exécution du code direct doit donc pouvoir être remonté (Fig. 2.3).

Ainsi, l'exécution du code adjoint nécessite, pour chaque sous programme (subroutine), une exécution de la subroutine directe puis une exécution de la partie transposée. Le temps d'exécution du code adjoint est alors plus long que celui du code direct.

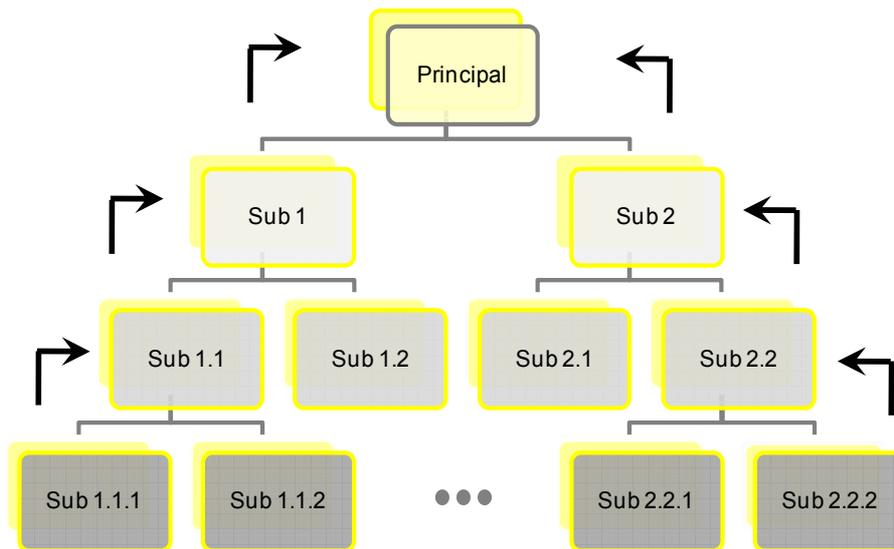


Fig. 2.4 : Arborescence d'un code de calcul nécessaire à la différentiation automatique.

Afin de diminuer le temps d'exécution du code adjoint, il est possible de remplacer les sauvegardes des trajectoires par leur re-calcul. Ceci a pour conséquence de diminuer l'espace mémoire nécessaire à l'exécution du code adjoint.

TAPENADE réalise un compromis entre re-calcul et stockage.

L'utilisation d'outil de différentiation automatique permet un calcul rapide du code adjoint. Toutefois, leur utilisation nécessite une connaissance approfondie du code utilisé car il est souvent nécessaire d'apporter des modifications au code source avant et après sa transformation.

Méthode de transposition

Afin d'illustrer la méthode de transposition employée, un exemple simple est proposé. Si, dans un sous programme du code source, une expression reliant les variables s'écrit :

$$x = yz + kx \tag{2.16}$$

où x, y et z sont des variables intermédiaires et k une constante. Dans le code linéaire tangent, l'expression correspondante est :

$$\delta x = y\delta z + z\delta y + k\delta x \tag{2.17}$$

L'équation (2.17) s'écrit sous forme matricielle :

$$\begin{pmatrix} \delta x \\ \delta y \\ \delta z \end{pmatrix} = \begin{pmatrix} k & z & y \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \\ \delta z \end{pmatrix} \tag{2.18}$$

En notant $\delta\alpha = \begin{pmatrix} \delta x \\ \delta y \\ \delta z \end{pmatrix}$ et $F = \begin{pmatrix} k & z & y \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, l'équation (2.18) devient $\delta\alpha = M\delta\alpha$.

En utilisant les propriétés du produit scalaire et le vecteur des variables adjointes $\alpha^* = (x^* \ y^* \ z^*)^T$, il vient :

$$(\alpha^*, \delta\alpha) = (\alpha^*, M\delta\alpha) = (M^T \alpha^*, \delta\alpha)$$

soit,

$$\alpha^* = M^T \alpha^* \text{ ou encore } \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} = \begin{pmatrix} k & 0 & 0 \\ z & 1 & 0 \\ y & 0 & 1 \end{pmatrix} \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix}$$

Ainsi, dans le code adjoint, l'expression (2.16) devient :

$$\begin{cases} z^* & = & z^* + yx^* \\ y^* & = & y^* + zx^* \\ x^* & = & kx^* \end{cases} \tag{2.19}$$

L'ordre des instructions est important, une inversion fausserait totalement le calcul. L'unique expression du code direct (2.16) se retrouve démultipliée en trois expressions (2.19) dans le code adjoint, ce qui a pour conséquence d'augmenter le temps de calcul du code adjoint.

2.2.3.6 Difficultés liées à l'utilisation de la méthode de l'état adjoint

Même si la méthode de l'état adjoint est une méthode très efficace pour le calcul du gradient quelle que soit la dimension de l'espace de contrôle, quelques difficultés peuvent se présenter lors de son implémentation numérique (Castaing, 2007) :

- des problèmes liés à la linéarisation d'une physique contenant des processus à seuil. Dans le cas d'une formulation continue différentiable mais fortement non-linéaire, le calcul des dérivées est correct mais se fait souvent au premier ordre.

Ainsi son utilisation peut donc être limitée s'il s'agit d'extrapoler une dynamique fortement non-linéaire. Les processus à seuil se traduisent par des instructions conditionnelles dans le code source du modèle.

Il arrive que la fonction résultante soit continue mais non différentiable ou encore qu'elle soit discontinue en certains points. Dans ce cas, il faut modifier le code direct en développant une physique continue et différentiable (Mahfouf, 1999 ; Laroche et al., 2002) ou conserver les seuils dans la formulation du modèle linéaire tangent et du modèle adjoint (Zou et al., 1993) comme le font les outils de différentiation automatique,

- des problèmes liés à l'utilisation de la différentiation automatique. La formulation algorithmique du modèle, implémentée par le code direct, peut contenir des fonctions ayant des points de non-différentiabilité (Griewank, 1995). Dans ce cas, une intervention de l'utilisateur s'avère nécessaire. Par ailleurs, comme précisé ci-dessus, le code direct peut contenir des instructions conditionnelles aboutissant à des fonctions différentiables par morceaux,
- des problèmes liés à une grande dimension de l'espace de contrôle. Lorsque l'assimilation variationnelle est utilisée pour des modèles de dimension importante, la dimension de l'espace de contrôle et la non-linéarité du modèle rendent l'optimisation délicate et très coûteuse en temps de calcul.

Dans le cas de fonctions non différentiables, la direction de descente estimée à partir du gradient calculé par la méthode de l'état adjoint donne encore de bons résultats avec un algorithme d'optimisation différentiable de type quasi-Newton (Zhang et al., 2000 ; Zhang et al., 2001 ; Zhang et al., 2002). Autrement, il est possible d'utiliser des méthodes d'optimisation non-différentiables comme la méthode des faisceaux de Lemaréchal (1976).

D'autre part, la formulation d'un critère de performance prenant en compte la réponse à l'exutoire du bassin versant permet, dans le cas de modèles distribués, de procéder à une intégration spatiale et temporelle, ce qui a pour conséquence de s'affranchir des discontinuités dans les dérivées de la fonction coût dues aux seuils de fonctionnement du modèle (Castaing, 2007).

2.2.4 Synthèse

Différentes techniques d'assimilation de données peuvent être utilisées pour la calibration des modèles. Elles permettent d'estimer un jeu de paramètres optimal en minimisant l'écart, suivant un critère choisi, entre les observations et les simulations du modèle.

Cependant, le concept d'équifinalité apporte quelques limites à la notion de paramètres optimaux car celui-ci stipule que plusieurs jeux de paramètres peuvent être également acceptables pour une modélisation donnée (Beven, 1989 ; Beven et Binley, 1992). En d'autres termes, différents jeux de paramètres peuvent reproduire de façon satisfaisante (au sens d'une mesure donnée) le comportement observé du système étudié. Cette équifinalité provient des incertitudes sur la modélisation du processus physique et sur les données (caractéristiques du bassin versant, conditions initiales...).

Deux objectifs distincts sont visés en assimilation de données.

- Le premier objectif concerne la calibration du modèle, il s'agit d'estimer certains paramètres. Les incertitudes sur les entrées du modèle peuvent être ainsi réduites.
- Le second objectif consiste, d'un point de vue opérationnel, à prévoir et estimer au mieux l'état du système à un instant donné. Le modèle sera utilisé par la suite en prévision.

De façon complémentaire à l'assimilation de données, une évaluation objective de la sensibilité de la réponse du modèle aux différents facteurs d'entrée s'avère indispensable afin de déterminer le jeu de paramètres à estimer. Cette analyse de sensibilité préliminaire au processus d'assimilation de données devrait contribuer à réduire les problèmes d'équifinalité.

Le modèle MARINE, modèle à base physique basée sur la résolution d'équations aux dérivées partielles correspond parfaitement au cadre d'application des méthodes variationnelles. Diverses applications de la méthode de l'adjoint au modèle MARINE sont présentées dans cette thèse :

une première application à l'estimation de paramètres est basée sur des expériences jumelles utilisant des observations synthétiques (chapitre 3 et 5),

une seconde application, au bassin versant Méditerranéen du Gardon d'Anduze, consiste à estimer les paramètres les plus sensibles du modèle MARINE en utilisant des observations réelles (chapitre 5),

la troisième application de la méthode s'oriente vers une optique de prévision des crues (chapitre 6),

finalement, une dernière application concerne l'analyse de sensibilité spatiale utilisant le calcul des dérivées de la fonction objectif (chapitre 7).

Chapitre 3

Procédures d'analyse des modèles

3.1	Analyse de sensibilité	48
3.1.1	Introduction	48
3.1.2	Méthode globale	49
3.1.3	Méthode locale	52
3.1.4	Conclusion	59
3.2	Optimisation	61
3.2.1	Introduction	61
3.2.2	Critère d'évaluation des performances d'un modèle	62
3.2.3	Equifinalité	64
3.2.4	Différentes stratégies d'optimisation des modèles hydrologiques	64
3.2.5	Algorithme d'optimisation utilisant la méthode de l'état adjoint	68
3.2.6	Méthodologie des expériences jumelles	69
3.2.7	Synthèse	73

CHAPITRE 3 Procédures d'analyse des modèles

3.1 Analyse de sensibilité

3.1.1 Introduction

L'analyse de sensibilité permet d'étudier la façon dont les perturbations sur les entrées du modèle engendrent des perturbations sur sa réponse. L'analyse de sensibilité peut être soit locale, soit globale ; elle consiste en une méthode d'analyse quantitative qui permet d'une part d'établir une hiérarchie au sein des variables d'entrée et d'autre part de donner un ordre de grandeur des écarts au sein de cette hiérarchie.

L'objectif de l'analyse de sensibilité locale est d'analyser le comportement de la réponse du système localement autour d'un point donné pour des systèmes statiques ou autour d'une trajectoire donnée pour des systèmes dynamiques dans l'espace des phases des paramètres et des variables d'état.

En revanche, l'analyse de sensibilité globale a pour but de déterminer tous les points critiques du système : les maxima et minima, points de selle, bifurcations, points de non retour toujours dans l'espace de phase des paramètres et des variables d'état. Elle montre comment la variabilité des entrées se répercute sur celle des sorties en déterminant la part de variance due à certains paramètres d'entrée ou à certains ensembles de paramètres d'entrée.

Selon Jacques (2005) la distinction entre les deux procédures d'analyse de sensibilité peut se résumer en ces termes : « l'analyse locale s'intéresse à la valeur de la réponse, tandis que l'analyse globale s'intéresse à sa variabilité ».

Les méthodes d'analyse de sensibilité sont basées sur des procédures statistiques ou déterministes. Bien qu'en théorie les deux procédures soient applicables à l'analyse de sensibilité locale et globale, en pratique, les méthodes déterministes sont plutôt utilisées pour des analyses locales alors que les méthodes statistiques sont employées dans les deux cas.

Les méthodes statistiques ne permettent pas le calcul exact des sensibilités, les méthodes déterministes sont alors utilisées dans ce but. La méthode directe (Forward Sensitivity Formalism) et la méthode par les équations adjointes (Adjoint Sensitivity Formalism) font partie des méthodes déterministes les plus fréquemment utilisées pour le calcul des sensibilités locales.

3.1.2 Méthode globale

« Global sensitivity analysis is a valuable tool in understanding flood inundation modes and deriving decisions on strategies to reduce model uncertainty » Pappenberger et al., 2008.

3.1.2.1 Généralités

L'objectif de l'analyse de sensibilité globale est de déterminer quelle part d'incertitude sur la réponse est due à l'incertitude sur chaque variable ou groupe de variables d'entrée. L'ensemble des méthodes existantes est présenté dans Saltelli et al. (2000). Quatre types de méthodes sont répertoriées : les méthodes fiabilistes pour l'analyse de risque, les méthodes bayésiennes, les méthodes graphiques et les méthodes basées sur l'étude de la variance. Une analyse de cette dernière méthode est exposée dans Jacques, 2005.

L'analyse de sensibilité globale permet d'apporter des éléments de réponse aux questions que se pose tout modélisateur :

- par la confrontation entre la connaissance et l'expérience que l'on a sur l'influence, l'importance de certains paramètres et les résultats de l'analyse de sensibilité, on pourra savoir si le modèle reflète correctement le processus modélisé.
- par la connaissance des variables les plus « importantes » ou encore les plus influentes sur la réponse modèle, on pourra contribuer à l'amélioration du modèle, soit en diminuant les erreurs sur ces variables, soit en modifiant la structure du modèle dans le but de réduire l'effet des erreurs sur ces entrées. Les erreurs sur la réponse du modèle seront alors atténuées.
- par la connaissance des variables les moins influentes, on pourra alléger le modèle en fixant ces variables comme des paramètres déterministes.
- par la connaissance des interactions entre les variables ou groupe de variables, on pourra améliorer la compréhension du phénomène modélisé.

L'analyse de sensibilité s'avère donc indispensable à la compréhension d'un modèle, à sa validation et à son développement.

L'analyse de sensibilité touche un large domaine d'applications : génie nucléaire, écologie, médecine, économie, hydraulique (Pappenberger et al., 2008 ; Hall et al., 2005) , hydrologie (Saltelli et al., 2000).

La méthode de sensibilité généralisée (GSA – Generalised Sensitivity Analysis) est la plus utilisée (3.1.2.2) en hydrologie. D'autres méthodes commencent à être utilisées comme celles basées sur la décomposition de la variance (méthode FAST Cukier et al., 1973), la méthode Sobol' (1993), la méthode de screening (Morris, 1991)

3.1.2.2 Un exemple : la méthode de sensibilité généralisée (GSA)

L'analyse de sensibilité généralisée a été développée par Hornberger et Spear, 1981 dans le contexte de la modélisation environnementale. L'objectif était de déterminer si le modèle est fidèle au comportement du système et si tel était le cas, de déterminer les variables du modèle les plus importantes.

En général, le modèle peut s'écrire de la façon suivante :

$$\frac{d\mathbf{x}}{dt} = F(\mathbf{x}, \boldsymbol{\alpha}) \quad (3.1)$$

où \mathbf{x} est le vecteur d'état et $\boldsymbol{\alpha}$ le vecteur des paramètres réels du modèle. A chaque jeu de paramètres $\boldsymbol{\alpha}$ correspond un unique vecteur d'état \mathbf{x} .

Deux types d'incertitudes interviennent dans cette modélisation : l'incertitude sur la structure du modèle (dans la fonction F), et l'incertitude sur les paramètres du modèle (dans le vecteur $\boldsymbol{\alpha}$).

Chaque composante du vecteur des paramètres $\boldsymbol{\alpha}$ est définie comme une variable aléatoire dont la distribution représente une mesure de l'incertitude sur la valeur réelle, mais inconnue, du paramètre en question. Ainsi, parmi l'ensemble des distributions des paramètres, un certain nombre va être capable de reproduire le comportement recherché du système.

A chaque jeu de paramètres $\boldsymbol{\alpha}^*$ choisi de manière aléatoire correspond un vecteur unique d'observations \mathbf{y}^* qui se compose en fait des mesures du système réel. Le comportement du système sera alors défini en fonction de ces observations. Il existe de nombreuses catégories de comportement possibles. Cependant, il est possible de définir ce comportement de façon binaire : le comportement se produit ou non pour un jeu de paramètres donné $\boldsymbol{\alpha}$. Les jeux de paramètres ayant permis de simuler le comportement désiré appartiennent à la catégorie « behavioural » (jeu de paramètres ayant un comportement correct) et les autres à la catégorie « non-behavioural » (jeu de paramètres ayant un comportement incorrect). Il s'ensuit qu'un critère doit être spécifié pour déterminer l'occurrence ou la non occurrence du comportement, ce critère dépendra de la finalité du modèle.

Dans ce contexte, une étude numérique basée sur des simulations de Monte Carlo permet de simuler le comportement du système à partir d'un jeu de paramètres choisi de manière aléatoire à partir des distributions initiales. Il s'agit ensuite de comparer les résultats de ces simulations au vecteur des observations \mathbf{y} à partir du critère défini. Il est alors possible de déterminer deux échantillons de vecteurs de paramètres pour lesquels le comportement recherché s'est produit ou non, par exemple le dépassement d'un seuil. Le but est d'une part de rechercher un ensemble de paramètres importants et déterminants pour le comportement recherché et d'autre part de classer ces paramètres en fonction de leur influence sur les résultats de la simulation.

Après analyse des résultats des simulations de Monte Carlo, la probabilité cumulative a priori $f(\alpha_k)$ (avant les simulations de Monte Carlo) est divisée en deux parties (Fig 3.1) :

- $f(\alpha_k / B)$ la densité de probabilité cumulative associée à la catégorie « behavioural » B
- $f(\alpha_k / \bar{B})$ la densité de probabilité cumulative associée a la catégorie « non-behavioural » \bar{B}

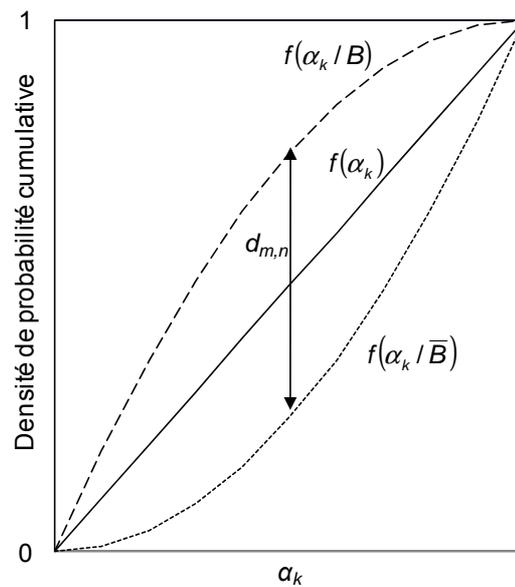


Fig. 3.1 : Principe de l'analyse de sensibilité généralisée, d'après Hornberger et Spear (1981)

Il en résulte une classification des paramètres vis-à-vis du comportement étudié. Si les densités de probabilités cumulatives $f(\alpha_k / B)$ et $f(\alpha_k / \bar{B})$ associées aux catégories « behavioural » et « non-behavioural » sont bien distinctes, alors le paramètre α_k est important. Dans le cas contraire, si la densité de probabilité cumulative n'est pas divisée par la classification comportementale, c'est-à-dire si les densités de probabilité $f(\alpha_k / B)$ et $f(\alpha_k / \bar{B})$ sont confondues avec la densité de probabilité a priori $f(\alpha_k)$, alors le paramètre α_k n'est pas important (Fig. 3.1).

Dans le cas où le vecteur des paramètres a une moyenne nulle et une matrice de variance-covariance égale à la matrice identité ; un classement de la sensibilité des paramètres peut être réalisé, par exemple, à partir de la statistique suivante :

$$d_{m,n} = \sup_x |S_n(x) - S_m(x)| \quad (3.2)$$

S_n et S_m sont les fonctions de distribution des échantillons correspondant à $f(\alpha_k / B)$ et $f(\alpha_k / \bar{B})$ pour n comportements corrects et m comportement incorrects. $d_{m,n}$ est la statistique utilisée dans le test de Kolmogorov-Smirnov bilatéral.

Cette mesure représente la distance verticale maximale entre les courbes associées aux fonctions $f(\alpha_k / B)$ et $f(\alpha_k / \bar{B})$.

Si la distance $d_{m,n}$ est grande, alors le paramètre considéré est important, malheureusement le contraire n'est pas toujours vrai. La distance $d_{m,n}$ peut être petite alors que le paramètre considéré, en raison de ses interactions avec d'autres paramètres, est indispensable à une bonne simulation du modèle. La sensibilité de ce paramètre doit alors être évaluée grâce aux covariances entre les paramètres.

L'implémentation pratique de la méthode est composée des étapes suivantes :

- choix des intervalles d'échantillonnage des valeurs à tester pour chaque paramètre,
- choix d'une méthodologie pour échantillonner l'espace des valeurs possibles des paramètres (généralement méthode de Monte-Carlo avec un échantillonnage uniforme),
- définition de la mesure de vraisemblance à utiliser,
- définition d'un critère de séparation (seuil behavioural / non-behavioural).

Cette méthode est recommandée comme étape préliminaire dans le développement d'un modèle pour détecter les incertitudes dans la connaissance du système.

3.1.3 Méthode locale

3.1.3.1 Introduction

L'utilisation des fonctions adjointes pour l'analyse de sensibilité date du début des années quarante. Les principales approches de l'analyse de sensibilité étaient jusqu'alors basées sur la théorie des perturbations (Wigner, 1945) première utilisation pour la théorie des réacteurs nucléaires) ou sur les approches variationnelles (Levine et Schwinger, 1949 ou Roussopolos, 1953). Le fort potentiel des méthodes adjointes pour l'analyse de sensibilité de nombreux problèmes rencontrés dans la théorie des réacteurs nucléaires a été démontré par Stacey (1974) et Greenspan (1976). Le succès de l'application de ces diverses méthodes adjointes à l'analyse de sensibilité a suscité l'intérêt des chercheurs et a motivé l'application et la généralisation de ces approches à d'autres types de problèmes non-linéaires rencontrés dans des champs d'application plus étendus. Ainsi, son application à la thermodynamique des réacteurs voit le jour à la fin des années soixante-dix où Oblow (1978) aborde l'analyse de sensibilité formellement d'un point de vue différentiel.

L'application de l'analyse de sensibilité par les équations adjointes s'étend aujourd'hui sur divers domaines (économie, économétrie, sciences pour l'ingénieur, industrie, génie nucléaire, physique des réacteurs, sciences environnementales. En particulier, les sensibilités adjointes sont utilisées dans les sciences environnementales (science de l'atmosphère, océanographie, hydraulique, hydrologie, chimie ...) pour :

- l'assimilation de données variationnelles (Le Dimet et Talagrand, 1986 ; Thépaut et Courtier, 1992 ; Navon et al., 1992),
- l'estimation de paramètres optimaux (Smestad et O'Brien, 1991 ; Zou et al., 1992),
- pour l'évaluation de l'augmentation optimale des perturbations initiales (Farrel, 1990 ; Tribbia, 1991 ; Barkmeijer, 1991).

Cacuci (1988) a été le premier à apporter une description détaillée des aspects mathématiques de ces méthodes.

La réponse du système est traduite par différentes fonctions réponses ou critères de sensibilité dépendant des paramètres du système. Les méthodes utilisées consistent en fait à dériver ces fonctions pour évaluer la sensibilité de la réponse du système à des perturbations sur les paramètres d'entrée. Les sensibilités sont alors les dérivées respectives de ces fonctions par rapport aux paramètres d'entrée.

Par exemple la sensibilité s_i de la fonction g par rapport au paramètre α_i s'exprime de la façon suivante :

$$s_i = \frac{\partial g}{\partial \alpha_i} \quad (3.3)$$

Dans les problèmes environnementaux, beaucoup d'auteurs se sont penchés sur cette problématique notamment en météorologie (Rabier et al., 1992 ; Errico et al., 1992 pour en citer quelques uns)

En 1981, Cacuci mit en place une théorie générale de l'analyse de sensibilité pour les systèmes non-linéaires comprenant deux formalismes : la méthode de sensibilité directe (« forward sensitivity formalism » ou FSF) et la méthode de sensibilité par les équations adjointes (« adjoint sensitivity formalism » ou ASF).

Le calcul exact des sensibilités locales contribue aux objectifs suivants :

- améliorer la compréhension du système en déterminant les paramètres importants,
- éliminer les données les moins importantes, réduire la « surparamétrisation »,
- déterminer l'influence des variations des paramètres sur le comportement du système,
- améliorer la représentation physique du système,
- cibler les améliorations à apporter à la modélisation du système,
- implémenter une méthode d'estimation de paramètres et d'assimilation de données.

Ce chapitre présente la théorie générale de l'analyse de sensibilité et de sa formulation dans un cadre mathématique. La dérivée de Gâteaux de la réponse est considérée comme la mesure la plus générale de la sensibilité d'une réponse à des variations des paramètres du système (Tomovic et Vukobratovic, 1972).

3.1.3.2 Théorie générale

Le principe général de l'analyse de sensibilité est représenté ici selon le schéma classique rencontré dans la littérature.

Le modèle est représenté selon la formulation suivante :

$$F(\mathbf{x}, \boldsymbol{\alpha}) = 0 \quad (3.4)$$

$$\begin{aligned} \mathbf{x} &: \mathfrak{R}^m \rightarrow \mathfrak{R}^n \\ \boldsymbol{\alpha} &\rightarrow \mathbf{x}(\boldsymbol{\alpha}) \end{aligned} \quad (3.5)$$

$$\begin{aligned} F &: H_x \times E_\alpha \rightarrow H \\ (\mathbf{x}, \boldsymbol{\alpha}) &\rightarrow F(\mathbf{x}, \boldsymbol{\alpha}) \end{aligned} \quad (3.6)$$

- \mathbf{x} représente le vecteur d'état du système, il appartient à un espace de Hilbert H_x
- $\boldsymbol{\alpha}$ est le vecteur des paramètres du modèle. On suppose que $\boldsymbol{\alpha}$ appartient à un espace vectoriel normé E_α et que les conditions aux limites du système ne dépendent pas de $\boldsymbol{\alpha}$.
- F est un opérateur différentiel, non linéaire en général. On suppose que l'expression des conditions aux limites est incluse dans F . H est un espace de Hilbert.

On suppose que le système (3.4) admet une unique solution $\mathbf{x}(\boldsymbol{\alpha})$ pour tout $\boldsymbol{\alpha}$ fixé dans E_α et que l'opérateur F est différentiable au sens de Gâteaux.

La dérivée de Gâteaux \hat{F} de F dans une direction α est donnée par :

$$\hat{F}(\mathbf{x}, \boldsymbol{\alpha}).\alpha = \lim_{\lambda \rightarrow 0} \frac{F(\mathbf{x}, \boldsymbol{\alpha} + \lambda.\alpha) - F(\mathbf{x}, \boldsymbol{\alpha})}{\lambda} \quad (3.7)$$

La réponse du modèle à étudiée a la forme intégrale suivante :

$$R(\mathbf{x}, \boldsymbol{\alpha}) = \int_{t_0}^{t_f} r(\mathbf{x}, \boldsymbol{\alpha}, t) dt \quad (3.8)$$

où r est une fonction à valeur réelles. D'autres types de réponses sont proposés par Cacuci (1981).

A partir d'un état de référence $(\alpha_0, \mathbf{x}_0) = (\alpha_0, \mathbf{x}(\alpha_0))$, il s'agit d'exprimer, au premier ordre, la variation totale de R due à une variation $(\delta_\alpha, \delta_x)$ autour de (α_0, \mathbf{x}_0) :

$$DR = \hat{R} = D_\alpha R(\alpha_0, \mathbf{x}_0).\delta_\alpha + D_x R(\alpha_0, \mathbf{x}_0).\delta_x \quad (3.9)$$

Avec :

$$D_\alpha R(\alpha_0, \mathbf{x}_0).\delta_\alpha = \int_{\Omega} D_\alpha r(\alpha_0, \mathbf{x}_0, t).\delta_\alpha dt \quad (3.10)$$

$$D_x R(\alpha_0, \mathbf{x}_0).\delta_x = \int_{\Omega} D_x r(\alpha_0, \mathbf{x}_0, t).\delta_x dt \quad (3.11)$$

Les variations $(\delta_\alpha, \delta_x)$ sont définies telles que $\mathbf{x}_0 + \delta\mathbf{x}$ soit solution de l'équation (3.4) pour $\alpha_0 + \delta\alpha$.

Dans l'expression (3.9) le premier terme est qualifié de terme à « effet direct », qui dépend explicitement de δ_α . Le second terme, à « effet indirect », ne dépend pas explicitement de δ_α et son calcul nécessite la connaissance de δ_x .

L'approximation au premier ordre de la différentiation de l'équation (3.4) donne une relation entre $\delta\alpha$ et $\delta\mathbf{x}$ (dérivée de Gâteaux de (3.4)).

$$D_\alpha F(x_0, \alpha_0).\delta_\alpha + D_x F(x_0, \alpha_0).\delta_x = 0 \quad (3.12)$$

En notant :

$$L = D_x F(x_0, \alpha_0) \text{ et } M = -D_\alpha F(x_0, \alpha_0) \quad (3.13)$$

La relation (3.12) devient :

$$L.\delta_x = M.\delta_\alpha \quad (3.14)$$

Ainsi, si δ_α est connu, il est possible de déterminer δ_x en résolvant le système linéaire (3.14), aussi appelé le système linéaire tangent. L'analyse de sensibilité consiste alors à calculer la variation DR de R autour de sa valeur de base $R(\alpha_0, x_0)$ à partir d'une variation δ_α de α autour de α_0 .

3.1.3.3 La méthode de sensibilité directe (FSF)

La méthode de sensibilité directe consiste à calculer DR à partir du système différentiel (3.14). La variation δ_x est déterminée à partir de δ_α donné, il est alors possible de calculer DR .

Comme le vecteur d'état du système dépend du vecteur des paramètres, il est possible de définir une fonction g telle que :

$$g : \alpha_0 \rightarrow g(\alpha_0) = R(\alpha_0, \mathbf{x}(\alpha_0)) \quad (3.15)$$

Si l'application (3.5) est différentiable, alors l'application (3.15) l'est aussi :

$$D_\alpha g(\alpha_0).\delta_\alpha = DR \quad (3.16)$$

Avec :

$$\delta_x = D_\alpha \mathbf{x}(\alpha_0).\delta_\alpha \quad (3.17)$$

Dans la pratique, le vecteur des paramètres α est souvent réel, donc $E_\alpha = \Re^p$ avec $p \geq 1$. Il est alors possible de calculer séparément les sensibilités de g par rapport aux différents paramètres du système :

$$\frac{\partial g}{\partial \alpha_i}(\alpha_0) = \frac{\partial R}{\partial \alpha_i}(\alpha_0, x_0) + D_x(\alpha_0, x_0) \cdot \frac{\partial \mathbf{x}}{\partial \alpha_i}(\alpha_0) \quad (3.18)$$

$\frac{\partial \mathbf{x}}{\partial \alpha_i}(\alpha_0)$ est déterminé par le système (3.14) avec :

$$\delta_\alpha = (0, \dots, \delta_{\alpha_i}, \dots, 0) \text{ et } \delta_x = \frac{\partial \mathbf{x}}{\partial \alpha_i} \delta_\alpha \quad (3.19)$$

Pour une variation δ_α autour de α_0 , il est alors nécessaire de résoudre le système linéaire (3.14) pour calculer δ_x (terme à « effet indirect ») et ainsi déterminer la sensibilité de g par rapport au paramètre α_i . Par conséquent pour n paramètres, la méthode directe requiert alors au moins n évaluations du modèle direct. D'un point de vue coût de calcul, la méthode directe n'est avantageuse uniquement lorsque l'objectif est de calculer simultanément la sensibilité d'un grand nombre de fonctions réponse par rapport à un nombre limité de paramètres d'entrée. Dans le cas contraire, pour un grand nombre de paramètres d'entrée, ce qui est le cas pour des modèles hydrologiques distribués, la méthode de sensibilité directe est trop coûteuse pour être utilisée.

3.1.3.4 La méthode de sensibilité par les équations adjointes (ASF)

La méthode de sensibilité utilisant les équations adjointes a été développée dans le but d'éliminer la dépendance explicite par rapport à δ_x (dans l'expression de DR). La formulation de la sensibilité adjointe nécessite l'introduction d'hypothèses restrictives supplémentaires à celles déjà introduites dans le cadre de l'analyse de sensibilité directe. Ce paragraphe décrit la formulation faible de la méthode adjointe d'analyse de sensibilité, pour un cadre mathématique rigoureux l'auteur peut se référer à Ngodock (1996) par exemple. Le formalisme faible s'adapte à la résolution des équations aux dérivées partielles modélisant les problèmes de physique classiques où la question de l'existence d'une solution est posée en termes de formulation faible du problème.

Les espaces H et H_x sont maintenant supposés espaces de Banach. Les applications G et F sont supposées continues et différentiables.

$$R : H_x \times E_\alpha \rightarrow \mathfrak{R}$$

$$F : H_x \times E_\alpha \rightarrow H$$

Il est admis que le système direct (3.4) admet une solution unique $\mathbf{x} = \mathbf{x}(\boldsymbol{\alpha})$ dans H_x pour tout $\boldsymbol{\alpha}$ fixé dans E_α (formulation faible). L'application g est alors définie de la façon suivante :

$$\begin{aligned} g : E_\alpha &\rightarrow \mathfrak{R} \\ \boldsymbol{\alpha} &\rightarrow R(\mathbf{x}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) \end{aligned} \quad (3.20)$$

Les variations autour du cas de référence $(\alpha_0, \mathbf{x}_0) = (\alpha_0, \mathbf{x}(\alpha_0))$ sont liées par la relation (3.14) et la variation Dg de g due à la perturbation δ_α s'écrit :

$$Dg = D_x R(\alpha_0, \mathbf{x}_0) \cdot \delta_x + D_\alpha R(\alpha_0, \mathbf{x}_0) \cdot \delta_\alpha \quad (3.21)$$

L'objectif de la formulation adjointe est d'exprimer Dg uniquement en fonction de la perturbation δ_α .

L'opérateur L dans (3.14) est linéaire continu entre les deux espaces de Banach H et H_x . Il est alors possible de définir un opérateur linéaire continu L^* , appelé opérateur adjoint tel que :

$$L^* : H' \rightarrow H'_x$$

$$\langle q, L.\delta_x \rangle_{H',H} = \langle L^*q, \delta_x \rangle_{H'_x, H'_x}, \text{ avec } \delta_x \in H_x \text{ et } q \in H'$$

L'isométrie canonique I de H dans H' est définie telle que pour tout $q \in H$:

$$\langle I(q), p \rangle_{H',H} = \langle p, q \rangle_{H,H}, \quad \forall p \in H'$$

Alors,

$$\begin{aligned} \langle L.\delta_x, q \rangle_{H,H} &= \langle I(q), L.\delta_x \rangle_{H',H} \\ &= \langle L^*I(q), \delta_x \rangle_{H'_x, H'_x} \end{aligned} \quad (3.22)$$

Comme R est différentiable, alors :

$$D_x R(x_0, \alpha_0) \in H'_x$$

Il est possible d'introduire le système adjoint :

$$\begin{cases} L^*.I(q) = D_x R(x_0, \alpha_0) \\ q \in H \end{cases} \quad (3.23)$$

Si q_0 est la solution du système adjoint, alors en utilisant la relation (3.22) :

$$\begin{aligned} D_x R(x_0, \alpha_0).\delta_x &= \langle L^*I(q_0), \delta_x \rangle_{H'_x, H'_x} \\ &= \langle L.\delta_x, q_0 \rangle_{H,H} \\ &= \langle M.\delta_\alpha, q_0 \rangle_{H,H} \end{aligned}$$

Et par conséquent :

$$Dg = \langle M.\delta_\alpha, q_0 \rangle_{H,H} + D_\alpha R(x_0, \alpha_0).\delta_\alpha \quad (3.24)$$

Dans l'expression de la sensibilité de g par rapport au paramètre α_i , il n'y a maintenant plus de dépendance explicite par rapport à δ_x .

3.1.4 Conclusion

La méthode de sensibilité directe (FSF) est un formalisme avantageux pour évaluer les effets de variations de relativement peu de paramètres sur plusieurs réponses du système. La méthode directe est plus simple et plus générale, cependant elle nécessite la résolution du modèle linéaire tangent qui demande autant d'évaluations du modèle que le nombre de paramètres du modèle. Finalement, elle peut se révéler beaucoup trop coûteuse pour des problèmes avec un grand nombre de paramètres d'entrée, comme dans le cas des modèles distribués.

Au contraire, pour étudier l'influence des variations d'un grand nombre de paramètres sur relativement peu de réponses, le formalisme de sensibilité adjoint (ASF) est plus avantageux au niveau coût de calcul. En effet la sensibilité d'une réponse à tous les paramètres du modèle et à tous les états précédents du modèle peut être évaluée par une simple intégration du modèle adjoint. Contrairement au formalisme direct, l'expression des équations adjointes ne change pas en fonction du nombre de paramètres considérés. Finalement, la méthode adjointe est la plus efficace pour calculer des sensibilités locales de systèmes à grande échelle où le nombre de paramètres est supérieur au nombre de fonctions réponses étudiées. Par exemple, la méthode de l'adjoint permet l'analyse des réponses du modèle spatialement distribuées à moindre coût. En outre, les outils de différentiation automatique (2.2.3.5) facilitent le calcul des dérivées de fonctions directement à partir du code source implémentant le modèle. L'examen des dérivées calculées améliore la connaissance du fonctionnement du modèle.

Par rapport aux méthodes globales qui apportent des mesures de sensibilité moyennes sur l'espace échantillonné, les méthodes d'analyse de sensibilité locales donnent des mesures plus précises en termes d'amplitude et de signe. Si l'indice de sensibilité d'un paramètre est positif, l'augmentation de la valeur de ce paramètre va contribuer à l'augmentation de la valeur de la fonction réponse étudiée.

Cependant, les méthodes locales sont restreintes à un point de l'espace de contrôle et dans le cas de modèles non-linéaires, les valeurs nominales des paramètres pour lesquelles sont calculées les dérivées ont une grande importance. L'influence des paramètres ne sera pas la même pour différents points de l'espace de contrôle.

Par ailleurs, le calcul des dérivées se fait souvent au premier ordre (2.2.3.6) à cause du coût de calcul important qu'engendre le calcul des dérivées d'ordre plus élevé. Ainsi, il n'est pas possible d'évaluer les interactions potentielles entre les paramètres avec les méthodes locales.

Par conséquent, deux types d'analyse de sensibilité (globale et locale) ont été menés pour le modèle MARINE. La première est une étude globale, réalisée avec la méthode de sensibilité généralisée (Le, 2008), ayant pour objectif la détermination des paramètres les plus sensibles du modèle. Afin de réduire l'espace de contrôle des paramètres, des coefficients multiplicatifs ont été utilisés pour ajuster l'amplitude d'une distribution spatiale fixée a priori des paramètres du modèle (5.2). La seconde analyse de sensibilité présentée dans ce mémoire est une analyse de sensibilité locale (7.2.1) post-calibration pour une étude spatiale de la réponse du modèle.

3.2 Optimisation

3.2.1 Introduction

Les modèles hydrologiques de tout type (à base physique, empiriques ou conceptuels) sont caractérisés par des fonctions mathématiques exprimant la relation entre la pluie et le débit. Généralement, ces fonctions contiennent des paramètres qui ne sont pas mesurables à l'échelle du bassin versant ou qui ne permettent pas une interprétation physique directe. Il est alors nécessaire de procéder au calage du modèle pour estimer ces paramètres afin d'en améliorer les prévisions ou les simulations. Le processus de calage consiste à minimiser, manuellement ou automatiquement, la distance entre l'hydrogramme simulé par le modèle et l'hydrogramme observé en cherchant un jeu de paramètres optimum (i.e. qui donne la meilleure simulation possible).

« The task of the modeller is to identify a model, i.e., a parameter set within a specific model structure, that is suitable for the given catchment characteristics, data and modelling propose" (Wagener, 1998).

Le calage d'un modèle est une tâche difficile. D'une part, les processus hydrologiques sont caractérisés par un grand nombre de paramètres dont les contributions sont mal connues. D'autre part, de nombreuses incertitudes existent dans le processus de modélisation du fonctionnement d'un bassin versant et sont inévitablement propagées dans les sorties du modèle. Ces incertitudes proviennent de différentes sources (Kitanidis et Bras, 1980) :

- Incertitudes dans les processus naturels dues à leurs fluctuations temporelles et spatiales aléatoires (précipitations par exemple),
- Incertitudes sur les données dues aux erreurs de mesures ou de leur prétraitement,
- incertitudes sur la spécification du modèle qui se traduisent par l'incapacité du modèle à converger vers un unique jeu de paramètres « optimal » à partir des données disponibles,
- incertitudes sur la structure du modèle introduites lors des simplifications et des approximations faites sur la description du processus réel.

Il est important de noter que même si ces incertitudes pouvaient être diminuées, il resterait toujours une part d'imprévisibilité dans les processus naturels (Melchin et al., 1990) dont les incertitudes associées ne peuvent pas être réduites.

Les incertitudes d'échelle des modèles distribués viennent également s'ajouter à la liste précédente, elles sont de trois types (Dehotin et Braud, 2007) :

- l'échelle à laquelle l'équation différentielle représente le système physique,
- l'échelle du modèle, c'est-à-dire la taille du maillage utilisé pour évaluer les équations,
- l'échelle des observations utilisées pour construire et calibrer le modèle.

Une méthode d'optimisation doit être efficace, efficiente et robuste. C'est-à-dire, qu'elle doit être capable de trouver l'optimum de la surface de réponse (ou le jeu de paramètres optimal) rapidement (ou avec un nombre d'itérations limité). Le jeu de paramètres trouvé doit être robuste, c'est-à-dire adapté à des conditions hydrologiques différentes de la période de calage. Les questions soulevées lors de l'utilisation des méthodes de d'optimisation sont nombreuses et difficiles à résoudre dans un contexte général. Elles concernent :

- le niveau correct d'adéquation (ajustement) entre les simulations du modèle et les observations du système physique,
- la mesure de ce niveau d'adéquation,
- l'existence d'un seul jeu de paramètres optimal,
- les informations qui peuvent être tirées du (des) jeu(x) de paramètres optimisé(s).

3.2.2 Critère d'évaluation des performances d'un modèle

Il existe un grand nombre de critères d'évaluation pour juger des performances d'un modèle. Selon Green et Stephenson (1986) ou Weglarczyk (1998), il n'y a pas de critère idéal ou unique qui permette cette évaluation. Le choix d'un critère particulier parmi la grande diversité des formulations existantes résulte en général du choix subjectif des utilisateurs du modèle.

Les deux principaux types de critères sont les critères qualitatifs et les critères quantitatifs. Les critères qualitatifs sont principalement basés sur l'interprétation visuelle des hydrogrammes. Ce type d'évaluation est fréquemment utilisé pour comparer les performances du modèle (Seibert, 2001 ; Perrin et al, 2005) ; en particulier pour comparer deux simulations l'une par rapport à l'autre. Pour la prévision de crue, où l'intérêt majeur est la bonne représentation de la phase de montée et la valeur du pic de crue, ce type de critère est très utile. Cependant, la comparaison graphique est très subjective et elle n'est pas pratique lorsqu'il s'agit de comparer un grand nombre de simulations d'un même modèle ou de modèles différents. Les critères quantitatifs sont généralement adimensionnels pour permettre de comparer différents modèles ou simulations sur plusieurs bassins versants ayant des caractéristiques différentes.

3.2.2.1 L'erreur quadratique moyenne

Un critère très utilisé dans les modèles pluie-débit est l'erreur quadratique moyenne (*RMSE*, root mean square error) qui a la même unité que la variable observée :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{obs,i} - Q_{sim,i})^2} \quad (3.25)$$

où $Q_{obs,i}$ et $Q_{sim,i}$ sont respectivement le débit observé et simulé au pas de temps i et N le nombre de pas de temps.

Plus l'erreur quadratique moyenne est proche de zéro, plus le modèle est performant.

3.2.2.2 Le critère de Nash et Sutcliffe

Nash et Sutcliffe (1970) ont construit une forme dérivée de la mesure précédente. Le critère de Nash (*NSE*) permet par exemple de comparer la performance du modèle sur différents bassins versants. Il s'écrit :

$$NSE = 1 - \frac{\sum_{i=1}^N (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^N (Q_{obs,i} - \overline{Q_{obs}})^2} = 1 - \frac{SE}{V} \quad (3.26)$$

où $Q_{obs,i}$ et $Q_{sim,i}$ sont respectivement le débit observé et simulé au pas de temps i et N le nombre de pas de temps. $\overline{Q_{obs}}$ est la moyenne des débits observés. SE représente la somme des erreurs quadratiques du modèle. V représente la variance du débit observé.

SE est fonction de la valeur des débits, ainsi pour comparer deux périodes, deux événements ou deux bassins versants entre eux, SE est normalisé par V . Le modèle étudié est ainsi comparé au modèle $Q = \overline{Q_{obs}}$. Le critère de Nash est donc adimensionnel, il varie dans l'intervalle $]-\infty, 1]$. Si NSE négatif, le modèle est moins bon que le modèle de référence ($Q = \overline{Q_{obs}}$), il est meilleur dans le cas contraire. La valeur 1 indique que le modèle est parfait.

3.2.2.3 L'analyse multi-critères

Des problèmes ont été soulevés avec l'utilisation des deux mesures précédentes : elles amplifient le poids de certains aspects de l'hydrogramme (période de hautes eaux par exemple) et apportent différents jeux de paramètres optimaux. Ceci est dû à l'utilisation du résidu au carré.

Les modèles ne sont pas capables de reproduire tous les aspects de l'hydrogramme avec un unique jeu de paramètres (Gupta et al., 1998 ; Wagener et al., 2001). Le modélisateur doit sacrifier l'adéquation à un aspect de l'hydrogramme au détriment d'un autre. Gupta et al. (1998) pensent alors que l'approche multi-critères pour la calibration automatique peut être avantageuse.

Il est possible également de combiner plusieurs fonctions objectif dans une seule valeur, ce qui permet par la suite d'utiliser des algorithmes d'optimisation classiques (Seibert, 2000).

Par exemple, dans son travail de thèse concernant l'analyse de sensibilité du modèle MARINE (5.2.3), Le (2008) a utilisé un critère défini de la façon suivante :

$$J = aNSE + b \left(1 - \frac{|\Delta T|}{T_c} \right) + c \left(1 - \frac{|\Delta Q|}{Q_{max}} \right) \quad (3.27)$$

où a , b , c sont des coefficients arbitraires dont la somme est égale à l'unité ; ΔT et ΔQ sont respectivement la différence entre la valeur observée et calculée de la date et du débit, Q_{max} est la valeur du débit maximal observé et T_c le temps de concentration. Ce critère a été choisi dans le but d'accorder plus de poids à la différence entre les pics de crue observé et calculé, caractéristique importante dans l'étude des crues éclair.

3.2.3 Equifinalité

Dans les années 90, Beven a introduit la notion d'équifinalité. Le fondement de cette approche réside dans le constat suivant : il n'y a pas en général un seul et unique jeu de paramètres optimal dans le calage des modèles distribués. Il existe donc une famille de solutions ou des régions d'une surface de réponse donnant des résultats similaires.

3.2.4 Différentes stratégies d'optimisation des modèles hydrologiques

3.2.4.1 Caractéristiques d'un algorithme d'optimisation

Un algorithme d'optimisation doit posséder les trois caractéristiques suivantes :

- une fonction objectif définie pour mesurer les écarts entre les débits simulés et les débits observés,
- une stratégie de recherche de l'optimum de la fonction objectif dans l'espace des paramètres,
- un critère d'arrêt qui stoppe l'algorithme de recherche.

Les interactions complexes entre la fonction objectif, la stratégie d'optimisation, les données et la structure du modèle testé conditionnent la réussite de l'optimisation.

3.2.4.2 L'optimisation manuelle ou « expertale »

La méthode la plus intuitive pour calibrer un modèle, bien que difficilement quantifiable est la calibration manuelle qui utilise comme critères d'évaluation de la performance du modèle « l'interprétation visuelle des hydrogrammes ». Le calage expertal des paramètres des modèles hydrologiques permet d'obtenir des jeux paramètres capables de produire de bonnes simulations. Cependant, le calage expertal est un exercice difficile et fastidieux dont la réussite dépend de l'expérience du modélisateur car elle nécessite une connaissance profonde du modèle et de son application. Il est en effet souvent difficile d'évaluer le « meilleur » jeu de paramètres. C'est donc une méthode subjective, difficilement formalisable.

Une procédure de calibration manuelle très avancée a été mise en place au National Weather Service (NWS) pour la calibration du modèle SAC-SMA (Sacramento Soil Moisture Accounting) (Boyle et al., 2000). Cependant, comme le soulignent Boyle et al. (2000), cette calibration est très longue (plusieurs centaines d'heures de travail) et nécessite une bonne compréhension du comportement du bassin et de la structure du modèle. Le choix de l'arrêt de la calibration est très subjectif et le transfert des compétences d'une personne à une autre est difficile.

3.2.4.3 L'optimisation automatique

Les méthodes d'optimisation automatiques sont quant à elles objectives et plus rapides que la méthode manuelle. Elles peuvent être de deux types : locale ou globale. A cause de la complexité de la surface de réponse des modèles hydrologiques, les stratégies automatiques d'optimisation ne sont pas parfaites et manquent soit d'efficacité (pour trouver le « meilleur » jeu de paramètre) soit d'efficience (un très grand nombre de vecteurs de paramètre sont testés avant d'en trouver un bon). En effet, comme le rapportent de nombreux auteurs (Johnston et Pilgrim, 1976 ; Duan et al., 1992 ; Sorooshian et Gupta, 1995, etc.), les surfaces de réponse des modèles hydrologiques sont complexes et peuvent présenter :

- de nombreuses régions d'attraction ou zones de convergence, ce qui rend le vecteur de paramètres optimal dépendant du point de départ de l'algorithme d'optimisation. La présence de ces optima locaux mineurs ou petits pics de la surface de réponse peut piéger l'algorithme d'optimisation,
- des points singuliers responsables de discontinuités dues à l'utilisation de seuils dans la structure du modèle et impliquant la non dérivabilité de la surface de réponse. Des discontinuités apparaissent également sur les dérivées de premier et second ordre de la surface de réponse,
- des zones de sensibilité différentes par rapport aux différents paramètres montrant alors des interactions entre les paramètres du modèle,
- des surfaces non convexes autour de l'optimum global.

Les méthodes locales

A partir d'un jeu de paramètres et d'une fonction objectif donnés, la méthode d'optimisation locale modifie de manière itérative la valeur du vecteur de paramètres jusqu'à ce qu'un critère d'arrêt spécifié au préalable soit satisfait. Les méthodes locales peuvent être de deux types : directe ou gradient (Tab. 3.1). La méthode directe utilise uniquement la valeur de la fonction objectif au cours de son itération. La méthode de type gradient utilise à la fois la valeur de la fonction objectif et son gradient pour calculer la direction de descente et le pas de convergence. La méthode de type gradient est plus puissante car le calcul des directions de convergence permet une convergence plus rapide.

Les caractéristiques de la surface de réponse listées ci-dessus impliquent que les méthodes d'optimisation locales soient particulièrement dépendantes du point de départ de l'algorithme de recherche, de sa distance à l'optimum et de la dimension de l'espace dans lequel elles évoluent. Si le point de départ est trop éloigné de l'optimum, l'algorithme aura du mal à converger vers celui-ci. Par contre, si le point de départ est bien choisi (par exemple si les gammes de paramètres optimaux des modèles sont connus), alors les méthodes locales peuvent s'avérer robustes et efficaces. Par ailleurs, ces méthodes sont souvent moins gourmandes en temps de calcul que les méthodes globales. Kuzmin et al. (2004) ont utilisé des relations de régionalisation donnant des valeurs du vecteur de paramètres a priori pour déterminer le point de départ d'une méthode de recherche locale. Par ce biais, le vecteur de paramètres a priori est situé dans la région d'attraction de l'optimum de la surface de réponse et l'algorithme de recherche local converge sans difficulté majeure vers l'optimum de la surface de réponse.

Tableau 3.1. Exemples de méthodes locales

Méthode	Type	Référence
Rosenbrock	directe	Rosenbrock, 1960
Simplex	directe	Nelder et Mead, 1965
Pas à pas	directe	Michel, 1989 Edijatno, 1991 Nascimento, 1995
Davidon-Fletcher-Powell	gradient	Fletcher et Powell, 1963
Levenberg-Marquardt	gradient	Levenberg, 1944 Marquardt, 1963
Type Newton (Newton, quasi-Newton, Gauss-Newton ...)	gradient	Lemarechal et Panier, 2000

Les méthodes globales

Les méthodes globales ont été développées dans le but de corriger les inconvénients des méthodes locales et de garantir l'indépendance du résultat par rapport au point de départ de l'algorithme et des optima locaux. Elles consistent alors à explorer de manière exhaustive la surface de réponse de façon déterministe ou stochastique (Tab. 3.2). Les méthodes déterministes ne font pas d'hypothèse a priori sur la localisation de l'extremum et discrétisent l'espace des paramètres de façon régulière en testant chaque nœud d'un maillage déterminé (Exhaustive Gridding, Kuczera, 1997). Théoriquement, ces méthodes localisent parfaitement l'optimum d'une surface de réponse mais elles nécessitent une infinité d'itérations.

Les méthodes stochastiques explorent l'espace des paramètres de manière aléatoire. Dans le cas le plus simple, la distribution est supposée uniforme.

Ces méthodes nécessitent un très grand nombre d'itérations pour atteindre l'optimum car elles ne tirent généralement pas profit de l'information acquise au cours des itérations précédentes. Par conséquent le temps de calcul est beaucoup plus important que pour les méthodes locales. Dans le but d'intégrer la mémorisation des résultats des itérations précédentes, des méthodes basées sur des concepts d'évolution de populations de systèmes physiques ou naturels ont été développées, comme les algorithmes génétiques (Wang, 1991 ; Wang, 1997). Enfin, la méthode SCE-UA (Shuffle Complex Evolution Algorithm, University of Arizona) proposée par Duan et al. (1992) combine plusieurs stratégies d'optimisation (stratégie de multi-départ, algorithmes génétiques, recherche locale et échange d'informations au cours de l'optimisation). C'est une méthode globale très utilisée dans la communauté des hydrologues qui a été comparée avec succès dans plusieurs études à d'autres méthodes d'optimisation : à la méthode simplex (Abdulla et al., 1999), à la méthode multi-start simplex et à une approche multi-start quasi-Newton (Kuczera, 1997), à une méthode d'algorithme génétique.

Tableau 3.2. Exemples de méthodes globales

Méthode	Type	Référence
Exhaustive Gridding	Déterministe	Kuczera, 1997
Multi-start simplex	Multi-départ	Johnston et Pilgrim, 1976
Genetic algorithm	Stochastique	Wang, 1991, 1997
Shuffle Complex Evolution Algorithm, University of Arizona	Stochastique	Duan et al., 1992

L'approche multi-critères

Les combinaisons de paramètres estimés sont dépendantes du critère d'estimation et de la série de données, il n'existe pas de critère universel pour juger de la qualité de l'ajustement des variables simulées par le modèle aux observations.

L'approche multi-critères est utilisée lorsque plusieurs fonctions objectives interviennent dans la procédure d'optimisation. Ces procédures utilisent des algorithmes d'évolution de population pour estimer le jeu de paramètres Pareto optimal. Un jeu de paramètres est Pareto optimal si la performance de l'une des fonctions objectives ne peut être améliorée qu'au détriment d'au moins une autre fonction objective. L'exploration du front de Pareto permet d'analyser les résultats obtenus pour différentes combinaisons de poids accordés à chaque fonction objective. Yapo et al. (1998) ont modifié l'algorithme SCE pour y introduire une analyse multi-critères : la méthode MOCEM (Multi-Objective Complex Evolution Method). Cette méthode minimise différentes fonctions objectives avec une seule exécution de l'algorithme d'optimisation. Les algorithmes génétiques peuvent également être utilisés pour mener des analyses multi-objectives (Liong et al., 2001).

La méthode GLUE (Generalized Likelihood Uncertainty Estimation)

Après avoir reconnu le concept d'équifinalité, la méthode GLUE de Beven et Binley (1992) a été développée dans ce sens puisqu'elle reconnaît la quasi-équivalence de différents jeux de paramètres dans le calage. Elle est basée sur la méthode de Monte Carlo qui consiste à faire de nombreuses simulations d'un modèle avec des jeux de paramètres choisis aléatoirement pour réaliser une exploration de la gamme des simulations possibles. Il s'agit d'attribuer à chaque jeu de paramètres une mesure de vraisemblance en comparant les simulations du modèle et les données d'observation. Les valeurs de vraisemblance obtenues sont utilisées pour calculer des intervalles de confiance et estimer l'incertitude associée aux prédictions du modèle. Cette méthode basée sur une approche bayésienne permet également d'évaluer l'apport de nouvelles observations en modifiant les valeurs de vraisemblance. Cependant, cette méthode nécessite un temps de calcul élevé. Cette méthode a été défendue (Freer et al., 1996 ; Beven et Freer, 2001 ; Beven, 2006) mais également critiquée (Montanari, 2005 ; Mantova et Todini, 2006).

3.2.5 Algorithme d'optimisation utilisant la méthode de l'état adjoint

La méthode de l'état adjoint issue de la théorie du contrôle optimal permet d'évaluer de façon précise et efficace le gradient de la fonction objective par rapport aux paramètres de contrôle, quelle que soit la dimension du vecteur de contrôle. La minimisation de la fonction objective est ensuite réalisée grâce à des méthodes de type gradient. L'efficacité de ces méthodes d'optimisation est améliorée par l'utilisation de la méthode de l'état adjoint (White et al., 2003).

La méthode employée dans cette thèse est une méthode de descente de type quasi-Newton (Lemaréchal et Panier, 2000). Les méthodes de quasi-Newton sont basées sur le même principe que les méthodes de Newton mais elles évitent le calcul de la Jacobienne ou de la Hessienne.

Le système d'optimisation peut se résumer de la façon suivante (2.2.3):

- définition du modèle direct
- définition du modèle adjoint
- calcul du gradient de la fonction coût J
- détermination de la condition nécessaire d'optimalité : $\nabla J = 0$

L'algorithme d'optimisation est schématisé de la façon suivante :

- initialisation du vecteur de contrôle
- itération k
 - o intégration du modèle directe de 0 à T pour le calcul de J
 - o intégration du modèle adjoint de T à 0
 - o calcul du gradient de la fonction coût
 - o calcul d'une direction de descente
 - o calcul d'un pas de descente optimal minimisant la fonction coût suivant la direction
- critère d'arrêt

Dans une optique de prévision en temps réel, le même formalisme peut être employé pour la mise à jour des conditions initiales et des coefficients correctifs sur les variables de forçage à travers l'assimilation des observations du système physique.

3.2.6 Méthodologie des expériences jumelles

3.2.6.1 Introduction

Le principe général de la méthodologie utilisée dans cette thèse est présenté dans ce paragraphe. Elle s'appuie sur la méthode de l'état adjoint avec l'algorithme de minimisation N2QN1 de type quasi-Newton avec contraintes de bornes (Lemaréchal et Panier, 2000) pour l'estimation de paramètres à l'aide d'expériences jumelles. L'objectif des expériences jumelles est d'évaluer les performances de cette méthode de descente afin de valider la procédure d'estimation de paramètre employée.

3.2.6.2 Principe général

Des paramètres de référence du modèle servent à générer des observations synthétiques. Soit α^{ref} ce jeu de paramètres de référence à estimer grâce à la résolution du système d'optimalité. Le modèle hydrologique est appelé P . La chronique de débits est obtenue par exécution du code direct P :

$$P(\alpha^{ref}) = Q^{ref}$$

Les débits Q^{ref} obtenus vont servir d'observations.

Après perturbation $\delta\alpha$ du jeu de paramètres α^{ref} , le code direct P apporte une chronique de débits Q :

$$P(\alpha^{ref} + \delta\alpha) = Q$$

Les jeux d'observations et de paramètres perturbé vont être utilisés afin de retrouver le jeu de paramètres de référence grâce à l'utilisation du code adjoint et à la minimisation de la fonctionnelle :

$$J(\alpha^{ref}) = \frac{1}{2} \|CQ - Q^{ref}\|_{obs}^2 \quad (3.28)$$

Les expériences jumelles constituent un cadre idéal où la seule source d'incertitude est celle liée aux paramètres du modèle. Cette méthodologie a l'avantage d'apporter autant d'observations désirées et d'éviter les erreurs dues à la structure du modèle. L'objectif est d'évaluer la stabilité du problème d'estimation et de valider les méthodes testées.

3.2.6.3 Estimation des paramètres avec la méthode de l'adjoint

Le problème de l'estimation des paramètres, souvent appelé problème inverse, consiste à assimiler les données relatives à la période de simulation pour effectuer un calage des paramètres représentatifs pour toute la période.

Dans le cadre des expériences jumelles, l'objectif est de retrouver le jeu de paramètres de référence à partir de conditions initiales différentes.

- La fonction objectif

Une fonction coût, mesurant les écarts entre les simulations et les observations, est définie et implémentée dans le modèle étudié. Une large gamme de fonction coût est disponible et son choix résulte d'une décision subjective qui dépend de l'objectif du modélisateur. Par exemple, dans cette étude, la fonction objectif utilisée pour l'algorithme d'optimisation est définie de la façon suivante :

$$J = 1 - NSE = \frac{\sum_{i=1}^N (Q_i^{ref} - Q_i)^2}{\sum_{i=1}^N (Q_i^{ref} - \overline{Q^{ref}})^2} \quad (3.29)$$

où Q_i^{ref} et Q_i sont respectivement le débit de référence (observé) et simulé au pas de temps i , N le nombre de pas de temps, $\overline{Q^{ref}}$ est la moyenne des débits observés. NSE est le critère de Nash (3.2.2.2).

- Le code adjoint

Dans cette thèse, le code adjoint du modèle est obtenu grâce à l'outil de différentiation automatique TAPENADE (Tangent and Adjoint PENultimate Automatic Differentiation Engine – Hacoët et Pascual, 2004), successeur d'ODYSEE (Faure et Papegay, 1998) et développé depuis 1999. Cet outil est basé sur la transformation de code source écrit en langage FORTRAN, comme le modèle MARINE (chapitre 4). Il est disponible en ligne à l'adresse <http://www-sop.inria.fr/tropics/>. Le code adjoint obtenu est ensuite validé par des tests du gradient et du produit scalaire.

A titre d'illustration, pour le modèle (MARINE) et la fonction coût étudiés, les résultats des tests du produit scalaire et du gradient effectués avec le code linéaire tangent et le code adjoint sont donnés sur la figure 3.2. On voit que le code adjoint obtenu est bien valide (les produits scalaires sont égaux et le test de Taylor vérifié : $\lim_{\alpha \rightarrow 0} (1 - \alpha) = 0$ - annexe A) et on obtient un gradient d'une très grande précision vers lequel converge l'approximation par les différences finies jusqu'à l'apparition des erreurs d'arrondi.

Le code adjoint obtenu permet alors de calculer les dérivées de la fonction objectif pour la détermination des directions de descente de l'algorithme d'optimisation

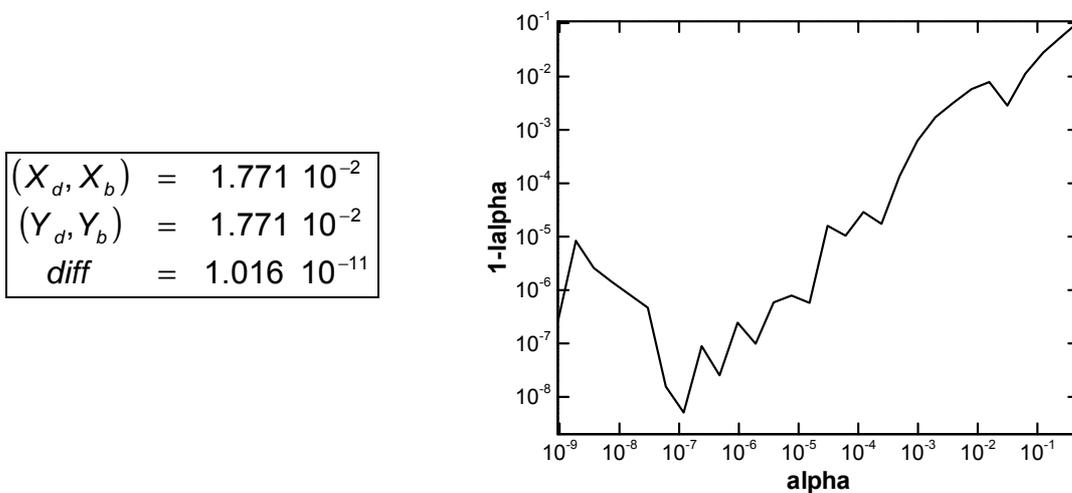


Fig. 3.2 : Tests du produit scalaire et test du gradient pour la validation du code adjoint de MARINE.

- L'algorithme d'optimisation

L'algorithme N2QN1 de la librairie MODULOPT (Lemarechal et Panier, 2004) est utilisé pour estimer les paramètres. C'est un algorithme de quasi-Newton qui résout les problèmes d'optimisation avec contraintes de bornes. Ce type d'algorithme présente deux principaux avantages : premièrement, il est robuste et convient aux modèles distribués, deuxièmement, il est adapté à la minimisation de fonction coût non lisse.

Si les deux bornes sont définies par deux vecteurs $\mathbf{b}_{inf} \in \mathfrak{R}^n$ et $\mathbf{b}_{sup} \in \mathfrak{R}^n$.

Alors le problème d'optimisation se formule de la façon suivante :

$$\left\{ \begin{array}{l} \min J(\boldsymbol{\alpha}) \\ \mathbf{b}_{inf} \leq \boldsymbol{\alpha} \leq \mathbf{b}_{sup} \end{array} \right. \Leftrightarrow \forall i \in [1, n] \left\{ \begin{array}{l} \min J(\boldsymbol{\alpha}) \\ (b_{inf})_i \leq \alpha_i \leq (b_{sup})_i \end{array} \right.$$

Si le vecteur $\boldsymbol{\alpha}$ est solution du problème de minimisation et qu'il est compris entre ces deux bornes alors il est admissible.

Cet algorithme utilise une méthode de quasi-Newton, nommée BFGS (Liu et Nocedal, 1989) ainsi que la stratégie de contraintes actives « à la Rosen-Goldfard ».

- Les calculs d'erreurs

Il est possible avec les expériences jumelles de comparer les paramètres estimés avec l'assimilation de données α^{assim} avec les paramètres de référence α^{ref} du modèle grâce au calcul de l'erreur relative suivant :

$$\delta_{\alpha}^{assim} = \left| \frac{\alpha^{assim} - \alpha^{ref}}{\alpha^{ref}} \right| \quad (3.30)$$

L'erreur relative de l'initialisation de l'assimilation α^{init} par rapport à l'état de référence α^{ref} s'écrit :

$$\delta_{\alpha}^{init} = \left| \frac{\alpha^{init} - \alpha^{ref}}{\alpha^{ref}} \right| \quad (3.31)$$

La figure 3.3 récapitule toutes les étapes de la méthodologie.

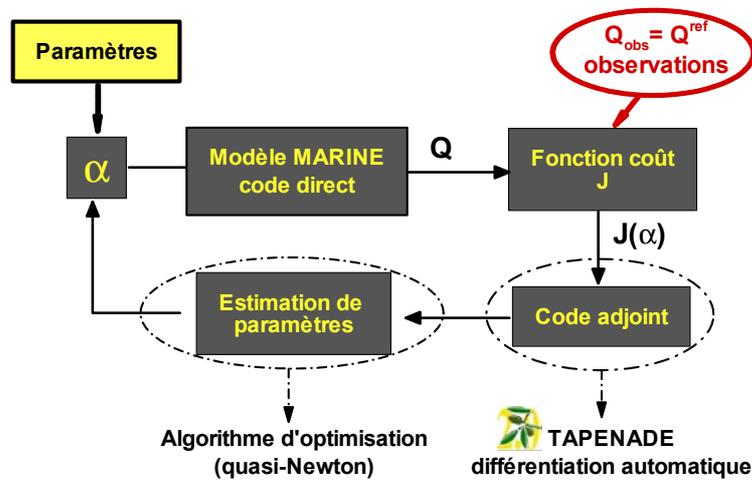


Fig. 3.3 : Schéma représentant le processus d'estimation des paramètres du modèle MARINE avec la méthode adjointe.

3.2.7 Synthèse

La calibration est une procédure subjective biaisée par le choix d'une fonction objectif. Toutefois, en supposant qu'il soit possible de mesurer parfaitement la capacité du modèle à reproduire les données observées on peut se demander comment le modèle devrait s'ajuster aux données. La réponse évidente serait de dire que le modèle doit parfaitement reproduire les données d'observation. Cependant, est-il réellement possible d'avoir un ajustement parfait entre les simulations du modèle et les observations ? Il est bon de rappeler que le modèle n'est qu'une simplification de la réalité, il représente au mieux les dynamiques simplifiées des processus physiques. De plus, il repose sur des données imparfaites, échantillonnées en temps et en espace. Il y a alors peu de chance qu'un jeu de paramètres reproduise exactement le comportement du système étudié. Ainsi, il est illusoire d'avoir un modèle parfait mais l'interprétation des imperfections permet un enrichissement raisonné lorsque le modèle est à base physique.

Indépendamment de l'incertitude associée aux observations et à la structure du modèle, la multi-modalité due à une paramétrisation trop complexe est la principale source d'équifinalité. Lorsque la fonction à minimiser comporte une zone unique d'attraction, les algorithmes d'optimisation basés sur le gradient sont plus efficaces que les méthodes globales. Dans la pratique, il est fréquent de coupler les méthodes d'analyse de sensibilité et d'optimisation globale et locale. Une analyse de sensibilité globale permet d'identifier les gammes de paramètres optimaux et la méthode locale utilise ses paramètres comme point de départ pour affiner la recherche. Ainsi l'estimation a priori des paramètres est améliorée et facilite l'utilisation des méthodes basées sur le gradient. De plus, les travaux réalisés par Kitanidis et Bras (1980) sur le lissage des seuils de fonctionnement éliminent des difficultés d'utilisation des méthodes de descente basée sur le gradient.

Dans ce travail de thèse, la calibration du modèle MARINE est réalisée. Une analyse de sensibilité globale préliminaire (GSA) a permis de déterminer les paramètres les plus sensibles du modèle, puis une méthode d'optimisation locale (méthode de type quasi-Newton en utilisant le calcul des dérivées avec la méthode de l'état adjoint) est utilisée par la détermination de la valeur de ces paramètres. Deux types de fonctions objectifs sont utilisées, l'une classique similaire au critère de Nash ; l'autre plus généralisée permettant d'intégrer des incertitudes sur les débits observés.

Chapitre 4

Présentation du modèle, du bassin versant, des observations et des critères de performance

4.1	MARINE : un modèle distribuée à base physique pour la modélisation des crues à cinétique rapide	76
4.1.1	Introduction	76
4.1.2	Données nécessaires au modèle MARINE	77
4.1.3	Equations représentatives du comportement des processus physiques	78
4.1.4	Prédétermination des paramètres du modèle	89
4.1.5	Synthèse	92
4.2	Bassin d'étude : le Gardon d'Anduze	93
4.2.1	Introduction	93
4.2.2	Description	93
4.2.3	Caractéristiques	94
4.2.4	Base de données disponibles	100
4.3	Les événements sélectionnés	108
4.3.1	Critères de sélection	108
4.3.2	Caractéristiques des événements sélectionnés	108
4.3.3	Critiques des données	109
4.3.4	Bilan des événements sélectionnés	113
4.4	Validité de l'approximation de l'onde cinématique sur le bassin versant	114
4.5	Formulation des critères de performance	115
4.5.1	Un critère classique	115
4.5.2	Un critère généralisé	115
4.6	Synthèse	117

CHAPITRE 4 Présentation du modèle, du bassin versant, des observations et des critères de performance

4.1 MARINE : un modèle distribuée à base physique pour la modélisation des crues à cinétique rapide

4.1.1 Introduction

La modélisation des crues éclair est très délicate à cause de la complexité du comportement du bassin versant lors de ces phénomènes et de l'importante variabilité spatio-temporelle des pluies responsables. De plus, la difficulté d'observation et de mesure de ces crues introduit un obstacle supplémentaire à leur modélisation, en particulier lors de la phase de calage.

Ainsi le modèle choisi parmi la multitude des modèles existants résulte d'un compromis entre le nombre de données, le coût et la rapidité de calcul (pour une utilisation temps réelle) et la précision des résultats attendus.

Le modèle MARINE (Modélisation de l'Anticipation du Ruissellement et des Inondations pour des Evénements Extrêmes) est développé au sein de l'équipe HydroEco de l'Institut de Mécanique des Fluides de Toulouse (IMFT) depuis la thèse de Estupina (2004). Il s'agit d'un modèle hydrologique événementiel, distribué à base physique, dédié à la prévention et à la prévision des crues éclair générées par des petits bassins versant (quelques centaines de km²) à cinétique rapide. Ce modèle de type pluie-débit peut intégrer des informations issues de l'imagerie satellitale (Ottlé et al., 2003) (Fig. 4.1).

Le modèle MARINE est un modèle perceptuel (Sivapalan, 2003) ayant pour objectif la modélisation des processus prépondérants (connus, dominants et influents) lors la genèse des crues à cinétique rapide. Il fait du ruissellement par dépassement de la capacité d'infiltration le processus prépondérant expliquant la genèse de ces crues. Seule l'infiltration dans la zone non-saturée est prise en compte, les processus d'interception et d'évapotranspiration sont supposés négligeables sur la genèse de la crue. Les sorties du modèle peuvent être des hydrogrammes en tout point du réseau de drainage (variables ponctuelles distribuées temporellement) ou des variables comme la hauteur d'eau sur le bassin versant, la hauteur d'eau dans le sol, la vitesse d'écoulement, ... (distribuées spatialement à tout instant de la simulation) (Fig. 4.1).

L'information fournie doit être pertinente au sens de la prévision des crues. Les temps de calcul doivent être compatibles avec les contraintes de prévision en temps réel : les simulations doivent être terminées avant que de nouvelles données de pluie ne soient disponibles.

Les résultats obtenus doivent apporter un certain nombre d'informations sur les caractéristiques de l'hydrogramme, en particulier la courbe de montée, la date et la valeur du pic de crue Estupina (2004).

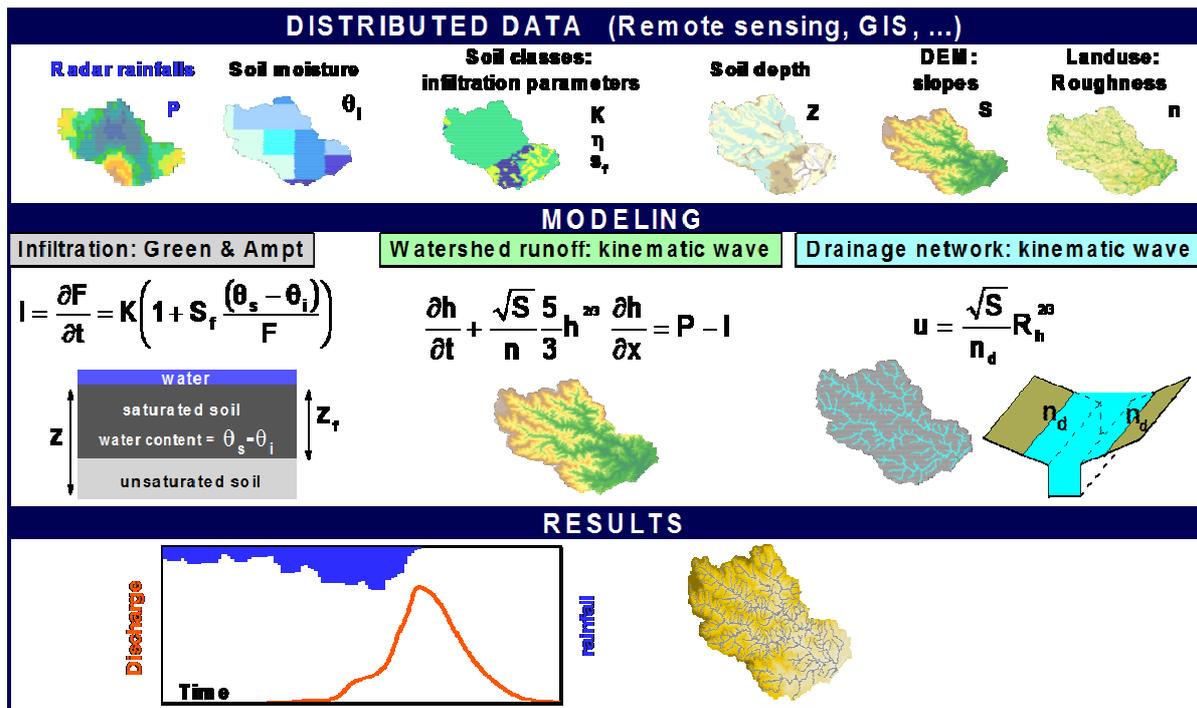


Fig. 4.1 : Schématisation du modèle MARINE

4.1.2 Données nécessaires au modèle MARINE

Afin d'obtenir une représentation fidèle de l'hétérogénéité de la pluie et des comportements variés du bassin versant, le modèle MARINE est distribué. Cette distribution spatiale contribue également à la compréhension des processus de surface. Le modèle MARINE intègre des informations issues de l'imagerie satellitale avec une résolution spatiale adaptée aux échelles de comportements hydrologiques. Le modèle nécessite un certain nombre de données hydrométéorologiques :

- les pluies. Issues des radars météorologiques et traitées par des modèles d'estimation et d'anticipation (Météo France, Rhéa), elles représentent le forçage atmosphérique de MARINE,
- l'humidité initiale du sol (issue du modèle SIM¹ de Météo France).

¹ Le modèle SIM (Safran-Isba-Modcou) a été développé au CNRM pour modéliser simultanément les bilans de surface et les débits des grands bassins. C'est un modèle qui lie des variables atmosphériques, des flux à la surface de la terre et des écoulements superficiels et souterrains.

et géographiques :

- le modèle numérique de terrain (MNT). Il contient la topographie du bassin versant, sa localisation (par ex : satellite SPOT) et permet de calculer les pentes à l'origine de l'écoulement superficiel de l'eau,
- la composition du sol (caractéristiques géologiques et pédologique) et les hauteurs maximales d'infiltration (qui renseignent sur l'infiltrabilité du sol),
- l'occupation du sol, la végétation (provenant des satellites SPOT ou LANDSAT), elles influencent les vitesses de ruissellement.

Les données issues de l'imagerie satellitale sont définies pour toute la surface du bassin versant par un maillage régulier (constitué de pixels carrés) qui définit l'échelle de support des données.

4.1.3 Equations représentatives du comportement des processus physiques

MARINE est un modèle perceptuel qui se concentre sur la modélisation des processus dominants, lesquels s'identifient, dans le cas des crues éclair, au ruissellement superficiel. Les phénomènes d'interception et d'évapotranspiration sont négligés dans la modélisation. En effet, Thirriot (1980) relève que la répartition de l'eau à travers le cycle hydrologique dépend des échelles spatio-temporelles étudiées et si l'échelle temporelle est la durée de la pluie, le phénomène d'évapotranspiration peut être négligé.

La fonction de production de MARINE est décrite par le modèle de Green qui a été choisi car il présente un compromis intéressant en termes de réalisme et de simplicité. La fonction de transfert de MARINE sur le bassin versant est approchée par l'approximation de l'onde cinématique.

4.1.3.1 Modèle d'infiltration de Green et Ampt

- Modèle physique

L'équation de Green et Ampt (1911) découle de l'équation de Darcy (Darcy, 1856) qui relie la vitesse de l'écoulement I (m.s^{-1}) au gradient hydraulique dh/dz par l'intermédiaire d'un paramètre caractéristique du milieu traversé : la conductivité hydraulique K (m.s^{-1}) :

$$I = K \frac{dh}{dz} \quad (4.1)$$

La loi de Darcy est définie pour des conditions d'écoulement laminaire dans un milieu homogène, isotrope et continu, le fluide n'interagissant pas avec le milieu. La conductivité hydraulique K est caractéristique des conditions d'écoulement dans un milieu poreux donné pour un fluide donné.

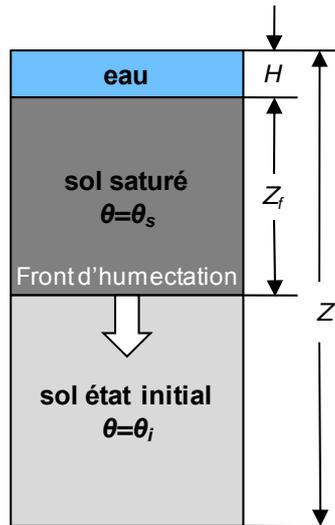


Fig. 4.2 : Schématisation de l'infiltration, d'après Green et Ampt (1911)

La schématisation proposée par Green et Ampt (1991) (Fig. 4.2), intègre un front d'humectation abrupt et horizontal qui sépare le sol non saturé du sol saturé. L'eau ayant une pression inférieure à la pression atmosphérique, est aspirée dans le milieu poreux par la force de succion S_f .

Selon cette schématisation, l'équation (4.1) devient :

$$I = K \frac{h_f - h_0}{Z_f - Z_0} = K \frac{S_f + Z_f}{Z_f} \quad (4.2)$$

H_f est la charge hydraulique au front d'humectation et h_0 la charge hydraulique à la surface ($h_0 = 0$ sauf si il y a de l'eau à la surface). S_f (m) est la force de succion au front d'humectation.

La profondeur au front d'humectation peut être reliée à la hauteur d'eau cumulée infiltrée F (m) par la relation suivante :

$$F = Z_f(\theta_s - \theta_i) \quad (4.3)$$

où θ_s (%) est l'humidité à saturation du sol ou la porosité et θ_i (%) l'humidité initiale du sol.

En combinant les équations (4.2) et (4.3), le taux d'infiltration $I(t)$ devient (Mein et Larson, 1973):

$$I(t) = \frac{dF(t)}{dt} = \begin{cases} P & t \leq t_p \\ K \left(1 + S_f \frac{\theta_s - \theta_i}{F(t)} \right) & t > t_p \end{cases} \quad (4.4)$$

où P est l'intensité de la pluie (m.s^{-1}), t_p est le temps nécessaire au « ponding » (à la formation d'une flaque à la surface du sol).

Le taux d'infiltration I est en fait égal à l'intensité de la pluie P tant que celle-ci est inférieure à la vitesse d'infiltration potentielle $I(t)$. Lorsque l'intensité de la pluie est supérieure à la vitesse d'infiltration, le sol devient saturé et il y a formation au temps t_p d'une lame d'eau non négligeable (Gourley et Vieux, 2006).

L'équation (4.4) ne fait pas explicitement apparaître de variable temporelle qui se trouve de façon implicite dans l'expression de $F(m)$: la hauteur d'eau infiltrée cumulée. En notant $\Delta\theta = \theta_s - \theta_i$, il vient après intégration que l'infiltration cumulée F s'écrit lorsque $t > t_p$:

$$F(t) = Kt + S_f \Delta\theta \ln \left(1 + \frac{F(t)}{S_f \Delta\theta} \right) \quad (4.5)$$

Mein et Larson (1973) ont également déduit une relation permettant de calculer t_p :

$$t = t_p + \frac{1}{K} \left[F - F_p + S_f \Delta\theta \ln \left(\frac{S_f \Delta\theta + F_p}{S_f \Delta\theta + F} \right) \right] \quad (4.6)$$

où F_p est le total d'eau infiltrée à l'instant de formation de la lame d'eau en surface. Cette quantité est obtenue pour $t = t_p$ et $P > K$, par conséquent elle s'exprime de la façon suivante :

$$F_p = \frac{S_f K \Delta\theta}{P - K} \quad (4.7)$$

$$t_p = \frac{F_p}{P} \quad (4.8)$$

Pour déterminer le volume total d'eau infiltrée à partir d'un événement pluvieux de durée t_r et d'intensité P , il faut d'abord déterminer le moment où la lame d'eau en surface se forme à partir des équations (4.6) et (4.7). Si $t_r < t_p$ alors l'infiltration cumulée vaut $F = Pt_r$ et le taux d'infiltration $I = P$. Si $t_r > t_p$, alors il faut utiliser l'équation (4.6) et trouver de manière itérative la valeur F pour laquelle $t = t_r$.

- Approche numérique

Le même raisonnement peut être appliqué à l'échelle du pas de temps de simulation Δt . En utilisant un schéma explicite amont, lorsque $t > t_p$ l'infiltration cumulée au pas de temps $t + \Delta t$ peut être reliée à celle du pas de temps précédent par une relation équivalente à l'équation (4.5) :

$$F_{t+\Delta t} - F_t - S_f \Delta\theta \ln \left(\frac{S_f \Delta\theta + F_{t+\Delta t}}{S_f \Delta\theta + F_t} \right) = K \Delta t \quad (4.9)$$

Pour un événement pluvieux donné, lors du calcul des pertes par infiltration sur une maille intégrant le ruissellement amont, différents régimes d'infiltration vont se succéder au cours du temps. L'équation précédente est résolue par une méthode de Newton quelque soit le régime d'infiltration.

Le détail de l'algorithme de résolution de l'infiltration est présenté dans la thèse de Castaing (2007). Cependant, dans la version actuelle du modèle une profondeur de sol Z est introduite, ce qui permet de limiter la capacité du réservoir sol égale à $Z\Delta\theta$. Et dans ce cas :

$$\forall t \text{ si } F(t) > Z\Delta\theta \text{ alors } I(t) = 0 \quad (4.10)$$

- Les paramètres d'infiltration

La conductivité hydraulique rend compte des forces de gravité, elle est caractéristique des conditions d'écoulement dans un milieu poreux donnée pour un fluide donné. La porosité est une propriété directement mesurable d'un aquifère. C'est une fraction comprise entre 0 et 1 qui indique la quantité d'espace vide entre des particules de sol libres ou dans une roche fracturée. Elle donne directement la valeur de l'humidité à saturation du sol. Elle agit sur la profondeur du front d'humectation et modifie alors la valeur du gradient de charge hydraulique. Plus la valeur de la porosité est importante, plus la capacité d'infiltration du sol est grande. L'humidité initiale du sol représente la fraction du milieu poreux initialement pleine d'eau liquide. C'est une fraction comprise entre 0 et 1 qui doit rester inférieure ou égale à la porosité totale. La succion effective au front d'humectation détermine la valeur du gradient de charge hydraulique de l'équation de Darcy. Cette valeur n'est pas définie parfaitement à un instant donné et s'étale sur une épaisseur de sol. Plus la valeur de la succion est importante, plus la capacité d'infiltration du sol est grande. Les paramètres du modèle de Green et Ampt ne sont pas directement mesurables. De nombreuses tables existent dans la littérature pour les estimer à partir de caractéristiques de sols (Tab. 4.2).

De nombreuses modélisations de la capacité d'infiltration du sol existent (Hillel, 1974 ; Mein, 1973). Le modèle de Green et Ampt réalise un compromis entre la physique du processus d'infiltration, la rapidité de calcul, le sens physique des paramètres et leur facilité d'acquisition.

4.1.3.2 Modèle de ruissellement : approximation de l'onde cinématique

Sur les versants

- Modèle physique

Ce modèle de l'onde cinématique a été utilisé dès 1965 (Wooding, 1965) pour simuler le ruissellement sur des bassins versants. Cette approximation constitue la base de la modélisation du mécanisme de ruissellement sur le bassin versant pour MARINE, comme pour la majorité des modèles pluie-débit distribués à base physique (Jain et al. 2004 ; Liu et al. 2004).

Les équations de Saint Venant à faible profondeur (shallow water) sont obtenues à partir des équations de Navier-Stokes en les intégrant verticalement sur une épaisseur relativement faible avec une hypothèse de non glissement sur le fond. Celles-ci se décomposent en termes inertiels et en termes de frottement. Ces derniers font apparaître les coefficients de Chézy, Manning-Strickler ou Darcy-Weisbach suivant la formulation retenue. Ils dépendent de la nature du fond, de sa forme, des diverses obstructions à l'écoulement et de la végétation.

Le passage aux équations d'ondes diffusantes se fait en considérant que les forces de frottement, de gravité et de pression gouvernent l'écoulement, alors que les termes inertiels d'accélération locale et convective sont négligés. Ceci est le cas de beaucoup de rivières naturelles à faibles vitesses d'écoulement (< 7.5 m/s (Tregarot 2001)) et à pentes douces (supérieures toutefois à 0.01 % (Woolhiser et al., 1996)). L'expression générale de l'équation des ondes diffusantes est ensuite obtenue en insérant l'expression du module de la vitesse dans l'équation de continuité.

Le passage à l'onde cinématique (Beven, 1982, Liu et al. 2004) suppose en sus que les termes de pression des équations sont négligeables (la pente de la ligne d'énergie est égale à la pente du sol localement $S_0 = S_f$), il n'y a donc pas d'atténuation de l'hydrogramme de crue. La hauteur de la lame d'eau devient alors égale à celle du régime uniforme. Ainsi, en considérant que les coefficients de frottement sont uniformes suivant x et y (isotropie) et que le problème peut se ramener à un écoulement monodimensionnel suivant la ligne de plus grande pente, le modèle se réduit à deux équations principales (4.11) et (4.12).

L'équation de continuité en régime non permanent s'exprime localement par :

$$\frac{\partial h}{\partial t} + \frac{\partial uh}{\partial x} = P - I \quad (4.11)$$

où h est la hauteur de lame d'eau (m), t le temps (s), u est la vitesse de ruissellement ($m.s^{-1}$), x est la variable d'espace (m), P représente l'intensité de pluie ($m.s^{-1}$) et I est le taux d'infiltration ($m.s^{-1}$).

L'équation de la dynamique est réduite à celle du régime uniforme :

$$u = a.h^{m-1}$$

où a et m sont des coefficients.

Lorsque ce ruissellement se fait suivant des drains, et donc à tirant d'eau plus important, il devient pleinement turbulent. La vitesse peut être exprimée au moyen de la formule de Manning (Carlier, 1986) en prenant h pour valeur du rayon hydraulique :

$$u = \frac{S_0^{1/2}}{n} h^{2/3} \quad (4.12)$$

où n est le coefficient de Manning ($m^{-1/3}.s$) et S_0 la pente ($m.m^{-1}$).

Ainsi avec l'équation de continuité (4.11), cette loi devient :

$$\frac{\partial h}{\partial t} + \frac{S_0^{1/2}}{n} \frac{5}{3} h^{2/3} \frac{\partial h}{\partial x} = P - I \quad (4.13)$$

- Approche numérique

Neveu et Perrot (2002) se sont intéressés à la formulation eulérienne du ruissellement superficiel. Pour chaque élément du maillage structuré de MARINE, quatre directions potentielles de l'écoulement sont considérées (directions cardinales) et l'unique direction utilisée est celle de la plus grande pente. La lame d'eau est supposée uniformément répartie sur toute la surface du pixel considéré. La résolution des équations de conservation est faite grâce à la méthode des volumes finis. La formulation de l'équation (4.11) peut s'écrire sur une surface de contrôle Ω_k délimitée par Γ_k :

$$\int_{\Omega_k} \frac{\partial h}{\partial t} d\Omega_k + \int_{\Omega_k} \text{div}(uh) d\Omega_k = \int_{\Omega_k} (P - I) d\Omega_k \quad (4.14)$$

Grâce au théorème d'Ostrogradsky, le second terme de l'équation (4.14) se transforme en intégrale de flux sur le contour Γ_k :

$$\int_{\Omega_k} \frac{\partial h_k}{\partial t} d\Omega_k + \int_{\Gamma_k} u_k d_k h_k d\Gamma_k = \int_{\Omega_k} (P_k - I_k) d\Omega_k$$

d_k est la direction de plus grande pente et la variable h_k est constante sur Ω_k . Un schéma d'Euler explicite amont pour la dérivée en temps donne :

$$\frac{h_k^{t+\Delta t} - h_k^t}{\Delta t} + \frac{u_k^t h_k^t}{\Delta x} = P_k^t - I_k^t$$

où t représente le temps et Δt le pas de temps. L'approximation de l'onde cinématique s'écrit :

$$u_k^t = \frac{S_k^{1/2}}{n_k} h_k^{t 2/3}$$

Finalement, la hauteur d'eau sur chaque pixel est obtenue par l'équation suivante :

$$h_k^{t+\Delta t} = h_k^t + \Delta t (P_k^t - I_k^t) - \frac{\Delta t S_k^{1/2}}{\Delta x n_k} h_k^{t 5/3}$$

Ce schéma explicite en temps s'implémente assez aisément et il est moins coûteux en temps de calcul. Cependant, un schéma décentré introduit de la diffusion numérique qui a tendance à atténuer les hydrogrammes de crue. Le pas de temps de calcul doit être faible pour assurer la stabilité de la solution. Une estimation de ce pas de temps peut être obtenue grâce à la condition de stabilité de Courant, Friedrichs, Lewy (CFL).

Dans le réseau de drainage

- Calcul des dimensions

L'intégration du réseau de drainage dans MARINE a été mise en œuvre par Bessière (2005). Lorsque l'aire drainée est supérieure à 1 km², l'écoulement de surface est structuré en réseau de drainage. L'écoulement dans les drains est simulé en utilisant l'approximation de l'onde cinématique en tenant compte de la largeur des drains.

La connaissance même approximative de la géométrie du réseau de drainage (Beer et Borgas 1993 ; Agnese et al., 1996) est nécessaire pour déterminer l'écoulement concentré dans la maille drain. Une hypothèse a donc été faite sur la géométrie du profil en travers du réseau de drainage (Fig. 4.3)

Les caractéristiques du réseau de drainage sont calculées en utilisant des lois géomorphologiques. Des concepts théoriques généraux sont déduits du concept de dépense d'énergie optimale et ils sont plus ou moins vérifiés par de nombreuses observations de terrain.

La valeur de la largeur du drain est essentielle car elle conditionne directement la valeur des vitesses : plus le drain est étroit, plus les vitesses sont importantes. En outre, plus le ruissellement est intense, plus il aura tendance à éroder le terrain et à creuser les drains et à les élargir. Il s'agit donc d'obtenir une estimation physique de ces largeurs. (Leopold et Maddock 1953) ont été les premiers à introduire une loi de type : $B = KQ^\beta$ reliant la largeur du drain B au débit Q et cette loi reste d'actualité. Il faut choisir judicieusement le débit Q pour être en adéquation avec les phénomènes physiques étudiés. Dans son article, Ibbitt (1997) propose différentes valeurs pour les coefficients K et β . Léopold et Langbein (1962) obtiennent respectivement $\beta = 0.53$ $\beta = 0.45$. Différentes formules reliant la largeur au débit sont citées pour des bassins versants en Nouvelle Zélande. Par exemple pour le bassin Hutt River (Nouvelle Zélande), la loi devient : $B = 0.25 Q^{0.49}$ pour un écoulement moyen et $B = 0.27 Q^{0.51}$ pour un dépassement à 50 %. Parallèlement, pour le bassin versant Ashley River, (Ibbitt 1997) l'auteur trouve : $B = 0.314 Q^{0.43}$.

Le débit est ensuite relié à l'aire drainée. Les formules varient en fonction de la pluviométrie sur le bassin versant. Par exemple, pour un bassin sec $Q = 1.2 \cdot 10^{-5} \times S$, alors que pour un bassin humide $Q = 1.38 \cdot 10^{-5} \times (S + 2.2 \cdot 10^6)$ où S est la surface du bassin versant (m²).

Dans le modèle MARINE, la largeur est calculée par la formule suivante :

$$B = 0.3(q_s a_d)^{0.44} \quad (4.15)$$

où q_s correspond au débit spécifique (l/s/m²) basé sur une période de retour 2 ans et a_d est l'aire drainée (m²) du drain considéré.

Le calcul des aires drainées permet d'obtenir une représentation du réseau de drainage (Meisels et al., 1995 ; Schoghorfer et Rothman, 2002), il se fait grâce au MNT (Modèle Numérique de Terrain).

La valeur de la hauteur H_D est estimée à partir de la largeur des drains selon une relation approximative : $B = 8 H_D$.

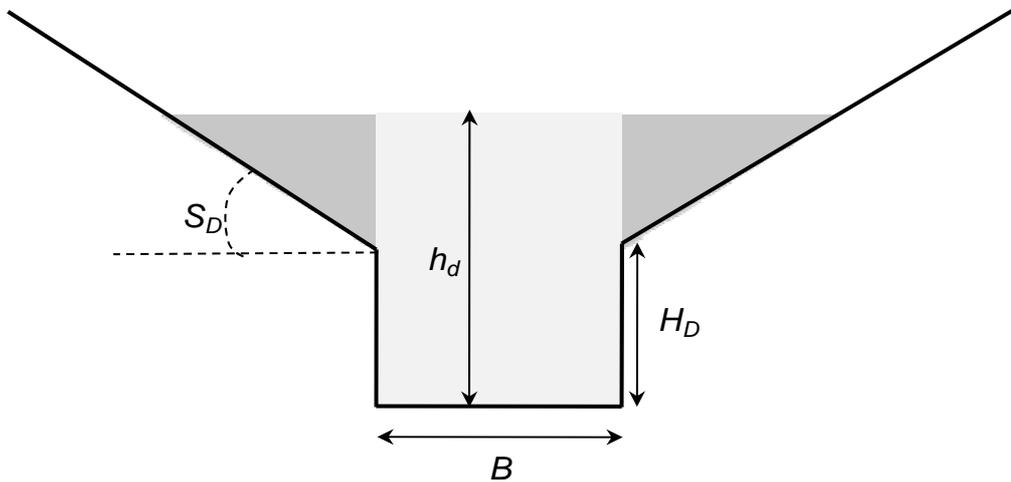


Fig. 4.3 : Géométrie du profil en travers du drain

- Expression de la vitesse dans les drains

Sur l'ensemble du bassin versant, le volume dans chaque maille peut s'écrire : $V = h \times dx^2$, où h est la hauteur d'eau dans chaque maille et dx la taille de la maille. Si la maille considérée fait partie du réseau de drainage, alors ce volume V est présent dans le drain correspondant. Il est alors possible d'en déduire la hauteur d'eau dans le drain h_d (Fig. 4.3). La vitesse dans le drain est ensuite calculée avec la loi de Manning :

$$u_d = \frac{S_0^{1/2}}{n} R_h^{2/3}$$

Deux cas de figure se présentent :

- Si $V \leq BH_D dx$, alors $h_d \leq H_D$

La hauteur d'eau h_d et la vitesse u_d dans la maille s'écrivent :

$$h_d = \frac{V}{Bdx}$$

$$u_d = \frac{S_0^{0.5}}{n_{D1}} \left(\frac{h_d B}{B + 2h_d} \right)^{2/3} \quad (4.16)$$

où n_{D1} est le coefficient de Manning du lit mineur.

- Si $V > BH_D dx$, alors $h_d > H_D$
La hauteur d'eau dans la maille devient :

$$h_d = H_D - \frac{BS_D}{2} + \sqrt{\left(H_D - \frac{B_D S_D}{2}\right)^2 - \left(H_D^2 - \frac{VS_D}{dx}\right)}$$

Les variables hydrauliques du lit mineur sont le rayon hydraulique $R_{H1} = \frac{Bh_d}{B+2h_d}$ et la surface mouillée $A_{D1} = Bh_d$. De même, les variables

hydrauliques du lit majeur sont le rayon hydraulique $R_{H2} = \frac{h_n - H_D}{2\sqrt{1+S_D^2}}$ et la

surface mouillée $A_{D2} = \frac{(h_d - H_D)^2}{S_D}$. La vitesse s'exprime dans ce cas :

$$u_d = \frac{S_0^{0.5} \left(\frac{R_{H1}^{2/3} A_{D1}}{n_{D1}} + \frac{R_{H2}^{2/3} A_{D2}}{n_{D2}} \right)}{A_{D1} + A_{D2}} \quad (4.17)$$

Où n_{D2} est le coefficient de Manning du lit majeur.

- Implémentation numérique

L'équation de conservation de la masse est résolue grâce à la méthode des volumes finis pour chaque maille du réseau de drainage (Fig. 4.4) entre deux instants t et $t+\Delta t$ en utilisant l'approximation de l'onde cinématique.

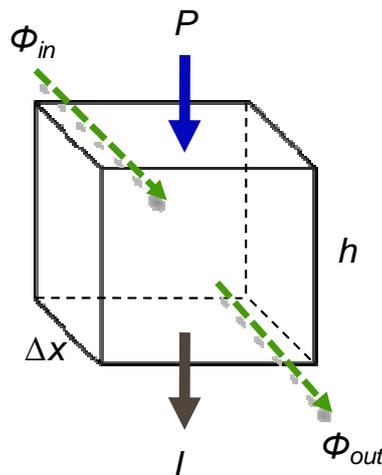


Fig. 4.4 : Flux à travers les mailles du réseau de drainage

Le bilan volumique dans la maille à chaque instant t s'écrit :

$$V_k^{t+\Delta t} = V_k^t + \Delta x^2 \Delta t (P_k^t - I_k^t) + \phi_{in}^t - \phi_{out}^t$$

où V_k^t et V_k^{t+1} les volumes d'eau de la maille k aux l'instant t et $t+\Delta t$ respectivement. P_k^t est l'intensité de la pluie et I_k^t la vitesse d'infiltration à l'instant t de la maille k . Les flux ϕ_{in}^t et ϕ_{out}^t sont les volumes d'eau ruisselés entrant et sortant dans la maille à l'instant t . La direction de ces flux (Nord, Sud, Est, Ouest) est déterminée par les pentes du bassin versant (calculés avec le MNT).

Le volume ruisselé pour la maille k est :

$$\phi = V \frac{u \Delta t}{\Delta x} = S_{mouillée} u \Delta t$$

où $S_{mouillée}$ est la surface mouillée.

Selon les deux cas de figure précédents ($h_d \leq H_D$ et $h_d > H_D$), les flux s'expriment en utilisant les mêmes notations :

$$h_d \leq H_D \text{ alors } \phi = h_d B u_d \Delta t = h_d B \frac{S_0^{0.5}}{n_{D1}} \left(\frac{h_d B}{B + 2h_d} \right)^{2/3} \Delta t$$

$$h_d > H_D \text{ alors } \phi = (A_{D1} + A_{D2}) u_d \Delta t = S_0^{0.5} \left(\frac{R_{H1}^{2/3} A_{D1}}{n_{D1}} + \frac{R_{H2}^{2/3} A_{D2}}{n_{D2}} \right) \Delta t$$

La hauteur d'eau ruisselée dans la maille h est égale à : $h = \frac{\phi}{\Delta x^2}$.

Validité de l'approximation de l'onde cinématique

La validité de l'approximation de l'onde cinématique a été étudiée à maintes reprises pour représenter des écoulements superficiels, notamment par Woolhiser et Liggett (1967) et Moore et Foster (1990) (Fig. 4.5).

Singh (2002, 1996) préconisent les conditions suivantes pour que l'approximation de l'onde cinématique soit applicable :

$$k_k = \frac{S_0 L_0}{F_0^2 h_0} > 20 \quad (4.18)$$

$$F_0^2 k_k = \frac{S_0 L_0}{h_0} > 5 \quad (4.19)$$

où k_k est le nombre d'onde cinématique, S_0 la pente de la surface du sol (m.m^{-1}), L_0 la longueur de ruissellement (m), h_0 le tirant d'eau (m) et F_0 le nombre de Froude. Le nombre de Froude s'exprime de la façon suivante :

$$F_0 = \frac{u_0}{\sqrt{gh_0}}$$

u_0 est la vitesse de l'écoulement (m.s^{-1}) et g l'accélération (m.s^{-2}) due à la pesanteur. L'approximation de l'onde cinématique est encore acceptable quand le nombre d'onde cinématique k_k est proche de 20 à condition que le nombre de Froude F_0 soit supérieur à 0.4. Pour des faibles valeurs de F_0 , l'approximation de l'onde cinématique ne peut être utilisée que si la condition (4.19) est satisfaite (Fig 4.6).

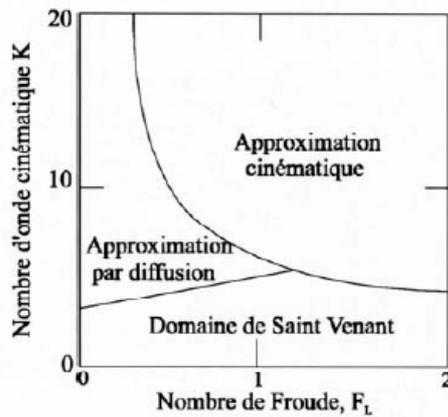


Fig. 4.5 : Domaine de validité de l'application de l'onde cinématique pour la modélisation du ruissellement de surface, d'après Moore et Foster (1990)

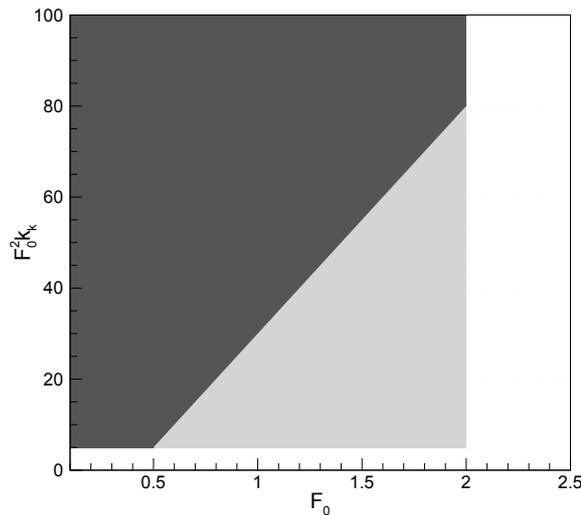


Fig. 4.6 : Critère de validité de l'approximation de l'onde cinématique (zone ombrée) (l'intersection des régions provident de (Lighthill and Whitham 1955, Woolhiser and Liggett 1967, Morris and Woolhiser 1980)); et (gris clair) de (Daluz Vieira 1983), d'après Singh (1996).

4.1.4 Prédétermination des paramètres du modèle

Le modèle est essentiellement contrôlé par les paramètres listés dans le tableau 4.1.

Tableau 4.1. Paramètres d'entrée du modèle MARINE

	Paramètre	Notation	Distribution spatiale	Sources des données
Versants	Humidité initiale du sol (%)	θ_i	Oui	SIM
	Hauteur de sol (m)	Z	Oui	BRGM ² - INRA ³
	Paramètres d'infiltration : - conductivité hydraulique (mm/h) - porosité (%) - force de succion (mm)	K θ_s S_f	Oui	BRGM - INRA
	Coefficient de Manning ($m^{-1/3}.s$)	n	Oui	SPOT ou LANDSAT
	Pente ($m.m^{-1}$)	S_0	Oui	MNT
Réseau de drainage	Largeur des drains (m)	B	Oui	Considérations géomorphologiques
	Pente des versants ($m.m^{-1}$)	S_D	Oui	MNT
	Manning du lit mineur ($m^{-1/3}.s$)	n_{D1}	Non	
	Manning du lit majeur ($m^{-1/3}.s$)	n_{D2}	Non	

Les coefficients de Green et Ampt sont reliés à la texture du sol (pourcentages de sable, d'argile et de limon) (Fig. 4.7) et les coefficients de rugosité de Manning à l'occupation du sol. Ces paramètres sont distribués selon des valeurs a priori déduites de tables provenant de la littérature (Maidment, 1993 ; Chow, 1959) (Tab. 4.2 et Tab. 4.3). La distribution spatiale ainsi obtenue permet de générer des cartes de référence.

² BRGM : Bureau de Recherche Géologiques et Minières.

³ INRA : Institut National de la Recherche Agronomique.

Tableau 4.2. Paramètres de Green et Ampt, d'après Maidment (1993)

Texture du sol	Porosité ϕ	Force de suction S_f (mm)	Conductivité hydraulique à saturation K_s (mm/h) ⁴
Sable	0.437 (0.374-0.500)	49.5 (9.7-253.6)	235.6
Sable limoneux	0.437 (0.363-0.506)	61.3 (13.5-279.4)	59.8
Limon sableux	0.453 (0.351-0.555)	110.1 (26.7-454.7)	21.8
Limon	0.463 (0.375-0.551)	88.9 (13.3-593.8)	13.2
Limon silteux	0.501 (0.420-0.582)	166.8 (29.2-953.9)	6.8
Limon argilo- sableux	0.398 (0.332-0.464)	218.5 (44.2-1080)	3.0
Limon argileux	0.464 (0.409-0.519)	208.8 (47.9-911.0)	2.0
Limon argilo-silteux	0.474 (0.418-0.524)	273.0 (56.7-1315.0)	2.0
Argile sableuse	0.430 (0.370-0.490)	239.0 (40.8-1402)	1.2
Argile silteuse	0.479 (0.425-0.533)	292.2 (61.3-1394)	1.0
Argile	0.475 (0.427-0.523)	316.3 (63.9-1565)	0.6

⁴ La conductivité hydraulique K intervenant dans l'équation de Green et Ampt du modèle MARINE est déduite de la conductivité hydraulique à saturation K_s par la relation suivante : $K = K_s / 2$.

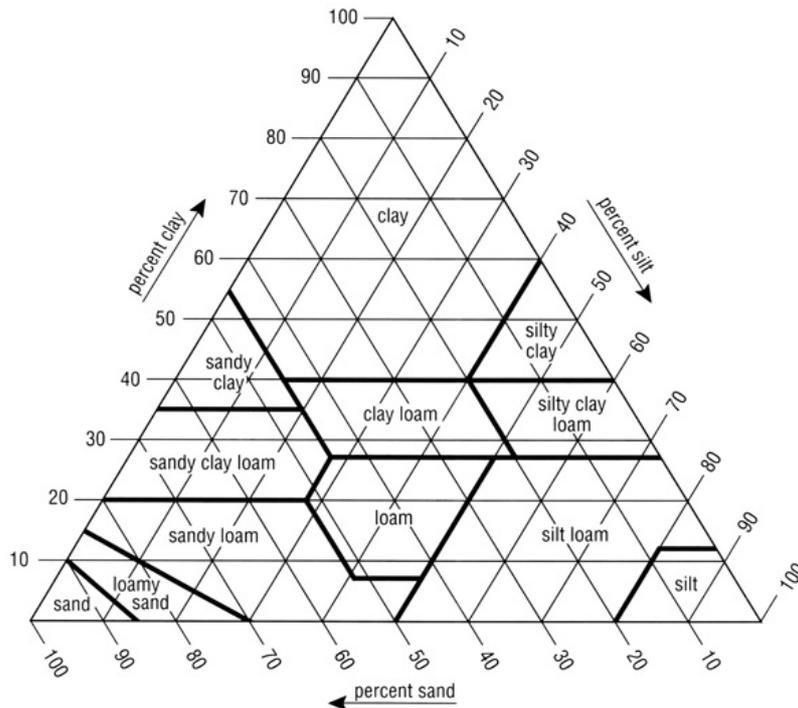


Fig. 4.7 : Pourcentage d'argile, de limon et de sable dans les textures classiques de sol.

Tableau 4.3. Valeurs du coefficient de Manning en fonction de l'occupation du sol, d'après Chow (1959)

Occupation du sol	Manning n ($m^{-1/3} \cdot s$)
Revêtement de béton	0.012
Sol recouvert de graviers et bords en durs	0.20 - 0.033
Prairie	0.035
Champ cultivé	0.040
Broussailles clairsemées	0.050
Broussailles denses	0.070
Bois dense	0.10

L'utilisation du MNT (avec des résolutions de 25 m à 200 m) pour la détermination des pentes des versants du réseau de drainage peut sembler abusive. Ce constat est tout à fait admis mais l'objectif est tout simplement d'obtenir un ordre de grandeur et d'en avoir une première estimation.

4.1.5 Synthèse

Le modèle MARINE utilise une description des processus physiques suffisamment simplifiée pour permettre l'utilisation des données en temps réel. Dans cette optique, le modèle a été implanté au SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) dans le cadre du projet BVNE (Bassins Versants Numériques Expérimentaux). Au moment de l'étude, les écoulements de subsurface n'étaient pas encore intégrés au modèle ; ils l'ont été par la suite (chapitre 7).

Les valeurs des tables issues de la littérature (pour les coefficients de Green et Ampt par exemple) doivent être considérées comme indicatives. En effet, les valeurs mesurées à l'échelle des travaux de Green et Ampt ne peuvent pas être directement utilisées à l'échelle du pixel (en général 200×200 m²). L'estimation a priori des paramètres doit donc être améliorée par la résolution d'un problème inverse. Cependant le ratio entre la dimension de l'espace de contrôle (ensemble des paramètres du modèle) et les observations disponibles (chronique de débits sur un ou plusieurs événements) est très défavorable. Cela conduit à un problème inverse mal posé. La méthodologie employée pour le calage des modèles distribués consiste alors à utiliser des facteurs multiplicatifs (ou coefficients correctifs) pour ajuster de manière relative une distribution spatiale fixée a priori des paramètres du modèle (Refsgaard, 1997 ; Senarath et al., 2000 ; Madsen, 2003). Il en résulte une paramétrisation d'ordre réduit souvent de faible dimension.

Le chapitre 5 de cette thèse est consacré à la calibration du modèle dans un cadre d'application réelle. Les données du bassin versant du Gardon d'Anduze (4.2) sont exploitées dans ce but.

4.2 Bassin d'étude : le Gardon d'Anduze

4.2.1 Introduction

Le bassin versant du Gardon d'Anduze est situé dans la partie Nord Est du Languedoc méditerranéen, principalement dans le département du Gard mais aussi dans le Sud de la Lozère (Fig. 4.8). Son exutoire se trouve au niveau de la ville d'Anduze.

L'ensemble de la population du département du Gard, soit environ 130 000 habitants vit en zone inondable. Le risque d'inondation constitue un enjeu majeur dans cette région. La commune d'Anduze se trouve à la sortie du massif des Cévennes et compte environ 3000 habitants. Cette commune est particulièrement affectée par les risques d'inondation du Gardon. Les constructions les plus récentes, installées en périphérie du vieux village, ainsi que les nombreux campings en bordure du Gardon sont particulièrement vulnérables. Quelques repères historiques sont disponibles dans les archives. En octobre 1795, l'eau est arrivée jusqu'au premier étage des habitations du vieux village. En octobre 1868, 1.5 m d'eau a été relevé dans les rues basses de la ville d'Anduze. Le camping du « Pradal » a été complètement inondé le 14 septembre 1976. Plus récemment, les crues de Septembre 1958 et Septembre 2002 ont été dévastatrices, elles ont causé de nombreux dégâts matériels ainsi que la mort de nombreuses personnes (24 pour celle de 2002). L'historique de ces crues a marqué les mémoires et façonné la vie des Cévenols.

Ces crues dévastatrices ont reçu l'appellation de "gardonnades".

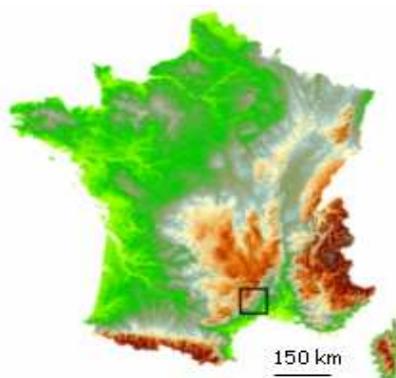


Fig. 4.8 : Situation géographique des Cévennes (source IGN)

4.2.2 Description

Le bassin versant du Gardon d'Anduze couvre une superficie de 545 km².

Il traverse les Cévennes et l'exutoire du bassin se situe en aval de la commune d'Anduze. Ce bassin s'intègre dans l'ensemble du bassin des Gardons qui prend sa source à la crête des Cévennes pour se jeter dans le Rhône en drainant une superficie de 2025 km².

Le Gardon d'Anduze fait partie du projet BVNE (Bassins Versants Numériques Expérimentaux) mené par le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) (Tanguy et al., 2005).

Ce bassin a fait l'objet de nombreuses études hydrologiques : Différence Première de la Fonction de Transfert (DPFT, Guillot et Duband, 1980), modélisation hydrologique spatialisée avec des modèles conceptuels (modèle ModSpa, Moussa, 2007), modélisation par TOPMODEL et comparaison de plusieurs modèles (Marchandise, 2007)

4.2.3 Caractéristiques

4.2.3.1 Caractéristiques physiques

Les caractéristiques physiques du bassin (la géomorphologie, l'occupation des sols, la pédologie ...) influencent directement les mécanismes de génération et de transfert de ruissellement le long du bassin jusqu'à l'exutoire.

- Le réseau de drainage

Le réseau de drainage est dense, en arêtes de poisson. Le bassin versant est divisé en deux sous-bassins : le Gardon de Mialet et le Gardon de Saint-Jean (Fig. 4.9). Leur confluence, deux kilomètres en amont de l'exutoire constitue l'ensemble du bassin.

Le Gardon est encaissé avec une pente longitudinale forte, supérieure à 1,5 %. Depuis la confluence du Gardon de Saint-Jean et du Gardon de Mialet, la vallée du Gardon s'évase un peu avant de rejoindre l'agglomération d'Anduze. Le Gardon s'écoule alors dans une plaine alluviale, relativement large par rapport aux secteurs plus en amont, encadrés par des massifs importants.

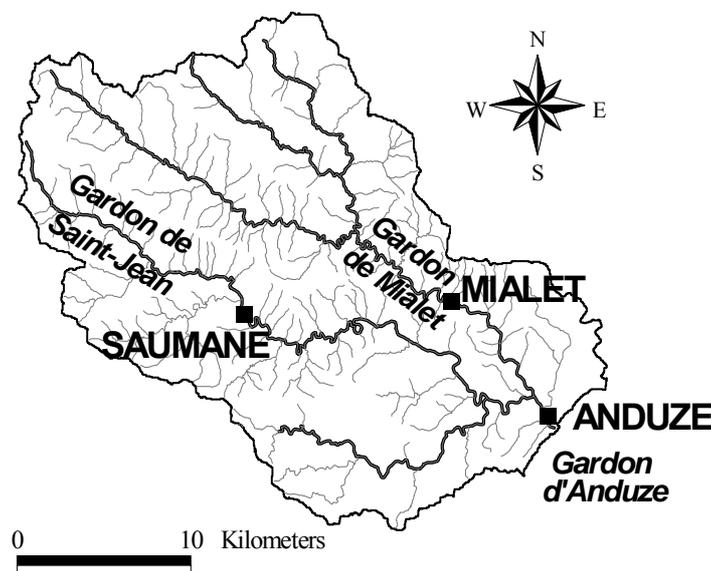


Fig. 4.9 : Réseau de drainage du Gardon d'Anduze

- La topographie

La partie amont du bassin versant correspond au haut bassin Cévenol, caractérisé par une région montagneuse (Fig. 4.10) avec des crêtes étroites (les « serres cévenols »), des vallées longues et très encaissées (les « valats »), des versants très pentus avec de nombreuses rivières au comportement de type torrentiel (Fig. 4.11).

Près de l'exutoire, au niveau du piémont cévenol, les pentes s'atténuent et valent 2% en moyenne ; le lit mineur s'élargit à l'intérieur d'une plaine alluviale (Fig. 4.12).

Dans les Cévennes, près du Mont Aigoual, les altitudes atteignent les 1200 m tandis que l'exutoire du bassin versant se situe aux environs de 130 m (Fig. 4.13).

La pente moyenne du bassin vaut approximativement 10% (Fig. 4.14).

Le sous-bassin versant du Gardon de Saint-Jean à Saumane occupe une superficie de 100 km² et des pentes moyennes de l'ordre de 35 %. Celui de Mialet s'étend sur 230 km² avec des pentes moyennes de 20 % (Ayrat, 2005)

- Les sols

Les épaisseurs de sol sont relativement petites, de 0.1 m à 1 m (Fig. 4.15). L'épaisseur moyenne est estimée à 29 cm.

Les sols sont en majorité des rankers, lithosols et des sols bruns (INRA). Un ranker est un type de sol peu épais sur sous-sol siliceux ; c'est le résultat de l'action de la végétation pionnière sur la roche mère. Les lithosols sont des sols peu épais, inférieurs à 10 cm, posés sur une roche mère dure, continue et peu fragmentée.

Les sols les plus épais sont situés au fond des vallées et dans les zones colluvionnaires. Les sous-bassins amont de Saumane et de Mialet ont des sols plus fins, 17.7 cm en moyenne pour Mialet et 20.7 cm pour Saumane.

La majeure partie du bassin est développée sur des formations cristallines et schisteuses (64%) généralement métamorphosées des Cévennes. Les sous-bassins de Mialet (230 km²) et Saumane (100 km²) sont constitués à 95 % de ces roches (Ayrat, 2005). Le substrat est composé de schiste et de roches cristalline recouvert à 83% par du limon argileux (silty clay loam) et à 17% par du limon sableux (silty loam) à la surface (Moussa et al., 2007).

- La végétation et l'occupation du sol

La végétation est dense et caractéristique des forêts Méditerranéennes. Elle est principalement composée de châtaigniers, de chênes verts, de conifères, de friches et de garrigues (Fig. 4.16).

Les châtaigniers se situent sur la partie amont du bassin et sur les versants exposés au sud et au soleil (les adrets). Les chênes verts et les garrigues se situent à l'aval, en dessous de 500 m et sur les pentes exposées au nord (ubacs).

Les surfaces boisées couvrent la partie amont du bassin versant. Sur la partie aval, l'occupation des sols est dominée par l'activité agricole (viticulture, cultures céréalières, cultures fruitières).

Hormis à Anduze qui représente un point de concentration de la population, l'urbanisation est assez diffuse. Cependant, de nombreuses communes sont situées dans des zones inondables (fiche BVNE Hansard et Louis, 2005).



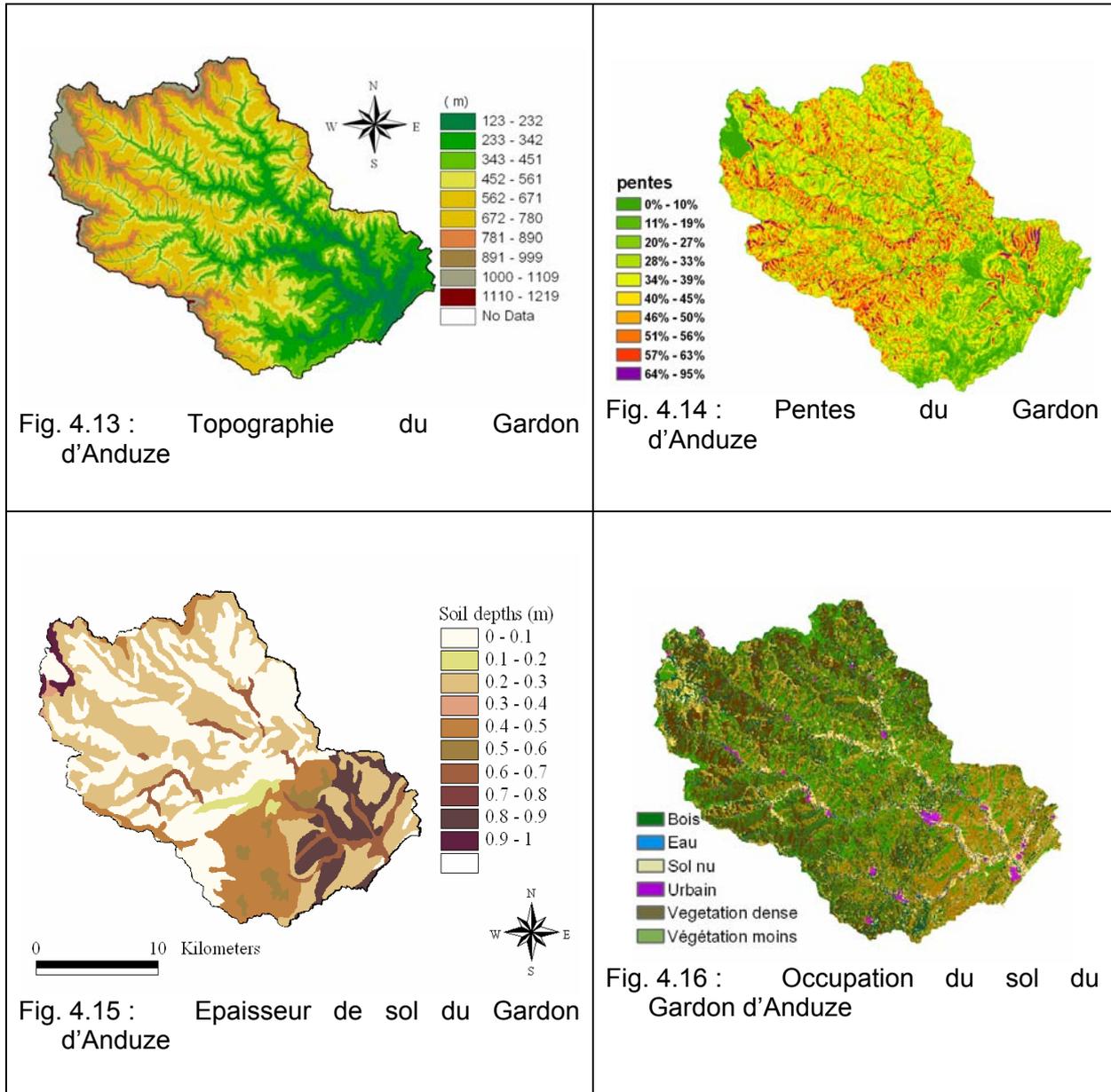
Fig. 4.10 : L'amont du Gardon d'Anduze est caractérisé par des hautes montagnes.



Fig. 4.11 : Les vallées encaissées du Gardon d'Anduze.



Fig. 4.12 : A l'aval le lit mineur s'élargie à l'intérieur d'une plaine alluviale



4.2.3.2 Caractéristiques climatiques

Le climat méditerranéen est caractérisé par des étés chauds et secs et des hivers frais et humides. C'est en automne que les intensités de pluie sont les plus importantes, il tombe alors environ les deux tiers des précipitations annuelles. Les mois d'été sont très secs, cependant certains orages isolés d'été peuvent engendrer des cumuls de pluie importants.

Au cours de l'automne, la rencontre d'une masse d'air chaud et humide (due à la présence de la mer Méditerranée) et d'une autre froide et sèche (d'origine arctique) associée aux effets locaux dus au relief sont à l'origine de ces fortes pluies. Le relief cévenol soulève l'air chaud et humide des basses couches qui se condense pour former des nuages.

Dans cette région, les précipitations qui génèrent les crues sont généralement de forte intensité, de courte durée (quelques heures) et d'une grande variabilité spatio-temporelle.

En France, les plus fortes intensités de pluie ont été enregistrées dans la région du Gard. Par exemple, il est tombé 608 mm en 24 h en Octobre 1963 sur le Mont Aigoual ou plus récemment, en Septembre 2002, 687 mm ont été relevés à Anduze. Pour le bassin versant du Gardon d'Anduze, Lebel (1984) a estimé une lame d'eau moyenne de 59 mm pendant 1h et 350 mm en 24h pour la période de retour centennale.

4.2.3.3 Caractéristiques hydrologiques

- Type de crue

La conjonction de ces pluies très intenses, des sols peu profonds et de fortes pentes provoque des crues très dévastatrices en automne essentiellement, même si les orages d'été présentent également un risque non négligeable de crue.

Les débits résultants de ces épisodes pluvieux intenses sont important et les temps de montée très rapides (Tab. 4.4), ce qui laisse peu de temps d'anticipation.

Tableau 4.4. Temps de réaction des différents bassins (source : fiche BVNE : Gardons, Hansard et Louis, 2005)

Station	Temps de réaction (h)
Saumane	1
Mialet	2
St Jean du Gard	2
Anduze	3

Au dessus d'un certain seuil de pluie cumulée, le ruissellement est caractérisé par une augmentation brutale du coefficient de ruissellement. Ce seuil est défini comme le total d'excès de précipitations cumulées sur une période donnée sur le bassin nécessaire à générer une crue (mineure) (Carpenter et al., 1999).

Pour les bassins Méditerranéens, la valeur de ce seuil est environ 100 mm à 200 mm de pluie cumulée. (Claudet et Bouvier, 2005).

Zhang et al. (2001) ont montré que selon la configuration du réseau de drainage, une production de pluie supérieure à 200 mm sur une superficie d'une centaine de kilomètre carrés pouvait générer des crues rapides. Le coefficient de ruissellement dépend également de l'humidité initiale du sol (due aux précipitations antérieures), des propriétés du sol (perméabilité du sol etc.), de la végétation et de la topographie. De ce fait, une différence de 10 mm sur le total des pluies cumulées peut causer une importante augmentation de débit et transformer un phénomène météorologique de moyenne intensité en un phénomène hydrologique exceptionnel.

Les débordements du lit mineur sont généralement limités à l'aval du bassin versant car une fois captées par le réseau de drainage, les eaux s'écoulent dans le lit mineur à l'amont.

- Principales crues historiques

Les principales crues sont celles de 1697, 1741, 1768, 1846, 1861, 1890, 1891, 1907, 1958 et 2002 (Tab. 4.5 et Fig 4.17).

Tableau 4.5. Données hydrologiques des principales crues

Date	3/11 1768	20/09 1846	11/11 1861	21/09 1890	21/10 1891	16/10 1907	30/09 1958	09/09 2002
Côtes max		7 m	7- 8.2 m	7.1 m	6.8 m	6.5 m	7.6 m	7.5 m

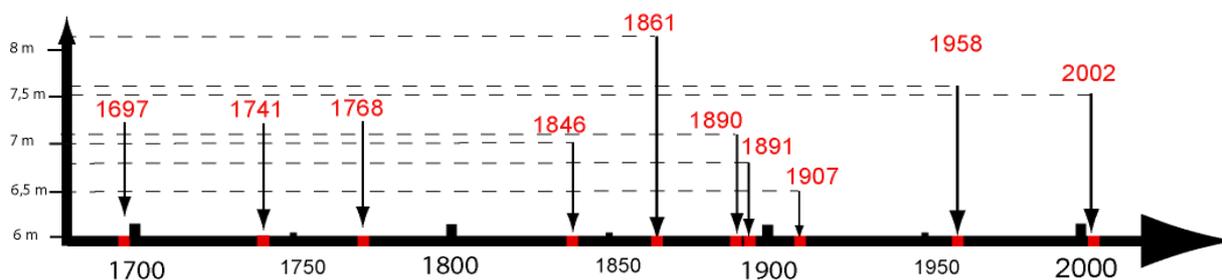


Fig. 4.17 : Chronologie des principales crues sur le Gardon à Anduze

- Débits caractéristiques

Les valeurs de référence des débits de pointe instantanés sont présentées dans le tableau 4.6.

Tableau 4.6. Débits caractéristiques en m^3/s à différentes stations. Qi_2 , Qi_{10} et Qi_{100} sont les débits instantanés de période retour 2, 10 et 100 ans. Les intervalles de confiance sont notés entre crochets.

Station	Aire (km ²)	Qi_2	Qi_{10}	Qi_{30}	Qi_{100}	Source
Généragues (Gardon de Mialet)	240	340 m^3/s [300-370]	600 [530-690]	-	820 [730-980]	Banque HYDRO
Saumane (Gardon de St Jean)	104	150 [130-170]	270 [240-320]	-	370 [320-460]	Banque HYDRO
Anduze (Gardon d'Anduze)	546	520	950	-	1750	CEMAGREF
		-	1340	1790	-	SAGE Gardons

Il est possible de remarquer qu'en fonction des différentes sources, les valeurs sont très différentes. Ceci démontre encore un fois les incertitudes existantes sur les estimations des débits.

4.2.4 Base de données disponibles

4.2.4.1 Origine des données événementielles

Le Service de Prévision des Crues (SPC) de la Direction Départementale de l'Équipement du Gard, basée à Nîmes, est responsable de la surveillance et de l'annonce des crues sur le bassin versant du Gardon d'Anduze. Ce service dispose d'un réseau de stations de mesure de la pluie au sol, d'un radar météorologique implanté à Nîmes et d'un réseau de stations limnimétriques (Fig. 4.18). Les courbes de tarage élaborées par ce service permettent de reconstituer les débits à partir des hauteurs d'eau mesurées.

La base de données événementielle couvre la période 1994 - 2006.



Fig. 4.18 : Echelles limnimétriques à Anduze (à gauche) et Mialet (à droite).

4.2.4.2 Données météorologiques

Les lames d'eau précipitées sont mesurées à partir du Radar de Nîmes exploité par Météo-France. Le logiciel CALAMAR (CALcul de LAMes d'eau à l'Aide du Radar), développé par la société Rhéa, est utilisé pour traiter les images brutes fournies par Météo-France. Le logiciel transforme les données brutes en un format adéquat et contrôle la qualité de ces données. C'est un outil standardisé et distribué pour le calcul et la production de données de lames d'eau précipitées sur des bassins versants à partir de l'exploitation croisée des données radar de pluie et des données pluviographiques.

Les données radar sont calibrées avec les données pluviométriques, et le facteur de calibration est réactualisé en permanence.

Ce facteur de calibration est utilisé pour corriger les erreurs liées à une mauvaise estimation du profil vertical de réflexivité. Il est défini comme le rapport entre le total des pluies cumulées aux postes pluviométriques et les cumuls aux pixels radar correspondants.

Les images radar permettent de mesurer l'intensité pluvieuse et d'analyser la dynamique des cellules orageuses en temps réel. Une information spatialisée des précipitations contribue à renforcer la compréhension et la connaissance des mécanismes de crue. Ces données constituent une information essentielle pour la prévision des crues en temps réel.

Sur le bassin d'étude, les données de pluies issues du radar météorologique sont disponibles au pas d'espace 1 km² et au pas de temps 6 min de 1994 à 2002 et 5 min depuis 2002.

Ces données sont répertoriées dans le tableau (4.7).

Tableau 4.7. Données radar météorologiques disponibles

Numéro événement	Date de début	Date de fin	Durée (h)	Remarques
1	21/09/1999 – 21h	24/09/1994 – 06h	57	
2	18/09/1995 – 04h	19/09/1995 – 08h	27	
3	03/10/1995 – 18h	06/10/1995 – 16h	70	
4	13/10/1995 – 07h	15/10/1995 – 06h	47	
5	26/11/1995 – 00h	27/11/1995 – 05h	29	Données aberrantes
6	10/01/1996 – 06h	10/01/1996 – 09h	3	Données manquantes
7	14/10/1996 – 10h	15/10/1996 – 14 h	28	Données manquantes en début d'épisode
8	10/11/1996 – 08h	14/11/1996 – 05h	93	
9	06/10/1997 – 09h	07/10/1997 – 00h	15	
10	04/09/1998 – 19h	05/09/1998 – 00h	5	
11	17/05/1999 – 03h	18/05/1999 – 14h	35	Données manquantes le 18 de 10h18 à 11h18

Numéro événement	Date de début	Date de fin	Durée (h)	Remarques
12	19/09/2000 – 06h	20/09/2000 – 01h	20	
13	28/09/2000 – 05h	29/09/2000 – 21h	41	
14	06/10/2001 – 17h	07/10/2001 – 06h	14	
15	14/03/2002 – 20h	15/03/2002 – 20h	25	Données manquantes le 15 de 12h à 13h
16	03/06/2002 – 23h	06/06/2002 – 19h	69	
17	02/09/2002 – 06h	02/09/2002 – 15h	10	
18	03/09/2002 – 17h	04/09/2002 – 00h	8	
19	08/09/2002 – 01h	09/09/2002 – 17h	41	
20	27/10/2004 – 10h	29/10/2004 – 07h30	45h30	
21	05/09/2005 – 16h	09/09/2005 – 08h	88h	
22	03/11/2005 – 11h	05/11/2005 – 11h	48h	
23	12/09/2006 – 22h	16/09/2006 – 22h	88h	
24	23/09/2006 – 00h	26/09/2006 – 00h	48h	
25	18/10/2006 – 10h	20/10/2006 – 11h20	49h20	
26	14/11/2006 – 23h	19/11/2006 – 23h	120	

4.2.4.3 Données hydrologiques

Les données d'observation disponibles sont des hydrogrammes de crue à Anduze (exutoire) et aux deux stations hydrométriques intermédiaires (Saumane et Mialet) de 1994 à 2007. Ces données sont répertoriées dans le tableau 4.8. De 2003 à 2006, les débits sont disponibles en continu pour les trois stations, cependant, seuls les événements correspondant aux épisodes pluvieux disponibles ont été relevés.

Il est possible de remarquer que de manière générale, le pic de crue à Mialet arrive en moyenne 1h avant le pic de crue à Anduze et celui à Saumane 2h à 3h avant.

Tableau 4.8. Hydrogrammes disponibles aux trois stations : Anduze, Mialet et Saumane

Événement	Anduze		Mialet		Saumane	
	Max (m ³ /s)	Heure	Max (m ³ /s)	Heure	Max (m ³ /s)	Heure
N°1 23/09/1994	775	23/09/1994 23h	263	23/09/1994 22h	155	23/09/1994 13h
N°2 18/09/1995	946	18/09/1995 16h	276	18/09/1995 15h	130	18/09/1995 13h
N°3 03/10/1995	1608	05/10/1995 07h	356	05/10/1995 06h	680	05/10/1995 05h
N°4 13/10/1995	1411	14/10/1995 07h	249	14/10/1995 06h	données manquantes	
N°5 25/11/1995	462	26/11/1995 20h	124	26/11/1995 19h	103	26/11/1995 17h
N°6 09/01/1996	650	12/01/1996 17h	données manquantes		données manquantes	
N°7 14/10/1996	164	14/10/1996 22h	données manquantes		données manquantes	
N°8 10/11/1996	693	13/11/1996 03h	222	13/11/1996 02h	170	13/11/1996 01h
N°9 06/10/1997	données manquantes		données manquantes		données manquantes	
N°10 04/09/1998	données manquantes		données manquantes		données manquantes	
N°11 17/05/1999	707	18/05/1999 01h	données manquantes		114	17/05/1999 21h
N°12 19/09/2000	338	19/09/2000 15h48	139	19/09/2000 14h36	11	19/09/2000 12h24
N°13 28/09/2000	1184	29/09/2000 13h	428	29/09/2000 12h12	240	29/09/2000 10h

Événement	Anduze		Mialet		Saumane	
	Max (m ³ /s)	Heure	Max (m ³ /s)	Heure	Max (m ³ /s)	Heure
N°14 06/10/2001	données manquantes		données manquantes		données manquantes	
N°15 14/03/2002	666	15/03/2002 9h10	70	15/03/2002 7h10	données manquantes	
N°16 03/06/2002	25	05/06/2002 10h40	20	06/06/2002 16h20	3	Pas de pic observé
N°17 02/09/2002	0 m ³ /s observé !		0 m ³ /s observé !		0 m ³ /s observé !	
N°18 02/09/2002	21	Pas de pic observé	4	Pas de pic observé	1	Pas de pic observé
N°19 08/09/2002	3634	09/09/2002 06h00	915	09/09/2002 05h50	833	09/09/2002 04h40
N°20 27/10/2004	248	28/10/2004 07h45	83	Pas de pic observé	53.54	28/10/2004 04h20
N°21 05/09/2005	154	07/09/2005 01h40	74	07/09/2005 00h40	13	07/09/2005 01h55
N°22 03/11/2005	207	04/11/2005 15h40	99	04/11/2005 14h45	57	04/11/2005 12h30
N°23 12/09/2006	40	14/09/2006 07h05	32	14/09/2006 02h25	9	Pas de pic observé
N°24 23/09/2006	193	25/09/2006 00h55	93	25/09/2006 23h45		Pas de pic observé
N°25 18/10/2006	1441	19/10/2006 23h55	105	19/10/2006 02h55	267	19/10/2006 21h10
N°26 14/11/2006	277	17/11/2006 07h25	125	17/11/2006 07h00	40	17/11/2006 05h20

4.2.4.4 Les données nécessaires à la paramétrisation du modèle MARINE : les caractéristiques du bassin versant

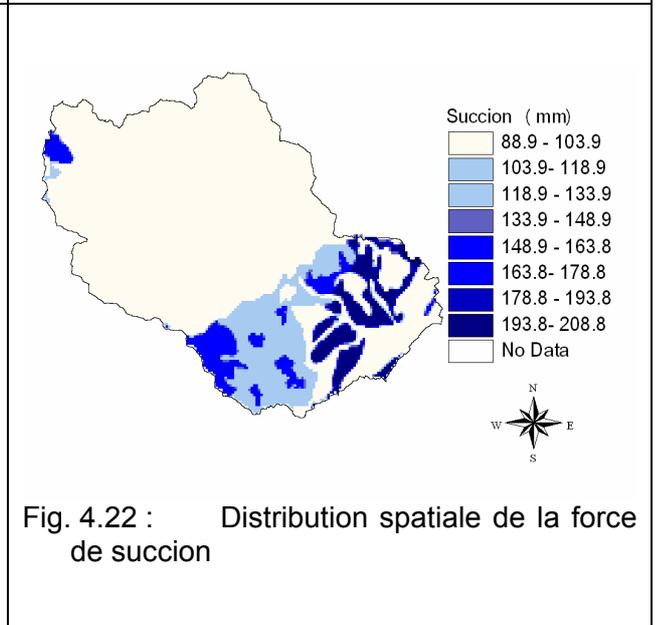
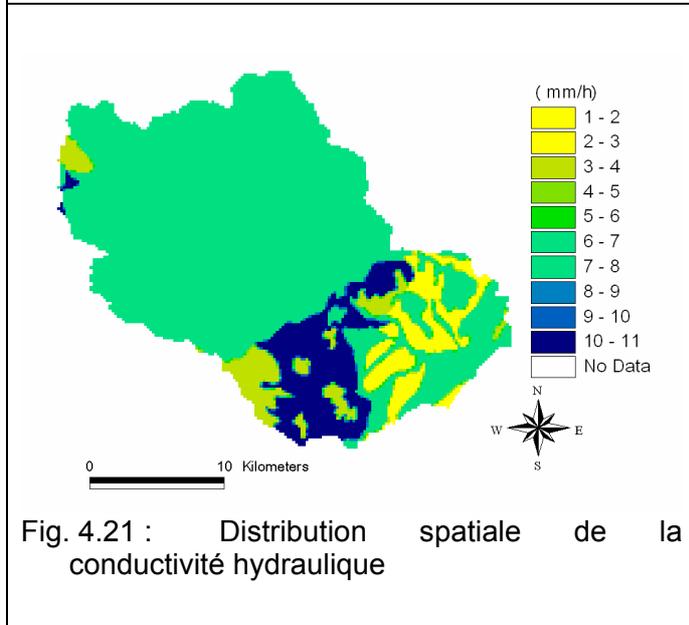
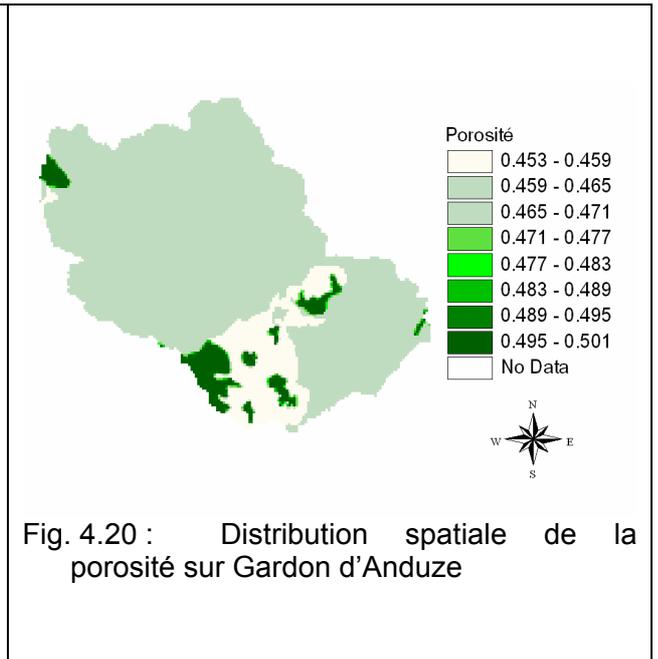
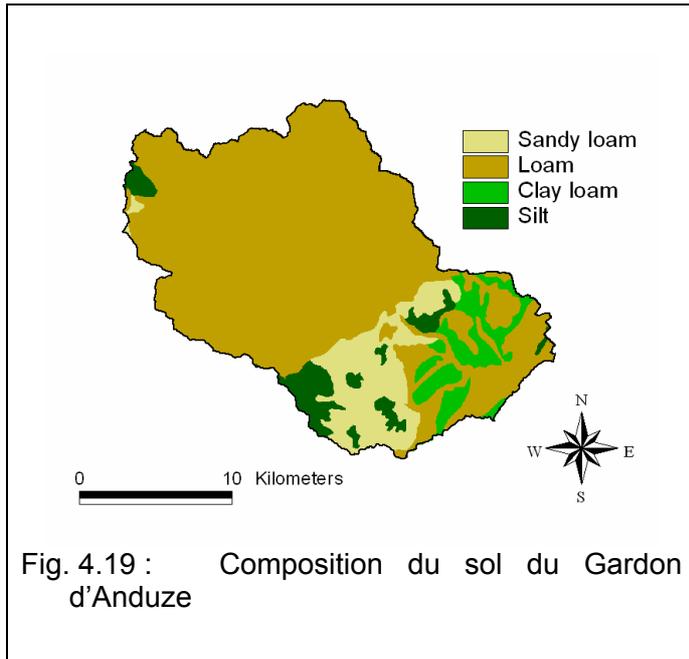
- Un modèle numérique de Terrain (MNT) au pas d'espace 50 m x 50 m est fourni par l'Institut National Géographique (IGN – BD TOPO).
- La composition du sol est donnée par le BRGM et l'INRA (Fig. 4.19). Les différents paramètres d'infiltration (la porosité Fig. 4.20, la conductivité hydraulique Fig. 4.21 et la force de succion Fig. 4.22) sont déduits de cette composition à partir du tableau 4.2.
- Les épaisseurs de sol proviennent également de l'INRA. La majorité des sols est d'épaisseur moyenne, i.e. entre 20 cm et 60 cm. Les sols très fins, d'épaisseur inférieure à 5 cm représente 53 % de la surface du bassin versant et les sols épais 12%. Les sols épais sont localisés en aval du bassin, à proximité des cours d'eau (Fig. 4.15)
- Une classification de l'occupation du sol issu de l'imagerie satellitale (satellite SPOT) est disponible. Les coefficients de frottement de Manning sur le bassin versant sont déduits de cette classification à partir du tableau 4.3.

Le tableau 4.9 résume les caractéristiques des paramètres distribués listés ci-dessus.

Tableau 4.9. Caractéristiques des paramètres d'infiltration, de l'épaisseur du sol et du coefficient de frottement de Manning sur le Gardon d'Anduze.

	Porosité θ_s	Conductivité hydraulique K (mm/h)	Force de succion S_f (mm)	Epaisseur de sol Z (m)	Manning du BV n ($m^{-1/3}.s$)
Min	0.45	1	88.9	0	0.02
Max	0.5	10.9	208.8	1	
Moyenne	0.46	6.5	103.4	0.29	0.0878
Std Dev⁵	0.009	2.1	31.8	0.24	

⁵Std Dev : déviation standard



4.2.4.5 Données nécessaires à l'initialisation du modèle MARINE : l'humidité initiale réduite du sol

L'humidité initiale réduite est issue des données SIM. C'est l'humidité de la couche racinaire (couche « HU2 » de SIM). Elle est donnée tous les jours à 6h sur le bassin des Gardons sur 25 ans de 1981 à 2006 (Fig. 4.23).

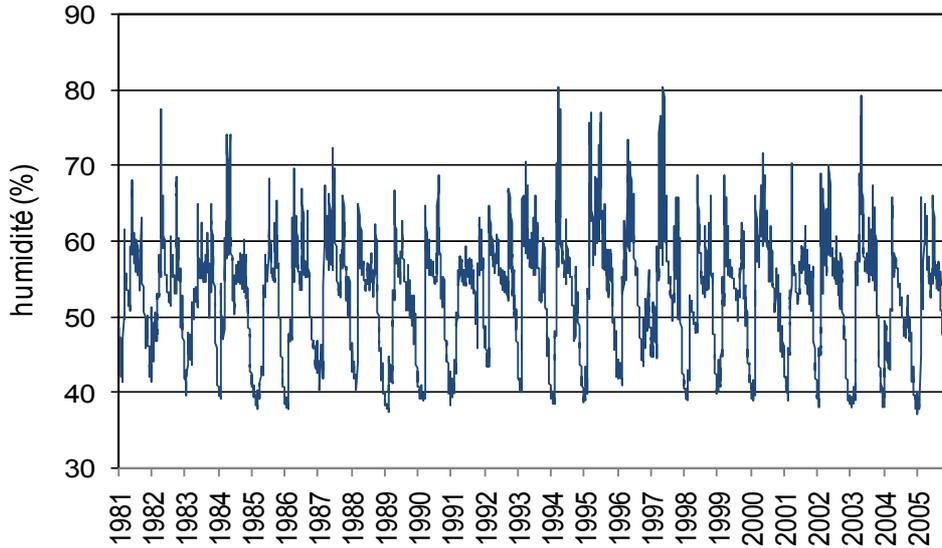


Fig. 4.23 : Evolution de l'humidité initiale réduite moyenne sur le Gardon d'Anduze de 1981 à 2006

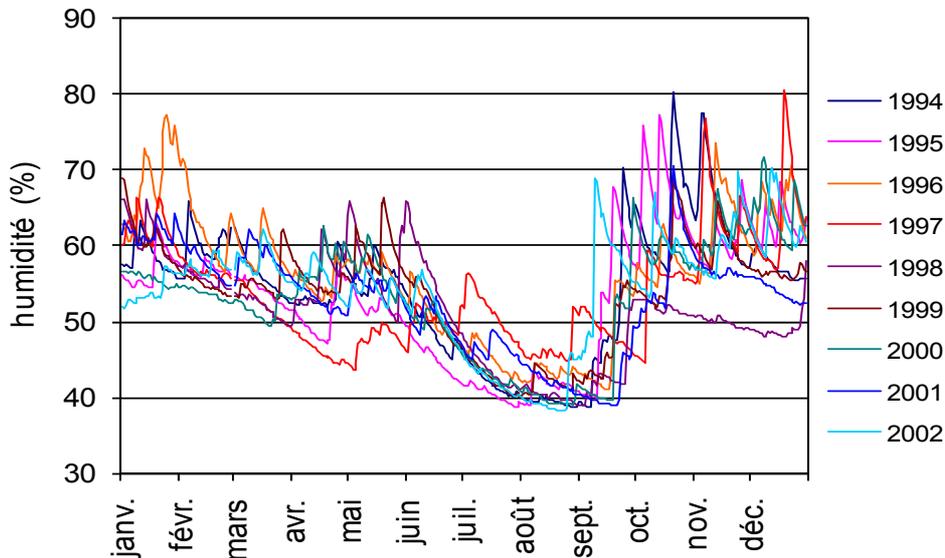


Fig. 4.24 : Evolution annuelle de la moyenne de l'humidité initiale réduite sur le Gardon d'Anduze de 1994 à 2002

L'humidité initiale réduite du sol suit une évolution périodique (Fig. 4.23 et Fig. 4.24). L'humidité décroît de Juin à Août, pendant l'été, elle atteint sa valeur minimum (environ 40 %) aux mois d'Août et de Septembre. puis augmente assez brutalement entre fin Septembre et début Novembre. L'augmentation est due à l'arrivée des épisodes pluvieux assez intenses sur un sol sec. Au cours des premiers événements pluvieux d'automne, les sols s'humidifient et même s'approchent de la saturation pour des événements intenses. Cependant, ils retrouvent assez vite (24 h) une valeur palier (autour de 60 %). Pendant les mois d'automne, l'humidité fluctue donc en fonction des épisodes pluvieux. De Janvier à Juin, la fluctuation est moins importante et l'humidité est comprise entre 50 % et 60 %.

4.3 Les événements sélectionnés

4.3.1 Critères de sélection

Deux critères de sélection ont naturellement été choisis en partant de la base de données disponible.

- Le premier critère de sélection a été attribué à la disponibilité des données de pluie.

Tous les événements dont les données de pluie sont aberrantes ou manquantes en début d'épisode ont été éliminés. Cela concerne les événements n°5, 6, 7.

Les événements n°11 et 15 ont été conservés car l'intervalle de temps d'une heure où les données de pluie sont manquantes se situe après le pic de crue et quasiment à la fin de l'épisode pluvieux. La plage de données manquante n'a cependant pas été remplacée et les simulations s'arrêtent après la dernière donnée disponible avant cette plage. Le pic de crue et la première descente peuvent tout de même bien être reconstitués avec les données fournies.

- Le second critère de sélection est un critère basé sur le débit de pointe.

En effet, le modèle MARINE a pour vocation la simulation des crues à cinétique rapide et des crues éclair. Le débit doit être assez élevé pour rentrer dans le cadre d'application du modèle. Un seuil arbitraire a été fixé au débit instantané de période de retour 2 ans pour Anduze, soit $520 \text{ m}^3/\text{s}$ selon le SAGE Gardons.

Au total 10 événements répondent aux critères utilisés. Il s'agit des événements n°1, 2, 3, 4, 8, 11, 13, 15, 19 et 25. La majorité de ces événements (6 sur 10) ont lieu à l'automne. De 2003 à 2006, un seul événement est présent.

4.3.2 Caractéristiques des événements sélectionnés

Le tableau 4.10 répertorie les caractéristiques des crues utilisées pour la calibration du modèle. Les valeurs de l'humidité initiale correspondent aux données SIM pour des mailles de 64 km^2 et les valeurs des pluies cumulées correspondent à des mailles de 1 km^2 .

Tableau 4.10. Caractéristiques des crues sélectionnées

Date de l'événement	Q _{max} (m ³ /s)	Humidité initiale				Pluie cumulée			
		Min	Max	Moy	Std dev	Min	Max	Moy	Std dev
N°1 21/09/1994	774.6	43	59	47.6	4.3	139	291	228.9	33
N°2 18/09/1995	946	48	70	56.5	5.5	40	291	142	44.4
N°3 03/10/1995	1608	51	69	56.7	4.6	32	591	217.5	111.6
N°4 13/10/1995	1411	50	79	61.8	6.9	49	397	189.2	88
N°8 10/11/1996	693	53	67	55.9					
N°11 17/05/1999	706.7	52	69	55.8	5	58	261	134	43
N°13 28/09/2000	1184.5	44	60	51.5	3.1	102	337	197.5	36.3
N°15 14/03/2002	666	51	70	56.5	4.9	6	256	63.5	70.3
N°19 08/09/2002	3634	43	60	47.6	4.6	65	792	302.8	159
N°25 18/10/2006	1440.6	34	49	37.7	4.2	74	360	219.2	49.4

4.3.3 Critiques des données

Comme il a été précisé en 1.2.3.5, il existe des incertitudes sur les données.

4.3.3.1 Erreurs de données de débits

Elles sont dues aux erreurs de mesures, de techniques d'acquisition des données et des méthodes d'interpolation et d'extrapolation. D'autre part, les crues éclair sont des phénomènes exceptionnels et l'acquisition des données est souvent difficile voire même dangereuse dans ces conditions extrêmes.

- Des incertitudes de mesures

Il existe des incertitudes sur la mesure des niveaux d'écoulement et des incertitudes sur la mesure des débits nécessaire à l'établissement de la courbe de tarage.

Lang et al., (2005) proposent dans leur article quelques ordres de grandeur des incertitudes dues à deux types d'erreurs de mesure.

En période de crue, les fortes vitesses d'écoulement peuvent entraîner du transport solide et modifier la mesure des hauteurs d'eau au niveau des échelles. Par exemple, un taux élevé de matière en suspension de la rivière peut augmenter la densité de l'eau et induire une incertitude sur la mesure de niveau de l'ordre de 3 % pour des capteurs à mesure de pression.

Pour les rivières de montagne à cinétique rapide, comme celles de la partie amont du Gardon d'Anduze, le niveau peut être très variable à cause des pentes et des blocs rocheux. Une erreur de 5 % sur la mesure du niveau, couramment observée, induit une incertitude de plus ou moins 10 % sur le débit.

Par ailleurs, le niveau de fond du lit peut être modifié, des pertes de charge locales au niveau des capteurs sont susceptibles de fausser la mesure ; enfin les capteurs peuvent être emportés par la crue.

La mesure des débits ou le jaugeage peut se faire soit par exploration des champs de vitesse à l'aide moulinets d'hydrométrie, soit par dilution. Lorsqu'elles sont réalisables, les méthodes par exploration du champ des vitesses sont assez robustes car leurs incertitudes ne dépassent pas 5 % en général, sauf pour des conditions d'écoulement plus difficile (piles de ponts, vagues, tourbillons) ou elles peuvent atteindre 10 %. Pour des petits bassins versants à cinétique rapide, il est nécessaire pour des conditions de sécurité de recourir à une autre méthode moins contraignante : le jaugeage par dilution. Sa mise en œuvre est délicate et les opérateurs n'ont pas l'accès visuel à certains paramètres.

- Des incertitudes dues aux extrapolations

Concrètement, les données de débit reposent sur l'établissement de courbes de tarage reliant le débit à la hauteur d'écoulement. Ces courbes sont construites à partir de jaugeages mesurant la correspondance entre la hauteur d'eau et le débit. Or, en période de crue, les données de jaugeage sont peu nombreuses du fait de la difficulté d'être sur le terrain au bon moment ou de l'impossibilité d'effectuer des mesures. En effet, du fait de la rapidité de ces phénomènes, plusieurs contraintes organisationnelles se succèdent pour l'équipe chargée d'effectuer les mesures. Dans un premier temps, elle doit être informée assez rapidement sur l'existence d'une crue, ensuite elle doit se rendre sur le site de mesure dans les plus brefs délais, lorsque les voies d'accès aux rivières ne se sont pas coupées, et finalement elle doit s'organiser lorsque ces phénomènes ont lieu pendant la nuit et le week-end. Il arrive aussi que les vitesses d'écoulement soient trop élevées pour le matériel.

Ainsi, des erreurs sont introduites dans l'extrapolation de la courbe de tarage en période de crue. Elles sont essentiellement dues à la non prise en compte des propriétés hydrauliques des écoulements (extrapolation linéaire de la courbe de tarage, relation non univoque à cause de l'influence aval...) ou aux discontinuités dans les écoulements (débordement lit mineur-majeur, contournement de la station de mesure, mise en charge d'ouvrages ...) (Lang et al., 2006).

Par ailleurs, plusieurs courbes de tarage doivent être définies à cause des modifications d'écoulement au cours du temps (modification ou envasement du lit du cours d'eau, travaux hydrauliques, etc.) et le raccordement de ces différentes courbes doit être validé.

Pour caler la courbe de tarage en hautes-eaux, il est nécessaire d'avoir des jaugeages au-delà du débordement du lit mineur car l'unique extrapolation graphique de la courbe apporte des erreurs beaucoup trop importantes.

Renouf et al. (2005) ont montré dans leur étude que pour la majorité des 325 stations hydrométriques étudiées, essentiellement sur la zone Rhône Méditerranée Corse, le débit correspondant à la hauteur maximale jaugée est de période de retour inférieur à 2 ans. Ainsi, l'incertitude associée à l'extrapolation des courbes de tarage peut être importante.

Finalement, le calcul complet des incertitudes sur les débits englobe toute la chaîne de traitement des données, depuis la mesure des hauteurs et des débits jusqu'à l'établissement de la courbe de tarage.

Au cours de ce travail de thèse, une incertitude de 20 % ($\pm 10\%$) a été arbitrairement retenue sur la mesure des débits.

4.3.3.2 Erreurs de données des pluies radar

La mesure de l'intensité de la pluie par le radar passe par deux étapes. La puissance rétrodiffusée par les hydrométéores (pluie, neige ou glace) vers l'antenne du radar est d'abord mesurée, puis convertie en intensité de pluie. La relation Z/R reliant la réflectivité radar aux taux de pluie est une loi empirique qui dépend du type de pluie.

Les problèmes liés à la mesure de la pluie par radar sont représentés sur la figure 4.25. Le relief produit des « échos fixes » ce qui se traduit par une zone masquée en aval. La zone de fonte, dont l'altitude est légèrement inférieure au niveau 0°, se traduit par un important renforcement de la réflectivité radar (bande brillante).

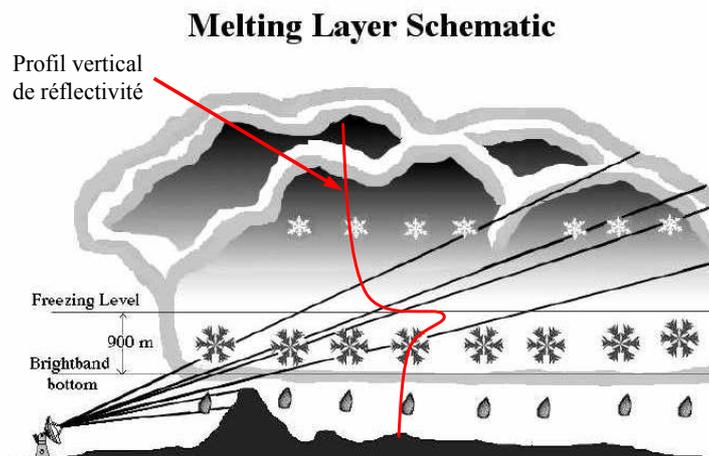


Fig. 4.25 : Principales sources d'erreurs pour les mesures des précipitations par radar : échos de sol, bande brillante et profil vertical de réflectivité, d'après Parent du Châtelet (2005)

D'autres sources d'erreur viennent s'ajouter à celles citées précédemment. Une propagation anormale du faisceau radar se traduit par des échos de sol inhabituels : par exemple, le rayon peut être rabattu vers le sol à cause d'un profil atypique de l'indice de réfraction. Des erreurs sont également introduites dans la calibration électronique du radar. Les problèmes d'atténuation des ondes électromagnétiques par la pluie peuvent également être mentionnés.

De plus, la loi empirique utilisée pour la relation Z/R dépend en réalité de la nature de l'hydrométéore, du type de pluie et de la distribution granulométrique des gouttes de pluies. Finalement, le problème principal de la mesure de la pluie par radar est l'identification imparfaite du profil vertical de réflectivité. Les informations ne sont fournies que pour les cibles visibles et le radar ne voit pas les zones cachées par les masques du relief ni les régions situées au dessous de l'altitude minimale observable. Cette altitude augmente avec la distance et une mesure quantitative de pluie n'est plus possible au delà de 100 km.

Des études ont été menées pour réduire ces différentes sources d'incertitude : prise en compte des échos fixes dans les algorithmes (Pellarin, 2001), utilisation de radars à double polarisation contre l'atténuation par la pluie (Testud et al., 2000), amélioration de l'identification du profil vertical de réflectivité par exploration volumique (balayage de la même zone de l'espace à des pas de temps rapprochés et des angles d'élévation différents pour obtenir une image volumique du système pluvieux) (Andrieu et Creutin, 1991 ; Guéguen et Tabary, 2004). Afin de lutter contre le problème de visibilité hydrologique, la densité de l'observation a été augmentée : acquisition de 8 nouveaux radars dans le projet PANTHERE (Parent et al., 2003).

4.3.3.3 Impact des erreurs de données sur les événements sélectionnés

Pour l'événement n°11, des interpolations de débits ont été réalisées sur certaines plages horaires (Fig. 4.26), cependant les pics de crue sont bien représentés. Cette crue sera tout de même utilisée pour la calibration mais les résultats seront discutés.

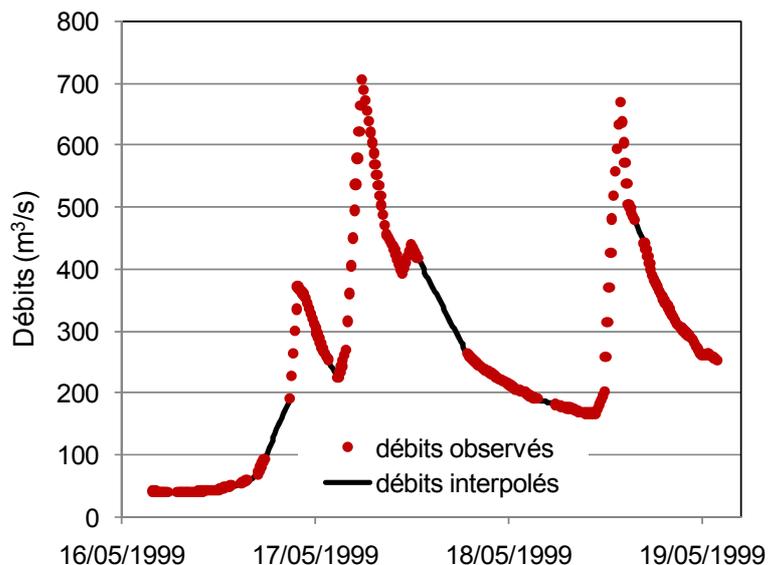


Fig. 4.26 : Interpolation des débits sur certaines plages horaires pour la crue du 17/05/1999 (événement n°11) à Anduze.

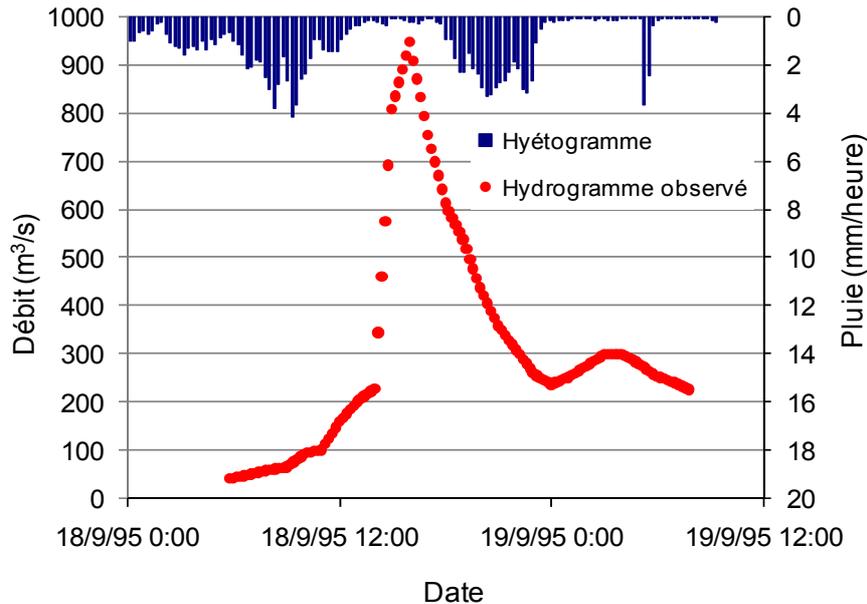


Fig. 4.27 : Pour l'événement n°2, l'hydrogramme observé ne présente aucune réaction au second pic de pluie, pourtant assez intense. Il y a probablement un problème au niveau de la mesure des pluies ou des débits.

Pour l'événement n°2, l'hydrogramme observé et les données de pluie ne semblent pas correspondre au niveau du second pic de pluie. En effet, le second pic de pluie, pourtant intense, ne provoque pas de remontée de débit significative (Fig. 4.27). Il est alors probable que les données de pluie ou de débit soient erronées. Cet événement ne sera pas retenu pour la calibration du modèle. Il sera utilisé uniquement au niveau de la validation pour la reconstitution du pic de crue et la première partie de l'hydrogramme.

4.3.4 Bilan des événements sélectionnés

Au final, 8 événements seront utilisés pour la calibration du modèle. Il s'agit des événements n°1, 3, 4, 11, 13, 15, 19 et 25.

L'événement n°8 est particulier car il dure très longtemps (93h). Sa dynamique est plus lente donc il ne rentre pas vraiment dans le même cadre d'application que les autres crues. De plus, il pourrait poser problème lors de l'application des méthodes de la calibration à cause de la lourde demande en temps de calcul et en espace mémoire.

Les événements n°2 et 8 serviront alors à la validation uniquement en tenant compte des diverses critiques émises précédemment.

4.4 Validité de l'approximation de l'onde cinématique sur le bassin versant

On rappelle que l'approximation de l'onde cinématique est valide si le nombre d'onde cinématique k_k est supérieur à 20 et lorsque k_k est proche de 20, le nombre de Froude F_0 doit être supérieur à 0.4. Pour des faibles valeurs de F_0 , l'approximation de l'onde cinématique ne peut être utilisée que si la condition $F_0^2 k_k$ est supérieur à 5.

Sur ce site d'étude, pour tous les événements étudiés, le nombre d'onde cinématique k_k est supérieur à 20, en revanche le nombre de Froude F_0 n'est pas toujours supérieur à 0.4. Cependant, la valeur de $F_0^2 k_k$ satisfait à la condition d'application de l'approximation comme il est possible de le constater sur la figure 4.28 pour la crue de Septembre 2002 jusque avant le pic de crue. L'approximation de l'onde cinématique a donc été considérée comme valide sur le bassin versant entier.

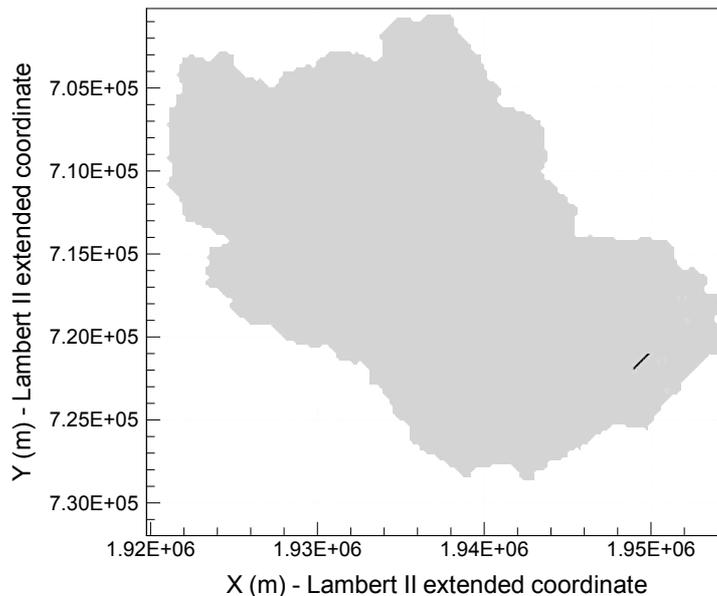


Fig. 4.28 : Validité de l'approximation de l'onde cinématique pour la crue de Septembre 2002, à 26h, d'après Singh (1996) : la zone en noir correspond aux pixels avec une valeur de $F_0^2 k_k$ inférieur à 5.

4.5 Formulation des critères de performance

Deux critères sont utilisés pour évaluer la performance du modèle.

4.5.1 Un critère classique

Le premier critère J_1 est défini de la façon suivante :

$$J_1 = 1 - NSE = \frac{\sum_{i=1}^N (Q_{obs}^i - Q_{sim}^i)^2}{\sum_{i=1}^N (Q_{obs}^i - \overline{Q_{obs}})^2} \quad (4.20)$$

où Q_{obs}^i et Q_{sim}^i sont respectivement le débit observé et simulé au pas de temps i , N le nombre de pas de temps, $\overline{Q_{obs}}$ est la moyenne des débits observés et NSE est le critère de Nash (3.2.2.2).

J_1 a une valeur nulle si les débits simulés sont confondus avec les débits observés (Fig. 4.29).

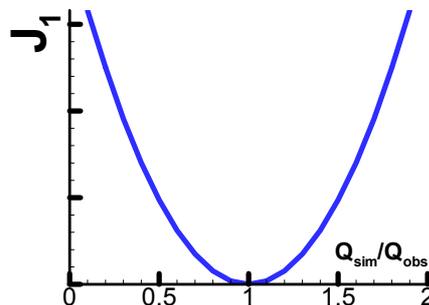


Fig. 4.29 : Représentation de J_1 en fonction de Q_{sim}/Q_{obs}

4.5.2 Un critère généralisé

Afin de prendre en compte les erreurs sur les débits observés (depuis la mesure des hauteurs et des débits jusqu'à l'établissement de la courbe de tarage), un second critère J_2 est utilisé. Dans ce travail, une incertitude de 20 % (± 10 %) est estimée sur la mesure des débits.

Le nouveau critère J_2 est alors défini comme suit :

$$\left\{ \begin{array}{l} \text{si } Q_{sim}^i < 0.8 Q_{obs}^i \text{ alors } J_2 = \frac{\sum_{i=1}^N (0.8 Q_{obs}^i - Q_{sim}^i)^2}{\sum_{i=1}^N (Q_{obs}^i - \overline{Q_{obs}})^2} \\ \text{si } Q_{sim}^i > 1.2 Q_{obs}^i \text{ alors } J_2 = \frac{\sum_{i=1}^N (1.2 Q_{obs}^i - Q_{sim}^i)^2}{\sum_{i=1}^N (Q_{obs}^i - \overline{Q_{obs}})^2} \end{array} \right. \quad (4.21)$$

en utilisant les mêmes notations que précédemment.

J_2 a une valeur nulle si les débits simulés correspondent à $\pm 20\%$ des débits observés (Fig. 4.30).

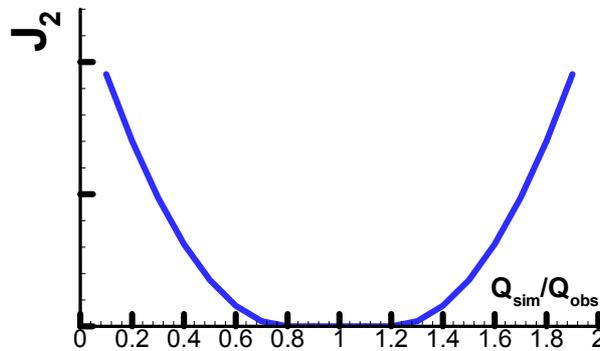


Fig. 4.30 : Représentation de J_2 en fonction de Q_{sim}/Q_{obs}

En utilisant ce nouveau critère de performance, une erreur de 20% sur les débits observés est donc admise (Fig. 4.31). Toutes simulations se situant dans la plage de débits définis par les traits discontinus de la figure 4.31 sont considérées comme « parfaites ».

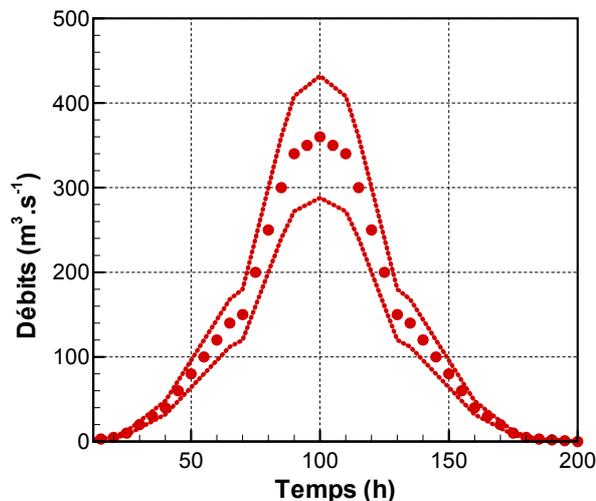


Fig. 4.31 : Les cercles correspondent aux débits observés et les traits discontinus représentent la plage des débits acceptables avec une erreur de $\pm 20\%$ sur les valeurs observées.

4.6 Synthèse

Le bassin versant du Gardon d'Anduze est soumis à un aléa pluviométrique intense en particulier au cours des mois d'automne. Les intensités de pluie varient entre quelques dizaines et plusieurs centaines de millimètres par heure sur plusieurs heures consécutives. D'autre part, la bordure des Cévennes est caractérisée par des reliefs accidentés. Ainsi, la combinaison de ces différents facteurs cause des crues violentes, voire meurtrières, dénommées crues « éclair ». Ce risque de crue est d'autant plus problématique que l'urbanisation à proximité des cours d'eau augmente.

La compréhension et la connaissance de la genèse de ces crues et de leurs paramètres hydrologiques sont nécessaires pour prévenir les risques associés. D'un point de vue modélisation hydrologique, il s'agit de d'obtenir une représentation la plus fidèle possible du phénomène physique.

A partir des caractéristiques du bassin versant et des données disponibles, la calibration du modèle MARINE est réalisée (chapitre 5). La base de données utilisée pour la calibration a été sélectionnée selon différents critères adaptés à l'application du modèle. La plupart des événements sélectionnés ont lieu en automne. Une attention particulière a été accordée à la qualité des données. L'incertitude sur les débits sera prise en compte à travers un critère de performance généralisé tolérant une erreur de 20 % sur le débit.

Chapitre 5

Calibration du modèle et prise en compte des incertitudes sur le débit

5.1	Introduction	120
5.2	Calibration du modèle	121
	5.2.1 Détermination des paramètres de calibration : analyse de sensibilité et expériences jumelles	121
	5.2.2 Résultats des simulations de Monte-Carlo	129
	5.2.3 Calibration avec la méthode d'optimisation	138
	5.2.4 Synthèse	161
5.3	Prise en compte des incertitudes sur le débit	162
	5.3.1 Résultats des simulations de Monte-Carlo	162
	5.3.2 Utilisation de la méthode d'optimisation	168
	5.3.3 Synthèse	172
5.4	Synthèse de la calibration et validation	173
	5.4.1 Agglomération des résultats	173
	5.4.2 Validation	174
5.5	Conclusion	178

CHAPITRE 5 Calibration du modèle et prise en compte des incertitudes sur le débit

5.1 Introduction

Un modèle performant en simulation a de grandes chances d'être performant en prévision. Il est alors nécessaire de rechercher dans une première étape un modèle performant en simulation.

Ce chapitre est ainsi consacré à la calibration du modèle MARINE par l'utilisation des deux méthodes : un processus d'estimation des paramètres basé sur la méthode de l'état adjoint et l'exploitation des simulations de Monte-Carlo. Les deux critères de performance présentés dans le chapitre précédent seront utilisés.

Il s'agit d'utiliser le mieux possible les observations disponibles pour améliorer la compréhension des processus de surface et cerner la réponse dominante du bassin en contraignant les incertitudes liées la caractérisation des systèmes hydrologiques, notamment les incertitudes paramétriques.

La méthodologie employée dans ce travail pour le calage du modèle distribué étudié consiste à utiliser des facteurs multiplicatifs (ou coefficients correctifs) pour ajuster de manière relative une distribution spatiale fixée a priori des paramètres du modèle (Refsgaard, 1997 ; Senarath et al., 2000 ; Madsen, 2003). Il en résulte une paramétrisation d'ordre réduit souvent de faible dimension.

L'objectif est bien sûr d'obtenir des simulations se rapprochant le plus possible des observations, toutefois une attention particulière sera accordée à la reproduction du pic de crue. La mesure de l'adéquation entre les simulations et les observations est faite selon les deux critères présentés en 4.5. Les huit crues sélectionnées en 4.3 ont été testées.

La démarche adoptée est la suivante :

- premièrement, les résultats d'une analyse de sensibilité globale des paramètres du modèle (les paramètres constants du modèle ou les coefficients multiplicateurs associés aux variables distribuées) et des expériences jumelles sont analysés pour déterminer les paramètres les plus sensibles,
- deuxièmement, les résultats des simulations de Monte-Carlo sont exposés et synthétisés,
- troisièmement, la méthode d'optimisation est employée pour déterminer les jeux de paramètres optimums pour chaque crue sélectionnée,
- finalement, une confrontation de ces deux approches est menée pour en déduire le jeu de paramètre issu de la calibration et procéder à la validation du modèle.

5.2 Calibration du modèle

5.2.1 Détermination des paramètres de calibration : analyse de sensibilité et expériences jumelles

5.2.1.1 Expériences jumelles – test I.

L'estimation des paramètres du modèle en utilisant la méthodologie des expériences jumelles décrite au chapitre 3 est effectuée dans cette partie (Fig. 5.1). Cette méthode est intéressante car elle permet de tester les performances de la méthode d'optimisation et d'en étudier la stabilité. En effet, dans ce cas le modèle est supposé parfait et la seule source d'incertitude est celle liée à la méconnaissance des paramètres.

L'approche employée consiste à utiliser les distributions spatiales fixées a priori et d'y affecter des facteurs multiplicatifs pour en ajuster l'amplitude (Senarath et al., 2000).

A partir de conditions initiales d'optimisation différentes (jeux de paramètres perturbés), l'objectif est de retrouver le jeu de paramètres de référence.

Parmi les nombreux tests effectués, deux d'entre eux pour lesquels des enseignements majeurs ont été tirés sont présentés dans ce paragraphe. Dans tous les cas, la méthodologie adoptée est toujours la même, celle exposée en 3.2.6.

Pour ce premier test, 6 paramètres sont testés : C_K , C_{Sf} , C_n , C_Z , C_θ , n_d (les coefficients multiplicatifs de la conductivité hydraulique, de la force de succion, du coefficient de Manning du bassin, de la profondeur du sol, de l'humidité initiale ainsi que le coefficient de Manning du lit majeur des drains).

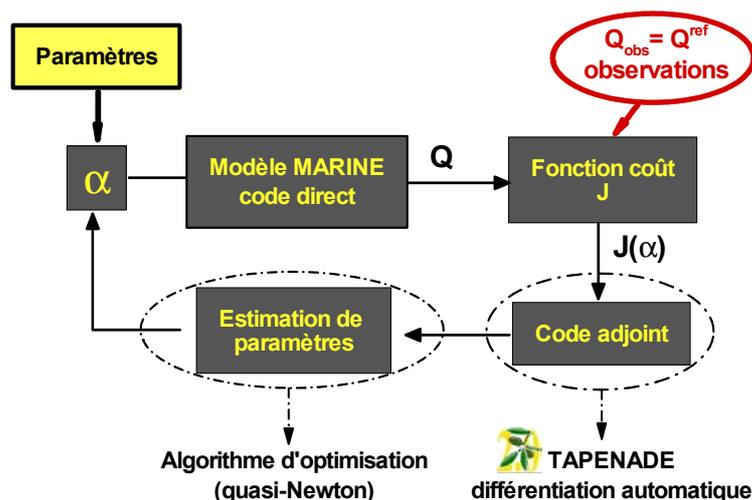


Fig. 5.1 : Schéma représentant le processus d'estimation des paramètres du modèle MARINE avec la méthode adjointe.

Le tableau 5.1 donne les valeurs de référence de ces paramètres (α^{ref}), leurs valeurs initiales (α^{init}) et les valeurs estimées (α^{assim}) ainsi que les statistiques d'erreur calculées (3.2.6). Les valeurs des fonctions coût sont aussi indiquées.

Le premier constat est que les valeurs de référence des paramètres ne sont retrouvées $\alpha^{assim} \neq \alpha^{ref}$ pour aucune des deux initialisations (Fig. 5.2). Certains paramètres n'en sont pas très éloignés (C_K , C_n), d'autres très proches (n_d), mais les résultats ne sont pas satisfaisants dans la mesure où la seule source d'incertitude est celle liée aux paramètres. Cela ne veut pas dire pour autant que la méthode n'est pas performante car les valeurs des fonctions coût obtenues après assimilation sont très faibles (quasiment nulles) (Tab. 5.1) et les hydrogrammes assimilés sont confondus avec celui observé (Fig. 5.2.h). Ceci traduit en fait que plusieurs jeux de paramètres différents apportent des résultats similaires (au sens de la fonction objectif utilisée), en d'autres termes qu'il y a équi-finalité. Des interactions se produisent entre certains paramètres et apportent des effets compensatoires.

Il est alors nécessaire de déterminer les paramètres les plus sensibles du modèle afin de réduire le nombre de paramètres à déterminer et de s'affranchir des problèmes d'interaction.

Tableau 5.1. Valeurs de référence, initiales et assimilées des six paramètres, de la fonction coût et statistiques d'erreur du test I.

	ref	initialisation 1				initialisation 2			
	α^{ref}	α^{init}	δ_{α}^{init}	α^{assim}	δ_{α}^{assim}	α^{init}	δ_{α}^{init}	α^{assim}	δ_{α}^{assim}
C_K	5	1	0.8	4.08	0.18	10	1	4.74	0.053
C_{Sf}	1.2	1	0.17	1.88	0.57	5	3.17	2.2	0.83
C_n	0.8	1	0.25	0.53	0.33	5	5.25	1.31	0.64
C_Z	5	1	0.8	3.23	0.35	10	1	11.82	1.36
C_θ	0.7	1	0.43	0.001	0.99	1.5	1.14	1.41	1
n_d	0.1	0.2	1	0.103	0.03	0.067	0.34	0.093	0.066
J		5.8		3.9 10 ⁻⁵		0.036		1.1 10 ⁻⁴	

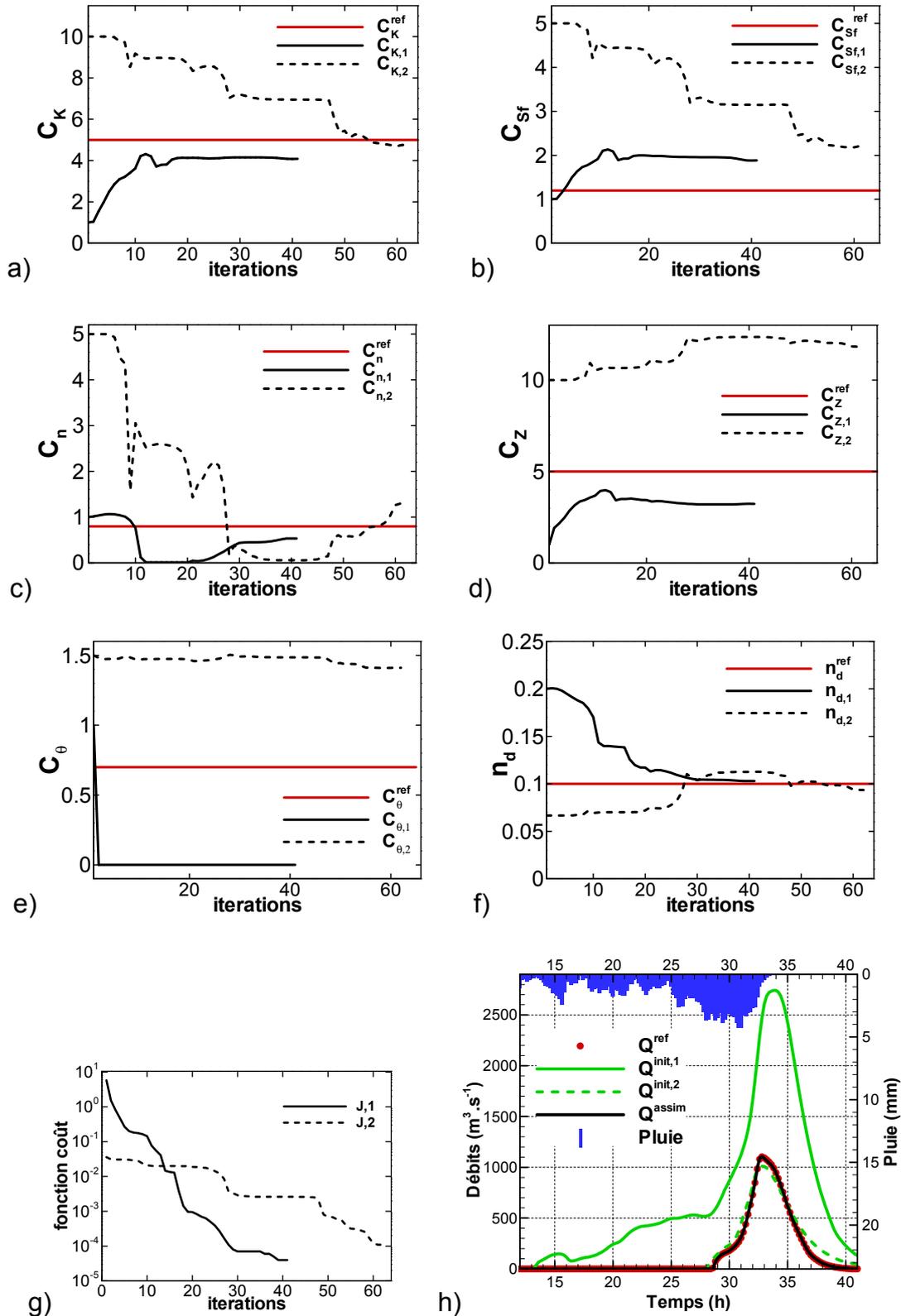


Fig. 5.2 : Convergence des paramètres C_K (a), C_{Sf} (b), C_n (c), C_Z (d), C_θ (e), n_d (f) et de la fonction coût (g) pour les deux initialisations. Hydrogrammes initiales (trait vert continu et discontinu) et assimilé (trait noir) comparés à l'hydrogramme de référence (points rouges) (h) – test I.

5.2.1.2 Analyse de sensibilité globale

Une analyse de sensibilité globale avec la méthode généralisée (GSA- 3.1.2.2) a été menée en parallèle de cette étude (Le, 2008).

On a pu conclure de l'analyse globale :

- qu'il était difficile de faire une distinction entre tous les paramètres du modèle lorsqu'ils étaient étudiés simultanément. Cette observation rejoint le constat établi lors du premier test réalisé avec les expériences jumelles,
- que le paramètre le plus influent parmi les coefficients de frottement intervenant dans la fonction de transfert (à savoir les coefficients de Manning du versant, du lit mineur et du lit majeur) est le coefficient de Manning du lit majeur des drains,

En conclusion, cette analyse de sensibilité a montré que le paramètre le plus important est, dans la plupart des cas, le coefficient de Manning-Strickler du lit majeur des drains. Ceci peut s'expliquer par le fait que les crues étudiées sont des crues rapides, importantes et que la plupart du temps elles s'accompagnent d'une invasion du lit majeur des drains. Ensuite, ce sont l'humidité initiale et la profondeur du sol, paramètres déterminant directement la capacité d'infiltration dans le sous-sol, qui influent sur les résultats des simulations. La conductivité hydraulique se positionne après et la force de succion ne joue qu'un rôle mineur.

En résumé de cette analyse, on peut écrire que : $C_{Sf} < C_K < C_{\theta_s} < C_Z \approx C_{\theta_i} < n_d$

5.2.1.3 Détermination des paramètres de calibration

Dans l'étude présente, il est admis que C_{θ_s} et C_Z sont corrélés. Le produit $Z \times \theta_s$ détermine en quelque sorte la capacité de stockage du sous-sol (en fait $Z\Delta\theta$). La figure 5.3 montre d'ailleurs que les hydrogrammes issus de simulations avec un produit $C_{\theta_s} \times C_Z$ constant sont confondus pour la crue de Septembre 1994. Ainsi, dans cette étude, seul le paramètre C_Z sera à déterminer pour ne pas inclure des phénomènes d'interaction entre les deux paramètres lors de la calibration du modèle. Par ailleurs, l'humidité initiale sera traitée à part car il s'agit d'un paramètre événementiel dépendant de la crue étudiée au même titre que le forçage par les pluies. Ce paramètre ne fera alors pas partie du processus de calibration mais sera considéré comme une donnée.

Le facteur C_{Sf} ne sera pas pris en compte car il est peu influent sur la réponse du modèle. La figure 5.4 montre effectivement que ce paramètre n'a que peu d'influence pour la crue de Septembre 1994, les hydrogrammes sont quasiment identiques pour des coefficients C_{Sf} allant de 0.5 à 20.

En conclusion, les paramètres à déterminer seront donc : C_k, C_Z et n_d .

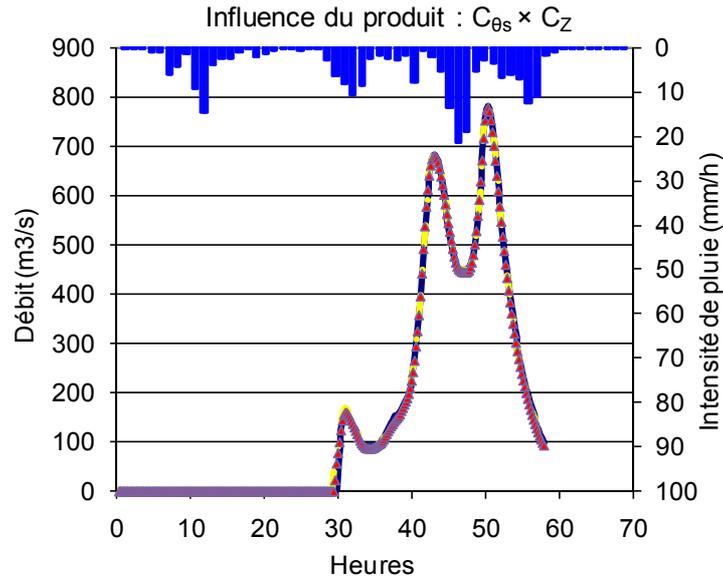


Fig. 5.3 : Hydrographes issus de trois simulations avec un produit $C_{\theta s} \times C_Z$ constant pour la crue de Septembre 1994.

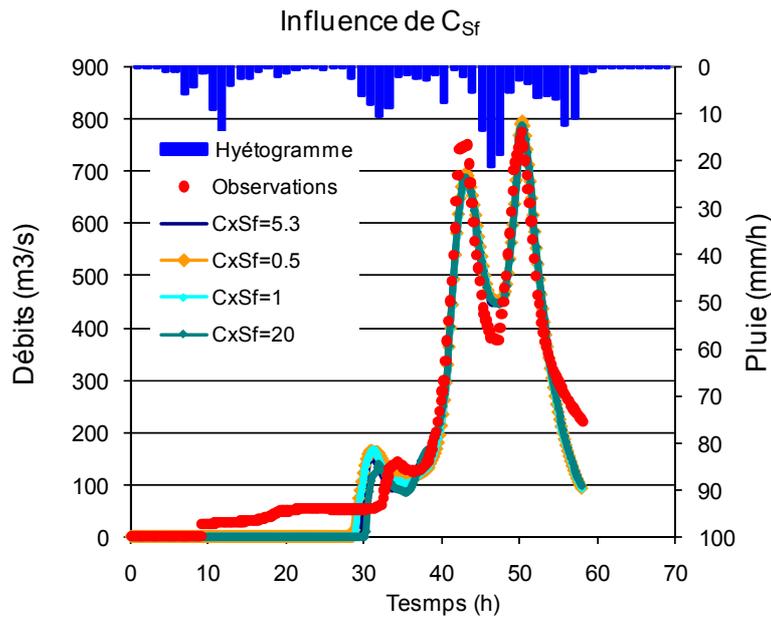


Fig. 5.4 : Influence directe de C_{Sf} pour la crue de Septembre 1994.

5.2.1.4 Expériences jumelles (test II) et validation du choix des paramètres à déterminer

L'estimation est dans ce cas réduite au calage de trois paramètres : (C_K, C_Z, n_d). Les résultats sont présentés dans le tableau 5.2 et sur la figure 5.5. On observe une vitesse de convergence à peu près similaire pour les deux initialisations, l'algorithme d'optimisation converge vers les paramètres de référence quelque soit le point de départ. Le coefficient de Manning du lit majeur des drains (n_d) s'ajuste un peu plus rapidement vers sa valeur de référence que les autres coefficients.

Comme les sensibilités de la fonction objectif aux paramètres (dérivées de la fonction par rapport aux paramètres) sont utilisées pour calculer les directions de descente de l'algorithme, les paramètres les plus influents sont retrouvés plus rapidement. Le paramètre n_d est donc plus sensible que les autres.

Les trois paramètres testés sont donc identifiables de façon stable à partir d'observations synthétiques des débits.

Tableau 5.2. Valeurs de référence, initiales et assimilées des trois paramètres, de la fonction coût et statistiques d'erreur pour le test II.

	ref	Test 1				Test 2			
	α^{ref}	α^{init}	δ_{α}^{init}	α^{assim}	δ_{α}^{assim}	α^{init}	δ_{α}^{init}	α^{assim}	δ_{α}^{assim}
C_K	5	1	0.8	5	0	10	1	5	10^{-3}
C_Z	5	1	0.8	4.99	$4 \cdot 10^{-5}$	10	1	4.99	$4.6 \cdot 10^{-4}$
n_d	0.1	0.2	1	0.1	0	0.05	0.5	0.1	0
J		2.83		$5.4 \cdot 10^{-10}$		0.26		$4.1 \cdot 10^{-8}$	

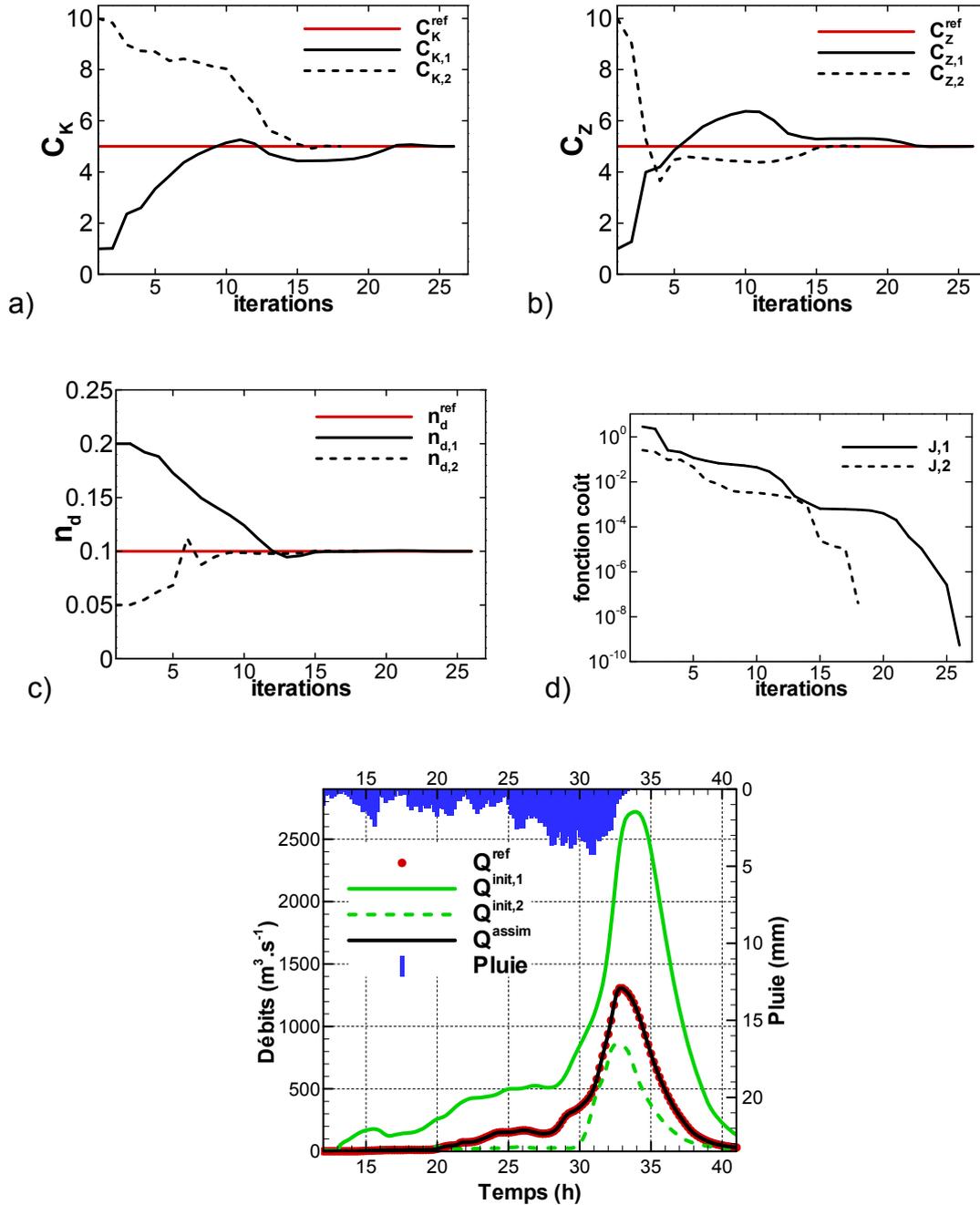


Fig. 5.5 : Convergence des paramètres C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) pour les deux initialisations. Hydrogrammes initiales (trait vert continu et discontinu) et assimilé (trait noir) comparés à l'hydrogramme de référence (points rouges) (e) – test II.

5.2.1.5 Bilan des enseignements

Les mêmes enseignements ressortent des expériences jumelles et de l'analyse de sensibilité globale :

- un nombre trop important de paramètres aboutit à des phénomènes d'interaction entre les paramètres et engendre des problèmes d'identification des paramètres,
- le modèle est surtout sensible aux trois paramètres C_k , C_z et n_d .

En outre, les résultats obtenus lors du test II des expériences jumelles sont très encourageants et mettent en évidence l'efficacité de l'approche variationnelle pour l'estimation de paramètres. Les gradients calculés par différentiation automatique peuvent guider de manière efficace un algorithme de descente de type quasi-Newton pour la minimisation d'une fonction objectif mesurant l'écart aux observations.

D'autre part, le test I des expériences jumelles, a permis de mettre en évidence que même dans le cas d'un modèle parfait et d'observations certaines, les interactions entre les paramètres pouvaient entraîner des phénomènes de compensation de l'erreur. Il est alors fort probable qu'un comportement similaire apparaisse pour la compensation d'incertitudes sur les observations, sur le forçage et sur la structure du modèle.

Après avoir déterminé le jeu de paramètres le plus sensible permettant une identification stable par la procédure d'estimation utilisée, le but est à présent de calibrer le modèle avec ce jeu de paramètres.

5.2.2 Résultats des simulations de Monte-Carlo

5.2.2.1 Introduction

Des simulations de Monte-Carlo (Freer et al., 1996 ; Kuczera and Parent, 1998) ont été réalisées pour chaque crue. Elles consistent à échantillonner l'espace des paramètres avec une distribution uniforme, aucune hypothèse préalable n'est faite sur les valeurs des paramètres mis à part les valeurs minimale et maximale de leurs intervalles de variation. Les paramètres testés ainsi que leurs intervalles de variation respectifs sont résumés dans le tableau 5.3.

Tableau 5.3. Intervalles de variation des paramètres utilisés dans les simulations des Monte-Carlo (* la valeur maximale dépend des crues)

Paramètre	Valeur minimale	Valeur maximale	Distribution
C_k	0	10 – 15 *	uniforme
C_z	0	10 – 15 *	uniforme
n_d	0.05	10	uniforme

Le but recherché dans un premier temps n'est pas de trouver immédiatement des valeurs des trois paramètres mais des intervalles de valeurs acceptables pour chaque paramètre et chaque crue, En effet, les erreurs dans la modélisation du processus physique, les incertitudes sur les caractéristiques du bassin versant et sur les conditions limites induisent des problèmes d'équifinalité. Grâce à ce concept très discuté en hydrologie par Beven (1989a), Beven et Binley (1992), Freer et al. (1996) on postule qu'à un modèle donné, il n'existe pas un jeu de paramètres optimal mais plusieurs jeux de paramètres acceptables dans le sens d'une mesure considérée. Par ailleurs, ces jeux de paramètres optimaux peuvent provenir de régions très différentes de l'espace des paramètres à estimer.

Cependant, selon Beven et Romanowicz (2003) le jeu de paramètre optimum d'un unique événement n'est pas assez robuste pour la calibration d'événements différents. L'objectif est alors de trouver pour toutes les crues une valeur unique pour ces coefficients à l'intérieur d'une fourchette de vraisemblance. C'est pourquoi, les plages de valeurs acceptables trouvées sont ensuite regroupées pour en déduire une unique plage de valeurs acceptables par paramètre valable pour toutes les crues.

La valeur de vraisemblance utilisée est le critère de Nash. Le nombre de tirages des simulations de Monte-Carlo dépend du nombre de paramètres à tester. Il est admis que si n paramètres sont à tester, un nombre de tirage de l'ordre de 10^n est nécessaire et suffisant.

5.2.2.2 Résultats

La figure 5.6 représente, pour toutes les crues, les valeurs du critère de Nash des simulations de Monte-Carlo pour les trois paramètres. Théoriquement, la valeur maximale de chaque paramètre est déterminée par la valeur maximale du critère de Nash. Pour une meilleure visibilité graphique, les valeurs de Nash ont été tracées pour $n_d \in [0.05, 1]$.

La figure 5.7 montre, pour toutes les crues, les hydrogrammes correspondant aux cent meilleures simulations.

La valeur moyenne du critère de Nash (NSE) obtenue pour toutes les crues vaut : $NSE = 0.87$. La plus faible valeur est obtenue pour l'événement n°3 pour lequel $NSE = 0.736$ et la plus forte pour l'événement n°13 avec $NSE = 0.954$.

Au vu de ces valeurs, un seuil arbitraire valant $NSE = 0.7$ a été choisi pour les plages des paramètres acceptables.

- Événement n°1 – crue de Septembre 1994

Pour cette crue, aucune valeur maximale n'apparaît clairement pour chaque paramètre. Cela signifie notamment que les paramètres ne sont pas indépendants. Il est possible néanmoins de déterminer des plages de variation de chaque paramètre aux performances équivalentes.

Le paramètre C_k est le moins sensible. Sa plage de valeur acceptable, correspondant à un Nash élevé, est très large. Le paramètre C_z ne doit pas être trop faible (inférieur à 3) pour obtenir des simulations satisfaisantes. Le paramètre n_d passe légèrement par une valeur maximale autour de 0.25 pour cette crue.

Toutes les simulations partent d'un débit nul et non du débit de base réel. Avant la montée du premier pic de crue, les simulations du modèle font apparaître un premier pic de faible intensité qui n'est pas présent sur les observations. Néanmoins, les deux pics de forte intensité sont bien encadrés par les différentes simulations du modèle. Finalement, la décrue pourrait être améliorée, par exemple par l'introduction des écoulements de subsurface.

En conclusion, eu égard au seuil choisi, on doit avoir : $C_k > 0.8$, $C_z > 2.8$ et $n_d \in [0.08, 7.7]$.

- Événement n°3 – crue du 03/10/1995

Le coefficient de Nash maximum pour cette crue ($NSE = 0.74$) est moins important que pour la crue précédente. La plage de valeur acceptable pour le paramètre C_k est toujours très étendue. Le paramètre C_z passe par un maximum aux alentours de 2.5 ; sa plage de valeur acceptable est assez localisée. Le coefficient C_z ne doit pas être trop élevé pour cette crue. Le paramètre n_d présente une valeur maximum de 0.14. Malgré une valeur du coefficient de Nash moins élevée pour cette crue que pour les autres, les simulations du modèle sont assez satisfaisantes.

Au final pour le seuil choisi, on a : $C_k \in [3, 13]$, $C_z \in [2.3, 4.5]$ et $n_d \in [0.13, 0.25]$.

- Événement n°4 - crue du 13/10/1995

Aucune valeur maximale ne ressort pour tous les paramètres. Le paramètre C_K s'étend sur une très large plage de valeurs.

Les simulations sont satisfaisantes dans l'ensemble sauf pour la deuxième partie de la décrue et la récession où les incertitudes sur le débit simulé sont importantes. Certaines simulations font même apparaître un second pic de faible intensité.

Pour le seuil choisi de $NSE = 0.7$, on obtient : $C_K > 0.8$, $C_Z > 2.8$ et $n_d \in [0.08, 7.7]$.

- Événement n°11 - crue du 17/05/1999

Le coefficient C_Z affiche une valeur maximale autour de 2.3. C_K s'étend à nouveau sur une plage très large.

Le premier pic de crue n'est pas bien représenté (sous estimé et en avance par rapport aux observations) sur les cent meilleures simulations.

Pour le seuil fixé, les conditions : $C_K > 0.4$, $C_Z \in [1.8, 10.5]$ et $n_d \in [0.175, 1.3]$ doivent être satisfaites.

- Événement n°13 - crue du 28/09/2000

Pour cette crue, les valeurs des paramètres C_K et C_Z ne doivent pas être trop faibles (surtout C_Z) pour obtenir des simulations correctes. Le paramètre affiche une valeur maximale autour de 0.12.

Les hydrogrammes représentatifs des cent meilleures simulations sont très proches des observations ; les marges d'incertitude sont assez faibles pour cette crue bien que la fin de la phase de décrue pourrait être améliorée.

Finalement, on doit avoir : $C_K > 1.5$, $C_Z > 2.8$ et $n_d \in [0.067, 0.5]$.

- Événement n°15 - crue du 14/03/2002

Le seul paramètre qui paraît bien déterminé pour cette crue est le coefficient C_Z qui présente une valeur maximale de 1.2. Contrairement aux autres crues, la plage de valeurs acceptable du paramètre n_d est étalée.

Les cent meilleures simulations obtenues pour cette crue font apparaître de fortes marges d'incertitude et l'hydrogramme observé n'est pas représenté de façon fidèle. En effet, les simulations donnent un premier pic de crue, inexistant sur l'hydrogramme observé et un second pic (celui des observations) sous-estimé et en avance par rapport aux observations.

On doit avoir pour cette crue : $C_K > 2$, $C_Z \in [0.5, 10]$ et $n_d \in [0.11, 6.6]$.

- Événement n°19 - crue du 08/09/2002

Les simulations de Monte-Carlo font clairement apparaître une valeur maximale autour de 0.09 pour le paramètre n_d .

L'hydrogramme observé est assez bien reproduit par les meilleures simulations. Le second pic est un peu sous-estimé mais il est bon de rappeler que cette crue présente un caractère exceptionnel. La prise en compte de la fin de la décrue est à améliorer.

On a : $C_K > 2.3$, $C_Z > 3$ et $n_d \in [0.07, 0.62]$

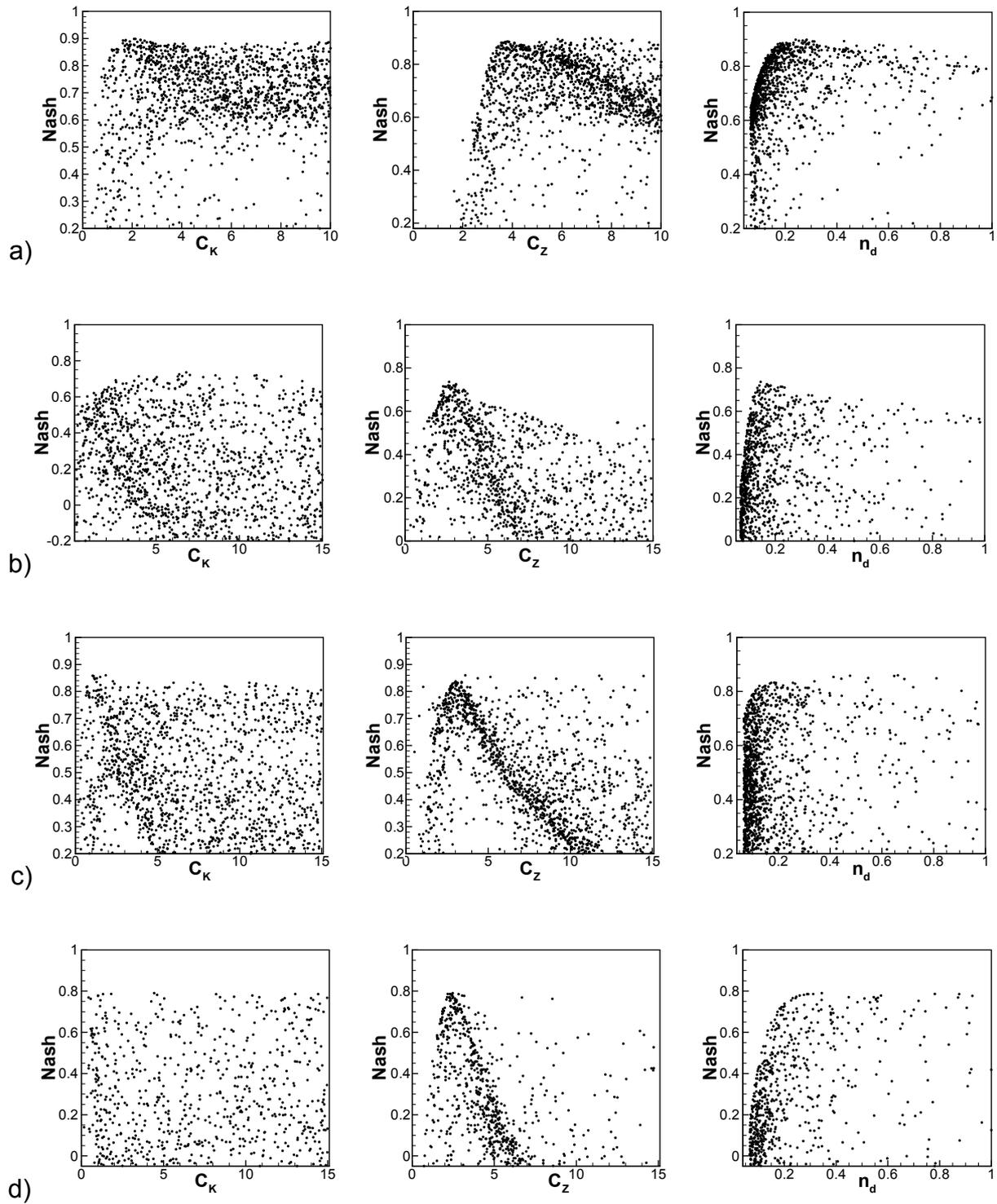
- Événement n°25 - crue du 18/10/2006

Les paramètres C_Z et n_d présentent deux valeurs maximales autour de 2.7 et 0.08 respectivement.

Les simulations sont bonnes, on peut simplement noter que la première partie de l'hydrogramme observé est sous-estimée.

Pour cette crue, le seuil est dépassé pour : $C_K > 0.4$, $C_Z > 1.7$ et $n_d \in [0.07, 0.14]$.

Un bilan général des résultats est dressé en (5.2.2.3).



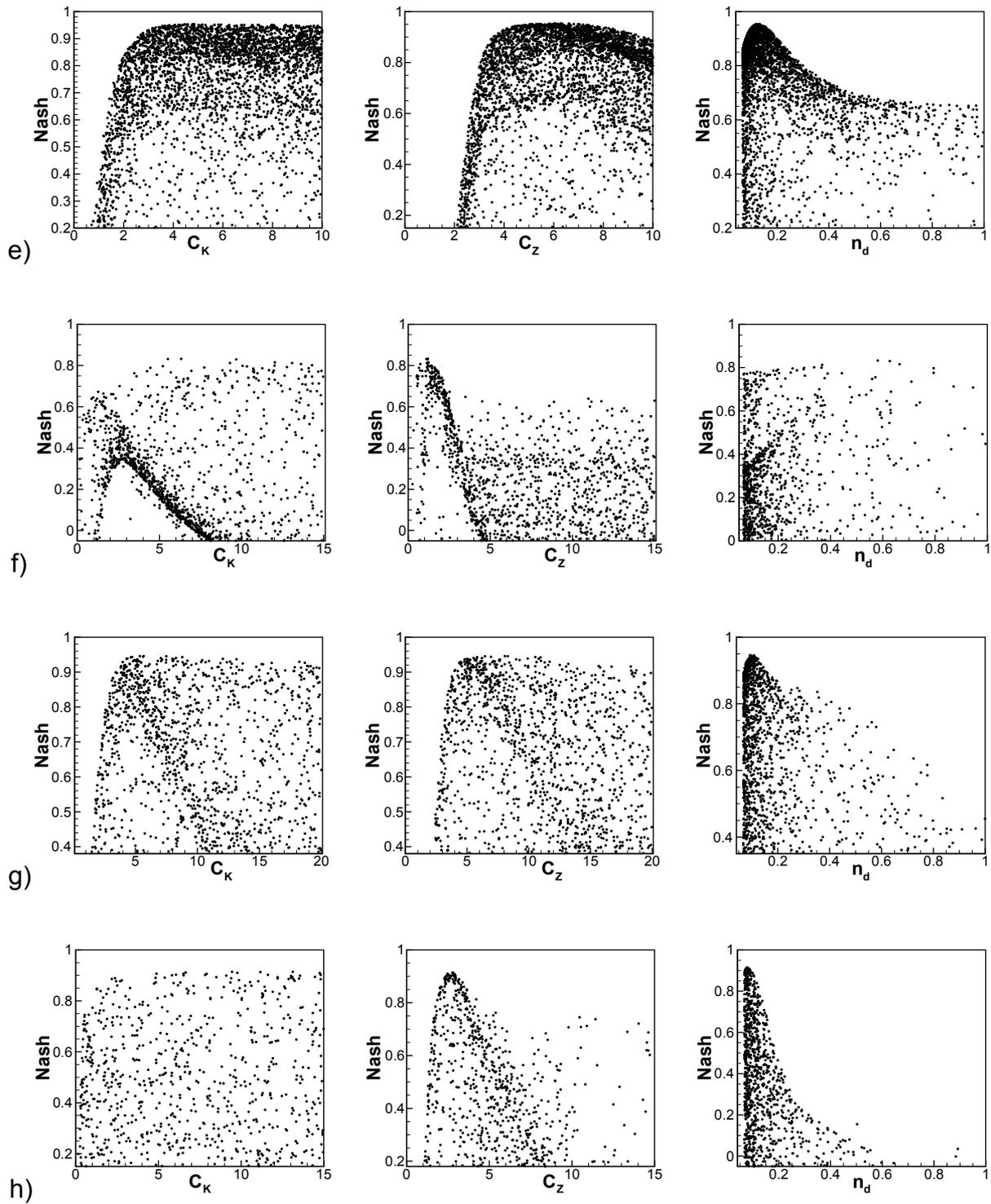
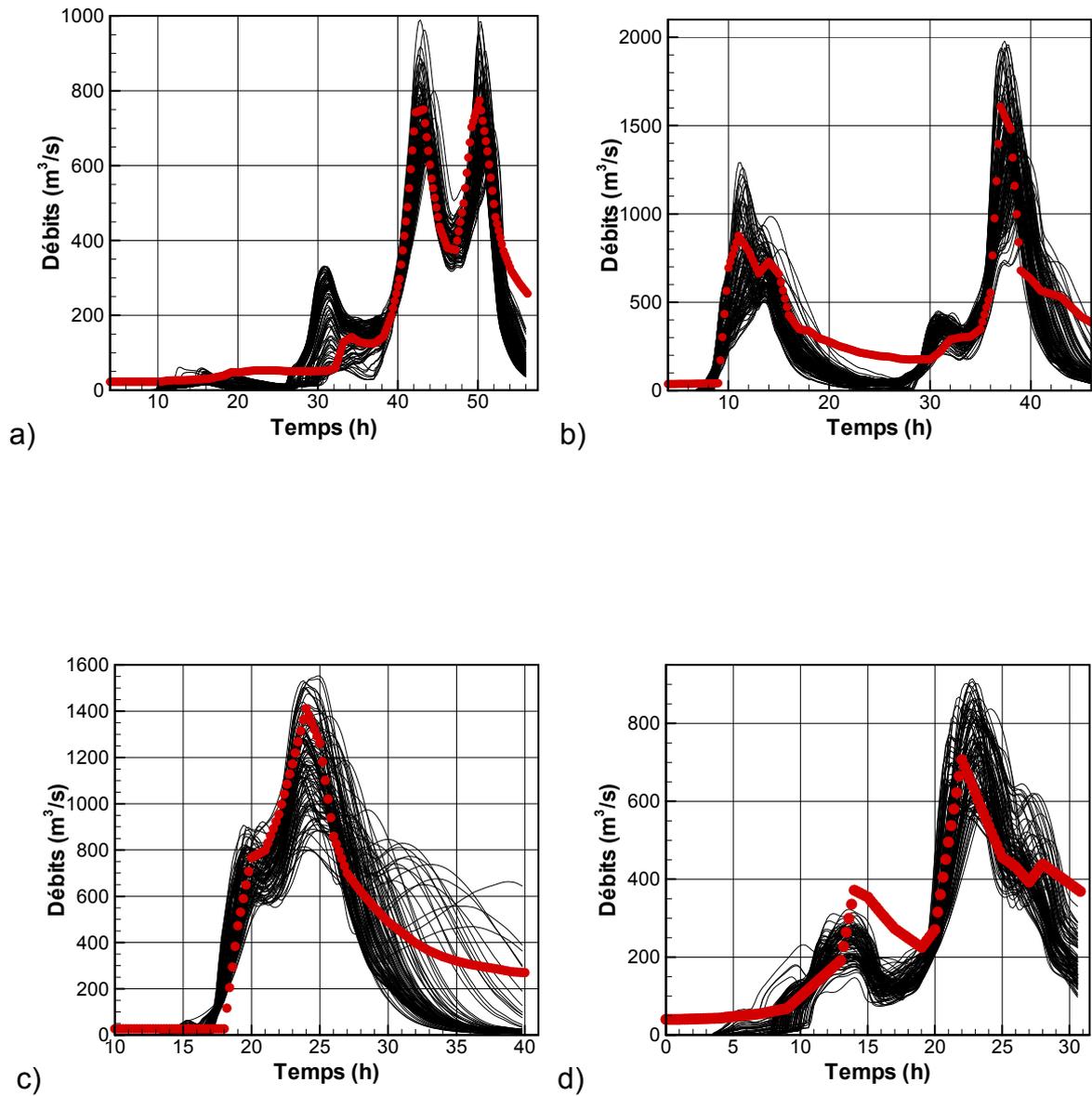


Fig. 5.6 : Valeurs du coefficient de Nash obtenues avec les simulations de Monte-Carlo pour les paramètres C_K , C_Z et n_d pour les événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g) et n°25 (h).



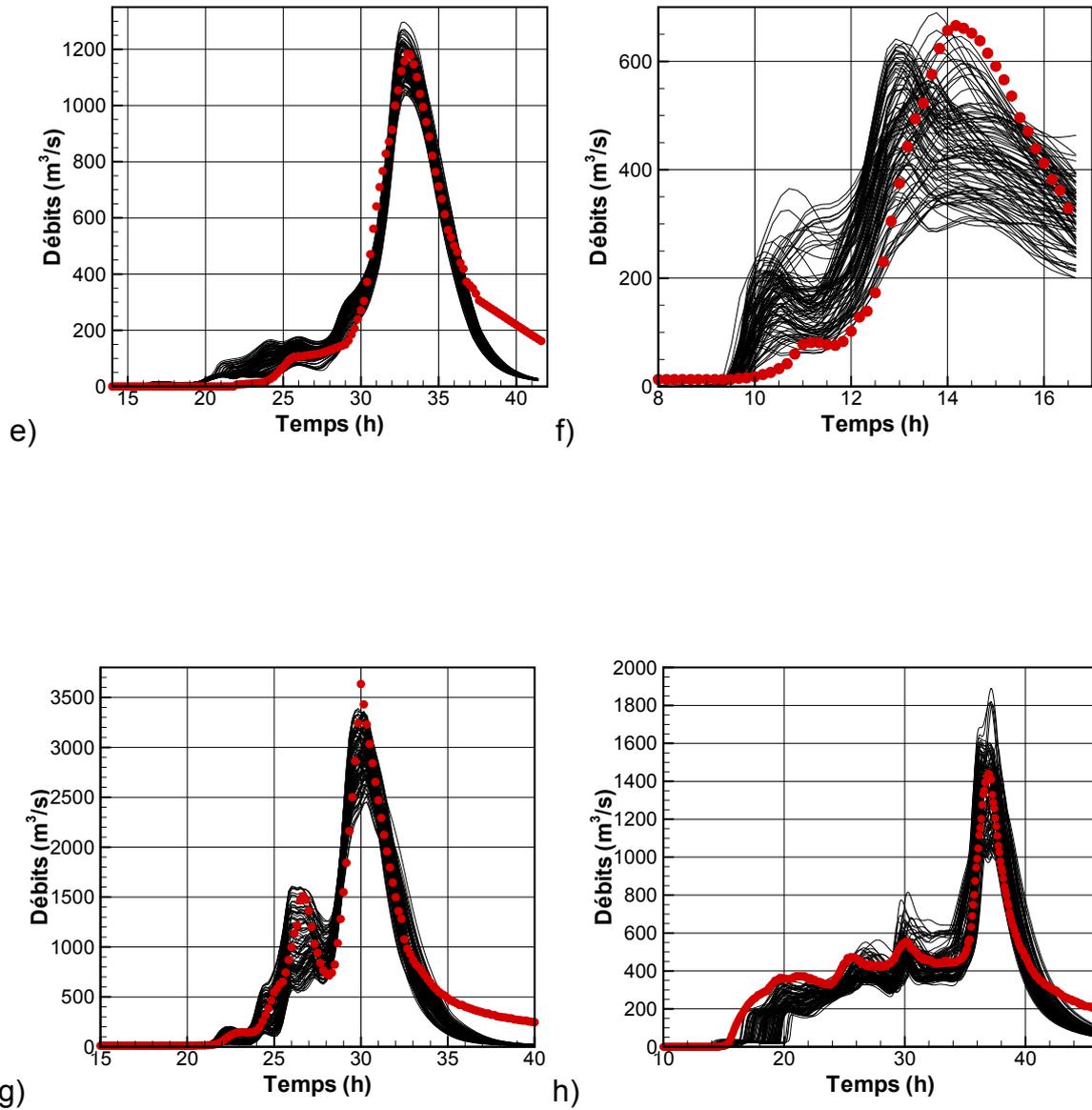


Fig. 5.7 : Tracé des cent meilleures simulations (traits noirs) comparés aux observations (points rouges) pour les événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g) et n°25 (h).

5.2.2.3 Synthèse des résultats des simulations de Monte-Carlo

En regroupant les informations obtenues pour chaque crue, deux intervalles de variation ont été trouvés pour les coefficients C_K et C_Z :

- $C_Z \in [3,10]$
- $C_Z \in [3,4.5]$

Les résultats sont synthétisés dans la figure 5.8. L'intervalle du coefficient C_K s'arrête à 10 car certains tests ont été effectués jusqu'à cette valeur. Il aurait été possible que l'intervalle de ce paramètre s'étende au-delà de la valeur 10 mais cette limitation ne posera pas de problème car, ainsi qu'on le verra plus loin, ce paramètre n'est plus sensible au-delà d'une certaine valeur pour une profondeur de sol donné.

En revanche, il n'a pas été possible d'en faire de même pour le paramètre n_d . En effet, pour les événements n°11 et n°25 par exemple, les intervalles ne se chevauchent pas. De plus, ce paramètre fait clairement apparaître des valeurs optimales en fonction des crues étudiées. En fait, l'utilisation de la loi de Manning pour la modélisation du ruissellement implique que le coefficient de Manning du lit majeur dépende du rayon hydraulique. Ce paramètre est moins important pour des crues de forte intensité à cause de la diminution du frottement occasionnée (Chow, 1959). C'est pourquoi, l'hypothèse selon laquelle ce paramètre dépend de l'intensité de la crue et de ses caractéristiques hydrodynamiques a été émise.

Des résultats similaires ont été obtenus avec d'autres valeurs de vraisemblance.

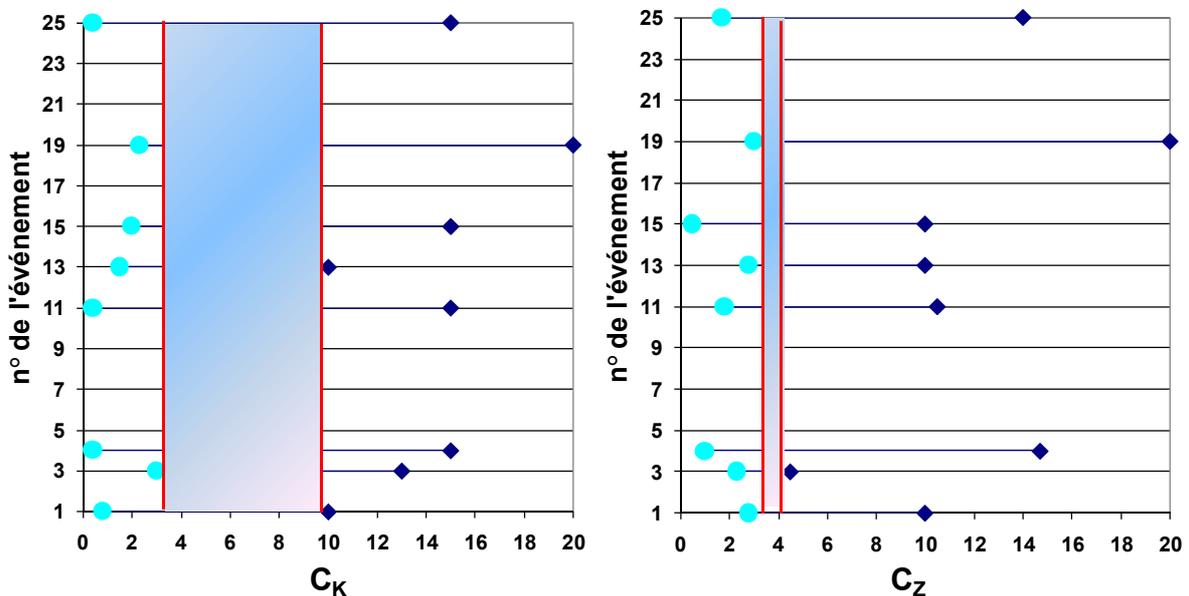


Fig. 5.8 : Bilan des intervalles de valeurs acceptables pour les paramètres C_K et C_Z avec le critère de Nash.

5.2.3 Calibration avec la méthode d'optimisation

5.2.3.1 Mise en œuvre de l'algorithme

- Fonction objectif

La fonction objectif utilisée dans cette partie est similaire au critère de Nash. Seuls les intervalles de variations diffèrent entre les deux critères.

$$J_1 = 1 - NSE \quad (5.1)$$

NSE est le critère de Nash (3.2.2.2). J_1 est une fonction décroissante (condition nécessaire à l'application de l'algorithme d'optimisation) qui prend une valeur nulle si les débits simulés sont confondus avec les débits observés, contrairement au critère de Nash qui est une fonction croissante qui prend la valeur maximale de 1.

- Les plages de variations des paramètres

L'algorithme d'optimisation N2QN1 (Lemarechal et Panier, 2000) de quasi-Newton utilisé est à contraintes de bornes. L'intervalle de variation des paramètres est défini dans le tableau 5.4. Ces plages de valeurs ont été choisies de telle sorte que les paramètres restent physiquement acceptables (selon les valeurs du tableau 4.2).

Tableau 5.4. Intervalles de variation des paramètres lors de l'optimisation

	C_k	C_z	n_d
Min	0.01	0.01	0.05
max	20	20	100

- Initialisation

Trois différentes conditions initiales du jeu de paramètres (C_k, C_z, n_d) (coefficient multiplicateur de la conductivité hydraulique, coefficient multiplicatif de la hauteur d'infiltration, coefficient de Manning du lit majeur des drains) ont été choisies comme points de départ de l'algorithme d'optimisation :

- test 1 : $(C_k, C_z, n_d) = (1, 1, 0.2)$
- test 2 : $(C_k, C_z, n_d) = (5, 5, 0.1)$
- test 3 : $(C_k, C_z, n_d) = (10, 10, 0.05)$

Il s'agit de comparer les résultats de l'algorithme d'optimisation pour trois différentes initialisations. Pour toutes les crues testées, les résultats sont présentés de la même manière. Le lecteur trouvera : les figures relatives aux tests de convergence des paramètres ; la comparaison des hydrogrammes simulé et observé et un tableau récapitulatif des résultats.

5.2.3.2 Résultats de l'optimisation

- Événement n°1 - crue du 21/09/1994

La crue de Septembre 1994 (Fig. 5.9) est une crue de moyenne intensité ($Q_{\max} = 775 \text{ m}^3/\text{s}$) à double pic.

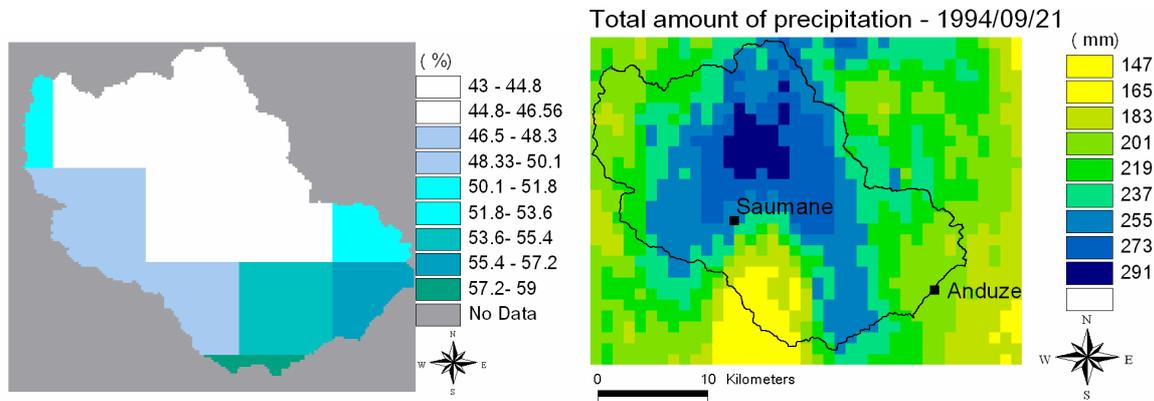


Fig. 5.9 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°1.

La Figure 5.10 représente l'évolution des paramètres C_k , C_z et n_d au cours de l'optimisation pour les différents tests d'initialisation.

Les trois tests d'initialisation aboutissent à trois jeux de paramètres (C_k , C_z , n_d) différents. L'algorithme converge donc vers trois minimums locaux avec des valeurs de fonction coût quasi identiques.

En fonction des différentes valeurs initiales des paramètres, le nombre d'itérations est plus ou moins grand. Pour cette crue, le nombre d'itérations est plus important lorsque les coefficients C_k et C_z initiaux sont unitaires, c'est-à-dire lorsque qu'aucune hypothèse de départ n'est émise sur leurs valeurs et pour un coefficient de Manning initial de 5. Pour les tests 2 et 3, la convergence des paramètres C_z et n_d est plus rapide (respectivement 7 et 12 itérations) que celle du coefficient C_k . Par ailleurs pour ces mêmes tests 2 et 3 les valeurs finales de C_z , n_d et J sont très proches alors que les valeurs de C_k sont différentes (respectivement 8.9 et 13.4). Il semble donc, pour cette crue, que le paramètre C_k n'ait plus d'influence à partir d'une certaine valeur. En d'autres termes, le coefficient C_k n'est plus sensible au-delà d'un certain seuil pour une profondeur de sol donnée, l'eau ne peut s'infiltrer davantage.

En comparant les tests 1 et 2, aux performances équivalentes pour des valeurs de n_d proches ($J = 0.151$ et $J = 0.156$), il est possible de noter qu'à un coefficient C_k plus faible est associé un coefficient C_z plus grand (test 1) et inversement (test 2) (Tab. 5.5). Les paramètres C_k et C_z ne sont donc pas complètement indépendants. Toutefois, la forme des hydrogrammes diffère légèrement en début de simulation ainsi que les écarts aux débits de pointe.

Sur la figure 5.11, on compare les hydrogrammes des différents tests, avant et après optimisation, à l'hydrogramme observé. L'hydrogramme de départ du test 1 est totalement surestimé, l'erreur par rapport au débit maximum observé est supérieure à 130%. Cependant, l'hydrogramme final est proche de l'hydrogramme observé, le coefficient de Nash ($1-J$) est supérieur à 0.85 pour cette crue. Les écarts des hydrogrammes initiaux des autres tests sont beaucoup moins importants.

Pour le test 1, le second pic de l'hydrogramme final, est moins bien estimé en temps que pour les autres tests (retard de 48 min par rapport à un retard de 24 min). En terme de prévision, il est important d'avoir une bonne synchronisation des pics de crue simulés et observés, même si le pic de crue simulé est supérieur en intensité au pic de crue observé. C'est pourquoi, même si le début de la simulation du test 1 est meilleur que celui des deux autres tests, les coefficients du test 1 ne seront pas retenus pour calibrer le modèle. Les résultats du test 2 seront alors sélectionnés pour la calibration car la plus petite valeur de C_k a été prise en compte.

Finalement, pour le jeu de paramètres choisi, l'hydrogramme simulé est satisfaisant au niveau de la détermination des pics de crue. La prise en compte du début de la simulation et celle de la décrue peuvent cependant être améliorés. La prise en compte des erreurs sur les pluies, sur les observations de débit et l'introduction d'un modèle de subsurface pourraient contribuer à améliorer les simulations du modèle.

Le tableau 5.5 récapitule les résultats des trois tests, le coefficient de ruissellement (CR) et les écarts en temps ($|\Delta T|$) et intensité ($|\Delta Q|$) des pics de crue simulés par rapport aux pics de crue observés sont également répertoriés.

Tableau 5.5. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux – événement n°1.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_k	1	2.21	5	8.9	10	13.4
C_z	1	6.7	5	3.72	10	3.6
n_d	0.2	0.21	0.1	0.175	0.05	0.171
J	3.27	0.151	0.225	0.156	0.52	0.155
CR (%)		25.1		30		30.6
Erreur premier pic						
$ \Delta T_1 $ (min)	12	0	48	12	84	0
$ \Delta Q_1 $ (%)	143.2	8.2	4	1.4	22.7	1.4
Erreur second pic						
$ \Delta T_2 $ (min)	24	48	24	24	60	24
$ \Delta Q_2 $ (%)	103.9	8.1	1.9	14.1	19.8	17.5

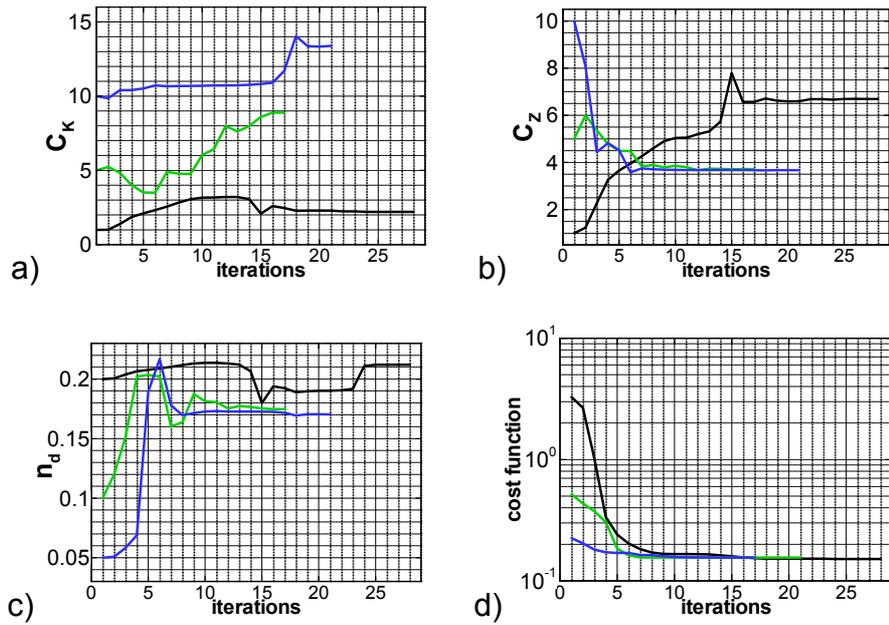


Fig. 5.10 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu – événement n°1.

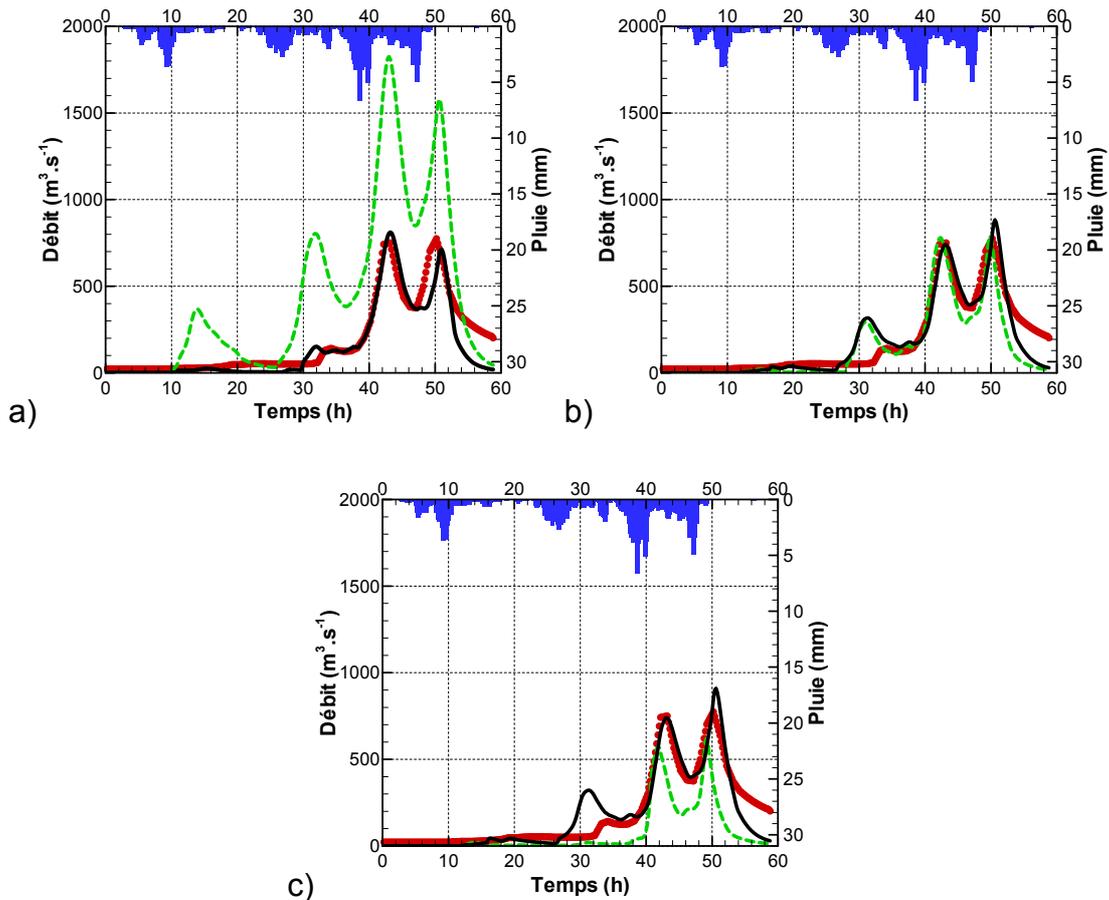


Fig. 5.11 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) – événement n°1 .

- Événement n°3 - crue du 03/10/1995

Il s'agit d'une crue de forte intensité ($Q_{\max} = 1608 \text{ m}^3/\text{s}$) à double pic (Fig. 5.12).

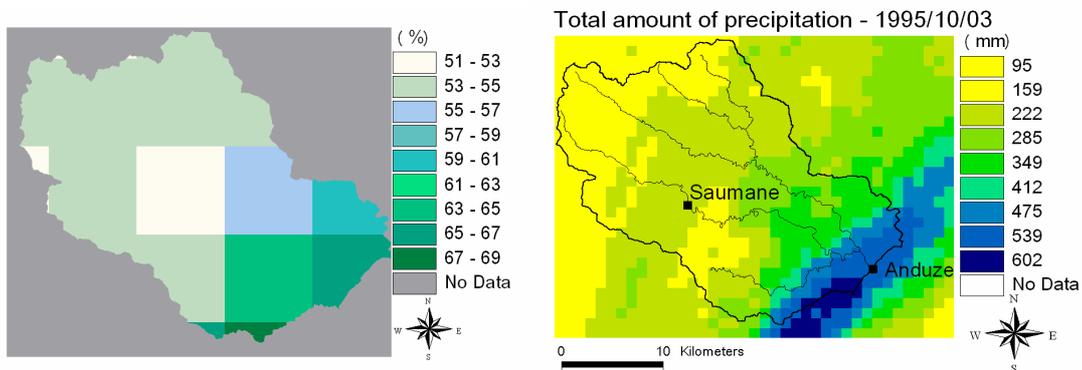


Fig. 5.12 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°3.

La Figure 5.13 représente l'évolution des paramètres C_k , C_z et n_d au cours de l'optimisation pour les différents tests d'initialisation.

Les trois tests d'initialisation convergent tous vers le même jeu de paramètres (C_k , C_z et n_d). L'algorithme semble alors converger vers un minimum global et ne se retrouve pas piégé au niveau des minimums locaux. L'algorithme d'optimisation conduit donc à une bonne identification des paramètres. Le test 2 est le plus rapide (13 itérations) car l'hydrogramme de départ est plus proche des observations.

L'hydrogramme final permet une estimation correcte des deux pics de crue avec un écart en intensité de 3% pour le premier et 5.5% pour le second avec cependant un léger retard d'environ 30 min (Tab. 5.6 et Fig 5.14)). Ces résultats sont satisfaisants compte tenu des erreurs potentielles d'observation. Comme précédemment, la simulation de la décrue peut être améliorée.

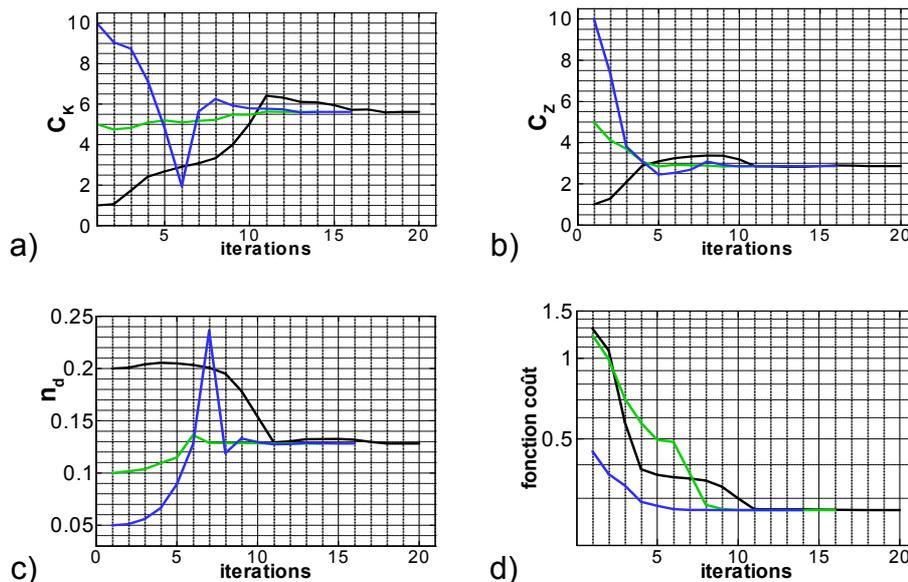


Fig. 5.13 : Convergence de C_k (a), C_z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°3.

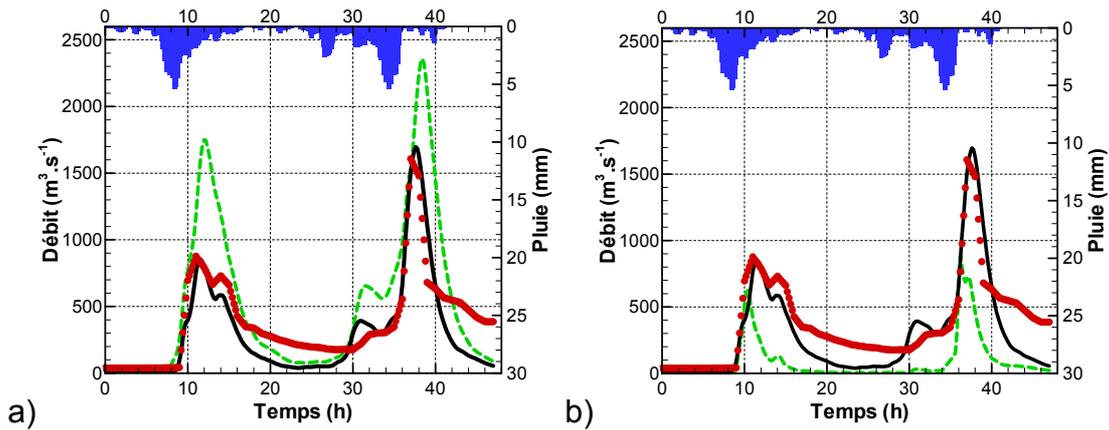


Fig. 5.14 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°3.

Tableau 5.6. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux – événement n°3.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	5.62	5	5.62	10	5.62
C_z	1	2.86	5	2.86	10	2.86
n_d	0.2	0.13	0.1	0.13	0.05	0.13
J	1.29	0.27	0.45	0.27	1.22	0.27
CR (%)		43.4		43.4		43.4
Erreur premier pic						
$ \Delta T_1 $ (min)	60	24	12	24	36	24
$ \Delta Q_1 $ (%)	99.7	3	3.9	3	28.8	3
Erreur second pic						
$ \Delta T_2 $ (min)	84	36	12	36	48	36
$ \Delta Q_2 $ (%)	46.6	5.5	16.8	5.5	45.2	5.5

- Événement n°4 - crue du 13/10/1995

La seconde crue d'octobre 1995 est une crue de forte intensité ($Q_{\max} = 1411 \text{ m}^3/\text{s}$).

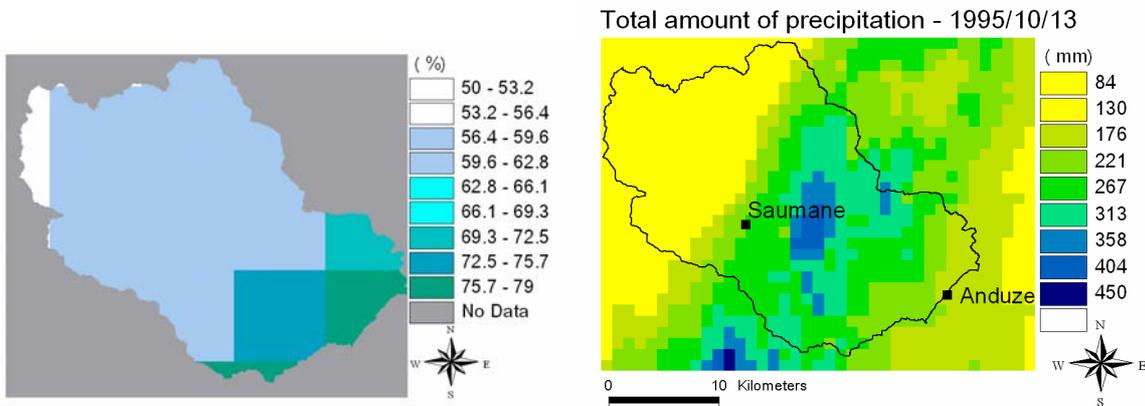


Fig. 5.15 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°4.

Les trois tests d'initialisation convergent vers deux jeux de paramètres différents, un premier jeu de paramètres pour le test 1 et un autre pour les tests 2 et 3 (Fig. 5.16). L'algorithme converge donc vers deux minimums locaux. La valeur de la fonction coût final du test 1 est plus faible ($J = 0.151$) que celle des tests 2 et 3 ($J = 0.167$). Cependant, les hydrogrammes après optimisation des tests 2 et 3 (Fig. 5.17) reproduisent plus fidèlement le pic de crue. L'erreur sur le débit de pointe est de 11% pour le test 1 contre 5% pour les tests 2 et 3. Le débit de pointe est mieux simulé notamment grâce à un coefficient de Manning plus faible pour les tests 2 et 3. Les résultats des tests 2 et 3 seront donc retenus pour la calibration. Le tableau 5.7 récapitule les résultats des trois tests d'initialisation.

Tableau 5.7. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux – événement n°4.

	Test 1		Test 2		Test 3	
C_K	1	1.93	5	11.15	10	11.15
C_z	1	3.75	5	3.05	10	3.05
n_d	0.2	0.38	0.1	0.124	0.05	0.124
J	1.17	0.151	0.31	0.167	0.724	0.167
CR (%)		43.7		39.3		39.3
Erreur pic de crue						
$ \Delta T $ (min)	48	12	0	24	48	24
$ \Delta Q $ (%)	58.5	10.9	24.8	4.7	48.4	4.7

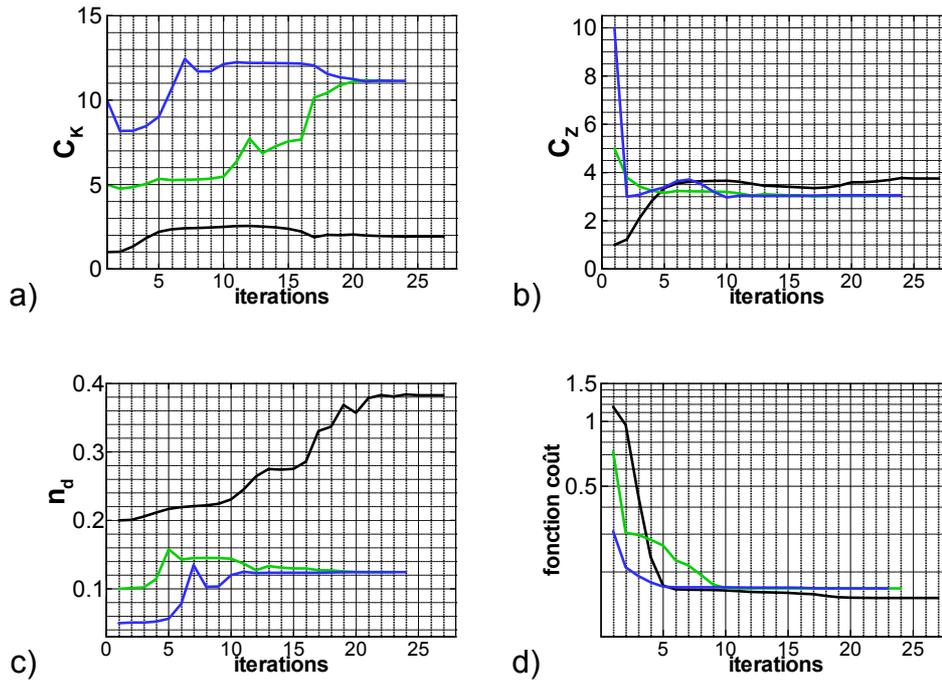


Fig. 5.16 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°4.

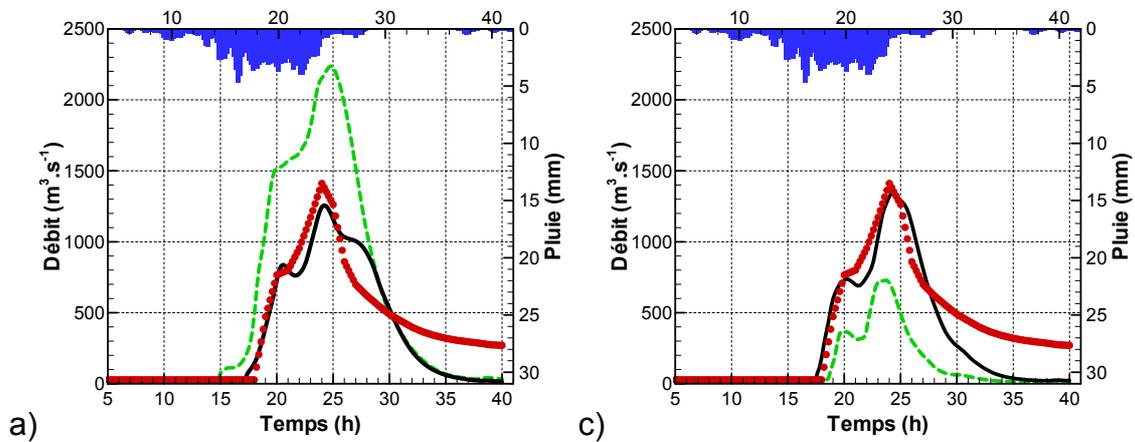


Fig. 5.17 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°4.

- Événement n°11 - crue du 17/05/1999

La crue de mai 1999 une crue de moyenne intensité ($Q_{\max} = 707 \text{ m}^3/\text{s}$).

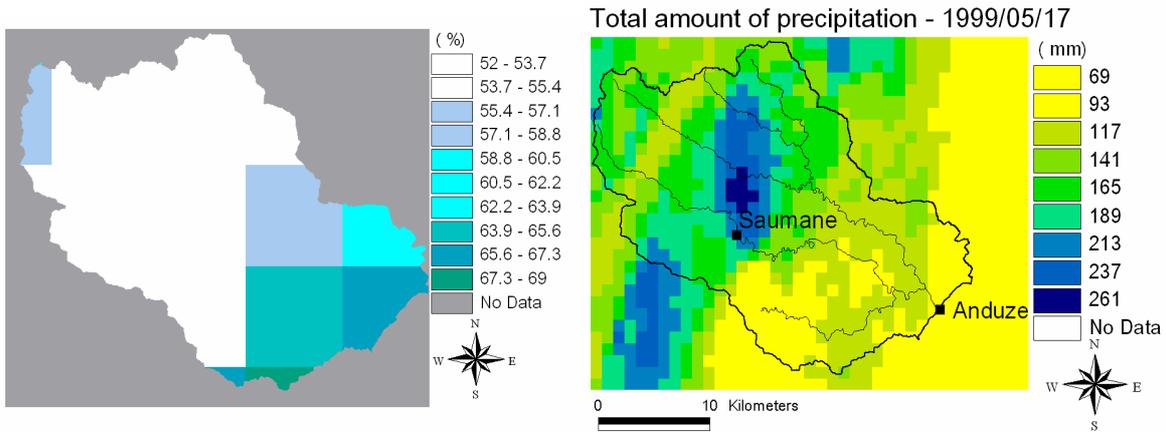


Fig. 5.18 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°11.

Contrairement à la majorité des événements sélectionnés, cette crue est crue s'est déroulé au printemps.

La simulation s'arrête au bout de 31h car au-delà les données de pluie sont erronées. L'événement a tout de même été retenu car la simulation ne s'arrête que bien après l'apparition du pic de crue.

Pour cette crue, le test 1 donne un premier jeu de paramètres et les tests 2 et 3 un autre jeu de paramètres avec une valeur identique de la fonction coût dans les 3 cas (Fig. 5.19 et Tab. 5.8). Le test 2 converge beaucoup plus vite que les deux autres tests, avec pratiquement deux fois moins d'itérations, car l'hydrogramme de départ est le plus proche des observations. Les valeurs finales du coefficient multiplicatif C_z et du coefficient de Manning n_d sont très proches pour les trois tests, seule la valeur du coefficient C_K varie de façon significative (Tab. 5.8). Les hydrogrammes finaux des trois tests, représentés sur la figure 5.19, sont également proches, l'intensité du pic de crue et la phase de montée sont bien reproduits. Bien que le test 1 aboutisse à une simulation du pic de crue avec 12 min de retard seulement (par rapport aux 2h des test 2 et 3), le jeu de paramètres des tests 2 et 3 ont été retenus car on peut penser qu'ils convergent tous les deux vers le minimum global.

Le début de l'événement, débit de base et premier pic, ainsi que la phase de décrue ne sont pas bien représentés par le modèle. Par rapport aux autres événements d'automne où les débits de base des crues sont très faibles, proches de zéro, cet événement de printemps a un débit de base non négligeable de $40 \text{ m}^3/\text{s}$. Les simulations du modèle n'intègrent pas cette information, le modèle étant événementiel par construction et en début d'événement il y a un déficit en volume. Des modifications sont en cours par remédier à cette faiblesse du modèle.

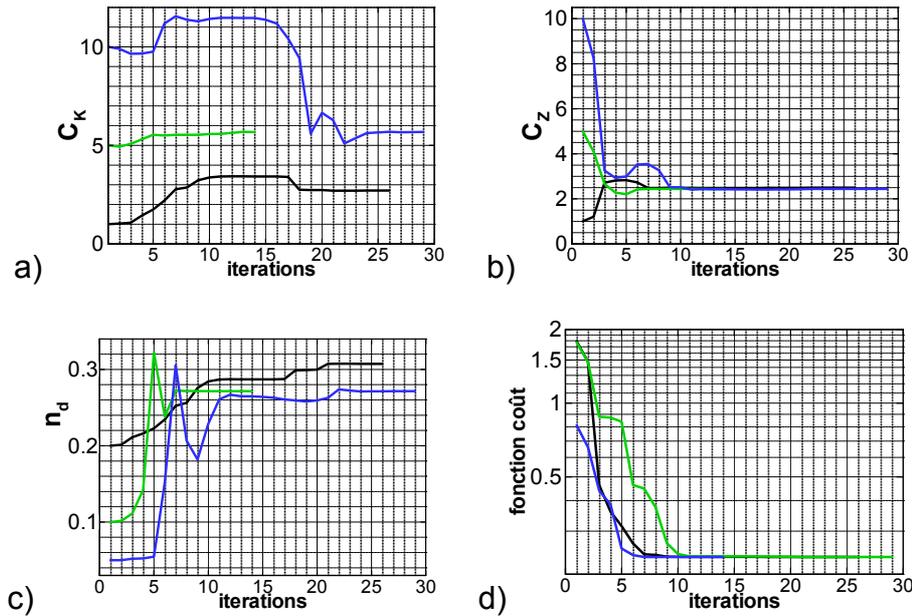


Fig. 5.19 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°11.

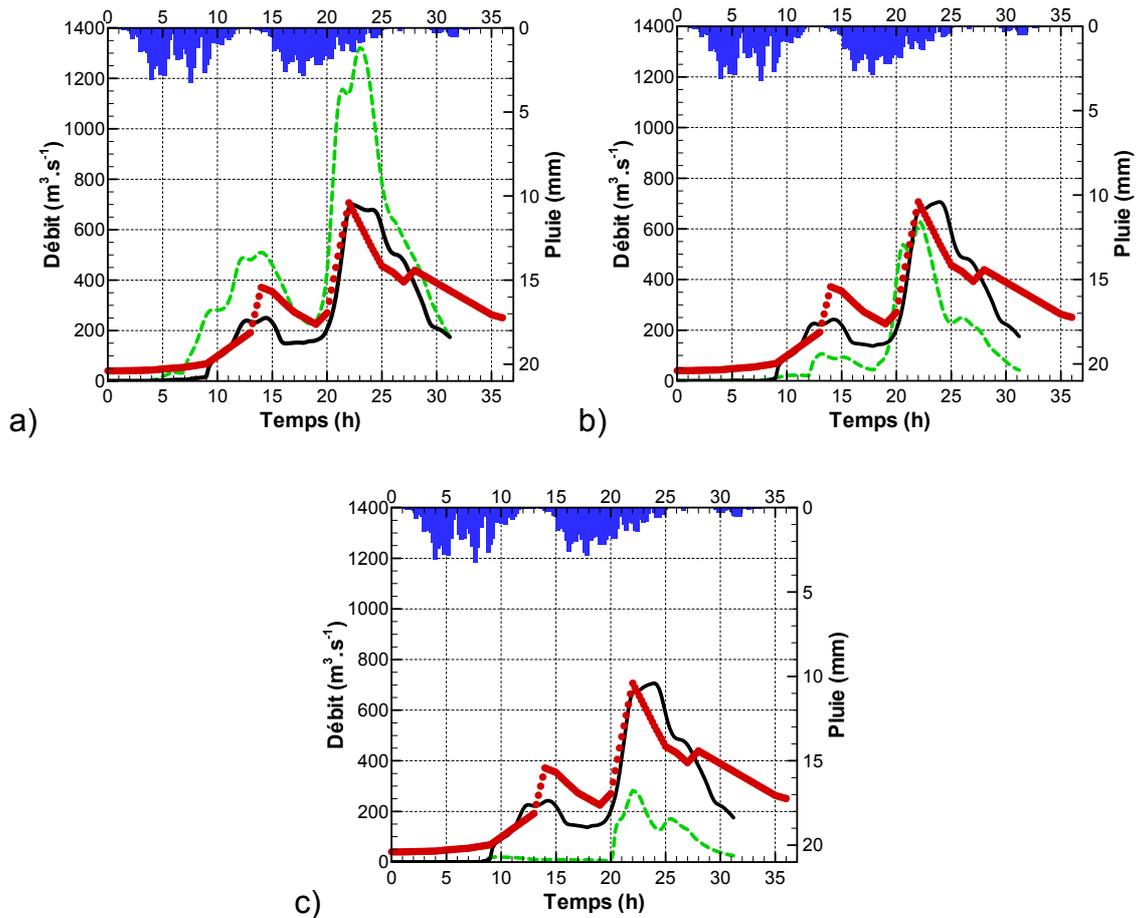


Fig. 5.20 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n° 11.

Tableau 5.8. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°11.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	2.70	5	5.67	10	5.67
C_Z	1	2.49	5	2.44	10	2.44
n_d	0.2	0.308	0.1	0.27	0.05	0.27
J	1.78	0.23	0.81	0.23	1.77	0.23
CR (%)		34.9		34.4		34.4
Erreur premier pic						
 \Delta T₁ (min)	0	24	48	12	264	12
 \Delta Q₁ (%)	37.3	32.6	71.2	34.8	94.6	34.8
Erreur second pic						
 \Delta T₂ (min)	60	12	12	120	0	120
 \Delta Q₂ (%)	87	0.9	11.3	0.01	60	0.01

- Événement n°13 - crue du 28/09/2000

La crue de septembre 2000 (Fig. 5.21) est une crue de forte intensité ($Q_{\max} = 1185 \text{ m}^3/\text{s}$).

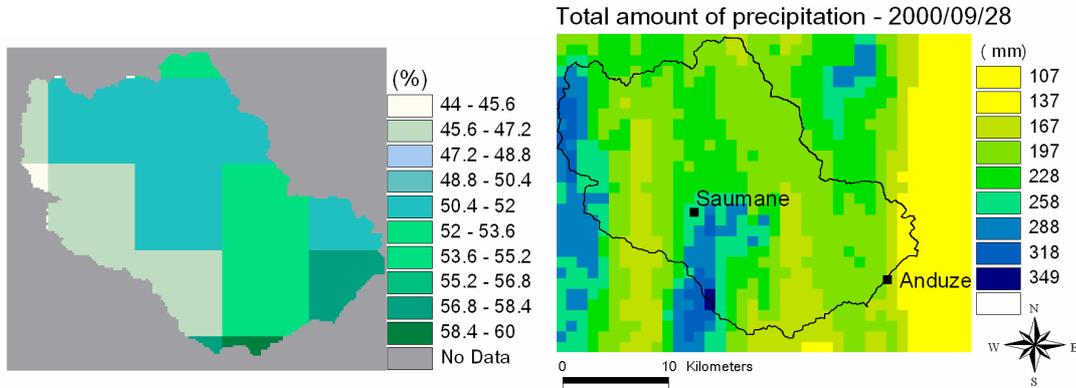


Fig. 5.21 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°13.

Les résultats des tests d'initialisation sont pratiquement identiques (Fig. 5.22). Le coefficient multiplicatif C_k est légèrement plus grand pour le test 3 ($C_k = 7.04$) que pour les tests 1 et 2 ($C_k = 5.88$). La valeur de la fonction coût est la même pour les trois tests, l'erreur sur le débit de pointe est légèrement supérieur pour le test 3 (Tab. 5.9). Les résultats des tests 1 et 2 sont sélectionnés pour cette crue.

Les résultats montrent l'efficacité de l'algorithme d'optimisation et l'hydrogramme obtenu après optimisation est très proche des observations pour cette crue (Fig. 5.22). Une fois encore, la phase de décrue pourrait être améliorée, notamment grâce à l'introduction des écoulements de subsurface.

Tableau 5.9. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°11.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_k	1	5.88	5	5.88	10	7.04
C_z	1	5.87	5	5.87	10	5.48
n_d	0.2	0.095	0.1	0.095	0.05	0.095
J	3.48	0.046	0.06	0.046	0.23	0.046
CR (%)		22.2		22.2		22.6
Erreur pic de crue						
$ \Delta T $ (min)	60	12	12	12	24	12
$ \Delta Q $ (%)	130	0.2	10.3	0.2	28.2	1.5

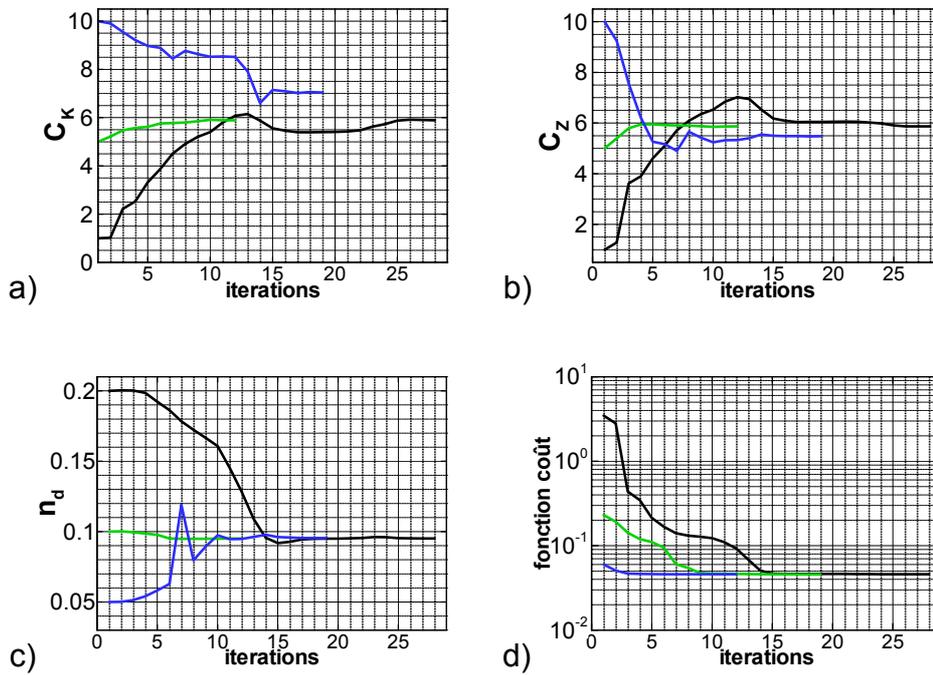


Fig. 5.22 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°11.

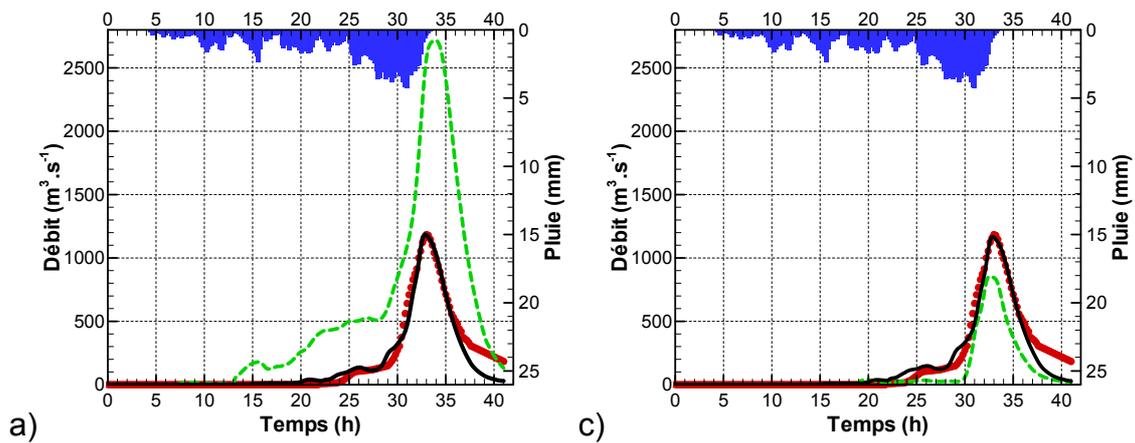


Fig. 5.23 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n°11.

Une analyse de l'évolution des paramètres a été menée durant la minimisation de la fonction coût.

Après 2 itérations, le coefficient de Manning reste à sa valeur initiale alors que la valeur du couple (C_K , C_Z) augmente (Fig 5.22). L'hydrogramme simulé après 2 itérations montre une diminution de l'intensité du pic de crue et du volume d'eau alors que la date du pic est la même (Fig. 5.24). Par conséquent, le couple (C_K , C_Z) réduit le volume d'eau disponible au ruissellement en forçant le taux d'infiltration.

Après 9 itérations, la valeur du couple (C_K , C_Z) continue d'augmenter pour se rapprocher de sa valeur finale estimée et n_d commence à diminuer (Fig 5.22). Le volume d'eau continue à décroître ; le pic de crue est sous-estimé mais l'écart temporel se réduit. De 9 à 14 itérations, la variation la plus importante correspond au paramètre n_d , elle contribue essentiellement à ajuster la position du pic de crue (en date et en intensité).

Finalement, l'hydrogramme simulé après 14 itérations est pratiquement confondu avec l'hydrogramme final (après optimisation).

Par conséquent, la méthodologie employée est performante et l'évolution des paramètres durant la minimisation de la fonction coût apporte des informations concernant leurs contributions respectives dans le processus de modélisation :

- le coefficient de Manning détermine la position du pic de crue. Candela et al. (2005) ont trouvé des résultats semblables et affirment que la meilleure reproduction de la position du pic n'est obtenue que pour une valeur particulière du coefficient de Manning. Ce paramètre contrôle donc le transfert des eaux sur le bassin (c'est le facteur le plus influent de la fonction de transfert comme l'a démontré l'analyse de sensibilité), sa valeur dépend de la crue étudiée.
- le couple (C_K , C_Z) joue sur la fonction de production en fixant le taux d'infiltration et la capacité de rétention d'eau du sol, il affecte donc le volume d'eau disponible pour la crue.

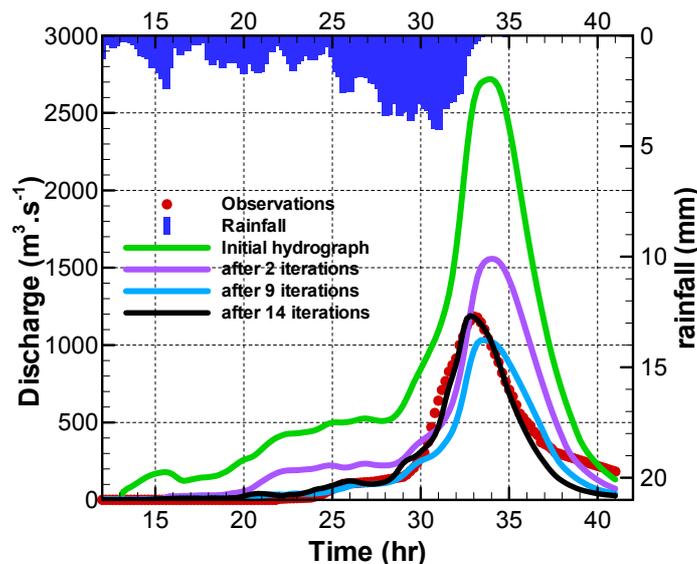


Fig. 5.24 : Hydrographes avant optimisation (trait vert), après 2 itérations (trait mauve), après 9 itérations (trait bleu) et après 14 itérations (trait noir) comparés aux observations (points rouges) pour le test 1- événement n°13.

- Événement n° 15 - crue du 14/03/2002

La crue de mai 2002 (Fig. 5.25) est une crue de moyenne intensité ($Q_{\max} = 666 \text{ m}^3/\text{s}$).

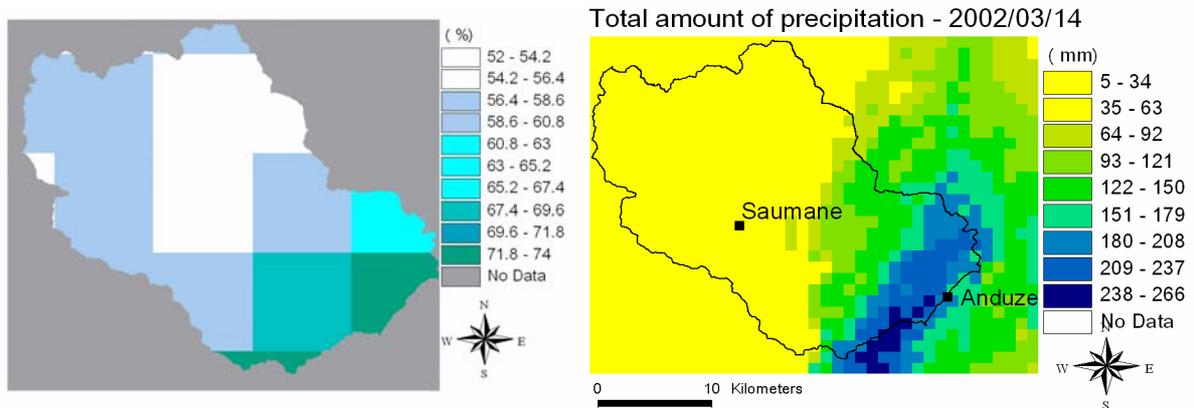


Fig. 5.25 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°15.

Il s'agit d'une crue de printemps. La simulation s'arrête au bout de 17h car au-delà les données de pluie sont erronées. L'événement a tout de même été retenu car la simulation ne s'arrête qu'après l'apparition du pic de crue.

Les tests d'initialisation aboutissent à deux jeux de paramètres distincts, un premier jeu de paramètres pour les tests 1 et 2, un second pour le test 3 (Fig. 5.26). Le premier jeu de paramètres affiche de meilleures performances, le critère de Nash est supérieur (0.855 par rapport à 0.72) et l'erreur sur le débit de pointe est inférieure (21% au lieu de 25%) (Tab. 5.10). Les résultats des tests 1 et 2 seront alors choisis pour la calibration.

Le coefficient de Manning est le même pour les trois tests, il semble donc indépendant du couple de paramètres (C_K , C_Z). Comme évoqué précédemment, ce paramètre contrôle essentiellement le transfert des eaux sur le bassin versant et détermine l'instant du pic de crue.

Cette crue n'est pas bien simulée par le modèle bien que le critère de Nash soit supérieur à 0.85 pour le jeu de paramètres sélectionné (Fig. 5.27). La montée de crue est trop rapide, l'hydrogramme final du test 3 représente même deux pics. Le débit de pointe n'est pas estimé correctement avec une erreur de plus de 20%.

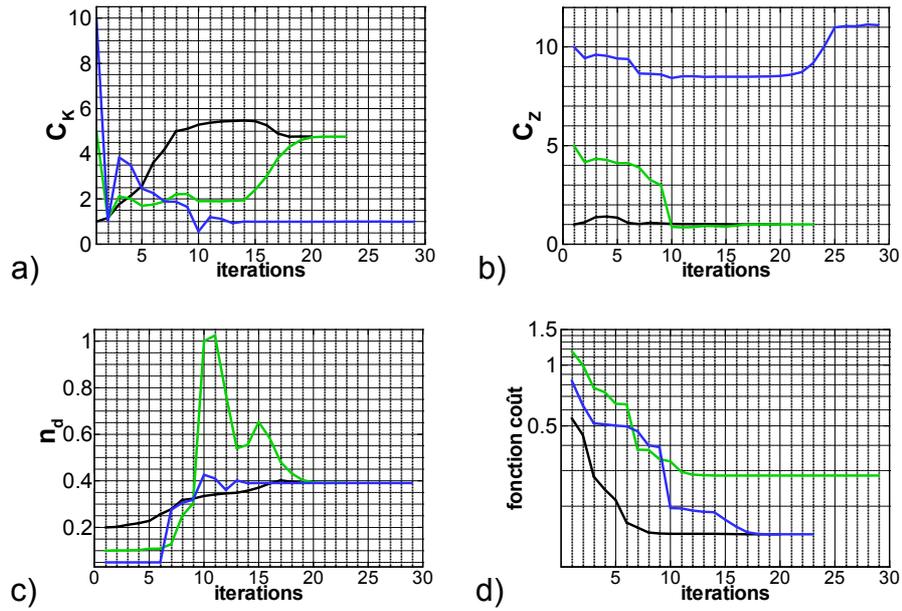


Fig. 5.26 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°15.

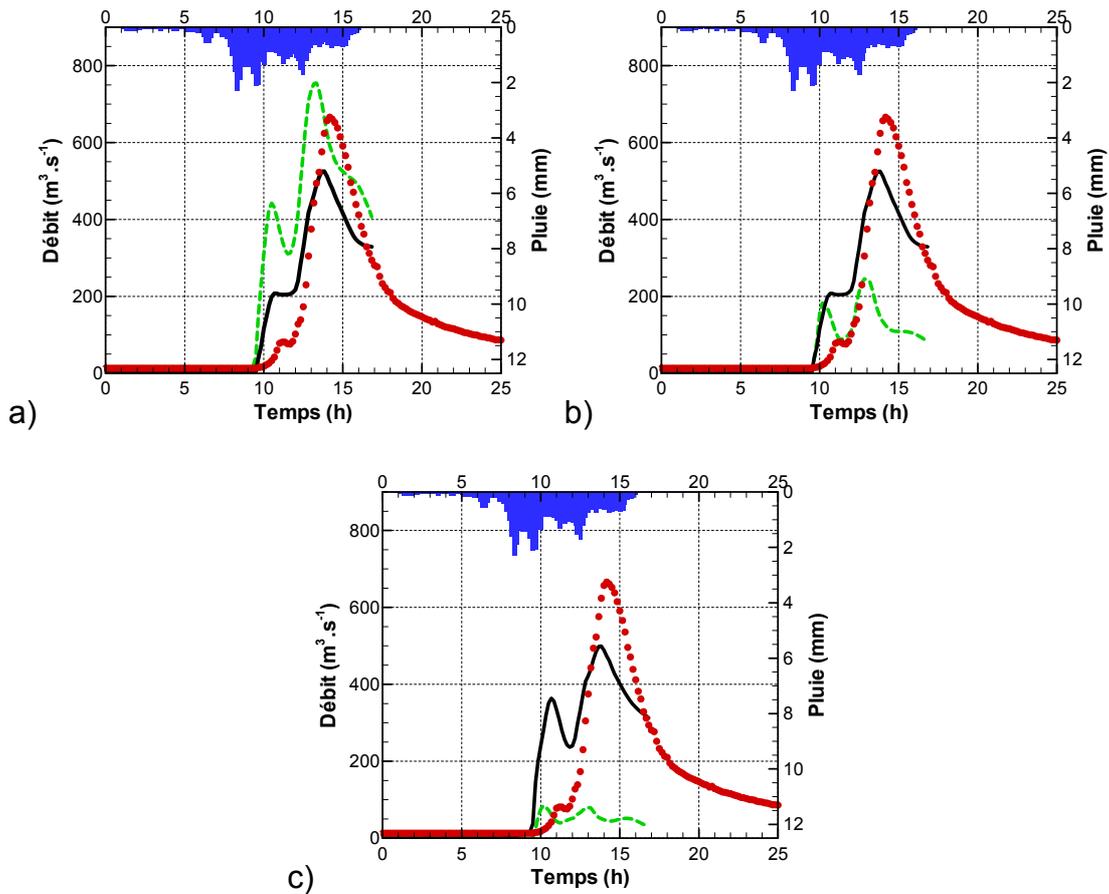


Fig. 5.27 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n° 5.

Tableau 5.10. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°15.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	4.76	5	4.76	10	0.99
C_z	1	1.01	5	1.01	10	11.1
n_d	0.2	0.39	0.1	0.39	0.05	0.39
J	0.545	0.145	0.84	0.145	1.17	0.28
CR (%)		27.9		27.9		30.5
Erreur pic de crue						
$ \Delta T $ (min)	50.4	20	80.4	20	240	20
$ \Delta Q $ (%)	13.4	21.1	63.1	21.1	87.3	25

- Événement n°19 - crue du 08/09/2002

La crue de septembre 2002 (Fig. 5.28) est une crue exceptionnelle de période de retour supérieure à cent ans ($Q_{max} = 3634 \text{ m}^3/\text{s}$). L'étude de ce genre d'événement contribue à l'évaluation de la performance du modèle dans des situations extrêmes.

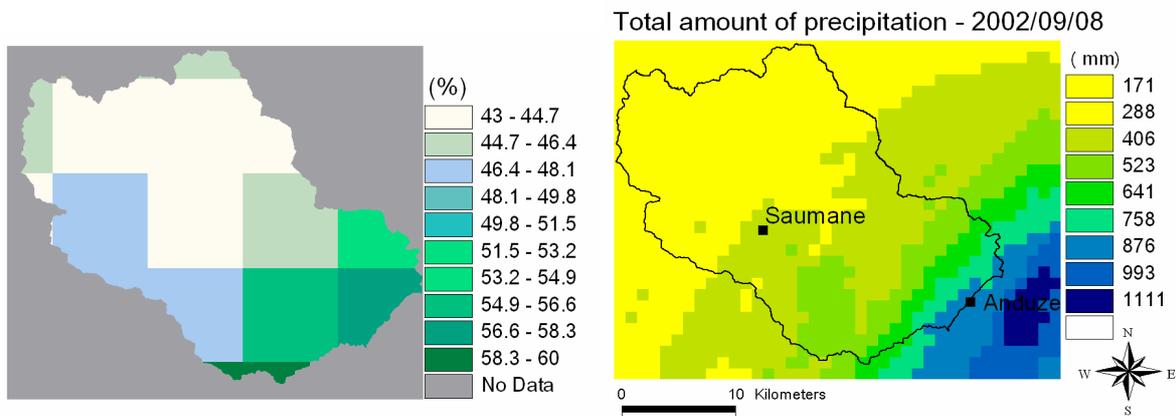


Fig. 5.28 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°19.

Les résultats des trois tests d'initialisation sont identiques. Les tests 2 et 3 convergent en moins de 15 itérations, le test 1 est beaucoup plus long et s'arrête au bout de 30 itérations.

L'hydrogramme final sous-estime les deux pics, néanmoins les montées de crue sont très bien représentées. Pour ce genre d'événement exceptionnel, les erreurs d'observation sont très probables surtout en ce qui concerne l'estimation des débits de pointe. Les simulations sont alors satisfaisantes pour cette crue, bien que la décrue puisse encore être améliorée.

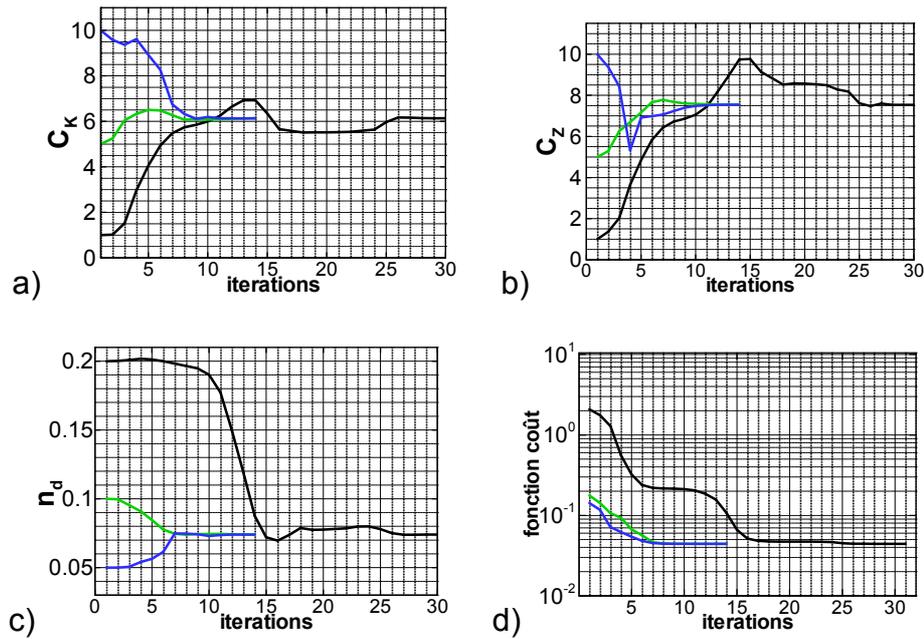


Fig. 5.29 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°19.

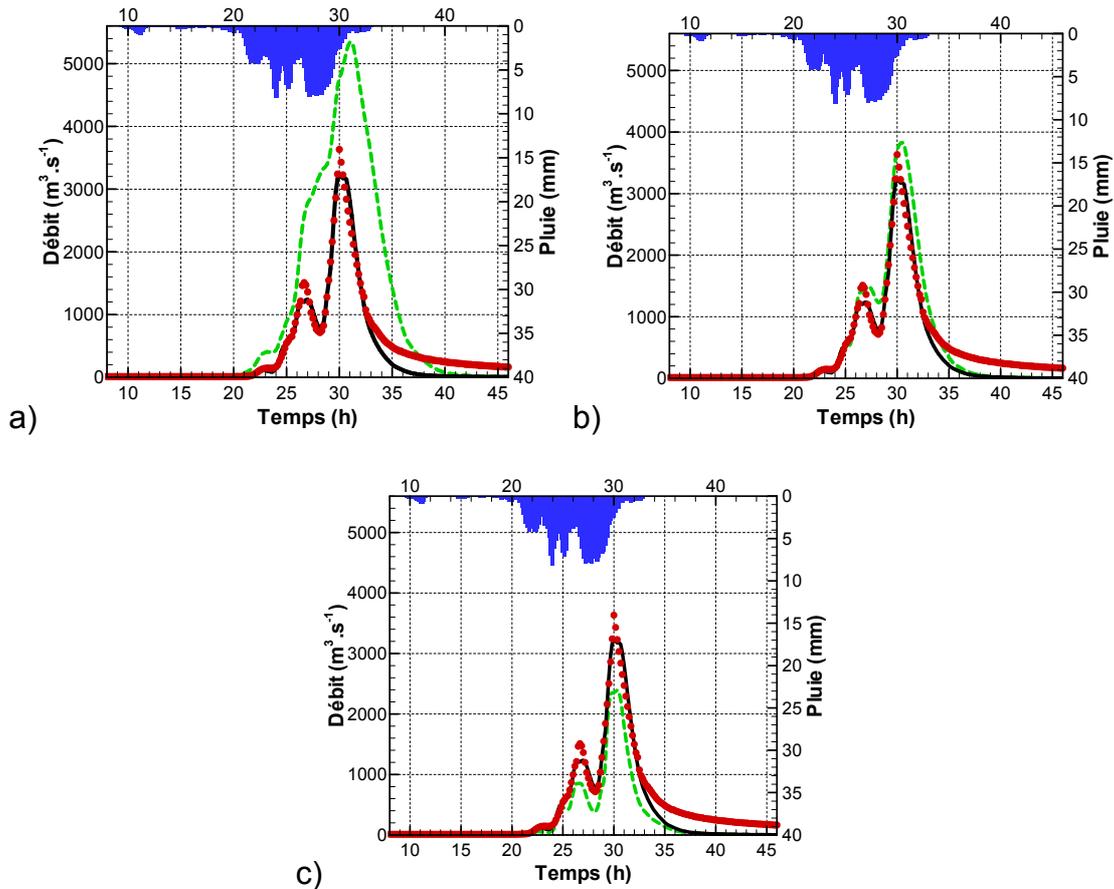


Fig. 5.30 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a), le test 2 (b) et le test 3 (c) - événement n° 9.

Tableau 5.11. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux- événement n°19.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	6.13	5	6.13	10	6.13
C_z	1	7.54	5	7.54	10	7.54
n_d	0.2	0.074	0.1	0.074	0.05	0.074
J	2.07	0.044	0.14	0.044	0.17	0.044
CR (%)		32.3		32.3		32.3
Erreur premier pic						
$ \Delta T_1 $ (min)	Pas de pic	20	39.6	20	0	20
$ \Delta Q_1 $ (%)		18.2	1.9	18.2	43.4	18.2
Erreur second pic						
$ \Delta T_2 $ (min)	70.2	0	30	0	19.8	0
$ \Delta Q_2 $ (%)	47	11.2	5.4	11.2	34.1	11.2

- Événement n°25 - crue du 18/10/2006

La crue d'octobre 2006 (Fig. 5.30) est une crue de forte intensité ($Q_{max} = 1441 \text{ m}^3/\text{s}$).

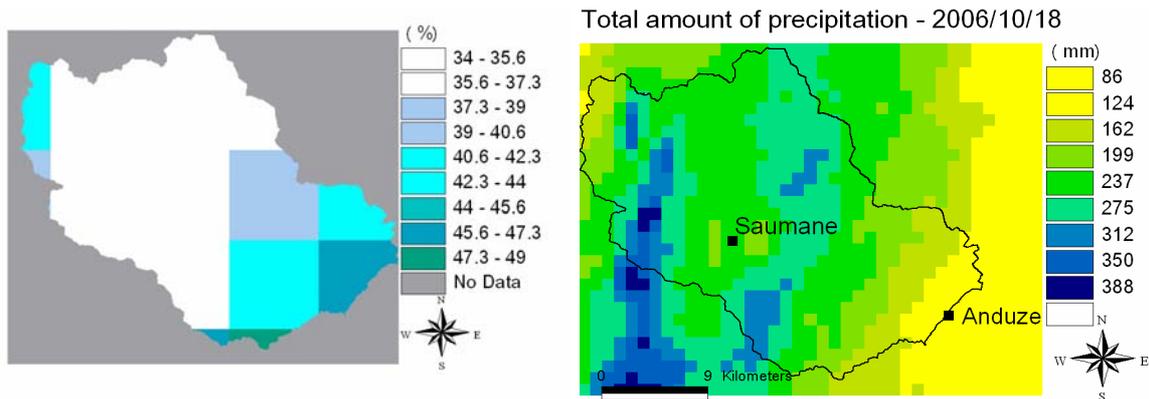


Fig. 5.31 : Humidité initiale relative (%) et total des pluies cumulées (mm) – événement n°25.

Le tableau 5.12 récapitule les résultats des trois tests. Ils convergent tous vers le même jeu de paramètres, vers le minimum global. Il est possible de noter la valeur élevée du coefficient C_K (Fig. 5.32). Le jeu de paramètres trouvé permet une très bonne simulation de l'hydrogramme observé (Fig. 5.33).

Tableau 5.12. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux - événement n°25.

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	20	5	20	10	20
C_z	1	3.1	5	3.1	10	3.1
n_d	0.2	0.099	0.1	0.099	0.05	0.099
J	1.63	0.103	0.81	0.103		0.103
CR (%)	70.8	34.8	43.7	34.8		34.8
Erreur pic de crue						
$ \Delta T $ (min)	105	15		15	6	15
$ \Delta Q $ (%)	43.6	5.5		5.5	43.8	5.5

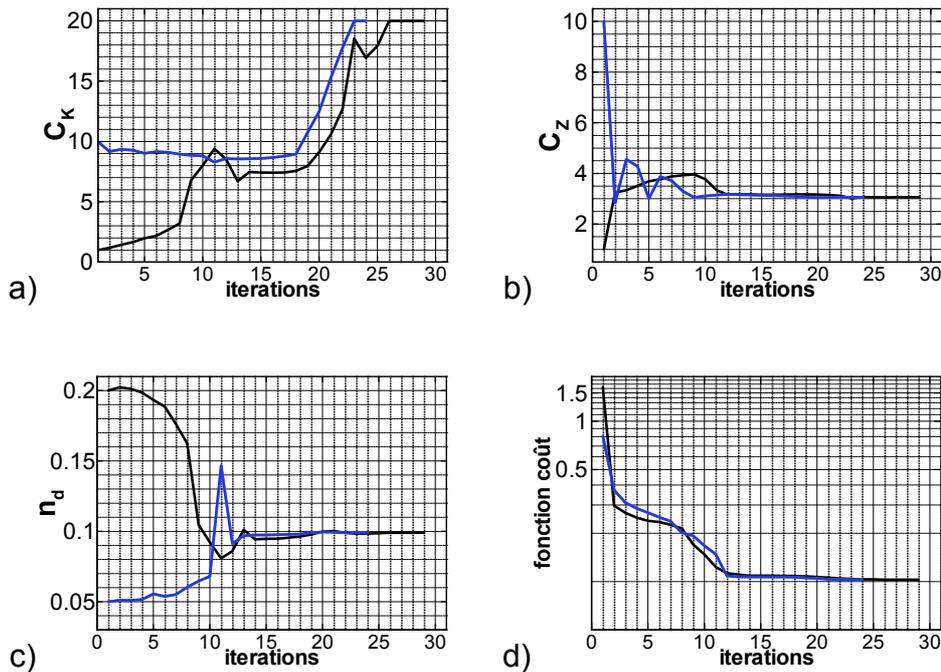


Fig. 5.32 : Convergence de C_K (a), C_z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu - événement n°25.

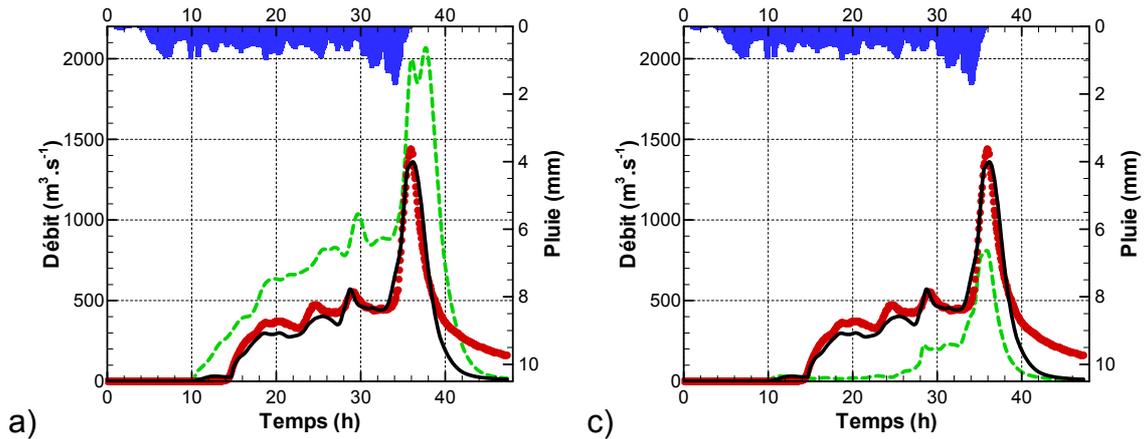


Fig. 5.33 : Hydrogrammes avant optimisation (en trait vert discontinu) et après optimisation (en trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°25.

5.2.3.3 Synthèse des résultats de l'optimisation

Pour certains événements (n°3, n°19 et n°25), l'algorithme converge toujours vers le même jeu de paramètres ; tandis que pour d'autres, le jeu de paramètres estimé est différent en fonction des valeurs initiales. Cependant, à l'exception de l'événement n°1, au moins deux des trois tests donnent les mêmes résultats. Le jeu sélectionné a donc été celui qui correspondait au résultat commun des deux tests. On peut alors supposer que l'algorithme a convergé vers le minimum global. En fait, pour pouvoir l'affirmer il aurait fallu tester davantage de conditions initiales d'optimisation, cependant cela aurait nécessité beaucoup plus de temps. Partant de cette hypothèse, on peut dire que pour que l'algorithme de minimisation converge vers le minimum global, il est important (crue du 14/03/2002, test 3) que l'hydrogramme initial ne soit pas trop éloigné des observations, ou encore qu'il est nécessaire d'avoir une bonne estimation a priori des paramètres.

Pour chaque crue, le test 2 avec comme jeu de paramètres initial $(C_K, C_Z, n_d) = (5, 5, 0.1)$ converge le plus rapidement vers le jeu de paramètres sélectionné car l'hydrogramme de départ est à chaque fois le plus proche des observations.

Le tableau 5.13 résume pour chaque crue les valeurs des paramètres C_K , C_Z et n_d sélectionnées. Ces valeurs correspondent respectivement au jeu de paramètres optimal pour chaque événement.

Il ne semble pas y avoir de répartition particulière pour le couple (C_K, C_Z) en fonction de l'intensité de la crue ou de la saison à laquelle a lieu l'événement. De toute façon, le nombre d'événements est insuffisant pour conduire à de telles conclusions.

Tableau 5.13. Tableau récapitulatif des paramètres sélectionnés

N° événement	Date	Intensité (m ³ /s)	C _k	C _z	n _d	Nash
n°1	21/09/1994	775	8.9	3.72	0.17	0.84
n°3	03/10/1995	1608	5.62	2.86	0.13	0.73
n°4	13/10/1995	1411	11.15	3.05	0.12	0.83
n°11	17/05/1999	707	5.67	2.44	0.27	0.77
n°13	28/09/2000	1185	5.88	5.87	0.095	0.95
n°15	14/03/2002	666	4.76	1.0	0.39	0.85
n°19	08/09/2002	3634	6.13	7.54	0.074	0.96
n°25	18/10/2006	1441	20	2.67	0.10	0.93

- Le couple (C_k, C_z)

Il joue principalement sur la fonction de production et détermine le volume d'eau qui s'infiltre au cours de l'événement.

- Le coefficient de Manning n_d

Il permet d'ajuster le temps et l'intensité du pic de crue. Pour chaque événement, le coefficient de Manning du lit majeur des drains n_d converge pratiquement vers la même valeur pour les trois tests d'initialisation. Ce paramètre semble donc indépendant du couple (C_k, C_z), il dépend de l'intensité de la crue, plus la crue est forte, plus le coefficient de Manning est faible. En effet, une crue de forte intensité implique une hauteur d'eau importante et par conséquent moins de frottements par unité de section. En fait, plus le débit est fort, plus les frottements sur le fond sont importants, mais plus le rayon hydraulique est important dans le même temps, ce qui entraîne une réduction du coefficient de Manning.

En fonction des valeurs estimées pour ce coefficient, il est possible de déterminer une loi du coefficient de Strickler (inverse du coefficient de Manning) en fonction de l'intensité du pic de crue (Fig. 5.34). Le tableau 5.14 dresse une comparaison entre les coefficients de Strickler estimé et ceux obtenus grâce à l'extrapolation logarithmique.

On peut supposer que dans l'ajustement du coefficient de Manning du lit majeur des drains, il y a l'ajustement des pentes des versants des drains et d'autres paramètres relatifs à la représentation du réseau de drainage.

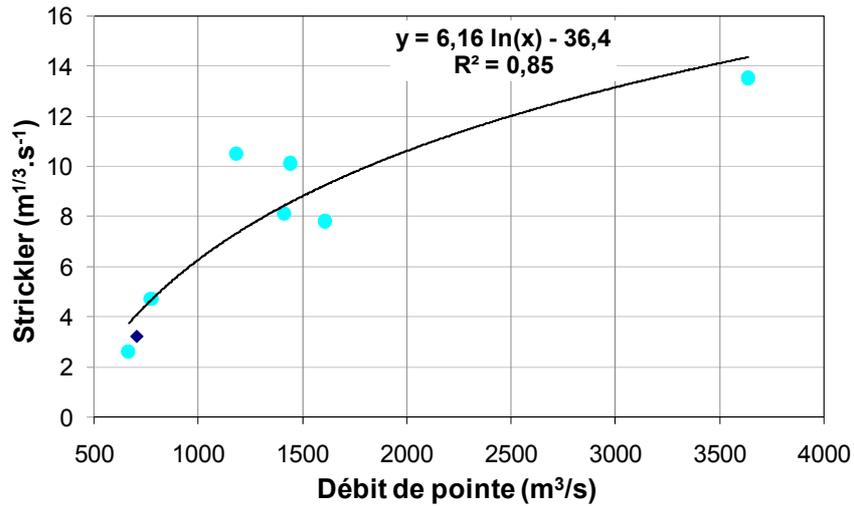


Fig. 5.34 : Loi du coefficient de Strickler en fonction du débit de pointe.

Tableau 5.14. Comparaison des coefficients de Strickler ($1/n_d$) estimé et obtenu par extrapolation logarithmique

	Intensité (m ³ /s)	Strickler estimé $1/n_d$ (m ^{1/3} .s ⁻¹)	Strickler extrapolé $1/n_d$ (m ^{1/3} .s ⁻¹)
14/03/2002	666	2.6	3.7
17/05/1999	707	3.7	4.0
21/09/1994	775	5.7	4.6
28/09/2000	1185	10.5	7.2
13/10/1995	1411	8.1	8.3
18/10/2006	1441	10.1	8.4
03/10/1995	1608	7.8	9.1
08/09/2002	3634	13.5	14.1

On pourrait désormais implémenter une loi de variation $n(Q)$ dans Marine du type : $1/n = 6,16 \ln(Q/Q_0)$ avec $Q_0 = 370 \text{ m}^3/\text{s}$.

5.2.4 Synthèse

La méthode d'optimisation est performante et apporte plusieurs contributions concernant :

- la validation des hypothèses physiques utilisées,
- la calibration du modèle,
- l'évaluation de ses performances,
- les perspectives d'amélioration de sa structure.

Les résultats de l'algorithme d'optimisation et des marges d'incertitude issues des meilleures simulations de Monte-Carlo démontrent une bonne performance du modèle pour la représentation des crues à cinétique rapide. La forme des hydrogrammes et les pics de crue sont simulés, cependant, la phase de décrue peut être améliorée. Par exemple, la mise en mouvement des écoulements de subsurface devrait contribuer à augmenter le volume d'eau à la fin de l'événement et légèrement au début par exfiltration dans les drains.

En comparant les résultats avec les simulations de Monte-Carlo, deux constats peuvent être faits :

- premièrement, les jeux de paramètres obtenus pour chaque crue sont en correspondance avec les différents intervalles de variation sélectionnés à l'issue des simulations de Monte-Carlo.
- deuxièmement, tous les coefficients C_K estimés entrent dans la plage de variation commune à toutes les crues : $C_z \in [3,10]$ (sauf la valeur 20 obtenue pour la crue de Novembre 2006 mais cette valeur élevée est due à l'insensibilité du coefficient au-delà d'un certain seuil). En revanche, cela n'est pas le cas pour le coefficient C_z ($C_z \in [3,4.5]$) qui est tantôt trop faible (événement n°11 et 15), tantôt trop élevé (événement n°13 et 19). Il faut alors réaliser un compromis entre toutes les simulations et adopter une sorte de valeur moyenne pour ces coefficients car les résultats obtenus individuellement pour chaque crue engendrent des marges d'incertitudes trop grandes.

Etant donné les erreurs possibles dans la mesure des débits, un critère plus souple tolérant une erreur relative de 20 % sur les débits, a été testé.

5.3 Prise en compte des incertitudes sur le débit

Une incertitude de 20 % sur le débit a été prise en compte dans cette partie par l'intermédiaire d'une seconde valeur de vraisemblance qu'on appellera le critère de Nash généralisé. Elle est définie comme suit :

$$\left\{ \begin{array}{l} \text{si } Q_{sim}^i < 0.8 Q_{obs}^i \text{ alors } NashG = 1 - \frac{\sum_{i=1}^N (0.8 Q_{obs}^i - Q_{sim}^i)^2}{\sum_{i=1}^N (Q_{obs}^i - \overline{Q_{obs}})^2} \\ \text{si } Q_{sim}^i > 1.2 Q_{obs}^i \text{ alors } NashG = 1 - \frac{\sum_{i=1}^N (1.2 Q_{obs}^i - Q_{sim}^i)^2}{\sum_{i=1}^N (Q_{obs}^i - \overline{Q_{obs}})^2} \end{array} \right. \quad (5.2)$$

où Q_{obs}^i et Q_{sim}^i sont respectivement le débit observé et simulé au pas de temps i , N le nombre de pas de temps, $\overline{Q_{obs}}$ est la moyenne des débits observés

$NashG$ a une valeur nulle si les débits simulés correspondent à ± 20 % des débits observés

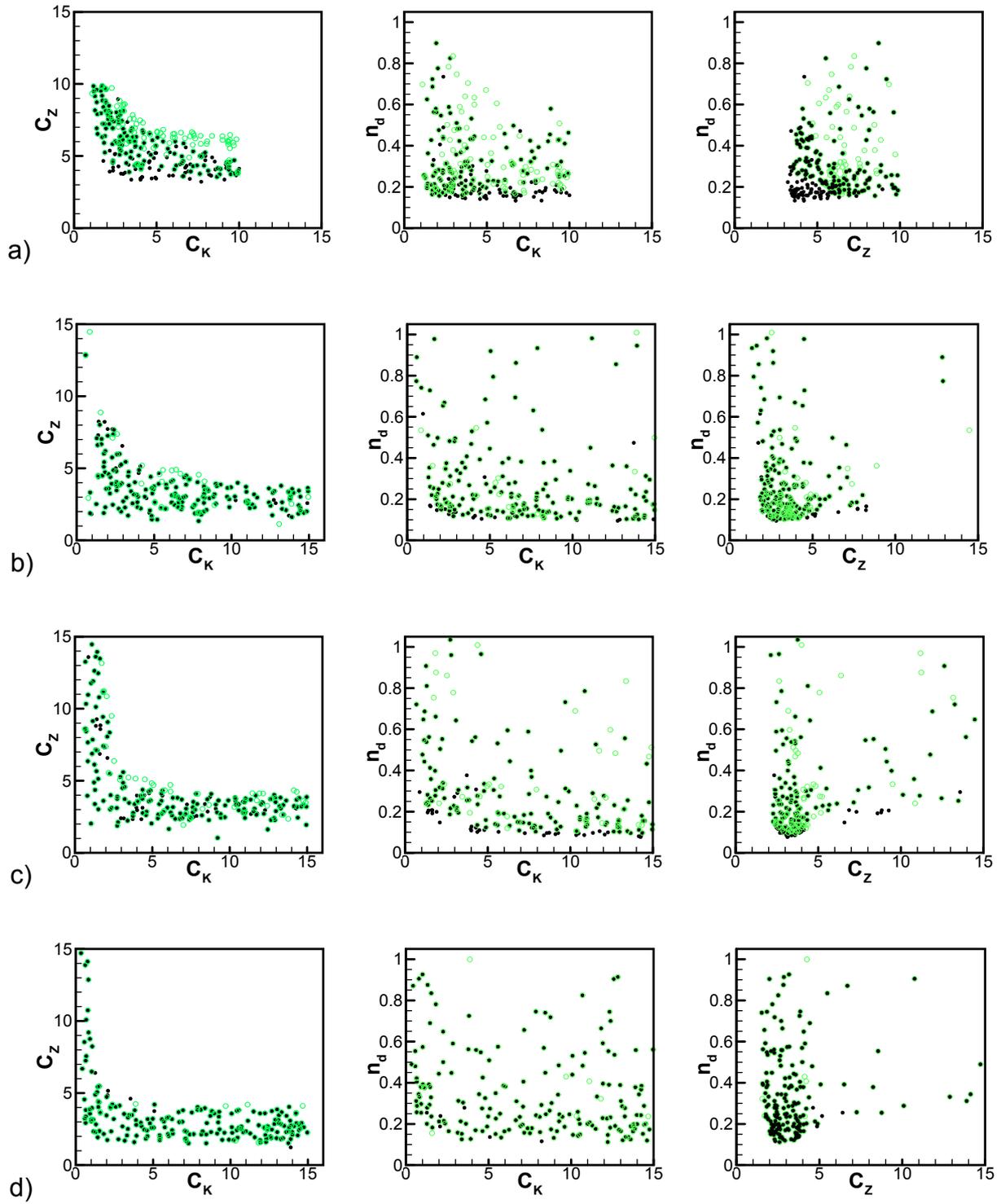
5.3.1 Résultats des simulations de Monte-Carlo

La figure 5.35 permet une comparaison des couples de paramètres (C_K, C_Z) , (C_Z, n_d) et (C_K, n_d) correspondant aux 200 meilleures simulations obtenues pour le critère de Nash et le Nash généralisé.

Il est possible de noter la très large répartition du paramètre C_K sur la majorité des événements, celle un peu plus restreinte du paramètre C_Z (événement n°13, 15 et 25) et finalement la localisation assez précise du paramètre n_d (événement n°2, 13 et 25 par exemple). Les graphiques correspondants à l'événement n°25 illustrent bien le contraste entre la distribution du coefficient C_K et la localisation de n_d .

La forme en demi arc de cercle du couple (C_K, C_Z) (événement n°3, 4, 11 et 15) met en avant les interactions qu'il y a lieu entre ces deux paramètres.

La distribution de ces couples de paramètres est proche (voire quasiment identique) pour les deux critères. Il semblerait que les résultats obtenus avec ce nouveau critère ne diffèrent pas beaucoup de ceux de l'ancien.



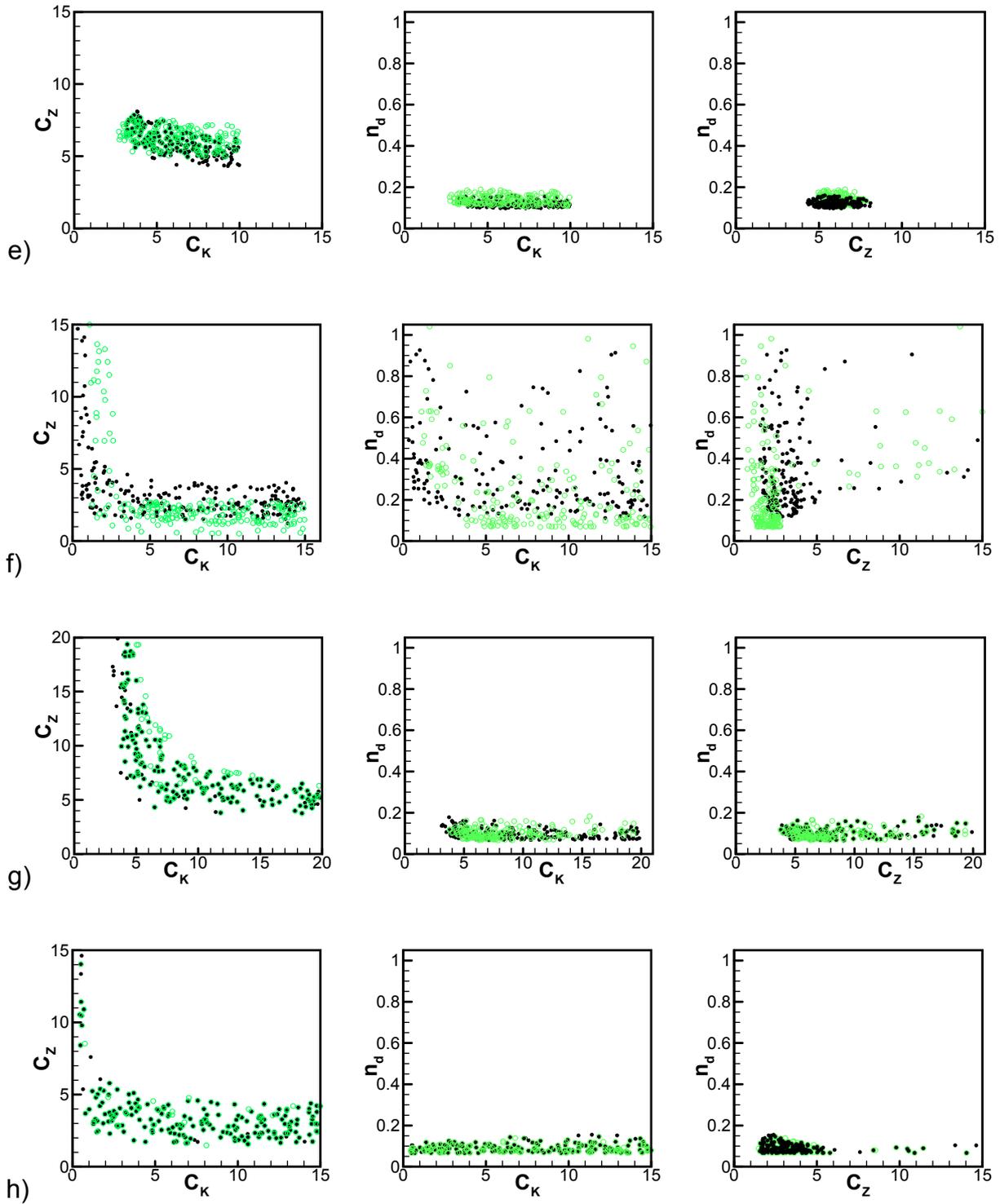


Fig. 5.35 : Tracé des couples (C_K, C_Z) , (C_K, n_d) et (C_Z, n_d) correspondant aux 200 meilleures simulations de Monte-Carlo pour le critère de Nash (points noirs) et le Nash généralisé (points verts) des événements n^1 (a), n^3 (b), n^4 (c), n^11 (d), n^13 (e), n^15 (f), n^19 (g) et n^25 (h).

La figure 5.36 montre une comparaison des simulations de Monte-Carlo pour les deux critères (Nash et Nash généralisé) pour les événements n°2 et n°19.

Les valeurs de vraisemblance obtenues avec le Nash généralisé sont supérieures à celles du critère de Nash classique. Ce résultat n'est pas étonnant dans le sens où le Nash généralisé est un critère plus souple intégrant un plus grand nombre de simulations « parfaites » (toutes celles comprises dans une marge d'incertitude de 20 % par rapport aux observations). Ce constat est généralisé pour toutes les autres crues.

La valeur moyenne de $NashG$ obtenue pour toutes les crues est de 0.96. La plus faible valeur est obtenue pour l'événement n°3 pour lequel $NashG = 0.89$ et la plus forte pour l'événement n°13 avec $NashG = 0.983$.

Cependant, afin de pouvoir comparer les résultats avec ceux obtenus pour le critère de Nash, la même valeur de seuil a été utilisée. Les intervalles des paramètres seront alors choisis de sorte que $NashG > 0.7$.

Concernant la possible identification des paramètres, des remarques similaires à celles établies pour le critère de Nash classique peuvent être établies.

Le paramètre C_K n'affiche pas une très forte sensibilité, il est réparti sur tout l'intervalle de variation (notamment pour les événements n°3, 4, 11 et 25); l'événement le plus contraignant pour ce paramètre est la crue exceptionnelle de Septembre 2002 (événement n°19).

Le paramètre C_Z est un peu plus contraint, il ne doit pas être trop faible (par exemple $C_Z > 2.5$ pour l'événement n°13) et présente des valeurs optimales pour les événements n°11 ($C_Z = 2.2$), n°15 ($C_Z = 1.3$) et n°25 ($C_Z = 2.2$).

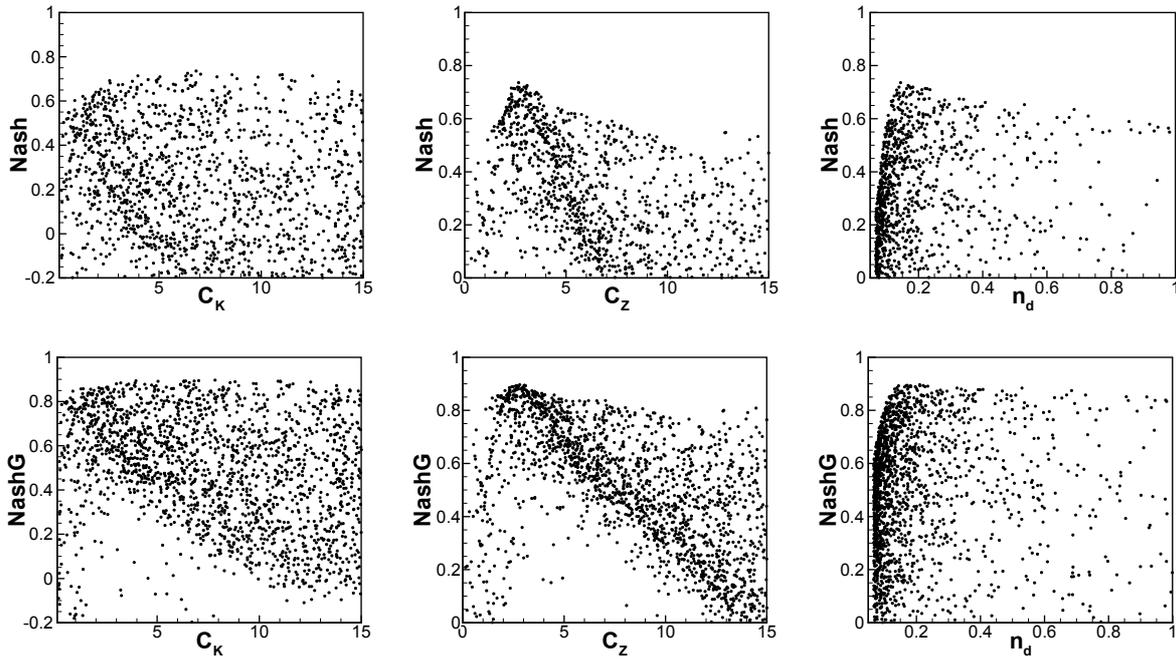
En revanche, le paramètre n_d est moins bien identifié que précédemment, les deux seuls événements pour lesquels des valeurs maximales apparaissent sont les événements n°19 (se situant autour de 0.11) et n°25 ($n_d = 0.09$). Il est toujours admis que ce paramètre dépend de l'intensité de la crue donc on ne cherchera pas à déterminer une plage de variation unique pour ce paramètre.

Les plages de valeurs acceptables pour les paramètres C_K et C_Z sont alors :

- pour l'événement n°1 : $C_K > 0.5$ et $C_Z > 2$
- pour l'événement n°3 : $C_K > 0.3$ et $C_Z > 0.7$
- pour l'événement n°4 : $C_K > 0.3$ et $C_Z > 0.5$
- pour l'événement n°11 : $C_K > 0.3$ et $C_Z \in [1.2, 14.7]$
- pour l'événement n°13 : $C_K > 1.2$ et $C_Z > 2.5$
- pour l'événement n°15 : $C_K > 0.7$ et $C_Z > 0.5$
- pour l'événement n°19 : $C_K > 1.8$ et $C_Z > 2.4$
- pour l'événement n°25 : $C_K > 0.3$ et $C_Z > 1.3$

Les marges d'incertitudes encadrent assez bien les observations sauf pour la phase de récession, elles ne diffèrent pas beaucoup par rapport au critère de Nash utilisé précédemment.

a)



b)

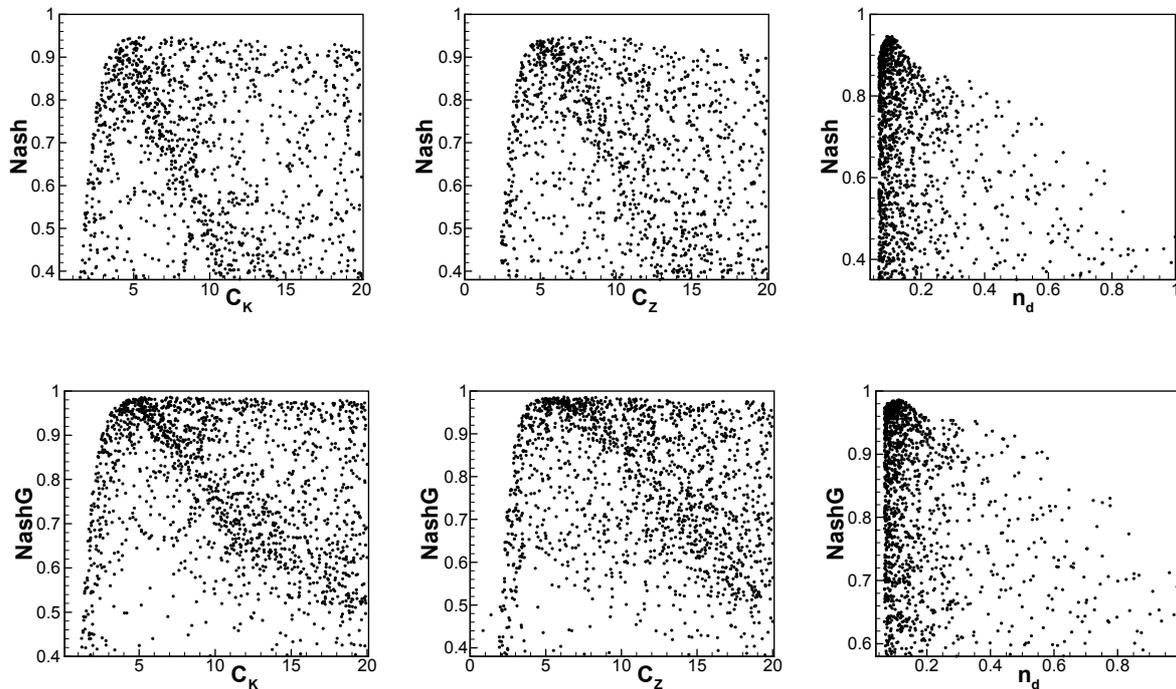


Fig. 5.36 : Comparaison des simulations de Monte-Carlo pour le critère de Nash et le Nash généralisé pour les paramètres C_K , C_Z et n_d pour les événements n°3 (a) et n°19 (b).

5.3.1.1 Synthèse des résultats

Les marges de valeurs acceptables pour les paramètres C_K et C_Z peuvent être résumées sur la figure 5.37.

En utilisant la même valeur de seuil pour les deux critères, Nash et Nash généralisé, les marges obtenues sont plus grandes pour le second :

- $C_Z \in [1.8, 10]$
- $C_Z \in [2.5, 10]$

Les résultats apparaissent plus diffus dans une plage plus large de paramètres acceptables.

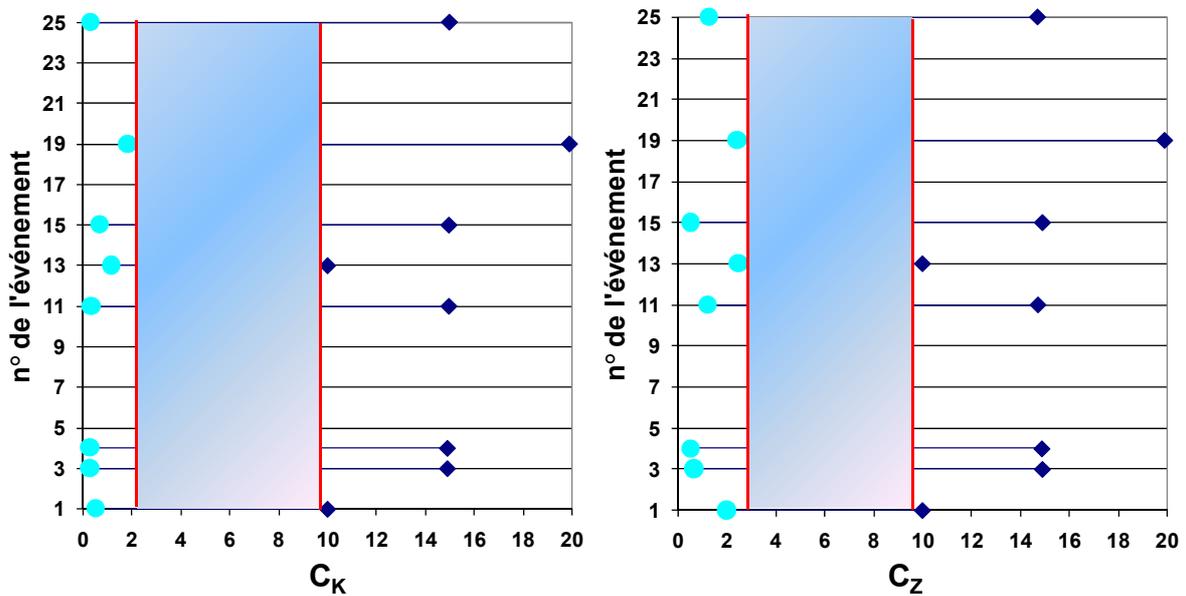


Fig. 5.37 : Bilan des intervalles de valeurs acceptables pour les paramètres C_K et C_Z avec le critère de Nash généralisé.

5.3.2 Utilisation de la méthode d'optimisation

5.3.2.1 Mise en œuvre de l'algorithme

La même procédure que précédemment a été mise en œuvre. Les intervalles de variation des paramètres sont les mêmes ainsi que les trois points d'initialisation correspondant aux tests 1, 2 et 3.

Une fonction objectif, similaire au critère de Nash généralisé utilisé pour les simulations de Monte-Carlo, a été implémentée :

$$J_2 = 1 - NashG \quad (5.3)$$

$NashG$ est le critère de Nash généralisé défini plus haut. J_2 est bien une fonction décroissante qui prend une valeur nulle si les débits simulés sont compris dans une gamme de $\pm 20\%$ des débits observés.

5.3.2.2 Résultats

Comme les résultats sont assez similaires à ceux obtenus pour le critère de Nash, seuls les résultats obtenus pour les deux crues utilisées dans la comparaison des simulations de Monte-Carlo sont présentés, il s'agit des événements n°3 et 19. Des commentaires très succincts sont donnés, basés essentiellement sur les différences remarquables entre ces deux critères.

Les résultats obtenus pour les autres crues sont renvoyés en annexe C.

- Événement n°3 - crue du 03/10/1995

Les résultats sont présentés dans le tableau 5.15 et sur la figure 5.39.

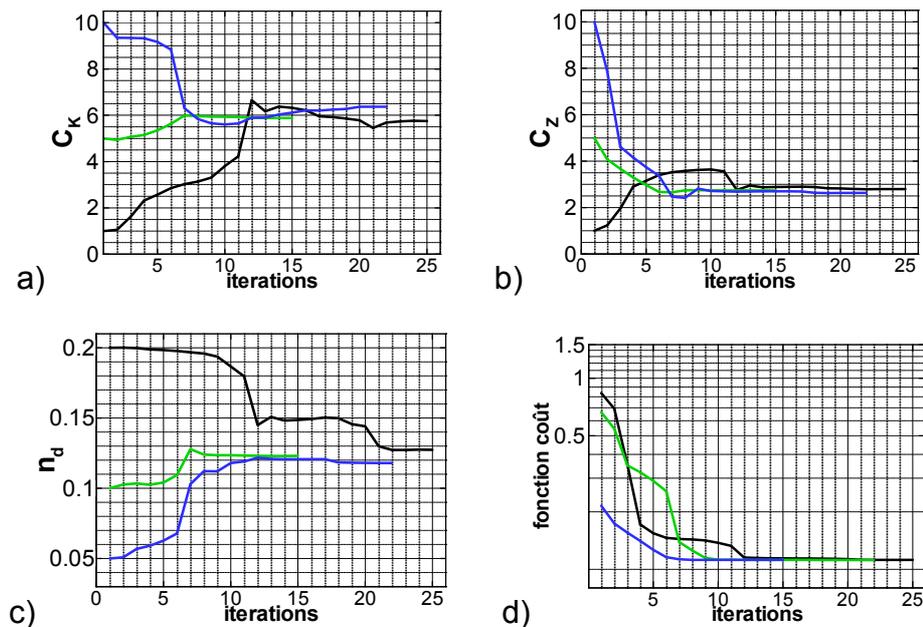


Fig. 5.38 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu.

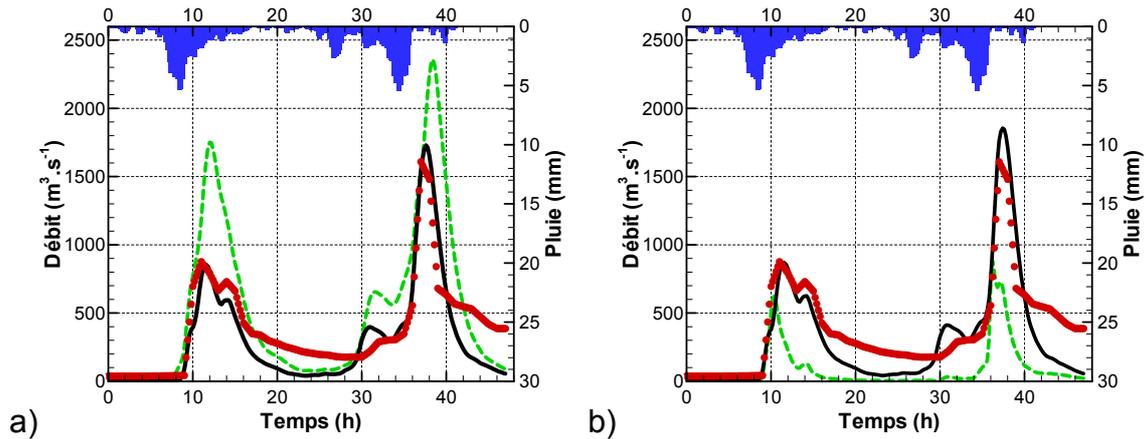


Fig. 5.39 : Hydrogrammes avant (trait vert discontinu) et après optimisation (trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°3.

Tableau 5.15. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	5.75	5	5.88	10	6.37
C_Z	1	2.8	5	2.74	10	2.63
n_d	0.2	0.13	0.1	0.12	0.05	0.12
J₂	0.84	0.11	0.22	0.11	0.66	0.11
CR (%)		44		44.6		45.6
Erreur premier pic						
ΔT₁ (min)	60	36	12	24	36	36
ΔQ₁ (%)	99.7	2.7	3.9	1.5	28.8	1
Erreur second pic						
ΔT₂ (min)	84	36	12	36	48	24
ΔQ₂ (%)	46.6	7.6	16.8	10.5	45.2	15.4

Les résultats obtenus pour les trois tests sont très proches et quasiment identique à ceux obtenus pour le critère de Nash.

- Événement n°19 - crue du 08/09/2002

Les résultats du test 1 et 2 sont très proches, en revanche, le test 3 diffère sensiblement. Il n'est pas étonnant que l'algorithme soit moins stable pour cette crue car l'incertitude est plus grande dans ce cas. Cependant, le test 3 apporte des résultats très satisfaisants avec une erreur de moins de 2 % sur le pic de crue (Fig. 5.41).

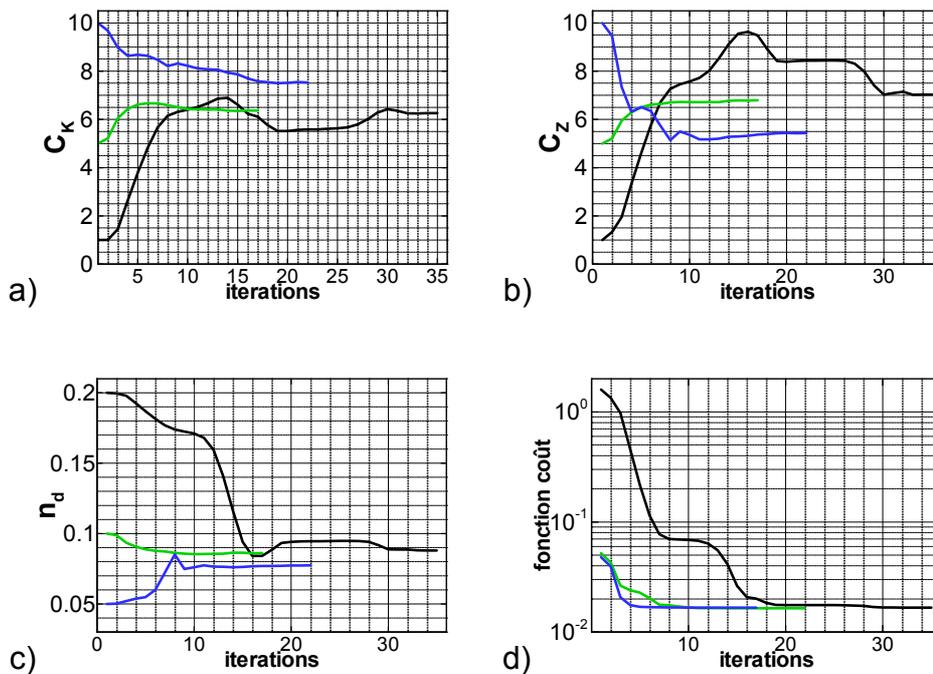


Fig. 5.40 : Convergence de C_K (a), C_Z (b), n_d (c) et de la fonction coût (d) lors de l'optimisation pour les trois différents tests d'initialisation : test 1 en noir, test 2 en vert, test 3 en bleu.

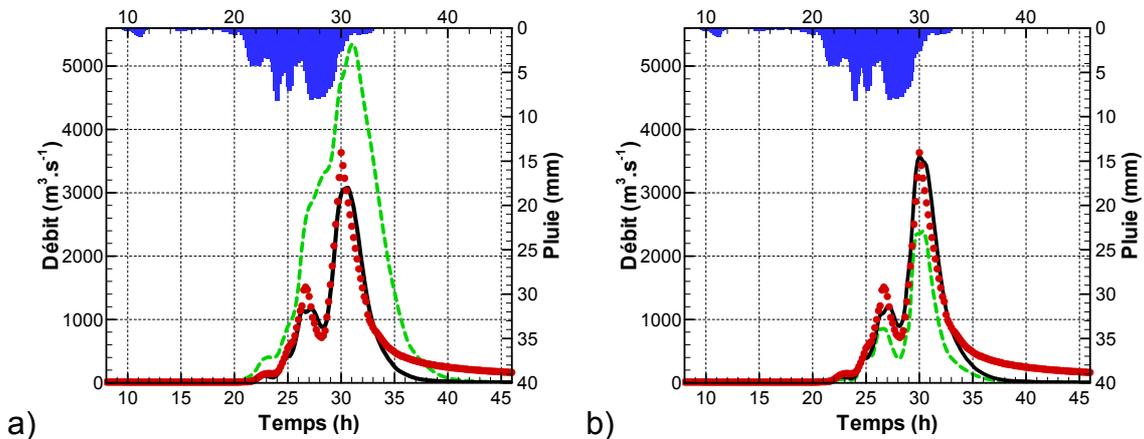


Fig. 5.41 : Hydrogrammes avant (trait vert discontinu) et après optimisation (trait noir plein) comparés aux observations (points rouge) pour le test 1 (a) et le test 3 (b) - événement n°19

Le tableau 5.16 résume les résultats des trois tests d'initialisation.

Tableau 5.16. Résultats de l'optimisation pour les 3 différents jeux de paramètres initiaux

	Test 1		Test 2		Test 3	
	Ini	Fin	Ini	Fin	Ini	Fin
C_K	1	6.26	5	6.36	10	7.53
C_z	1	7.03	5	6.79	10	5.43
n_d	0.2	0.88	0.1	0.086	0.05	0.077
J_2	1.59	0.017	0.048	0.017	0.052	0.016
CR (%)		33		33.2		34.7
Erreur premier pic						
ΔT_1 (min)	Pas de pic	30	39.6	30	0	30
ΔQ_1 (%)		22	1.9	22.4	43.4	19.4
Erreur second pic						
ΔT_2 (min)	47	40.2	30	40.2	19.8	0
ΔQ_2 (%)	81.9	15.3	5.4	13.3	34.1	1.8

5.3.2.3 Synthèse des résultats de l'optimisation

Le 5.17 récapitule la sélection des paramètres pour toutes les crues pour le Nash généralisé.

Tableau 5.17. Tableau récapitulatif des coefficients sélectionnés

	Intensité (m ³ /s)	C _k	C _z	n _d	1-J ₂
14/03/2002	666	7.48	1.05	0.57	0.95
17/05/1999	707	1.52	2.33	0.57	0.96
28/09/2000	1185	3.92	6.13	0.012	0.98
13/10/1995	1411	1.82	3.83	0.70	0.95
03/10/1995	1608	5.75	2.8	0.13	0.89
08/09/2002	3634	7.53	5.43	0.077	0.98

Les valeurs du coefficient de Manning estimées pour le critère de Nash généralisé sont trop faibles (moins de 2 m^{-1/3}.s pour les crues de Mars 2002 et Mai 1999).

5.3.3 Synthèse

Avec ce nouveau critère de performance, les résultats obtenus ont été assez proches de ceux obtenus avec le critère de Nash. On a tout de même pu constater que les pics de crue et les hydrogrammes avaient tendance à être sous-estimés. En effet l'algorithme d'optimisation converge plutôt vers la partie inférieure de la plage des débits acceptables (compris dans une incertitude de 20%) car il est plus facile de simuler un phénomène moins hydrodynamique, plus lissé. Ce qui a pour conséquence de surestimer les profondeurs de sol (et donc l'infiltration) et les coefficients de Manning (ce qui entraîne retard et sous-estimation du pic de crue).

On conclut donc que l'utilisation d'un critère classique, sans intégration des incertitudes sur le débit, suffit à une bonne identification des paramètres.

5.4 Synthèse de la calibration et validation

5.4.1 Agglomération des résultats

Les résultats de la calibration obtenus avec les deux critères ont été regroupés (agglomération) pour déduire un jeu de paramètre.

- Choix du couple (C_K , C_Z)

Le choix du couple (C_K , C_Z) a été réalisé en prenant une valeur moyenne de tous les coefficients estimés pour chaque événement. On trouve : (C_K , C_Z) = (6.5, 3.5). Cette valeur moyenne correspond à peu près au centre de l'intervalle des valeurs acceptables des simulations de Monte-Carlo pour le critère de Nash. Pour le critère de Nash généralisé, une valeur de 3.5 pour le coefficient C_Z est proche de la borne inférieure de l'intervalle des valeurs acceptables. Cependant, on a vu que pour ce critère l'hydrogramme avait tendance à être sous-estimé donc ce choix est justifiable.

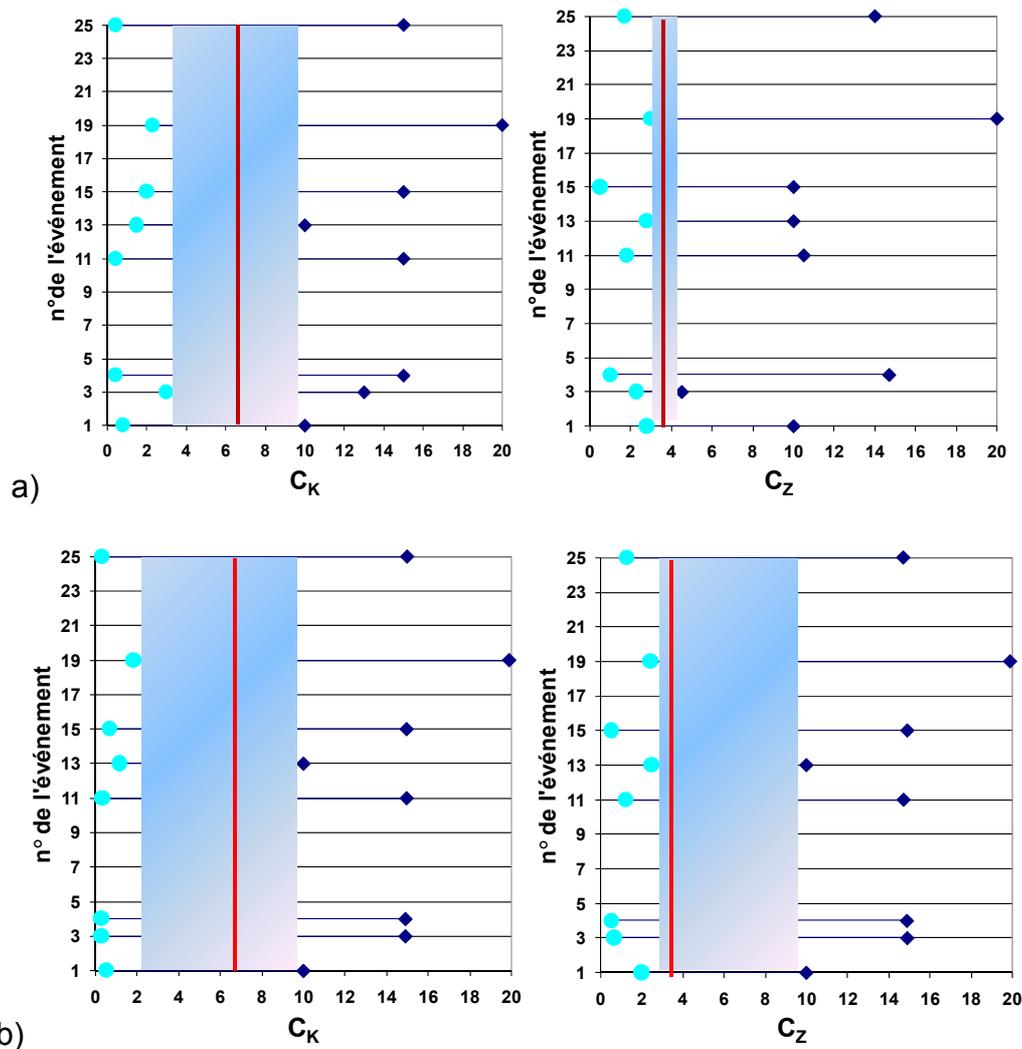


Fig. 5.42 : Choix du couple (C_K , C_Z) par rapport aux simulations de Monte-Carlo avec le critère de Nash (a) et le Nash généralisé (b).

La valeur moyenne de la conductivité hydraulique, initialement de 0.65 cm/h, devient égale à 4.2 cm/h après correction et la valeur maximale passe de 1.1 cm/h à 7 cm/h. Des valeurs de ce paramètre peuvent être trouvées dans les tables du Maidment, elles varient entre 0.06 cm/h à 23.6 cm/h. Bouvier et al. (2006) ont mesuré des valeurs de conductivité hydraulique avec un infiltromètre à disque sur le bassin du Gardon d'Anduze, ils ont trouvé des valeurs très élevées, entre 20 et 60 cm/h.

- Choix du coefficient de Manning

Le coefficient de Manning est lui dépendant de la crue, donc une valeur différente sera attribuée à chaque crue.

Les coefficients de Manning estimés avec le critère de Nash ont été retenus car avec le critère généralisé ils ont été sous-estimés. Les tests de validation seront effectués avec les coefficients de Manning estimés et ceux issus de l'extrapolation.

Les valeurs du coefficient de Manning pour différents types de rivières peuvent être issus de tables (Chow, 1959), de valeurs de la littérature (Hessel et al., 2003) ou de courbes de régression telles que celle que nous avons obtenue au paragraphe précédent pour ce bassin versant.

5.4.2 Validation

Des simulations ont été effectuées avec le jeu de paramètres déduit de la calibration. Les crues ayant servi à la calibration sont réutilisées pour la validation car très peu d'événements sont disponibles, les événements n°3 et n°8 ont été rajoutés.

Afin de valider les coefficients de Manning extrapolés, une comparaison des hydrogrammes simulés avec les Manning estimé et extrapolé est établie (Fig. 5.43).

Le tableau (5.18) donne les résultats de la validation, le critère de performance utilisé est J_1 .

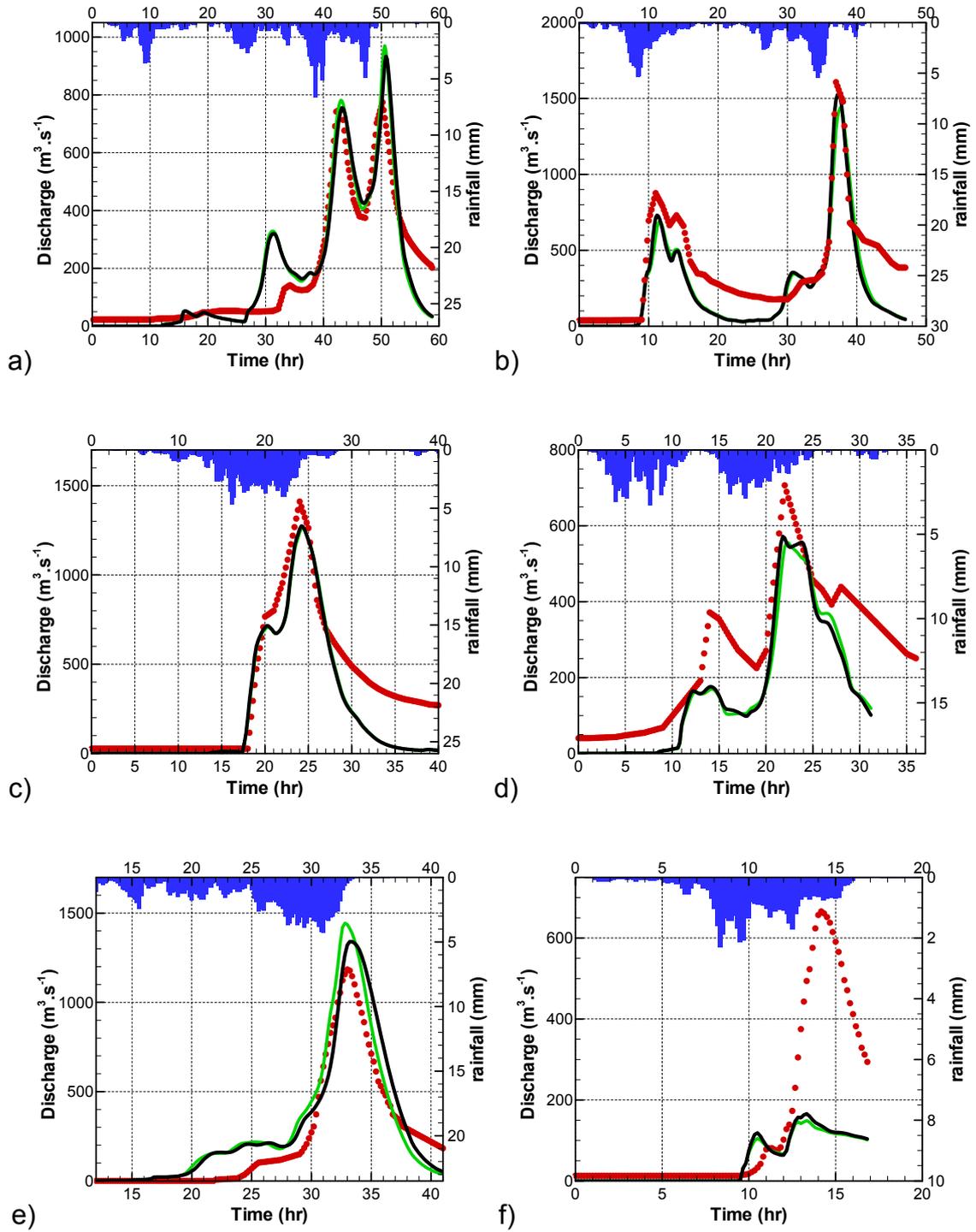
Les résultats montrent que :

- les valeurs des coefficients de Manning extrapolés sont correctes (le pic de crue arrive au bon moment),
- le couple de valeur utilisé pour (C_K , C_Z) est correct également,
- l'événement n°2 est difficile à reproduire, ceci est dû au problème évoqué dans le chapitre 4 concernant les erreurs possibles de mesure des pluies ou des débits.
- la simulation n'est pas satisfaisante sur toute plage temporelle de l'événement n°8 (crue de Novembre 1996) car celui-ci est très long (100 h) et présente deux pics séparés à plus de 35 h heures d'intervalle ; ce qui entraîne une saturation du sol et cadre mal avec les critères d'application du modèle.
- la valeur moyenne de J_1 pour tous les événements est $J_{1\text{moyen}} = 0.36$ (Nash = 0.64) et $|\Delta Q_{\text{moyen}}| = 26.6 \%$ alors que sans les événements n°2 et 8 on obtient : $J_{1\text{moyen}} = 0.26$ (Nash = 0.74) et $|\Delta Q_{\text{moyen}}| = 18.6 \%$.

En conclusion, on peut dire que le jeu de paramètres trouvé donne des résultats satisfaisants.

Tableau 5.18. Résultats de la validation du modèle avec le jeu de paramètres $(C_K, C_Z) = (6.5, 3.5)$ et le coefficient de Manning estimé ou extrapolé.

Événement	n_d	J_1	$ \Delta Q $ (%)	CR (%)
N°1 21/09/1994	0.18	0.17	25	32.3
	0.22	0.17	20.5	32.3
N°3 03/10/1995	0.13	0.32	10.5	37
	0.11	0.32	4.8	37
N°4 13/10/1995	0.124	0.18	10.5	36.6
	0.12	0.18	9.7	36.6
N°11 17/05/1999	0.31	0.42	20.5	25.5
	0.25	0.42	19	25.6
N°13 28/09/2000	0.095	0.13	21.9	30.7
	0.14	0.15	13.1	30.6
N°15 14/03/2002	0.39	0.85	77.6	9
	0.27	0.82	75.1	9.5
N°19 08/09/2002	0.074	0.27	24.6	45.4
	0.071	0.28	25.6	45.3
N°25 18/10/2006	0.10	0.13	11.9	31.8
	0.12	0.14	17.2	31.8
N°3 18/09/1995	0.17	0.66	42.4	25.6
N°8 10/11/1996	0.26	0.44	38.8	30.5



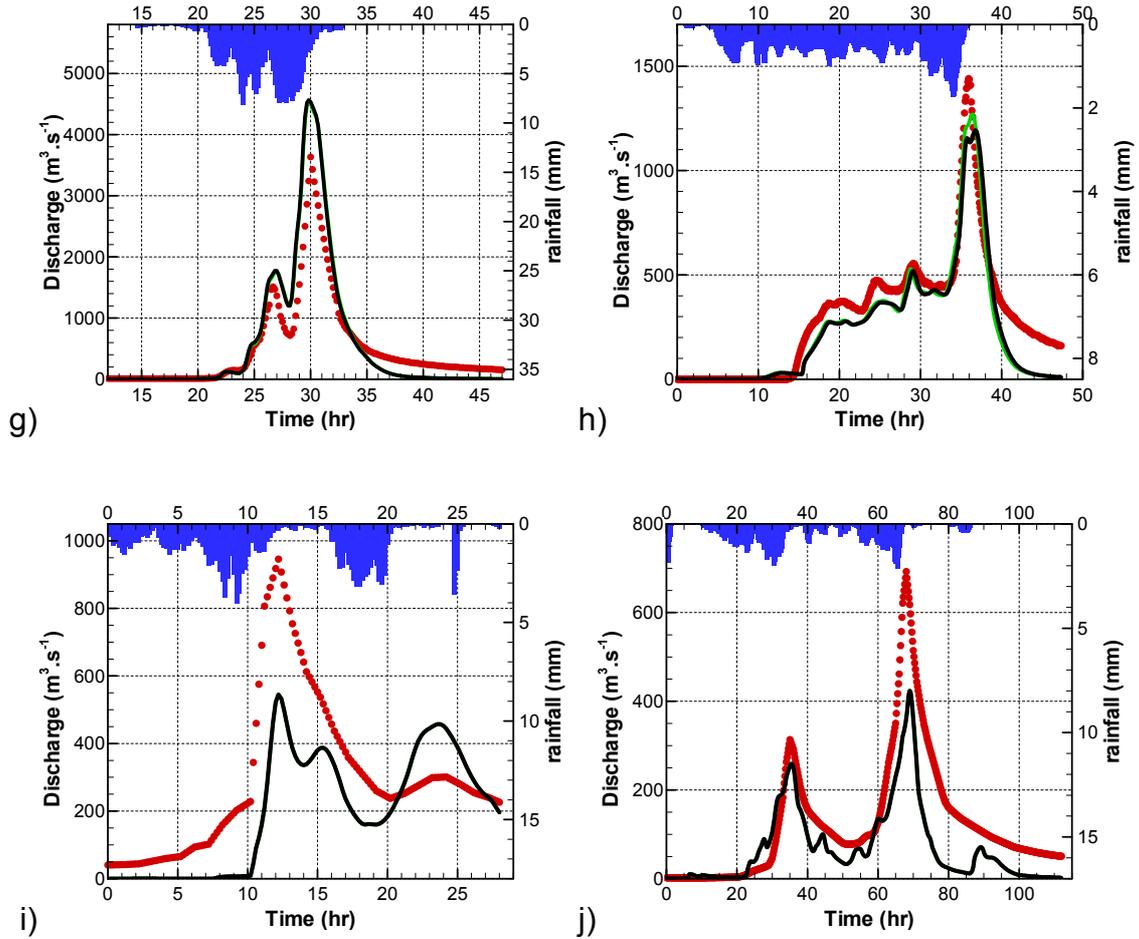


Fig. 5.43 : Validation du couple de paramètres $(C_K, C_Z) = (6.5, 3.5)$ et comparaison des coefficients de Manning estimés (trait vert) et extrapolés (trait noir) pour les événements n°1 (a), n°3 (b), n°4 (c), n°11 (d), n°13 (e), n°15 (f), n°19 (g), n°25 (h), n°2 (i) et n°8 (j).

5.5 Conclusion

La méthode d'estimation des paramètres utilisée a apporté une précieuse contribution à la calibration du modèle et a permis d'améliorer la compréhension de ses mécanismes de modélisation.

En effet, l'évolution des paramètres pendant la minimisation de la fonction coût a mené à des conclusions intéressantes à propos de la contribution respective de chaque paramètre vis-à-vis du processus de modélisation. Elle a également permis de valider les hypothèses physiques sur lesquelles est basé le modèle. La conductivité hydraulique et la profondeur de sol contrôlent l'infiltration (vitesse et capacité de stockage du sol) et déterminent donc le volume d'eau disponible au ruissellement. La valeur du coefficient de Manning du lit majeur influence essentiellement la position du pic de crue.

A cause des problèmes de non linéarité des phénomènes hydrologiques et donc de la présence possible de plusieurs minima locaux de la fonction objectif, certains auteurs préfèrent des méthodes globales d'optimisation. Duan et al. (1992) ont utilisé les stratégies de recherche basées sur l'évolution de la population et Wang (1991) des algorithmes génétiques. Cependant les méthodes globales sont très coûteuses en temps de calcul. Par conséquent, lorsque les valeurs a priori des paramètres sont cohérentes et physiquement acceptables, les méthodes locales sont plus avantageuses que les méthodes globales. C'est précisément ce qu'on a pu constater dans cette étude par l'utilisation d'une méthode locale de type gradient. La méthode s'est avérée efficace et robuste, elle a permis de contraindre les incertitudes inhérentes aux incertitudes paramétriques.

En outre, l'étude a permis d'évaluer les performances du modèle en montrant quelques points faibles de sa représentation physique. Pendant la phase de décrue, l'introduction d'un modèle de subsurface devrait contribuer à améliorer les simulations puisque les débits sont systématiquement sous-estimés pendant cette phase. Ainsi, un modèle basé sur les hypothèses de TOPMODEL (Beven et Kirkby, 1979) est en cours d'implémentation et de validation à la suite de ces observations. Une autre perspective d'amélioration du modèle est la modification du calcul du coefficient de Manning et on en a calé une loi d'évolution qu'il conviendra de tester (chapitre 7).

La confrontation des résultats du processus d'estimation et des simulations de Monte-Carlo avec les deux critères de performance utilisés, l'un classique et l'autre plus généralisé prenant en compte les incertitudes sur les débits, a permis de déterminer une valeur unique pour le couple (C_K, C_Z) , ainsi que des valeurs du coefficient de Manning en fonction de l'intensité de la crue. Les jeux de paramètres issus de cette calibration se sont avérés consistants lors de la validation.

Par conséquent, on peut dire que le modèle MARINE est robuste et constitue un outil d'aide à la décision en matière de gestion de crise d'inondation.

A partir du jeu de paramètre issu de la calibration, il sera toujours possible d'utiliser la méthode d'assimilation pour réajuster en temps réel, par exemple, la valeur de ces paramètres et améliorer l'estimation de l'hydrogramme. L'objectif du chapitre 6 concerne justement l'ajustement de ces paramètres en début d'événement pour prédire le pic de crue.

Lire
la seconde partie
de la thèse