



HAL
open science

Adjoint-based approach for estimation and sensor location on 1D hyperbolic systems with applications in hydrology and traffic

van Tri Nguyen

► **To cite this version:**

van Tri Nguyen. Adjoint-based approach for estimation and sensor location on 1D hyperbolic systems with applications in hydrology and traffic. Automatic. Université Grenoble Alpes, 2016. English. NNT: . tel-01420000v1

HAL Id: tel-01420000

<https://hal.science/tel-01420000v1>

Submitted on 20 Dec 2016 (v1), last revised 16 Feb 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

**DOCTEUR DE la Communauté UNIVERSITÉ
GRENOBLE ALPES**

Spécialité : **Automatique - Productique**

Arrêté ministériel : 7 Août 2006

Présentée par

Van Tri NGUYEN

Thèse dirigée par **Didier GEORGES**
et codirigée par **Gildas BESANÇON**

préparée au sein du **Laboratoire GIPSA-LAB**
et de l'école doctorale **EEATS**

Adjoint-based approach for estimation & sensor location on 1D hyperbolic systems with applications in hydrology & traffic

Thèse soutenue publiquement le **03 Novembre 2016**,
devant le jury composé de :

Mme Anne Catherine FAVRE

Professeur, Grenoble INP, Présidente

M. Laurent AUTRIQUE

Professeur, Université d'Angers, Rapporteur

M. Pierre-Olivier MALATERRE

Chercheur HDR, IRSTEA, Rapporteur

M. Vicenç PUIG

Professeur, Universitat Politècnica de Catalunya, Examineur

M. Didier GEORGES

Professeur, Grenoble INP, Directeur de thèse

M. Gildas BESANÇON

Professeur, Grenoble INP, Co-Directeur de thèse



In loving memory of my grandmother

*This thesis is dedicated to
my beloved parents Thi Suu Nguyen, Van Ly Nguyen
my darling wife Thi Vinh Vu
my lovely sister Thi Minh Trinh Nguyen
for their endless love, support and encouragement*

Acknowledgments

After three-year-work (10/201-10/2016) as a Ph.D. student at GIPSA-lab, today is the day, i am writing this finishing touch on my thesis. I would like to reflect on the people who have supported and helped me so much during three wonderful years of my life.

First and foremost, i would like to express my deepest gratitude to my supervisors Prof. Didier Georges and Prof. Gildas Besançon for the continuous support of my study and research, for your motivation, patience, and immense knowledge. Your sense of humor, joy and enthusiasm you have for your research was motivational for me, when i got lost during the tough times in my the Ph.D. pursuit. It has been an honor for me, to be yours student.

My sincere thanks also goes to Prof. Laurent Autrique and Chercheur HDR Pierre-Olivier Malaterre for having accepted to review my manuscript and for their time, precious and insightful comments to improve this work. I would also like to thank to Prof. Anne Catherine Favre, Prof. Vicenç Puig for their valuable comments and constructive feedbacks which helped me to widen my research from various perspectives.

Many thanks to all my colleagues and to the staff of GIPSA-lab for their help and support during my thesis. Special thanks go to Sonia Nogueira for administrative helps and to Patricia Reynier for service of documents and also for the photo contests. A very warm thank to my sworn brother Hong Nam Huynh for brotherhood and continuous support in work and life.

My deepest thanks and love goes to my parents who gave me this life and continuously sacrifice for my dreams, to my darling wife who has given a wonderful meaning to this empty world of mine and to my little sister. I dedicate this thesis to them.

Van Tri Nguyen
Grenoble, France

Contents

Acknowledgments	iii
Acknowledgements	iii
List of Figures	ix
List of Tables	xiii
I General introduction	1
Introduction	1
1 Motivation	1
2 Thesis objectives and contribution summary	3
3 Thesis organization	5
A State and parameters estimation	7
II Estimation methodology	9
Estimation methodology	9
1 State-of-the-art	10
2 Optimal estimation problem formulation	13
2.a 1-D hyperbolic system dynamics	13
2.b Optimization-based estimation approach	14
2.c Extension to switched 1-D hyperbolic systems	15
3 Calculus of variations	16
3.a Introduction	16
3.b Basic concept of calculus of variations	18
3.b.1 Gâteaux derivative	18
3.b.2 First-order necessary condition for minimization of a functional	19
3.c Application of calculus of variations	20
4 Numerical approach	26
4.a Introduction	26

4.b	Midpoint scheme	28
4.c	Preissmann scheme	29
4.d	Lax-Wendroff scheme	31
4.e	Discrete gradients and remarks	34
4.e.1	Discrete gradients	34
4.e.2	Remarks	34
5	Optimization algorithm	36
5.a	Introduction	36
5.b	Steepest descent method	38
5.c	Newton and Quasi-Newton method	40
6	Conclusion	43
III Estimation application		45
Estimation application		45
1	Traffic flow	46
1.a	Overview of free traffic flow	46
1.b	Estimation problems using synthetic measurements	49
1.b.1	Estimation of parameters	49
1.b.2	Estimation of state and parameter	54
1.b.3	Estimation of state and parameter in switched system	56
2	Overland flow	61
2.a	Overview of overland flow	61
2.b	Estimation problems using synthetic measurements	64
2.b.1	Estimation of initial condition	64
2.b.2	Estimation of distributed Manning coefficient	67
2.b.3	Estimation of initial state and distributed Manning roughness	72
2.b.4	Estimation of initial state, distributed Manning roughness and infiltration parameters	73
2.c	Estimation problems using real measurement of Tondi Kiboro catchment	77
2.c.1	Overview of Tondi Kiboro catchment and real measurements	77
2.c.2	Estimation of Manning coefficient with Green-Ampt model for unsteady rain	81
2.c.3	Estimation of Manning coefficient and infiltration parameter with Horton model	85
3	Estimation results with different configurations	89
4	Optimization algorithm with adjoint-gradient and without adjoint-gradient	92
5	Estimation with noisy measurements	94
6	Conclusion	96
B	Optimal sensor locations for state estimation	99

IV Optimal sensor location methodology	101
Optimal sensor location methodology	101
1 State-of-the-art	101
2 Preliminaries	103
2.a Observability	103
2.a.1 Linear time-invariant systems	103
2.a.2 Link between the Fisher information matrix and the observability gramian	104
2.a.3 Nonlinear systems	105
2.b Criteria for optimal sensor locations	107
2.b.1 Smallest eigenvalue of the observability Gramian	108
2.b.2 Trace of the inverse observability Gramian	108
2.b.3 Determinant of observability Gramian	109
3 Two approaches for the optimal location of sensors	109
3.a A Fisher information matrix approach	110
3.b An adjoint based approach	113
4 Conclusion	114
V Optimal sensor location application	117
Optimal sensor application	117
1 Sensor placement in a traffic flow	117
1.a Using the adjoint-based criterion	117
1.b State estimation using optimal and non optimal locations of a single sensor	118
2 One-sensor placement in overland flows	119
2.a One-sensor placement in a single overland flow	119
2.a.1 Using the Fisher Information Matrix criterion	119
2.a.2 Using the adjoint-based criterion	120
2.a.3 State estimation using optimal and non optimal locations	121
2.b Sensor placement in an overland flow network	122
2.b.1 Dynamics of overland flow network	123
2.b.2 Generalized adjoint-based criterion for the water flow network . .	124
2.b.3 State estimation using optimal locations	132
3 Conclusion	134
VI Conclusions and perspectives	135
Conclusion and perspectives	135
1 Conclusions	135
2 Perspectives	136

Bibliographie

149

List of Figures

II.1	Stencil of midpoint scheme	28
II.2	Stencil of forward Preissmann scheme	29
II.3	Stencil of backward Preissmann scheme	30
II.4	Stencil of forward Lax-Wendroff scheme	32
II.5	Stencil of backward Lax-Wendroff scheme	33
III.1	Interval $[x_1, x_2]$ of highway.	47
III.2	Relation density-velocity of Greenshields model.	48
III.3	Parameters estimation results in LWR traffic flow (I).	52
III.4	Parameters estimation results in LWR traffic flow (II).	52
III.5	Parameters estimation results in LWR traffic flow (III).	53
III.6	State and parameters estimation in LWR traffic flow (I).	55
III.7	State and parameters estimation in LWR traffic flow (II).	56
III.8	State and parameters estimation results in switched LWR traffic flow (I).	59
III.9	State and parameters estimation results in switched LWR traffic flow (II).	60
III.10	State and parameters estimation results in switched LWR traffic flow (III).	60
III.11	Illustration of overland flow model.	61
III.12	Estimation of $f_0^i(x)$	66
III.13	Estimation results of state in overland flow (II).	67
III.14	Overview of Radial Basic Function Network to approximate $K_{fh}(n, S_0)$	68
III.15	Estimation results of distributed Manning coefficient in overland flow (I).	70
III.16	Estimation results of distributed Manning coefficient in overland flow (II).	71
III.17	Estimation results of distributed Manning coefficient and initial condition in overland flow (III).	73
III.18	State and parameter estimation result in switched overland flow (II).	76
III.19	State and parameter estimation result in switched overland flow (I).	76
III.20	State and parameter estimation result in switched overland flow (III): the evolution of cost function \mathcal{J}	77
III.21	Location of Tondi Kiboro catchment.	79
III.22	Overview of Tondi Kiboro basins.	80

III.23	Observation stations on Tondi Kiboro catchment.	80
III.24	Manning coefficient estimation results using real data of Tondi-Kiboro catchment (I).	83
III.25	Manning coefficient estimation results using real data of Tondi-Kiboro catchment (II): The comparison between simulated discharge and measured discharge.	84
III.26	Manning coefficient estimation results using real data of Tondi-Kiboro catchment (III).	85
III.27	Manning coefficient and Horton infiltration parameters estimation results using real data of Tondi-Kiboro catchment (I).	87
III.28	Manning coefficient and Horton infiltration parameters estimation results using real data of Tondi-Kiboro catchment (II).	88
III.29	Manning coefficient and Horton infiltration parameters estimation results using real data of Tondi-Kiboro catchment (III).	88
III.30	Real values and 6 initial conditions of initial state and parameter $\alpha(x)$ in the case of constant initial state and parameter.	89
III.31	Real values and 6 initial conditions of initial state and parameter $\alpha(x)$ in the case of Gaussian initial state and parameter.	90
III.32	Estimation result of state and parameter estimation with initial condition 5 in Table III.10.	91
III.33	Real values and 6 initial conditions of initial state and parameter α in the case of random initial state and parameter.	92
III.34	Comparison of minimum of cost function and estimation time in two cases: with adjoint-gradient and without adjoint-gradient.	93
III.35	Comparison between noise-free measurement and noisy measurement at sensor location 10	94
III.36	Estimation of $\alpha(x)$ in two cases $SNR_{dB} = 15 dB$ and $SNR_{dB} = 30 dB$	95
V.2	State estimation in traffic flow with sensor position 10 and position 8.	118
V.1	One-sensor placement using the adjoint-based criterion.	118
V.3	Sensor placement used Fisher information matrix criterion in an overland flow: trace-based criterion and E-optimality criterion.	119
V.4	Sensor placement used Fisher information matrix criterion in an overland flow: D-optimality criterion.	120
V.5	One-sensor placement ubased on the adjoint-based criterion.	121
V.6	State estimation in an overland flow with a single sensor at position 10 and position 5.	122
V.7	State estimation in an overland with a single sensor at position 6.	122
V.8	Overland flow network configuration	123
V.9	Real initial state and initialization of systems 1 and 2.	129
V.10	Real initial state and initialization of system 3.	129

V.11	Values of the adjoint-based criterion for different locations of a single sensor	130
V.12	Index function with different location of 2 sensors	131
V.13	Adjoint-based criterion w.r.t. locations of 3 sensors	132

List of Tables

III.1	Parameters of highway and numerical values of optimization parameters used for estimating of v_{max} and ρ_{max}	51
III.2	Parameters of highway and numerical value of optimization parameter used for estimating of $\alpha_1 = v_{max}$ and $\alpha_2 = v_{max}/\rho_{max}$ and initial condition $\rho_0^i(x)$	55
III.3	Parameters of highway and numerical values of optimization parameters used for estimating of $\alpha_1 = v_{max}$ and $\alpha_2 = v_{max}/\rho_{max}$, initial condition $\rho_0^i(x)$ and relief characteristics $p = [\zeta \ \beta]$	58
III.4	Parameters of overland flow used in example of estimation the initial condition of flow per unit width $f(x, t)$	66
III.5	Parameters of overland flow and numerical value of optimization parameter used for estimating of $\alpha_1 = S_0^{1/2}/n(x)$, initial condition $h_0^i(x)$ and infiltration characteristics $p = [K_i \ K_i\Psi\eta(1 - \theta)]$	75
III.6	Geometric characteristic of Tondi Kiboro catchment.	81
III.7	Infiltration parameter of modified Green-Ampt model on Tondi Kiboro catchment.	82
III.8	Estimation errors with 6 different initial conditions in the case constant initial state and parameter.	90
III.9	Estimation errors with 6 different initial conditions in the case Gaussian initial state and parameter.	91
III.10	Estimation errors with 6 different initial conditions in the case random initial state and parameter.	92
III.11	Comparison of average relative error in 6 different initial conditions between two cases: with adjoint-gradient and without adjoint-gradient.	93
III.12	Average relative error with different values of the SNR_{dB}	95
V.1	Parameters used for the numerical simulations of the three overland flows.	129
V.2	Estimation errors with different sensor configurations	133

I | General introduction

1 Motivation

In both nature and ordinary life, one can find a lot of natural and social processes that can be described by *hyperbolic partial differential equations of order one*. They provide an effective way to mathematically model various phenomena such as water flow, traffic flow, gas dynamics or electrical lines for instance. For those reasons, hyperbolic systems of balance laws in one space dimension have been the subject of various researches in different domains, such as mathematics, physics, or automatic control. The present thesis is about this class of systems, with a special attention to application on the cases of overland flow on the one hand, which is an hydrological system, and traffic flow on the other hand, which is a more artificial phenomenon.

The motivation for those examples can be related to more general environmental issues: in recent decades, scientists have observed a serious evolution of global warming. It is one of the most important environmental challenges nowadays. Being driven directly or indirectly by human emissions of greenhouse gases, it causes not only bad effect on Earth's ecological systems but also on human life. More details and further discussions about global warming can be found in [Mat+10] and [Pac+14]. The climate change leads to problems of hydrological changes, and in the last four decades for example, strong rainfall decrements together with a significant change in the land use have been noted all around the world (see [Pie+02] and [ZC11]). This happens even in regions which seem to be insensitive to climate change like Sahelian desert area (with the case of *Tondi Kiboro* catchment for instance, which will be considered in this thesis). This phenomenon leads to strong adjustments in water cycle in such locations. Moreover, the increment of runoff factor and flow channel infiltration also causes vegetation clearing and topsoil debasement. One of the important changes which create a lot of difficulties for population in this area is changes in agricultural productivity and crops. The increment of temperature also creates glacier melting not only at Earth's poles, but also on iced tops of some high mountains. Along

with the heavy open flow water generated by rapid ice melting, the strong rain fall caused by weather change can create flood. It is a natural phenomenon that can in turn generate a lot of effects on many aspects of human life. It can for instance give rise to a broken dam. Flood usually happens when the water discharges of a river, stream or drainage ditch exceed the maximum of its banks. Moreover, in some special situations such as very excessive rainfall, hurricane or tropical storm, a flood can become a more dangerous disaster, called "flash flood" with duration less than six hours but with large destructive power, very high speed and unpredictability. One of the main components of these bad phenomena caused by global warming is the *overland flow*, whose dynamics can be modelled by hyperbolic partial differential equations, or at least one, chosen as the continuity equation of the well-known Saint-Venant model.

Beside the global warming, and being also part of the reasons for it, the population boom and the rise of motorized personal transport (car or motorbikes typically), specially in developing countries, have created heavy pressure in transport system (see [LA90] and [MW90]). The traffic jams in big and crowded cities are more and more serious, and car exhausts in the congestions make the environment more polluted as well as create potential threat on the health of urban residents. The *traffic flow*, which is considered as the interactions between travelers (including pedestrians, cyclists, drivers and their vehicles) and infrastructure, becomes a noticeable subject in many scientific branches, including automatic control. In the macroscopic point of view, the dynamics of traffic flow of vehicle can mathematically be again represented by a hyperbolic equation.

From the above analysis, hyperbolic partial differential equations of order one are an interesting subject for many scientific domains, and can in particular be used to reduce the bad effects of hydrological changes and traffic congestion. Automatic control, an application of control theory, can play an essential role in this problem. Two main issues can be considered: monitor, and control, the overland flow, or the traffic flow. Both of these purposes need information on the system, which is very often only provided by few sensors, and thus requires additional estimation processes.

In such estimation (or observer) issues, beside the prior knowledge about the system dynamics, the measurements or outputs of systems are essential. Normally they are provided by sensors placed at some points of the system spatial domain. For instance in the flood mechanism previously mentioned, an extreme rainfall causing huge overland flow is pretty common, and difficult to observe. Similarly, the vehicles characteristics (density and average velocity for instance) are necessary to formulate the control law in order to manage the vehicle flow and avoid traffic congestion in the traffic system. Therefore, monitoring in overland flow and traffic systems is a key process to detect, predict and provide warning when there is a possibility of flood or traffic jams. These duties are realized by a sensor

network. Because of technical and economic reasons, the number of sensors is limited, and as a result, state and parameter estimation over the whole overland flow or traffic systems from few measurements only should be achieved. Moreover, the sensor number and sensor location affect on the accuracy of estimation. An investigation of the optimal sensor placement is thus also needed before estimation step.

2 Thesis objectives and contribution summary

The *main objective* of this thesis is to develop a systematic technique for state and parameter estimation in overland flow and traffic flow, or more generally for a class of 1-D hyperbolic system. The proposed method is applied directly on the nonlinear system model in infinite dimensional way. A second challenge is the optimal sensor placement, which is investigated with the purpose to propose a new one in accordance with the proposed estimation method. The thesis was achieved within a Grenoble INP project on water management MEPIERA ("*Méthodologies innovantes pour l'ingénierie de l'eau et des risques associés*"), partly in collaboration with Mme I. ZIN from LTHE (Laboratoire d'étude des Transferts en Hydrologie et Environnement) for the validation on a real hydrological system. It was also included in the PERSYVAL-lab context (project ANR-11-LABX-0025-01).

In this context, the contributions of the thesis are twofold:

- Firstly, a *state and/or constant/distributed parameters estimation method for smooth/switched 1-D hyperbolic systems is developed*, and its effectiveness and applicability are *validated by simulation and real data examples*. In short, the adjoint method based on calculus of variations is used, formulated and solved directly on nonlinear infinite dimensional models. It relies on the sensitivity of measurements with respect to variables needing to be estimated, which can be distributed initial state variables, as well as parameters. The presented approach is also validated in the more simple case of constant parameters. It is generalized to the problem of state and distributed parameter estimation for a class of *switched hyperbolic partial differential equation* coupled with an ordinary differential equation. Overland flow and traffic flow are considered as illustrative examples. In the first case, the unknown infiltration model provides the switched system, and in the second case, the consideration of a relief route connected to the main traffic road also generates a switched dynamical model.
- Secondly, the problem of *optimal sensor placement* is considered. Beside the formulation of *Fisher information matrix* on an infinite dimensional hyperbolic system (the effectiveness of which is also illustrated in overland and traffic flows), a *new adjoint-gradient based criterion* to find optimal sensor position is introduced for an overland flow. This method

is also extended to the case of a overland flow network and validated by a simulation illustrative example.

Publications

This work gave rise to the following publications:

Journal papers:

- 1 V. T. Nguyen, D. Georges, G. Besançon, "State and parameter estimation in 1-D hyperbolic PDEs based on an adjoint method", *Automatica*, Volume 67, May 2016, Pages 185-191, ISSN 0005-1098.
- 2 V. T. Nguyen, D. Georges, G. Besançon, "Calculus of variation approach for state and parameter estimation in switched 1-D hyperbolic PDEs", *ESAIM - Control Optimisation and Calculus of Variations* (*under review*).

International conference proceedings:

1. V. T. Nguyen, D. Georges and G. Besançon, "Optimal state estimation in an overland flow model using the adjoint method", *IEEE Conf on Control Applications (CCA)*, Juan Les Antibes, France, 2014, pp. 2034-2039.
2. V. T. Nguyen, D. Georges, G. Besançon and I. Zin, "Application of adjoint method for estimating Manning-Strickler coefficient in Tondi Kiboro catchment", *IEEE Conf on Control Applications (CCA)*, Sydney, Australia, 2015, pp. 551-556.
3. V. T. Nguyen, D. Georges and G. Besançon, "Traffic flow parameter estimation based on an adjoint method", *European Control Conference (ECC)*, Linz, Austria, 2015, pp. 1333-1338.
4. V. T. Nguyen, D. Georges and G. Besançon, "Adjoint-method-based estimation of Manning roughness coefficient in an overland flow model", *American Control Conference (ACC)*, Chicago, IL, USA, 2015, pp. 1977-1982.
5. V. T. Nguyen, D. Georges, G. Besançon and I. Zin, "Parameter estimation of a real hydrological system using an adjoint method", *IFAC International Workshop on Adaptation and Learning in Control and Signal Processing (ALCOSP)*, Eindhoven, The Netherlands, 2016, pp. 300-305.
6. V. T. Nguyen, D. Georges and G. Besançon, "Adjoint-based state and distributed parameter estimation in a switched hyperbolic overland flow model", *IFAC Symposium on Nonlinear Control Systems (NOLCOS)*, Monterey, CA, USA, 2016.

7. V. T. Nguyen, D. Georges and G. Besançon, "Optimal sensor location for overland flow network monitoring", 3rd International Conference on Control and Fault-Tolerant Systems (SYSTOL), Barcelona, Spain, 2016.

3 Thesis organization

The rest of this thesis is organized as follows:

Chapter 2: In this chapter the estimation methodology is presented. A brief state-of-the-art on control design and estimation method for hyperbolic systems is first given. Then the optimal estimation problem is formulated with the presentation of the dynamics of 1-D hyperbolic systems under consideration. An extension to *switched* hyperbolic systems is then discussed, with the introduction of a smooth activation function to obtain an approximated continuous case. The optimal estimation is transformed into an optimization problem by defining a cost function as least square error between simulated values and measurement of system variable. The idea to solve this optimization problem is to combine the Lagrangian method and calculus of variation. Some recalls on the calculus of variations and Gâteaux derivatives are also given, and the first order optimality condition applied to the augmented objective functional allows to get the gradients of this functional with respect to the state and parameters needing to be estimated. These gradients are the fundamental inputs of gradient-based optimization algorithm which are steepest descent method and Quasi-Newton method. Some numerical schemes for numerically solving the partial differential equation and ordinary differential equations are also presented in this chapter.

Chapter 3: Here the application of the proposed optimal estimation method is considered in two main systems: overland flow modeled by continuity equation of Saint-Venant equations and traffic flow modeled by Lighthill–Whitham–Richards equation. For each system, the illustration simulation examples are presented from the less difficult cases (initial state estimation, constant parameters estimation) to the more difficult ones (state and distributed parameters estimation at the same time). The system dynamics of both systems also evolve from non-switched into switched. Moreover, in overland flow system, the parameters estimation approach is validated with the real data on Tondi-Kiboro catchment.

Chapter 4: This chapter is devoted to the optimal sensor placement for state/parameters estimation in non-switched hyperbolic systems. After a brief state-of-the-art on sensor positioning in various systems and applications, some preliminaries about observability for linear and nonlinear systems and criteria for finding sensor location are recalled. The application of Fisher information matrix approach to the problem of state and parameter

estimation in hyperbolic systems is first considered, and a new method based on adjoint approach to get the optimal sensor locations for state estimation is then proposed.

Chapter 5: In this chapter, applications of optimal sensor placement are presented, again divided in two cases: traffic flow and overland flow. In the first one, the Fisher information technique and gradient based criterion are applied on the same system to find the optimal locations. The found solutions are tested by a state estimation process. This is also the content of the first application in the case of overland flow. The second one is an extension of the adjoint gradient based criterion for an overland flow network consisting of three flows. The optimal location configurations must allow to estimate the initial conditions of all three flows of the network.

Chapter 6: This chapter concludes the manuscript, and gives some perspectives or possible extensions.

Part A

State and parameters estimation

II | Estimation methodology

This chapter is devoted to the methodology of state and parameter estimation which is considered in this thesis for a class of 1-D hyperbolic systems. It is basically an optimal estimation strategy, based on the calculus of variations. Some references related to the considered hyperbolic systems and the use of calculus of variations are firstly mentioned in the state-of-the-art section. The class of systems is then specified (section 2), made of a partial differential equation (PDE) coupled with an ordinary differential equation (ODE), and the estimation problem is stated as a non-convex optimization one, the solution of which provides the optimal values of the variables needing to be identified. It is also emphasized how the approach can be extended to a larger class of systems, including switches in the PDE dynamics, by means of a smooth activation function. The calculus of variations, a typical tool to deal with the type of optimization we are facing, is then recalled in section 3 with its historical development and basic concepts of Gâteaux derivative or first order necessary condition. Based on these concepts, the details of calculation and application of adjoint method on the presented optimization problem are given. The Lagrangian multiplier method is used to connect the system equation with the cost function and create a new unconstrained optimization problem; the first variation calculation is applied on this Lagrangian objective functional to derive the gradient of the functional with respect to the parameter and state. In this procedure, the adjoint equations (one PDE and one ODE) being conjugated with the original one are also obtained. Finally, numerical approaches to solve all the PDEs and ODEs which are met are discussed in section 4, and two gradient-based minimization algorithms including steepest descent and quasi-Newton as optimization solver are given in section 5. Section 6 gives the conclusions of the chapter.

1 State-of-the-art

The control design and state/parameter estimation problems for hyperbolic systems define a vast field of research for the automatic control community. Over the last decades, along with the developments in computer power, various advanced control and estimation methods which demand high computational power, have been developed in a lot of applications. Focusing only on the class of *non switched hyperbolic systems* of conservation laws, various researchers in control have investigated a large number of control issues with different purposes. One can cite here some typical ones such as: P.O Malaterre combined the multivariable predictive control method with Kalman filter (for state estimation) to control a 2-pool open canal [MR97]; X. Litrico et al. proposed the model and control law for a dam-river system [LG99a], [LG99b]; X. Sun et al. applied the predictive control method on gas jet flames distribution [SW05]; V. Puig et al. [Pui+05] and Y. Bolea et al. [BBP07] investigated linear parameter varying approaches for modeling and identification of open channel system; D. Georges [Geo09] designed an infinite-dimensional nonlinear predictive control for open-channel flow; V.T. Pham et al. investigated the problem of shock wave avoiding by using the receding horizon boundary control applied on the macroscopic Lighthill-Whitham-Richards traffic flow model [PGB12]; more recently Y. Li et al. used linear programming to optimally control traffic in highway network [LCC14].

Uncertainty in the initial condition and empirical parameters are ones of the important issues arising in the simulation and control of this type of systems. They can cause large errors and inconsistency between the output of control system and the real one. Related to the problem of constant parameter estimation in hyperbolic systems, one can also find studies such as: the method of influence coefficient has been used by Becker et al. to estimate Manning roughness coefficient in an unsteady open-channel flow [BY72]; an adjoint analysis method has been proposed by Y. Ding et al. to find out roughness coefficient in shallow water [DJW04]. In the case of distributed parameters of hyperbolic system, Bagchi et al. [BB90] developed an estimation method based on the maximization of a likelihood function for the parameter and system state of a discrete-time hyperbolic system with noisy boundary condition; the same adjoint gradient approach considered by Y. Ding et al. was also used on the multi-reaches channel flow network [DW05]; Richard et al. investigated a numerical scheme used to solve parameter identification issue for 1-D hyperbolic system [EL88]; whereas Wenhuan proposed a quasi-Newton method to deal with the same problem [Yu99]; the estimation problem of Manning roughness coefficient of a channel network was solved by H. Longxi by a complex method [Lon08]; a state estimation method based sequential Monte Carlo for an irrigation canal was developed in the work of J. Sau et al. [SMB10]; P.O. Malaterre et al. proposed the sequential Kalman filter and particle filter state observer applied on a 1-D hydrodynamic model to determine the water level, discharges and some parameters such as Manning-Strickler [Mal+09]; a moving hori-

zon technique was employed by Hasan et al. to estimate state and constant parameter in a 2×2 linear hyperbolic system based on a distributed model for drilling application [HI14].

In the case of *switched hyperbolic systems* (used with a general sense including switched characteristic in dynamic part or source part) some studies can also be found. One can cite here some works concerning the optimal control for this type of system: X. Xu et al [XA02] using nonlinear optimization techniques to find the optimal switching instants for switched autonomous systems; dealing with the same problem and system M. Egerstedt et al. [EWD03] used a different optimization technique; M. Rinehart et al. [RDK07] designed discrete-time state-feedback controllers for continuous-time switched homogeneous systems; further references can be found in a survey paper of F. Zhu et al. for optimal control of hybrid switched systems [ZA15]. Some fruitful results on infinite dimensional stabilization of switched hyperbolic system have also been reported, see e.g, the papers of S. Amin et al. [AHB08], C. Prieur et al. [PGW14]. More details on this subject can be found in the survey paper of H. Yang et al. [YJC14]. However, there are few available studies investigating the problem of state and parameters estimation in switched infinite dimensional PDE. These brief references are of course very far from the full summary of control application in hyperbolic systems but through them readers can find basic ideas.

In this context, the present chapter proposes the use of calculus of variations or adjoint method to solve the problem of state and parameter estimation for 1-D hyperbolic systems. The proposed approach is extended to a case of switched system where one PDE is coupled with an ODE. The term "switched" here refers to a system changing its characteristics on the source term in the dynamic equation, according to a switching rule which may depend on time, parameters of the system and/or state of the system.

The adjoint method being the main tool considered in this thesis to deal with the estimation problem, it is worth including some references on this concept hereafter. Based on the calculus of variations, a mature concept with more than 250 years (more details about its brief history are recalled in section 3), the adjoint method is employed in various problems in the field of control. An application survey of variational calculus can be found in the work of J. Ferguson [Fer04]. Two formalisms including forward sensitivity and the adjoint sensitivity for the investigation of mathematical foundations in nonlinear equations and functional analysis are provided in two works of Cacuci [Cac81a]; [Cac81b]. A large number of studies have investigated variational calculus to analyze the characteristics or control the behaviors of some dynamical systems in various domains including engineering systems, biological systems, economic systems and hydrological systems etc whose dynamics are described by differential equations. Some typical works concerning the application of this concept are: to design the controller for the contaminant releases in rivers [PK97a]; [PK97b]; to control the canal flow or reduce the effect of hazardous flood in rivers by using

depression waves [SK99b]; [SK99a]; to investigate the applicability of adjoint method to nonlinear control of irrigation canals or dam-river systems via the one-dimensional collocation control model [DGB04]; to monitor and control the traffic flow with the presence of congestion waves [JDWK05]; to plan the trajectory of aircraft and control air traffic flow [RT05]; to solve the boundary optimal control issue for open channel flow network [DW06] or alluvial channels [DW12b] or a watershed consisting of complex channel network with irregular topography [DW12a] or for Navier-Stokes equations [GC06]; to find the control law for an Eulerian network model of air traffic flow through solving an optimization problem of a log-barrier augmented objective functional [BRT06] or to regulate the optimal flow in the same system [WB08]; to formulate the nonlinear predictive control in infinite dimensional for open-channels system [Geo08]; and to solve a coordinated ramp metering problem for freeways [Rei+15].

A lot of studies have also been carried out to investigate the application of adjoint method on estimation problem. S. K. Das and al. estimated the bottom friction coefficient and the water depth in the tidal flow model [DL91]. I. M. Navon firstly provided some aspect of stability to validate the estimated result after providing a survey on the use of calculus of variations for inverse problems in oceanography and meteorology [Nav98]. More recently, V. M. Calo et al. proposed a gradient-based approach to identify the spatially distributed Manning coefficient in the shallow water equations [Cal+13]. L. Perez et al. investigated a conjugate gradient algorithm based on calculus of variations to identify the thermal diffusivity in moving boundaries system [PAG08]. L. Autrique [ARR05] performed a parametric identification by using the same conjugate gradient method as [PAG08] on a thermal system which is described by partial differential equations. K. Yoshida et al. proposed the same approach for estimating also the Manning roughness but on a real hydrological system (the Asahi River in Japan during flooding in 2011) [YM14] and for flood hydrograph [YI15]. The dynamics of flood of this work is modeled by shallow-water equations. The variational method applied on a Saint-Venant hydraulic network to estimated its water discharges is presented in the work of I. Gejadze et al. [GM16]. More specially, the work of W. Castaings et al., where the applicability of variational calculus for parameter estimation in distributed overland flows is analyzed and illustrated by a numerical example [Cas+09], shares the same idea as the one of parameter estimation in an overland flow investigated in the present thesis. But [Cas+09] does not consider the problem of state estimation, and the parameter needing to be estimated in this latter work is supposed to be constant.

The state and parameters estimation problem considered in this thesis consists in solving an inverse problem. But the inverse problem is often ill-posed if it is not well-posed in the sense given by Jacques Hadamard [Had02]. This means that at least one of the following properties is not satisfied: a solution exists, the existing solution is unique, the behavior of the solution depends continuously on the initial conditions. The degree of

ill-posedness which represents the characteristic of an inverse problem can be found in the works of B. Hofmann et al. [Hof94]; [HW05]; [HK+10]. The noises in the observation process are among the main reasons which cause the ill-posedness. For the class of descent methods used for solving inverse problem, the regulation iterative technique based on conjugate gradient proposed in the former works of Y. Jarny et al. [JOB91] and of O.M. Alifanov [Ali94] for parabolic systems can be used as potential approaches. In this thesis, an example of estimation problem with noisy measurement along with a regulation method is presented at the end of chapter III.

The next part of this chapter is devoted to describe the state and parameter optimal estimation in 1-D hyperbolic system. The calculation of variations and adjoint-based approach along with the optimization algorithm are then reformulated.

2 Optimal estimation problem formulation

2.a 1-D hyperbolic system dynamics

Being the subject of various researches in a lot of domains, such as mathematics, physics and automatic control, hyperbolic systems of balance laws in one space dimension have been used to describe a lot of phenomena in nature and ordinary life such as water flow, macroscopic traffic flow, gas pipelines, or electrical lines for instance.

Denoting by x the spatial variable and t the time variable, with $x \in [0, L]$, $t \in [0, T]$ for a spatial length L and a time horizon T , the system dynamical representation typically reads as follows:

$$\begin{cases} \frac{\partial u(x, t)}{\partial t} + \frac{\partial f(u(x, t), x)}{\partial x} & = g(x, t, p, y(t), u(x, t)) \\ u(x, 0) = u_0^i(x) \quad \text{and} \quad u(0, t) = u_0^b(t) \\ \frac{dy}{dt}(t) = h(y(t), p) \\ y(0) = y_0 \end{cases} \quad (\text{II.1})$$

where $u(x, t)$ and $y(t)$ are system variables of the PDE and ODE, respectively; $f(u(x, t), x)$ is the flow term depending on $u(x, t)$ and x , and $g(x, t, p, y, u)$ is a source term possibly depending on space and time, but also system variables y, u and a set of parameters p . Finally h is a function defining the dynamics of y , and possibly depending on p as well. The system dynamics are complemented by the Cauchy initial and boundary conditions described two functions $u(x, 0) = u_0^i(x)$ and $u(0, t) = u_0^b(t)$, while y_0 is the initial condition of the coupled ODE. The boundary condition of PDE and initial condition of ODE are known functions and fixed beforehand. The flow $f(u(x, t), x)$ can be written as a

product of some coefficient vector $\alpha = [\alpha_1(x)\dots\alpha_i(x)\dots\alpha_K(x)]^T$ and sub-function vector $\varphi = [\varphi_1(u)\dots\varphi_i(u)\dots\varphi_K(u)]^T$,

$$f(u(x, t), x) = \sum_{i=1}^K \alpha_i(x) \varphi_i(u) \quad (\text{II.2})$$

2.b Optimization-based estimation approach

The general estimation problem which is considered is that the initial condition $u_0^i(x)$ together with parameters p and α in equations (II.1) are undefined and needing to be estimated from some available measurements.

Instead of the estimation techniques mentioned in the introduction chapter, the estimation issue considered in this work is formulated as a nonlinear optimization problem. This means that a cost function or observation function is considered, being a function of initial state and parameters, and defined based on the measurements. This function needs to be minimized to get the desired variables. There are several ways to characterize this function or define the relationship between observations and simulated values. Typical functions which can be considered are the maximal error, the absolute error summation and the sum of least square errors, which will be the one considered in this thesis.

The measurements are provided by sensors normally placed at fixed locations on the spatial domain (except mobile sensors in some special cases). The observation process considered hereby can thus be modelled by the Delta-Dirac operator at a predefined set of discrete points. But to guarantee the smoothness of observation functions, a Gaussian approximation of the Dirac function is used. And with a variance σ^2 and center point x_j , this smooth operator is described in equation (II.3):

$$\delta_A(x - x_j) = \frac{1}{\Lambda(x_j)} e^{-\frac{(x - x_j)^2}{\sigma^2}} \quad (\text{II.3})$$

where

$$\Lambda(x_j) = \int_0^L e^{-\frac{(x - x_j)^2}{\sigma^2}} dx$$

with spatial domain length L . Notice that the reconstructed measurement values are conserved due to

$$\int_0^L \delta_A(x - x_j) dx = 1$$

The center point x_j is also the observation position. The size of observation window can be adjusted by tuning the variance σ . A small value of σ allows to have an accurate observed

data but causes some peaking in the gradients. Moreover, a large σ creates a wider observation window of sensor operator which propagates the information around observation point x_j but adds also a bias to the measured values. The value of σ is usually chosen "small enough", and manually in each specific case because its value also depends on the scale of spatial variable to keep a good compromise between exact measurement and information propagation. The measurements are rebuilt by the spatial Fredholm first kind integral of the convolution between the simulated value $u(x, t)$ with the shifted approximation of Delta Dirac function.

The cost function called \mathcal{J} finally takes the following form:

$$\begin{aligned} \mathcal{J} = & \frac{1}{2} \sum_{j=1}^N \int_0^T \left\{ \int_0^L \delta_A(x - x_j) u \, dx - u_j^{meas}(x_j, t) \right\}^2 dt + \frac{1}{2} \varepsilon_1 \int_0^L \|u_0^i(x) - u_{0F}^i(x)\|^2 dx \\ & + \frac{1}{2} \varepsilon_2 \sum_{i=1}^K \int_0^L \|\alpha_i(x) - \alpha_{iF}(x)\|^2 dx + \frac{1}{2} \varepsilon_3 \sum_{l=1}^H \|p(l) - p_F(l)\|^2 \end{aligned} \quad (\text{II.4})$$

where T is the observation horizon (*hours*); L is the spatial domain length (*m*); N is the number of observation values; $u_j^{meas}(x_j, t)$ is the measurement at observation position x_j with $x_j \in [0, L]$; K and H are respectively the size of parameter vector α and p ; α_{iF} , p_F and u_{0F}^i are the first guessed values of coefficients α_i , parameters p and initial state u_0^i respectively ; $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are weighting factors to calibrate the estimated values and the guessed ones, adjust the scale of cost function and improve the convergence of optimization algorithm. The first guessed values of state and parameters are selected in a reasonable range around the expected real ones.

2.c Extension to switched 1-D hyperbolic systems

Before entering into the process of minimizing the above criterion (II.4), let us discuss how the proposed formulation can be extended to systems (II.1) also including possible switching dynamics: here only switching in the source term g will be considered, assuming that it can change between 2 forms $\{g_1(x, t, p, y, u), g_2(x, t, p, y, u)\}$ when a switching condition is satisfied (more general cases could also be considered, but are left for future works). More precisely, let us focus only on a simple case when the system goes from $g_1(x, t, p, y, u)$ to $g_2(x, t, p, y, u)$ under the switching condition $\xi(t, p, u) > 0$. This condition can be a time-evolving or state-evolving inequality. With the same switching condition, the source of ODE, $h(y, p)$ takes one of its sub-functions $\{h_1(y, p), h_2(y, p)\}$. Equation

extrema (minima or maxima) of a nonlinear optimization problem on infinite dimensional function space. From the beginning of this branch of mathematics in 814 BC with the so-called "Dido problem" up to many modern time optimization issues, it is a very lively domain which has attracted the attention of many scientists. Let us go upstream the history to have an overview of its development and its typical problems. The recalled Dido problem, corresponding to finding the closed curve having the maximum area for a given perimeter, can be considered as the earliest isoperimetric inequality. This is also one of the oldest problems in the calculus of variations. Its first solution for polygons was given by Zenodorus, an ancient Greek mathematician around 200 BC with the contribution and improvement of Archimedes. Weierstrass made some significant advances in this isoperimetric inequality problem by his proof with new complete reformulation and opening the way to modern study of the calculus of variations by proving the necessary condition for the existence of strong extrema. In the seventeenth century, a lot of variational problems were proposed and considered such as: geometrical optics by Fermat, ship design to minimize the resistance of Newton and the well-known Brachistochrone curve problem formulated and solved by Johann Bernoulli in 1696. During this period, Euler-Lagrange equation, a systematic way allowing to solve systematically variational issue was found by Euler and Lagrange. It was extended later in the studies of Hamilton, Hilbert, Weierstrass etc. At the end of the nineteenth century and the beginning of twentieth century, a problem of finding extrema of multiple integrals or Dirichlet integral, one of the most famous problem of the calculus of variations, was solved by Hilbert by using the direct method with the previous contributions of Dirichlet, Thompson and Riemann. Being presented around 1900, the direct method relies on functional analysis and topology to prove the existence of a minimum for a given functional. The problem of minimal surface (that locally minimizes its area) which was presented in 1762, was completely solved by Douglas and Rado in 1930. Further information about calculus of variations history can be found in [Fer04] and [Gol12].

Nowadays, the minimization techniques of calculus of variations provide a powerful tool to characterize the functional equilibrium configuration of various systems in a lot of areas of mathematics, physics, classical mechanics, quantum mechanics, engineering etc. With the same spirit of multi-variable function optimization where the gradients of function with respect to variable point out the candidate of the minima, the functional gradients allow to distinguish the critical functions that might be the minimizer of the functional. The gradient of a function or the generalization of the usual concept of derivative of one variable function is given by its partial derivative. But in the case of a functional defined on an infinite dimensional function space, the ordinary derivative is replaced by Gâteaux differential or Gâteaux derivative, named after René Gâteaux, a French mathematician. This new concept along with the associated first order of necessary condition for optimality are presented in details in the next subsections.

3.b Basic concept of calculus of variations

3.b.1 Gâteaux derivative

Considering two locally convex topological vector spaces denoted by Υ and Ψ , an open sub vector space $\Theta \in \Upsilon$ and a functional called $\mathcal{F} : \Upsilon \rightarrow \Psi$, the Gâteaux derivative denoted by $d\mathcal{F}(u; v)$ at $u \in \Theta$ in the direction $v \in \Upsilon$ of \mathcal{F} is defined by the limit below:

$$d\mathcal{F}(u; v) = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{F}(u + \epsilon v) - \mathcal{F}(u)}{\epsilon} \quad (\text{II.9})$$

if the limit exists for all $\epsilon \in \mathbb{R}$. Classically, the direction v can be considered as the variation of function u (maybe written as δu) perturbed by the real factor ϵ . This is also the reason why this technique called as "calculus of variations". For an unknown function u which will be determined to minimize the functional $\mathcal{F}(u)$, the functional value $\mathcal{F}(u)$ perturbed by the perturbation function h and small real number ϵ is $\mathcal{F}(u + v\epsilon)$. If this limit exists for all vector $v \in \Upsilon$, the functional \mathcal{F} is Gâteaux differentiable at u . The Gâteaux derivative can be also rewritten under the form of partial derivative of a function if the functional \mathcal{F} is "considered" as a simple function of ϵ .

$$d\mathcal{F}(u; v) = \left. \frac{\partial}{\partial \epsilon} \mathcal{F}(u; v) \right|_{\epsilon=0} \quad (\text{II.10})$$

This formulation allows to calculate the Gâteaux derivative without evaluating the limit as in the previous formal definition. The concept of first variation of functional \mathcal{G} usually corresponds to the Gâteaux derivative. A linear functional denoted by $\delta\mathcal{G}|_u$ (linearity with the sense: $\delta\mathcal{G}|_u(v_1\epsilon_1 + v_2\epsilon_2) = \epsilon_1\mathcal{G}|_u(v_1) + \epsilon_2\mathcal{G}|_u(v_2)$) is called the first variation of \mathcal{G} at function u if for all function v and all real number ϵ , the following equality holds

$$\mathcal{G}(u + v\epsilon) = \mathcal{G}(u) + \delta\mathcal{G}|_u\epsilon + o(\epsilon) \quad (\text{II.11})$$

where the function $o(\epsilon)$ satisfies the condition below:

$$\lim_{\epsilon \rightarrow 0} \frac{o(\epsilon)}{\epsilon} = 0 \quad (\text{II.12})$$

or the function $o(\epsilon)$ goes to zero faster than ϵ when it tends to zero. This mean one can reformulate the formal definition of Gâteaux derivative in (II.9) by dividing the two sides of equation (II.11) by ϵ and taking the limit while ϵ tends to zero. For convenience and to avoid the confusion between the notation of Gâteaux derivative $d\mathcal{F}(u; v)$ in (II.9) and the ordinary derivative, the notation $\delta\mathcal{F}|_u(v)$ is used for both the first variation and the Gâteaux derivative. The first variation provides a necessary condition for a minimum of a functional and will be discussed in more details in the next subsection via a simple proof.

3.b.2 First-order necessary condition for minimization of a functional

Consider the functional $\mathcal{F}(u)$ defined on the two vector spaces Υ and Ψ , $\mathcal{F} : \Upsilon \rightarrow \Psi$. If the function u^* is a minimum of $\mathcal{F}(u)$, then the following equation will hold for all small real value ϵ because the functional value cannot be smaller in all possible direction:

$$\mathcal{F}(u^* + \epsilon v) - \mathcal{F}(u^*) \geq 0 \quad (\text{II.13})$$

As a result, if $\epsilon > 0$

$$\frac{\mathcal{F}(u^* + \epsilon v) - \mathcal{F}(u^*)}{\epsilon} \geq 0$$

and conversely $\epsilon < 0$

$$\frac{\mathcal{F}(u^* + \epsilon v) - \mathcal{F}(u^*)}{\epsilon} \leq 0$$

Take the limit of the two previous inequalities while ϵ goes to zero from both left and right sides:

$$\lim_{\substack{\epsilon \rightarrow 0 \\ \epsilon > 0}} \frac{\mathcal{F}(u^* + \epsilon v) - \mathcal{F}(u^*)}{\epsilon} \geq 0 \quad ; \quad \lim_{\substack{\epsilon \rightarrow 0 \\ \epsilon < 0}} \frac{\mathcal{F}(u^* + \epsilon v) - \mathcal{F}(u^*)}{\epsilon} \leq 0 \quad (\text{II.14})$$

This leads to the limit defining the first variation:

$$\lim_{\epsilon \rightarrow 0} \frac{\mathcal{F}(u^* + \epsilon v) - \mathcal{F}(u^*)}{\epsilon} = 0$$

If the Gâteaux derivative of \mathcal{F} exists, the previous limits also exists or the variation of \mathcal{F} at function u^* is equal to zero in all direction v . By reversing the inequality signs in the above equations, the same result can be obtained for the case of maximum. The first order of optimality of functional therefore is presented as follows:

$$\delta\mathcal{F}|_{u^*}(v) = 0 \quad \forall v \in \Upsilon \quad (\text{II.15})$$

On the function space Υ , an inner product \langle, \rangle is defined, as the standard L^2 inner product for two functions f and g

$$\langle h, g \rangle = \int g(x)h(x)dx \quad (\text{II.16})$$

The Gâteaux derivative of \mathcal{F} at point u in direction v can be written under the form of the inner product between the gradient of functional $\mathcal{F}(u)$ and vector v :

$$\delta\mathcal{F}(u, v) = \langle \nabla\mathcal{F}(u), v \rangle \quad (\text{II.17})$$

Comparing with equation of the previous optimality condition, the gradient of functional $\mathcal{F}(u^*)$ satisfies

$$\langle \nabla\mathcal{F}(u^*), v \rangle = 0 \quad \forall v \in \Upsilon \quad (\text{II.18})$$

which is known as the weak form of variational principle.

3.c Application of calculus of variations

The previous concepts of calculus of variations provide the idea to deal with the optimization problem presented in section 2. The application is proposed here for the more general case of system (II.7)-(II.8). The case of smooth system (II.1) is a reduced version and the details can be found in [NGB16b]. Here, the first variation calculation or the Gâteaux derivative of the cost function with respect to all possible variation directions of initial state and parameters is applied. But the considered problem is a constrained one with the system dynamics (II.5) as an equality constraint, along with the cost functional \mathcal{J} in equation (II.4). The Lagrangian multiplier method can then be used to combine the cost function and the constraint together. By introducing two multipliers $\lambda(x, t)$ for the switched PDE and $\gamma(t)$ for the ODE, the augmented objective functional called \mathcal{L} can be written as follows:

$$\mathcal{L}(u(x, t), u_0^i(x), y(t), \alpha(x), p) = \mathcal{J} + \int_0^T \int_0^L \lambda(x, t) \Sigma_1 dx dt + \int_0^T \gamma(t) \Sigma_2 dt \quad (\text{II.19})$$

where Σ_1 , Σ_2 are the left parts of equations (II.7) and (II.8), respectively. The constrained optimization becomes an unconstrained one. The functional $\mathcal{L}(u(x, t), u_0^i(x), y(t), \alpha(x), p)$ is defined from the set $L_2([0, L], [0, T]) \times L_2([0, L]) \times L_2([0, T]) \times L_2([0, L])^K \times \mathbb{R}^H$ into \mathbb{R} . The idea to solve this optimization problem is to rely on the first order necessary condition for functional optimality established by the calculus of variations (presented in equation (II.15)) that is to cancel the first order variation of $\mathcal{L}(u(x, t), u_0^i(x), y(t), \alpha(x), p)$. The candidates for optimal solutions of the cost functional are found at the point where its variations with respect to the variations of state and parameters vanish. The first variation of functional is considered here in the form of Gâteaux derivative which is applied on equation (II.19). The stationary point of the cost functional is at the point where its gradients with respect to the state and parameters vanish (by using the condition (II.18) for all vector v). The local minimum of \mathcal{J} corresponds to a stationary point of objective functional \mathcal{L} with a specific values of λ^* and γ^* .

For the reason of simplicity, notations $u(x, t)$, $f(u(x, t), t)$, $\lambda(x, t)$, $\gamma(t)$, $\varphi_a(t, p, u)$, $g_1(x, t, p, y)$, $g_2(x, t, p, y)$, $h_1(y, p)$, $h_2(y, p)$, $y(t)$ and $\xi(t, p, u)$ will be reduced to u , f , λ , γ , φ_a , g_1 , g_2 , h_1 , h_2 , y and ξ , respectively, except special case. The variation of the first part \mathcal{J} of \mathcal{L}

is given as follows:

$$\begin{aligned} \delta \mathcal{J} = & \sum_{j=1}^N \int_0^T \int_0^L \delta_A(x-x_j) \left[\int_0^L \delta_A(x-x_j) u dx - u_j^{meas} \right] \delta u dx dt + \varepsilon_1 \int_0^L [u_0^i(x) - u_{0F}^i(x)] \delta u_0^i(x) dx \\ & + \varepsilon_2 \sum_{i=1}^K \int_0^L [\alpha_i(x) - \alpha_{iF}(x)] \delta \alpha_i(x) dx + \varepsilon_3 \sum_{l=1}^H [p(l) - p_F(l)] \delta p(l) \end{aligned} \quad (\text{II.20})$$

The variation of second part, called \mathcal{A} , of equation (II.19) is obtained by the integration by parts technique and given in equation (II.21).

$$\begin{aligned} \delta \mathcal{A} = & \int_0^L \lambda \delta u \Big|_0^T dx - \int_0^T \int_0^L \frac{\partial \lambda}{\partial t} \delta u dx dt + \int_0^T \left[\lambda \frac{\partial f}{\partial u} \right] \delta u \Big|_0^L dt + \sum_{i=1}^K \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \delta \alpha_i(x) \Big|_0^L dt \\ & - \int_0^T \int_0^L \frac{\partial f}{\partial u} \frac{\partial \lambda}{\partial x} \delta u dx dt - \int_0^T \int_0^L \lambda \left[\frac{\partial g_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} g_1 - \varphi_a \frac{\partial g_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} g_2 + \varphi_a \frac{\partial g_2}{\partial p} \right] \delta p dx dt \\ & - \sum_{i=1}^K \int_0^T \int_0^L \frac{\partial f}{\partial \alpha_i(x)} \frac{\partial \lambda}{\partial x} \delta \alpha_i(x) dx dt - \int_0^T \int_0^L \lambda \left[\frac{\partial g_1}{\partial y} - \varphi_a \frac{\partial g_1}{\partial y} + \varphi_a \frac{\partial g_2}{\partial y} \right] \delta y dx dt \\ & - \int_0^T \int_0^L \lambda \left[\frac{\partial g_1}{\partial u} - \varphi_a \frac{\partial g_1}{\partial u} - g_1 \frac{\partial \varphi_a}{\partial u} + \varphi_a \frac{\partial g_2}{\partial u} + g_2 \frac{\partial \varphi_a}{\partial u} \right] \delta u dx dt \end{aligned} \quad (\text{II.21})$$

The same result of the variation of term \mathcal{A} can be derived by using the Green's theorem for double integrals on 2-D domain $[0, T] \times [0, L]$. By using also the integration by parts, the first variation of third part, called \mathcal{B} , is thus written as

$$\begin{aligned} \delta \mathcal{B} = & \left[\gamma \delta y(t) \right] \Big|_0^T - \int_0^T \frac{\partial \gamma}{\partial t} \delta y(t) dt - \int_0^T \gamma \left[\frac{\partial h_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} h_1 - \varphi_a \frac{\partial h_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} h_2 + \varphi_a \frac{\partial h_2}{\partial p} \right] \delta p dx dt \\ & - \int_0^T \gamma \left[\frac{\partial h_1}{\partial y} - \varphi_a \frac{\partial h_1}{\partial y} + \varphi_a \frac{\partial h_2}{\partial y} \right] \delta y dt + \int_0^T \gamma \left[h_1 \frac{\partial \varphi_a}{\partial u} - h_2 \frac{\partial \varphi_a}{\partial u} \right] \delta u dt \end{aligned} \quad (\text{II.22})$$

In order to establish the formulation of the gradients of Lagrangian functional, its variations in equations (II.20), (II.21) and (II.22) are written under the form of Gâteaux derivative of \mathcal{L} with respect to the directions $\delta u(x, t)$, $\delta u_0^i(x)$, $\delta y(t)$, $\delta \alpha_i(x)$ and δp . Denoting the new

vector of all variations directions $\pi = [\delta u(x, t) \quad \delta u_0^i(x) \quad \delta y(t) \quad \delta \alpha_i(x) \quad \delta p]^T$, the Gâteaux derivative of \mathcal{L} in all directions can be rewritten by the following equation:

$$\delta \mathcal{L} = \langle \nabla \mathcal{L}, \pi \rangle \quad (\text{II.23})$$

where $\delta \mathcal{L}$ and $\nabla \mathcal{L}$ are respectively the first variation and the weak form of gradient of \mathcal{L} with respect to $u(x, t)$, $u_0^i(x)$, $y(t)$, $\alpha_i(x)$ and p . The functional space where \mathcal{L} is defined on, $L_2([0, L], [0, T]) \times L_2([0, L]) \times L_2([0, T]) \times L_2([0, L])^K \times \mathbb{R}^H \rightarrow \mathbb{R}$ is equipped with an inner product \langle, \rangle defined as

$$\begin{aligned} \langle \nabla \mathcal{L}, \pi \rangle = & \int_0^T \int_0^L \nabla \mathcal{L}_{u(x,t)} \delta u(x, t) dx dt + \int_0^L \nabla \mathcal{L}_{u_0^i(x)} \delta u_0^i(x) dx + \int_0^T \nabla \mathcal{L}_{y(t)} \delta y(t) dt + \nabla \mathcal{L}_p \delta p \\ & + \int_0^L \nabla \mathcal{L}_{\alpha_i(x)} \delta \alpha_i(x) dx \end{aligned} \quad (\text{II.24})$$

where $\nabla \mathcal{L}$ is a vector of 5 elements $\nabla \mathcal{L} = [\nabla \mathcal{L}_{u(x,t)} \quad \nabla \mathcal{L}_{u_0^i(x)} \quad \nabla \mathcal{L}_{y(t)} \quad \nabla \mathcal{L}_{\alpha_i(x)} \quad \nabla \mathcal{L}_p]$. According to the first-order necessary condition for optimality, for all admissible variation vector π , $\delta \mathcal{L}$ must be equal to zero, or the weak form of gradients $\nabla \mathcal{L}$ must satisfy

$$\nabla \mathcal{L} = 0 \quad (\text{II.25})$$

to eliminate the variations in all directions. Based on the of first variation calculations in equations (II.20), (II.21) and (II.22), to establish the optimization condition of the weak gradient in direction $\delta u(x, t)$ in equation (II.25), all the terms related to $\delta u(x, t)$ are collected as shown in equation (II.26) below:

$$\begin{aligned} \delta \mathcal{L}_{u(x,t)} = & \int_0^T \int_0^L \left\{ -\frac{\partial \lambda}{\partial t} - \frac{\partial f}{\partial u} \frac{\partial \lambda}{\partial x} - \lambda \left[\frac{\partial g_1}{\partial u} - \varphi_a \frac{\partial g_1}{\partial u} + \varphi_a \frac{\partial g_2}{\partial u} - g_1 \frac{\partial \varphi_a}{\partial u} + g_2 \frac{\partial \varphi_a}{\partial u} \right] \right. \\ & \left. + \sum_{j=1}^N \delta_A(x - x_j) \times \left[\int_0^L \delta_A(x - x_j) u dx - u_j^{meas} \right] \right\} \delta u dx dt + \int_0^T \gamma \left[h_1 \frac{\partial \varphi_a}{\partial u} - h_2 \frac{\partial \varphi_a}{\partial u} \right] \delta u dt \end{aligned} \quad (\text{II.26})$$

The second integration concerning the Lagrangian multiplier $\gamma(t)$ can be rewritten under the form of double integration in time and space via an approximation of Dirac delta function $\delta_A(x)$.

$$\int_0^T \gamma \left[h_1 \frac{\partial \varphi_a}{\partial u} - h_2 \frac{\partial \varphi_a}{\partial u} \right] \delta u dt = \int_0^T \int_0^L \gamma \delta_A(x) \left[h_1 \frac{\partial \varphi_a}{\partial u} - h_2 \frac{\partial \varphi_a}{\partial u} \right] \delta u dx dt \quad (\text{II.27})$$

The weak form of gradient of cost functional \mathcal{L} with respect to system variable $u(x, t)$ can be found by merging two double integrations together and setting the term under integration to zero. This is also the adjoint system of Lagrangian variable $\lambda(x, t)$. Its dynamic is finally described by the PDE in equation (II.28) below.

$$-\frac{\partial \lambda}{\partial t} - \frac{\partial f}{\partial u} \frac{\partial \lambda}{\partial x} - \lambda \left[\frac{\partial g_1}{\partial u} - \varphi_a \frac{\partial g_1}{\partial u} + \varphi_a \frac{\partial g_2}{\partial u} - g_1 \frac{\partial \varphi_a}{\partial u} + g_2 \frac{\partial \varphi_a}{\partial u} \right] + \gamma \delta_A(x) \left[h_1 \frac{\partial \varphi_a}{\partial u} - h_2 \frac{\partial \varphi_a}{\partial u} \right] + \sum_{j=1}^N \delta_A(x - x_j) \times \left[\int_0^L \delta_A(x - x_j) u dx - u_j^{meas} \right] = 0 \quad (\text{II.28})$$

which describes the dynamics of the so-called adjoint system of variable $\lambda(x, t)$. This PDE is dual to the direct system. The minus signs in system (II.28) being contrary to the plus signs in system (Σ) represents this duality. The adjoint system of variable $\gamma(t)$ is similarly found by gathering related terms into one equation and removing the time integration, written readily as:

$$-\frac{\partial \gamma}{\partial t} - \gamma \left[\frac{\partial h_1}{\partial y} - \varphi_a \frac{\partial h_1}{\partial y} + \varphi_a \frac{\partial h_2}{\partial y} \right] - \int_0^L \lambda \left[\frac{\partial g_1}{\partial y} - \varphi_a \frac{\partial g_1}{\partial y} + \varphi_a \frac{\partial g_2}{\partial y} \right] dx = 0 \quad (\text{II.29})$$

The first variations of cost functional \mathcal{L} with respect to the variation of state and parameters are found by collecting the related variation terms of corresponding variables $\delta u_0^i(x)$, δp in equations (II.20), (II.21) and (II.22). They are expressed in equations (II.30), (II.31) hereafter, respectively.

$$\delta \mathcal{L}_{u_0^i(x)} = - \int_0^L \lambda(x, 0) \delta u_0^i(x) dx + \varepsilon_1 \int_0^L [u_0^i(x) - u_{0F}^i(x)] \delta u_0^i(x) dx \quad (\text{II.30})$$

$$\delta \mathcal{L}_p = - \int_0^T \int_0^L \lambda \left[\frac{\partial g_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} g_1 - \varphi_a \frac{\partial g_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} g_2 + \varphi_a \frac{\partial g_2}{\partial p} \right] \delta p dx dt - \int_0^T \gamma \left[\frac{\partial h_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} h_1 - \varphi_a \frac{\partial h_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} h_2 + \varphi_a \frac{\partial h_2}{\partial p} \right] \delta p dt + \varepsilon_3 [p - p_F] \delta p \quad (\text{II.31})$$

By using the same method, the variations of parameters $\alpha_i(x)$ is presented in equation (II.32). It is not identical all over the spatial domain.

$$\begin{aligned}\delta\mathcal{L}_{\alpha_i(0)} &= - \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=0} \delta\alpha_i(0) dt ; & \delta\mathcal{L}_{\alpha_i(L)} &= \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=L} \delta\alpha_i(L) dt \\ \delta\mathcal{L}_{\alpha_i(x)} &= - \int_0^T \int_0^L \frac{\partial f}{\partial \alpha_i(x)} \frac{\partial \lambda}{\partial x} \delta\alpha_i(x) dx dt + \varepsilon_2 \int_0^L [\alpha_i(x) - \alpha_{iF}(x)] \delta\alpha_i(x) dx \\ & & & \text{where } x \in]0, L[\end{aligned} \quad (\text{II.32})$$

But if parameter α_i is constant in some special cases, its variation is

$$\delta\mathcal{L}_{\alpha_i} = \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_0^L \delta\alpha_i dt - \int_0^T \int_0^L \frac{\partial f}{\partial \alpha_i} \frac{\partial \lambda}{\partial x} \delta\alpha_i + \varepsilon_2 [\alpha_i - \alpha_{iF}] dx dt \quad (\text{II.33})$$

To solve these equations of Lagrangian multipliers, the boundary and transversality conditions associated with them must be present. These conditions can be found by setting the remaining terms after forming the adjoint systems and the variation of state and parameters to zero to satisfy the first order optimality condition.

$$\int_0^L \lambda(x, T) \delta u(x, T) dx + \int_0^T \left[\lambda \frac{\partial f}{\partial u} \right] \delta u \Big|_0^L dt + \left[\gamma \delta y(t) \right] \Big|_0^T = 0 \quad (\text{II.34})$$

Because the boundary condition of the direct system is fixed beforehand, its variation is thus equal to zero or $\delta u(0, T) = 0$. The variation of initial condition of variable y of the ODE is also equal to zero because of its predetermination. To establish equation (II.34), the following equalities must hold $\lambda(x, T) = 0$; $\lambda(L, t) = 0$ and $\gamma(T) = 0$ which are as well the limit time condition and boundary condition for the two adjoint models (II.28) and (II.29). The initial and boundary conditions of the adjoint equation are dual to these conditions of the direct systems.

The formulation of each element of $\delta\mathcal{L}$ in equations (II.30), (II.31) and (II.32) is rewritten according to the previous definition of inner product \langle, \rangle in (II.24), and presented as

follow:

$$\begin{aligned} \delta \mathcal{L}_p &= \left\{ - \int_0^T \int_0^L \lambda \left[\frac{\partial g_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} g_1 - \varphi_a \frac{\partial g_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} g_2 + \varphi_a \frac{\partial g_2}{\partial p} \right] dx dt \right. \\ &\quad \left. - \int_0^T \gamma \left[\frac{\partial h_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} h_1 - \varphi_a \frac{\partial h_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} h_2 + \varphi_a \frac{\partial h_2}{\partial p} \right] dt + \varepsilon_3 [p - p_F] \right\} \delta p \\ \delta \mathcal{L}_{u_0^i(x)} &= \int_0^L \left\{ - \lambda(x, 0) dx + \varepsilon_1 [u_0^i(x) - u_{0F}^i(x)] \right\} \delta u_0^i(x) dx \end{aligned}$$

$$\begin{cases} \delta \mathcal{L}_{\alpha_i(0)} &= \int_0^L \frac{1}{L} \left\{ - \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=0} dt \right\} \delta \alpha_i(0) dx \\ \delta \mathcal{L}_{\alpha_i(L)} &= \int_0^L \frac{1}{L} \left\{ \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=L} dt \right\} \delta \alpha_i(L) dx \\ \delta \mathcal{L}_{\alpha_i(x)} &= \int_0^L \left\{ - \int_0^T \frac{\partial f}{\partial \alpha_i(x)} \frac{\partial \lambda}{\partial x} dt + \varepsilon_2 [\alpha_i(x) - \alpha_{iF}(x)] \right\} \delta \alpha_i(x) dx \quad \text{where } x \in]0, L[\end{cases}$$

As a result, the weak form of gradients of the cost functional with respect to the state and parameters are written readily as:

$$\begin{aligned} \nabla \mathcal{L}_p &= - \int_0^T \int_0^L \lambda \left[\frac{\partial g_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} g_1 - \varphi_a \frac{\partial g_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} g_2 + \varphi_a \frac{\partial g_2}{\partial p} \right] dx dt + \varepsilon_3 [p - p_F] \\ &\quad - \int_0^T \gamma \left[\frac{\partial h_1}{\partial p} - \frac{\partial \varphi_a}{\partial p} h_1 - \varphi_a \frac{\partial h_1}{\partial p} + \frac{\partial \varphi_a}{\partial p} h_2 + \varphi_a \frac{\partial h_2}{\partial p} \right] dt \end{aligned} \quad (\text{II.35})$$

$$\nabla \mathcal{L}_{u_0^i(x)} = -\lambda(x, 0) + \varepsilon_1 [u_0^i(x) - u_{0F}^i(x)] \quad (\text{II.36})$$

$$\begin{cases} \nabla \mathcal{L}_{\alpha_i(0)} &= -\frac{1}{L} \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=0} dt ; \quad \mathcal{L}_{\alpha_i(L)} = \frac{1}{L} \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=L} dt \\ \nabla \mathcal{L}_{\alpha_i(x)} &= - \int_0^T \frac{\partial f}{\partial \alpha_i(x)} \frac{\partial \lambda}{\partial x} dt + \varepsilon_2 [\alpha_i(x) - \alpha_{iF}(x)] \quad \text{where } x \in]0, L[\end{cases} \quad (\text{II.37})$$

And for the case of constant α_i

$$\nabla \mathcal{L}_{\alpha_i} = \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_0^L dt - \int_0^T \int_0^L \frac{\partial f}{\partial \alpha_i} \frac{\partial \lambda}{\partial x} dx dt + \varepsilon_2 [\alpha_i - \alpha_{iF}] \quad (\text{II.38})$$

4 Numerical approach

4.a Introduction

PDEs can describe the dynamics of a vast amount of mathematical models for physical, biological, hydrological and chemical phenomena. They have been extended in order to deal with the systems in more recent fields including economic, weather forecasting etc. The majority of PDEs falls into three types: the hyperbolic equations usefully associated to the advection problems, the parabolic equations related to the diffusion phenomena and the elliptic equations used to characterize the steady state of two above equation types. Most of the research related to the PDEs needs the solution of the equation to carry out their study. Except in some simple system model or simplified cases, an exact solution of PDE is difficult to be obtained. Specially, in the case of nonlinear and complex system, the demand of approximation of analytical solution by the numerical one is critical. But in other simpler cases, they can be combined together to give a better analysis of system characteristics. Because of the non-linearity and complexity of PDE governing both the system model and the adjoint systems in this work, the only way to solve them is using numerical technique due to the impossibility of finding analytical one. Through several decades of development, there is a large number of methods allowing to approximate the solutions of the three types of PDE which can be classified into some main approaches: the finite-difference approach uses approximation technique to approximate the derivatives, the method of lines discretizes the PDE in one dimension into a continuous and interdependent set of ordinary differential equations (ODE) which are then solved by simple numerical method, the finite-element method divides the original equation into a set of finite elements (smaller parts) and applies the calculus of variations to minimize the error function. For complex geometry equation with high demand of accuracy, some advanced methods can be applied such as multi-grid method, finite volumes, spectral method, domain decomposition method. In this large number of possible choice of numerical method, the chosen one not only depends on the type of governing equation but also the number of physical dimensions, the involved coordinates, and the type of associated boundary conditions. Among them, the finite difference method is the simplest one and easiest to implement with a good accuracy. This is the reason why this numerical method is used in this study.

The choice of finite difference scheme depends on the type of PDE needing to be solved. Some methods used to approximate the solution of hyperbolic PDE are upwind method of characteristics, Lax–Friedrichs, Lax–Wendroff, MacCormack, Preissmann and etc. Moreover, a finite difference method can be basically arranged into two mains classes: implicit and explicit methods. While the explicit one allows to get all unknown values at a point

on discretized grid directly from its previous value, the implicit one provides only a set of equations all over the grid which must be solved to get the solutions. But the implicit schemes have the strong characteristic to be always stable for all step sizes. The stability of explicit schemes must be guaranteed by some condition like the Courant condition for wave equation. Both of these two numerical scheme families will be presented and analyzed in the next sections with their typical representatives: Preissmann scheme and Lax-Wendroff scheme, respectively for implicit and explicit methods.

In order to make the following development of Preissmann scheme and Lax-Wendroff scheme be suitable for both the direct system equation and adjoint model, their PDEs are generalized and described by the following form of general hyperbolic equation:

$$\frac{\partial \vartheta(x, t)}{\partial t} + \frac{\partial \chi(\vartheta(x, t), x)}{\partial x} = \varrho(x, t) \quad (\text{II.39})$$

where $\vartheta(x, t)$ is the system variable, $\chi(\vartheta(x, t), x)$ is the flow and $\varrho(x, t)$ is the non homogeneous source term. This simple equation can be considered as the representation of all the PDE of system (Σ) and the adjoint equation of $\lambda(x, t)$ (II.28).

The numerical solution for the ODE which is the governing equation of the adjoint variable $\gamma(t)$ (II.29) is much more simple than the partial one. The ordinary differential equation (ODE) is the differential equation used to describe the relation between the functions of only one dependent variable ("ordinary" instead of "partial") with their derivatives. It arises in many contexts of social and also natural domains. The dynamics of various phenomena can be described by this type of equation such as weather modeling reaction rates, population modeling and the market equilibrium price changes. The exact analytic solution can be usually derived in the case linear ODE (under the series or integral form) while the numerical techniques are necessary to approximate the exact solution by a numerical one in the case of nonlinear ODE. The numerical methods for the ODE can be arranged into three classes according to its order. The first order numerical schemes contain the family of Euler method (forward, backward, semi-implicit and exponential). The second order method provides more accurate schemes: trapezoidal rule, midpoint method, Newmark-beta method. The last class is the higher order method including exponential integrator, Runge-Kutta methods and linear multi-step method. The ODE in equation (II.29) is a first order one and can be reformulated as follows:

$$\begin{cases} \frac{d\gamma(t)}{dt} = f(\gamma(t)) \\ \gamma(T) = \gamma_T \end{cases} \quad (\text{II.40})$$

Among the mentioned numerical scheme for ODE, the midpoint method (see [SM03]) is chosen to solve equation (II.40) because of its balance between simplicity and accuracy.

4.b Midpoint scheme

The system (II.40) is discretized on the time line $[0, T]$ with time step Δt . The time derivative, in Euler method, is approximated by the finite difference at time t as

$$\frac{d\gamma(t)}{dt} \approx \frac{\gamma(t) - \gamma(t - \Delta t)}{\Delta t} \quad (\text{II.41})$$

Replacing this equation by the derivative at a half time step $t - \Delta t/2$ gives more accuracy:

$$\frac{d\gamma\left(t - \frac{\Delta t}{2}\right)}{dt} \approx \frac{\gamma(t) - \gamma(t - \Delta t)}{\Delta t} \quad (\text{II.42})$$

and thus

$$\gamma(t - \Delta t) \approx \gamma(t) - \Delta t f(\gamma(t - \Delta t/2)) \quad (\text{II.43})$$

The value of γ at a half backward time step $t - \Delta t/2$ is obtained by using the Taylor expansion:

$$\gamma(t - \Delta t/2) \approx \gamma(t) - \frac{\Delta t}{2} \frac{d\gamma(t)}{dt} = \gamma(t) - \frac{\Delta t}{2} f(\gamma(t)) \quad (\text{II.44})$$

Plugging this equation to (II.43) results in the final equation allowing to calculate backwardly the value $\gamma(t - \Delta t)$:

$$\gamma(t - \Delta t) \approx \gamma(t) - \Delta t f\left(\gamma(t) - \frac{\Delta t}{2} f(\gamma(t))\right) \quad (\text{II.45})$$

By using the notation $\gamma(i)$ to shorten $\gamma(t + i \Delta t)$, equation (II.45) is rewritten as follow:

$$\gamma(i - 1) \approx \gamma(i) - \Delta t f\left(\gamma(i) - \frac{\Delta t}{2} f(\gamma(i))\right) \quad (\text{II.46})$$



Figure II.1: Stencil of midpoint scheme

From the final time condition at time T (being equal to γ_T) denoted by the solid circle, the value of $\gamma(t)$ is calculated backwardly from the right of time line to the left. Figure II.1 presents the stencil of this scheme. Through its development, the term "midpoint" in the name refers to the mediate evaluation of function value at position $t - \Delta t/2$, which is located between the current point t and the point needing to be evaluated $t - \Delta t$. This one step method has the local error of order 2 and converges faster than the Euler method with a small value of Δt .

4.c Preissmann scheme

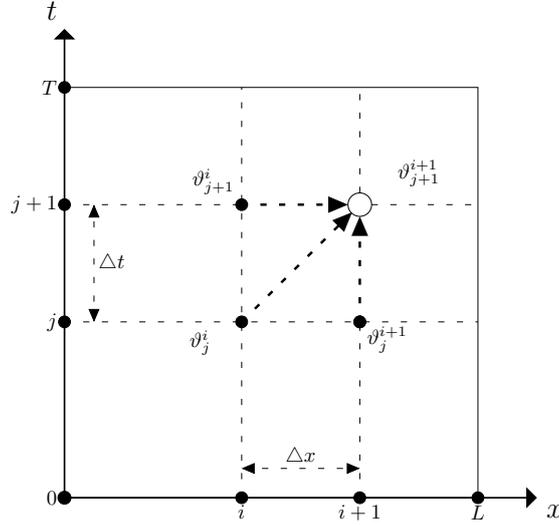


Figure II.2: Stencil of forward Preissmann scheme

Being introduced in 1960s by Preissmann [Pre61], this implicit scheme has been widely used to numerically discretize the one-dimension unsteady flows and overland flow because of its simplicity and effectiveness. Moreover, this weighted four-point implicit finite difference can be extended to the case of general hyperbolic equation (II.39). The main idea of this scheme is that a the value of variable $\vartheta(x, t)$ and its partial derivative with respect to x and t is approximated by four discretization points in space and time with suitable and chosen weighting factors α, β as presented in equation (II.47).

$$\begin{aligned} \frac{\partial \vartheta(x, t)}{\partial x} &= \alpha \frac{\vartheta_{j+1}^{i+1} - \vartheta_{j+1}^i}{\Delta x} + (1 - \alpha) \frac{\vartheta_j^{i+1} - \vartheta_j^i}{\Delta x} \\ \frac{\partial \vartheta(x, t)}{\partial t} &= \beta \frac{\vartheta_{j+1}^{i+1} - \vartheta_j^{i+1}}{\Delta t} + (1 - \beta) \frac{\vartheta_{j+1}^i - \vartheta_j^i}{\Delta t} \\ \vartheta(x, t) &= \alpha \frac{\vartheta_{j+1}^{i+1} + \vartheta_j^{i+1}}{2} + (1 - \alpha) \frac{\vartheta_j^{i+1} + \vartheta_j^i}{2} \end{aligned} \quad (\text{II.47})$$

where $\Delta t =$ time step; $\Delta x =$ space step; $\alpha, \beta =$ weighting factors with $0 \leq \alpha, \beta \leq 1$; $\vartheta_j^i = \vartheta(x + i \Delta x, t + j \Delta t)$. Figure II.2 describes the stencil of Preissmann forward difference. The initial and boundary conditions are fixed at the limits of the $x - t$ plan and depicted by the solid circle at two lines $x = 0 \forall t \in [0, T]$ and $t = 0 \forall x \in [0, L]$. From these initial values, the system is simulated forward in space then forward in time. The known values on the $x - t$ grid are presented by the solid circle while the white circle represents the unknown one, value of variable $\vartheta(x, t)$ at position (x_{i+1}, t_{n+1}) . It will be calculated

based on its values at three known neighbors ϑ_j^i , ϑ_{j+1}^i and ϑ_j^{i+1} . By replacing the value of $\vartheta(x, t)$ and its partial derivatives in equation (II.47) into the general system, an implicit equation of variable ϑ_{j+1}^{i+1} can be obtained. This nonlinear equation is then solved by the iterative Newton–Raphson method to get the solution of $\vartheta(x, t)$ at grid point (x_{i+1}, t_{n+1}) .

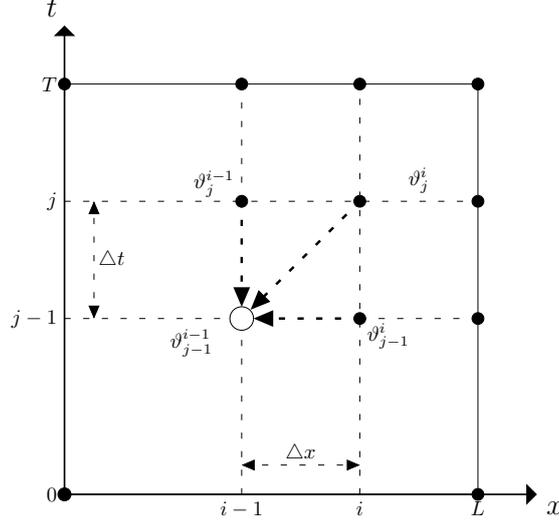


Figure II.3: Stencil of backward Preissmann scheme

But one can notice that this scheme is used for forward simulation only. Due to its boundary and final time condition, the adjoint equation of variable $\lambda(x, t)$ must be solved numerically backwardly in time and space. This means the above Preissmann scheme has to be modified to adapt to this situation. The stencil of backward Preissmann scheme is given in Figure II.3. And the equations of backward Preissmann are as follows:

$$\begin{aligned} \frac{\partial \vartheta(x, t)}{\partial x} &= \alpha \frac{\vartheta_j^i - \vartheta_j^{i-1}}{\Delta x} + (1 - \alpha) \frac{\vartheta_{j-1}^i - \vartheta_{j-1}^{i-1}}{\Delta x} \\ \frac{\partial \vartheta(x, t)}{\partial t} &= \beta \frac{\vartheta_{j-1}^i - \vartheta_{j-1}^{i-1}}{\Delta t} + (1 - \beta) \frac{\vartheta_j^i - \vartheta_{j-1}^i}{\Delta t} \\ \vartheta(x, t) &= \alpha \frac{\vartheta_j^i + \vartheta_j^{i-1}}{2} + (1 - \alpha) \frac{\vartheta_{j-1}^i + \vartheta_{j-1}^{i-1}}{2} \end{aligned} \quad (\text{II.48})$$

More details and discussion about this implicit backward scheme can be found in the studies of M. T. Shamaa et al. [SK11] and F. Liu et al. [LFB92]. Another important reference for the Preissmann scheme is the book of J. Cunge et al. [CHV80].

Through the stability analysis, this scheme is always stable for any time step Δt if coefficient α is chosen in the interval $[0.5, 1]$. A good accuracy of approximation can be achieved if the value of α belongs to the specified interval $[0.6, 0.75]$ according to the study of

H. Chanson [Cha04]. This characteristic allows to use a large time step Δt without any stability problem. This is a good point because the iterative Newton–Raphson method for each equation at all the grid mesh is time consuming (one of its largest disadvantage) with large number of space and time discretization points.

4.d Lax-Wendroff scheme

Instead of the implicit Preissmann scheme, which was used in the beginning of the work of this thesis, the explicit Lax-Wendroff scheme is employed to overcome the disadvantages of the firstly used one. Being presented in 1960 by P.D. Lax and B. Wendroff [LW60], this conservative scheme is one of the most used numerical method for solving the systems of one dimension hyperbolic conservation laws with second order of accuracy in both space and time. There are various ways to find the formula of this scheme. The Taylor series expansion is among of them. Because the dynamics of direct system are well-defined by fixing the initial condition at time $t = 0$ and boundary condition at position $x = 0$, the system equation will be solved forward in time and space. On the contrary, the adjoint equation will be simulated backward in both time and space due to its conditions are specified at the limits of $x-t$ plane. In the case of forward simulation, the Taylor expansion in time of $\vartheta(x, t)$ is written as

$$\vartheta(x, t + \Delta t) \approx \vartheta(x, t) + \Delta t \frac{\partial \vartheta}{\partial t} + \frac{\Delta t^2}{2} \frac{\partial^2 \vartheta}{\partial t^2} + \mathcal{O}(\Delta t^3) \quad (\text{II.49})$$

The first and second order partial derivative with respect to t of $\varphi(x, t)$ can be found and replaced by the space derivatives by taking the time derivative of both left and right part of equation (II.39) and changing the derivative order.

$$\begin{aligned} \frac{\partial \vartheta}{\partial t} &= -\frac{\partial \chi}{\partial x} + \varrho \\ \frac{\partial^2 \vartheta}{\partial t^2} &= -\frac{\partial}{\partial t} \left[-\frac{\partial \chi}{\partial x} + \varrho \right] + \frac{\partial \varrho}{\partial t} = -\frac{\partial}{\partial x} \left[-\frac{\partial \chi}{\partial x} + \varrho \right] - \frac{\partial \chi}{\partial \varphi} \left[-\frac{\partial^2 \chi}{\partial x^2} + \frac{\partial \varrho}{\partial x} \right] + \frac{\partial \varrho}{\partial t} \end{aligned} \quad (\text{II.50})$$

After neglecting the higher order term $\mathcal{O}(\Delta t^3)$, the Taylor expansion can be rewritten by inserting the terms in equation (II.50) into equation (II.49). Its final form can be found in equation (II.51)

$$\vartheta(x, t + \Delta t) \approx \vartheta(x, t) + \Delta t \left[-\frac{\partial \chi}{\partial x} + \varrho \right] + \frac{\Delta t^2}{2} \left\{ -\frac{\partial}{\partial x} \left[-\frac{\partial \chi}{\partial x} + \varrho \right] - \frac{\partial \chi}{\partial \varphi} \left[-\frac{\partial^2 \chi}{\partial x^2} + \frac{\partial \varrho}{\partial x} \right] + \frac{\partial \varrho}{\partial t} \right\} \quad (\text{II.51})$$

In previous equation, the time and space derivatives are replaced by their central finite difference approximations to get the second accuracy order to get the computational equation of Lax-Wenroff scheme for each point on the discretization mesh, as in equation(II.52).

$$\begin{aligned} \vartheta(i, n+1) \approx & \vartheta(i, n) + \Delta t \left[-\frac{\chi(i+1, n) - \chi(i-1, n)}{2 \Delta x} + \varrho(i, n) \right] \\ & + \frac{\Delta t^2}{2} \left\{ -\frac{\varsigma(i+1, n) - \varsigma(i-1, n)}{2 \Delta x} \left[-\frac{\chi(i+1, n) - \chi(i-1, n)}{2 \Delta x} + \varrho(i, n) \right] \right. \\ & - \varsigma(i, n) \left[-\frac{\chi(i+1, n) - 2\chi(i, n) + \chi(i-1, n)}{\Delta x^2} + \frac{\varrho(i+1, n) - \varrho(i-1, n)}{2 \Delta x} \right] \\ & \left. + \frac{\varrho(i, n+1) - \varrho(i, n-1)}{2 \Delta t} \right\} \end{aligned} \quad (\text{II.52})$$

The partial derivative of flow χ with respect to the variable ϑ is denoted by ς . The notation $\vartheta(x+i\Delta x, t+n\Delta t)$ is shortened to $\vartheta(i, j)$. The notations of other terms are used with the same meaning. The value of ϑ at the next time step are approximated from three values of its neighbor at current time. The stencil of this scheme is depicted in figure II.4 where the initial and boundary conditions along with the 3 known points are presented by the solid circles. The value of $\vartheta(i, j+1)$ needs to be updated is the white bigger one.

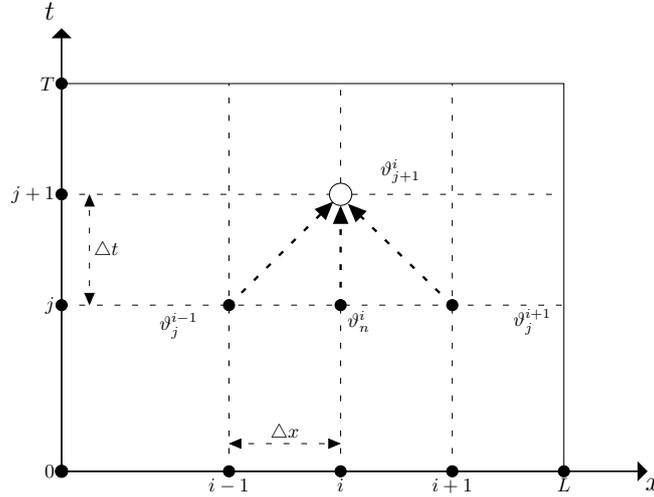


Figure II.4: Stencil of forward Lax-Wendroff scheme

Since the adjoint system in equation (II.28) of variable $\lambda(x, t)$ is conjugated with the direct one, the similar approach can be used for discretization. Again, it must be simulated backwardly. Consequently, the Taylor series expansion is modified to adapt to this new

situation:

$$\vartheta(x, t - \Delta t) \approx \vartheta(x, t) - \Delta t \frac{\partial \vartheta}{\partial t} + \frac{\Delta t^2}{2} \frac{\partial^2 \vartheta}{\partial t^2} + \mathcal{O}(\Delta t^2) \quad (\text{II.53})$$

Through the same previous technique for direct system, the final computational of backward Lax-Wendroff is given by equation (II.54) with its associated stencil depicted in figure II.5.

$$\begin{aligned} \vartheta(i, n-1) \approx & \vartheta(i, n) - \Delta t \left[-\frac{\chi(i+1, n) - \chi(i-1, n)}{2 \Delta x} + \varrho(i, n) \right] \\ & + \frac{\Delta t^2}{2} \left\{ -\frac{\varsigma(i+1, n) - \varsigma(i-1, n)}{2 \Delta x} \left[-\frac{\chi(i+1, n) - \chi(i-1, n)}{2 \Delta x} + \varrho(i, n) \right] \right. \\ & - \varsigma(i, n) \left[-\frac{\chi(i+1, n) - 2\chi(i, n) + \chi(i-1, n)}{\Delta x^2} + \frac{\varrho(i+1, n) - \varrho(i-1, n)}{2 \Delta x} \right] \\ & \left. + \frac{\varrho(i, n+1) - \varrho(i, n-1)}{2 \Delta t} \right\} \end{aligned} \quad (\text{II.54})$$

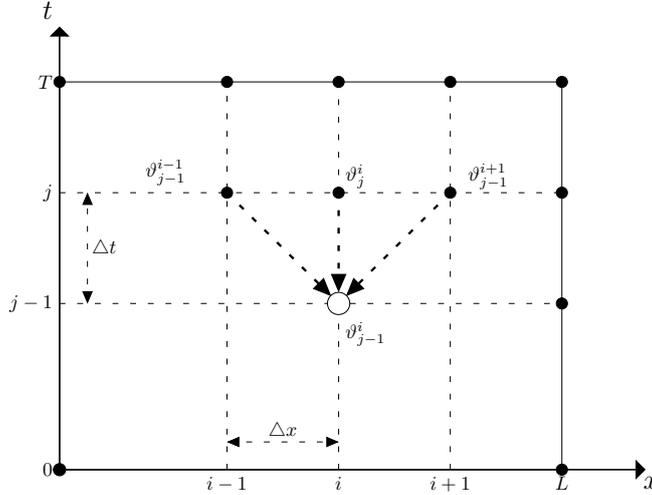


Figure II.5: Stencil of backward Lax-Wendroff scheme

The von Neumann stability analysis shows that the linear stability of this second order accuracy in both time and space is guaranteed by the Courant-Friedrichs-Lewy condition [CFL67], which is given by

$$|\vartheta_{max}| \frac{\Delta t}{\Delta x} \leq C_{max} \quad (\text{II.55})$$

where the coefficient C_{max} for this scheme is 1.

4.e Discrete gradients and remarks

4.e.1 Discrete gradients

Because both of the switched direct systems and adjoint systems are discretized by the previous numerical methods on the space and time grid, in order to be solved numerically, the continuous gradients of cost functional presented in equations (II.32), (II.30) and (II.31) must be consequently discretized for each discrete spatial point x_j with $\forall x_j = j \Delta x$ where $0 \leq j \leq M$ with $\Delta x = L/M$, M being the number of point. As a result, the gradient of initial state becomes

$$\nabla \mathcal{L}_{u_0^i(x_j)} = -\lambda(x_j, 0) + \varepsilon_1 [u_0^i(x_j) - u_{0F}^i(x_j)] \quad (\text{II.56})$$

However, the gradients of parameter α_i are not identical on the spatial grid. Their value at the limits (the beginning and the end of section $[0, L]$ in equations (II.57) and (II.58)) are different from the interior of section $[0, L]$, in equation (II.59).

$$\nabla \mathcal{L}_{\alpha_i(x_j)} = -\frac{1}{L} \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=x_j} dt \quad \text{with } x_j = 0 \quad (\text{II.57})$$

$$\nabla \mathcal{L}_{\alpha_i(x_j)} = \frac{1}{L} \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_i(x)} \right] \Big|_{x=x_j} dt \quad \text{with } x_j = L \quad (\text{II.58})$$

$$\nabla \mathcal{L}_{\alpha_i(x_j)} = - \int_0^T \frac{\partial f}{\partial \alpha_i(x)} \frac{\partial \lambda}{\partial x} \Big|_{x=x_j} dt + \varepsilon_2 [\alpha_i(x) - \alpha_{iF}(x)] \quad (\text{II.59})$$

with $\forall x_j = j \Delta x$ where $1 \leq j \leq M - 1$

On the other hand, the gradients of parameter p are the same with equation (II.35). They do not change the form because of their independence with x .

4.e.2 Remarks

The distributed parameters and state are discretized on the spatial domain into $M + 1$ sections. As previously recalled, the boundary condition of adjoint equation is fixed to zero or $\lambda(L, t) = 0$. This null value of $\lambda(x, t)$ at position L for all time domain leads to

$$\nabla \mathcal{L}_{\alpha_i(L)} = 0$$

and

$$\nabla \mathcal{L}_{u_0^i(x_j)} = \varepsilon_1 [u_0^i(L) - u_{0F}^i(L)]$$

The gradients of $\alpha_i(L)$ are equal to zero which prevents the algorithm from estimating the optimal value of $\alpha_i(L)$. For the same reason, the gradient of cost functional with respect to the last value of initial condition, $\nabla \mathcal{L}_{u_0^i(L)}$ depends only on first guessed $u_{0F}^i(L)$ value and calibration term ε_1 . As a result, this gradient is small when being compared to other positions that causes difficulties for the operation of optimization process, specially the gradient-based methods. This issue is caused by the use of the Gaussian approximation of Dirac function as a continuous measurement function in equation (II.3). The present formulation based on the approximated Dirac always gives the same boundary condition for PDE of $\lambda(x, t)$. This condition does not depend on the position of measurement. Or there is no sensitivity of the gradients with the measurement at the end of spatial domain, $u_j^{meas}(L, t)$. By replacing this approximation with the real Dirac function or the discrete measurements, the sensitivity of $\lambda(L, t)$ with the measurements can be recovered but the present formulation of the objective function \mathcal{J} and adjoint equations must be changed to adapt to the new sensor operator. The cost function \mathcal{J} is thus separately defined for each case of measurement at the interior locations or at the limit point L of the spatial domain $[0, L]$. For example, if there are measurements at position L denote $u_j^{meas}(L, t)$, the form of cost function is written as

$$\mathcal{J} = \frac{1}{2} \int_0^T \left\{ u(L, t) - u_j^{meas}(x_j, t) \right\}^2 dt \quad (\text{II.60})$$

All the calibration terms and first guessed values are neglected for simplicity without any loss of generality. By using the same calculus of variations approach, the dynamics of multiplier $\lambda(x, t)$ satisfy the equation below

$$\begin{aligned} -\frac{\partial \lambda}{\partial t} - \frac{\partial f}{\partial u} \frac{\partial \lambda}{\partial x} - \lambda \left[\frac{\partial g_1}{\partial u} - \varphi_a \frac{\partial g_1}{\partial u} + \varphi_a \frac{\partial g_2}{\partial u} - g_1 \frac{\partial \varphi_a}{\partial u} + g_2 \frac{\partial \varphi_a}{\partial u} \right] \\ + \gamma \delta_A(x) \left[h_1 \frac{\partial \varphi_a}{\partial u} - h_2 \frac{\partial \varphi_a}{\partial u} \right] = 0 \end{aligned} \quad (\text{II.61})$$

The equation of second Lagrangian multiplier $\gamma(t)$ and the gradient of initial state and parameters remain unchanged except the boundary condition of $\lambda(x, t)$. Its value is directly equal to the discrepancies between simulated and measured values

$$\lambda(L, t) = u(L, t) - u_j^{meas}(L, t)$$

which allows to have the effect of the last measurement of the domain on the adjoint values. The observation information from the boundary condition is propagated all over the grid

of $\lambda(x, t)$ due to the backward simulation that brings the sensitivity to all the positions before the last position L .

The present work only focuses on the application of continuous δ_A to investigate the adjoint based estimation method, and the use of discrete measurement will be an open topic for our future studies. The sensor locations are limited in the interior section of spatial domain $[0, L]$ and the corresponding discrete position for state and distributed parameter estimation are reduced to the set $\{x_j = j \Delta x\}$ where $2 \leq j \leq M - 1$.

5 Optimization algorithm

5.a Introduction

In order to select the most appropriate algorithm to solve the mentioned optimization issue, it is worth classifying them according to their characteristics. This is an important and necessary step because a particular type of problem has a tailored algorithm. There are many points of view for classification. Depending on the variables set which can be discrete or continuous, an optimization problem can be a discrete or continuous one which need different algorithms. The smoothness of the model will provide useful information for optimization solver because the function value at one point can be deduced from its neighbors. The continuous optimization issue is thus easier to solve than the discrete one. The optimization problem can be arranged by considering its number of objective functions. Most of them have only one objective. There are also optimization issues with none or many objectives. In the first case, it plays the role of a feasibility problem where the optimal solution will satisfy only the constraints. The multi-objective optimization problem is also a common problem in many domains including economics, engineering. The found solutions must satisfy all the objectives and the constraints. To reduce the complexity of optimizing process, some objectives functions are transformed into constraints or use weighted combinations to connect the objectives. Beside the objective function, the optimization algorithm must deal also with the constraints on the optimizing variable if they exist. This is called as constrained optimization problems. The constraints which describe the relationships between the variable, can be simple bounds (constrained optimization), linear functions (linear constraint programming) or maybe nonlinear functions (nonlinear constraint programming). The unconstrained optimization issues is not as frequent as the constrained one because it usually arises when the constraints are transformed to penalties in the objective function. The last framework to classify an optimization problem is the deterministic or probabilistic (stochastic) point of view. The deterministic optimization means that the models used for optimization are fully determined by the values of initial condition and parameters and no noised or disturbance present in the system or the

data contains some information about the future. In contrary, stochastic optimization concerned the case when the data possess some inherent randomness or uncertainties. Some advanced algorithms can be used to explicitly incorporate the uncertainty such as robust optimization or stochastic programming.

By comparing with previous analysis, the optimal estimation considered hereby can be put to the class of nonlinear continuous constrained deterministic optimization problem with one objective function. Augmented Lagrangian method is a typical method that can be used to solve this problem. The idea behind this method is that the original constrained issue is replaced by an unconstrained one with only bound constraints through using of the Lagrangian multiplier. It is also called the method of multipliers. The added virtual variable is iteratively updated along with the augmented Lagrangian functional in a successive minimization process. There are several issues involved in solving the optimization of new objective functional. One of the important key to be considered is whether to use gradient-free optimization approach or gradient-based approach. Both of them have its own advantages and disadvantages. The choice between them is problem-dependent. One of the strengths of gradient oriented algorithm is its fast convergence property compared to other methods thanks to the exploitation of gradients which give the direction to the optimal solution. However, this optimization approach contains also some weaknesses. It usually gives the a optimal local solution rather than the global one. It is also time consuming while developing the code to calculate the gradient of a complicated system. Moreover, the code and calculation must be changed to fit to new configuration if there is any change. This approach is also very sensitive to the uncertainty or the noisy data and disturbances added into the cost function. The gradient can be calculated by several way including finite-different method or through adjoint state.

In this study, after the step of adjoint based gradient method, the gradient of cost functional with respect to the state and parameters are found and will be served as the inputs of the optimization algorithm. While the gradients of objective functional with respect to the state and parameters are found, the remaining question is how to use them correctly. The gradient of the cost functional which belongs to the class of twice continuously differentiable functional is perpendicular to the contour of this functional and represents the direction of maximal increase. This gives the idea that the minimum (maybe local) of this functional can be found if one goes in the opposite direction compared to the gradient. This is the basic idea of the steepest descent method which is an important subclass of the direct search algorithm, a more general optimization approach which has been developed over several decades for constrained optimization. In a wider situation, the direct search, from an initial estimate of optimum point, generates a sequence of the next estimate points by seeking directly from each previous point in the descent directions which ensures the next function value will be lower than the previous one. The stopping condition usually

concerns the norm of first necessary condition of optimality (null value of gradient vector) or norm of two consecutive function values. Depending on the way to choose the descent direction, there are many sub classes of direct search method mainly including the first order line search descent methods and the second order ones. The first order term in its name reflects the first partial derivative of function value which is used to determine the search direction at each iteration. The mentioned steepest descent method belongs to the first order algorithm with the conjugate gradient method. The modified Newton method and Quasi-Newton method which belong to the second order class, use the Hessian matrix, second order partial derivative to improve the convergence.

Two of these methods, the steepest descent and Quasi-Newton which represent the characteristics of their classes will be presented in the next subsections. Both of them have been used to solve the optimization problems presented in the next chapter of estimation application. Their advantages and disadvantages are also discussed.

5.b Steepest descent method

At the beginning of this work, the optimization problem was solved by using the steepest descent method, one of the simplest and most famous first order line search method, proposed by Cauchy in 1847. This method is used in some publications [NGB14]; [NGB15a]; [NGB15b]. The search directions to update the optimal point from the initial guessed one are calculated based on the gradient of objective functional with respect to the estimated state and parameters, which are presented in the calculation of adjoint method in section 3.c. The general structure of a line search descent method is firstly considered and it is given in details in Algorithm 1 where the notations vector π contains the state and parameters needing to be estimated and $\nabla\mathcal{L}$ denotes the gradient vector of \mathcal{L} . As the discussion in the beginning of subsection 5, there are two important points in the general algorithm of gradient descent: how to choose the gradient based descent direction at each iteration u^i and how to perform the line search technique or solve the sub-optimization problem (II.62).

The method of steepest descent uses the negative value of normalized gradient as the search direction. The expression of this direction is derived via using the Schwartz's equality. At the point π , by supposing that there exists a unit descent direction u such that the directional derivative

$$d\mathcal{L}(\pi; u)|_{\eta} = \frac{d\Upsilon(0)}{d\eta} = \nabla^T \mathcal{L}(\pi)u \quad (\text{II.64})$$

where the function Υ is defined as $\Upsilon(\eta) = \mathcal{L}(\pi + \eta u)$. Via the Schwartz's equality:

$$\nabla^T \mathcal{L}(\pi)u \geq -\|\nabla^T \mathcal{L}(\pi)\| \|u\| = -\|\nabla^T \mathcal{L}(\pi)\| \quad (\text{II.65})$$

Goal The optimal values of π^* .

- $STOP = \text{false}$.
- Define the positive tolerances ξ_1 , ξ_2 and ξ_3 .
- Set iteration variable $i = 1$.

while $STOP$ is false **do**

- Simulate forwardly in time and in space the direct systems.
- Simulate backwardly in time and in space the adjoint models.
- Select descent direction u^i based on gradient $\nabla\mathcal{L}(\pi^{i-1})$.
- Calculate the descent step at iteration i η^i by perform a one-dimensional line search in direction u^i . Or solve the sub-optimization problem of new function Υ as follows:

$$\min_{\eta} \Upsilon(\eta) = \min_{\eta} \mathcal{L}(\pi^{i-1} + \eta u^i) \quad (\text{II.62})$$

The solution of the optimization problem (II.62) is denoted η^i

- Update the new value of π^{i+1} in direction u^i with step length η^i

$$\pi^i = \pi^{i-1} + \eta^i u^i \quad (\text{II.63})$$

- **if** $\|\pi^i - \pi^{i-1}\| \leq \xi_1$ or $\|\nabla\mathcal{L}(\pi^{i-1})\| \leq \xi_2$ or $\|\mathcal{J}(\pi^i) - \mathcal{J}(\pi^{i-1})\| \leq \xi_3$. **then**

- | • $STOP = \text{true}$.

else

- | • Set $i = i + 1$.

end

end

Algorithm 1: General gradient descent method algorithm

where $-\|\nabla^T\mathcal{L}(\pi)\|$ is the least value. By choosing the unit direction $u = \frac{-\nabla^T\mathcal{L}(\pi)}{\|\nabla^T\mathcal{L}(\pi)\|}$, the directional derivative at u satisfies the below equality:

$$\frac{d\Upsilon(0)}{d\eta} = -\nabla^T\mathcal{L}(\pi) \frac{-\nabla^T\mathcal{L}(\pi)}{\|\nabla^T\mathcal{L}(\pi)\|} = -\|\nabla^T\mathcal{L}(\pi)\| \quad (\text{II.66})$$

As a result, at the iteration i , the chosen $u^i = \frac{-\nabla^T\mathcal{L}(\pi^{i-1})}{\|\nabla^T\mathcal{L}(\pi^{i-1})\|}$ is the normalized direction of steepest descent used to updated the current value of π .

The step size η^i is the optimal solution of (II.62) which allows to minimize the cost functional in the chosen descent direction. In fact, the step size η^i can be selected to be constant for all iteration i , but the convergence will be very slow with a small η^i or the optimization does not converge with a large one. At each iteration, the best value of step size that can be achieved is the exact solution of the sub-optimization issue (II.62). But it is usually not possible to obtain this solution or it is time consuming due to the fact that sub-optimization issue must be solved for all iteration. Moreover, an exact optimal values

of step size are often not more efficient than its approximated value via backtracking line search, which is presented as follows.

Goal The values of η^i .

- $STOP_{backtracking} = \text{false}$.
- Define maximum candidate step size value η_0^i and parameters $\tau \in (0, 1)$ and $\sigma \in (0, 1)$.

• Set inner iteration variable $j = 0$.

while $STOP_{backtracking}$ is false **do**

 • Set $\varpi = -\sigma \|\nabla^T \mathcal{L}(\pi^{i-1})\|^T u^i$

 • Calculate the Armijo–Goldstein, stop condition

$$\Xi_j = \mathcal{J}(\pi^{i-1}) - \mathcal{J}(\pi^{i-1} + \eta_j^i u^i) - \varpi \eta_j^i.$$

 • **if** $\Xi_j \geq 0$ **then**

 • Set $j = j + 1$.

 • Set $\eta_j^i = \tau \eta_j^i$.

else

 • $STOP_{backtracking} = \text{true}$.

end

end

Algorithm 2: Backtracking line search method

Based on the Armijo–Goldstein condition, the backtracking line search allows to determine the maximal step size or how far one can move along the chosen search direction. From the beginning positive maximal candidate step size η_0^i (at iteration i of algorithm 1), this method decreases iteratively the candidate η_0^i until getting an expected decrease of the objective functional. This is the reason while it is call as "backtracking". By defining new search control parameters $\tau \in (0, 1)$ and $\sigma \in (0, 1)$, the Algorithm 2 describes how to determine the step size η^i . This is a very simple and quite effective inexact line search method. The constant τ can be interpreted as the fraction used to prevent long steps relative to the decrease in \mathcal{J} . In another work, the present of τ ensures that the achieved reduction of \mathcal{J} will be at least the amount of τ times the reduction promised by the Taylor approximation of \mathcal{J} at π^i . Typical values of τ and σ are both 0.5 which are presented in the study of Armijo [Arm66].

5.c Newton and Quasi-Newton method

As discussed previously, the steepest descent method, despite its simplicity, is one of the slowest methods with very long convergence time. While in the opposite extreme, the pure Newton and Quasi-Newton method give a good convergence characteristics but require more computational effort and a complex formation. Being also iterative methods, both of

original and modified Newton methods consist of the iteration to update the next values of variables from current values by seeking along the chosen direction with an appropriate step size. The only different thing compared with the first order method is the direction of descent, second order gradient based direction. However, while the pure Newton method relies on the Hessian matrix of function at each iteration, the Quasi-Newton one tries to approximate this matrix by various methods. Due to almost Quasi Newton method based on the basis idea of the original one, it is worth briefly presenting it and its variation here.

The algorithm of Newton method can be shared with the one of steepest descent method but the descent direction at iteration i is $u^i = -\mathcal{H}^{-1}(\mathcal{L}(\pi^{i-1}))\nabla\mathcal{L}(\pi^{i-1})$ where $\mathcal{H}(\mathcal{L}(\pi^{i-1}))$ is the second-order functional derivative of \mathcal{L} at point π^{i-1} . The line search method to calculate the step size is the same as in the preceding descent method. Some analysis of pure Newton method including local convergence and global convergence shows its advantages and reduce the disadvantages. When starting close enough to a non-singular minimum, the original Newton method converges super-linearly with order at least two. If the algorithm begin at a point which is very far from the local minimum, the Hessian matrix is possibly singular, as a result, its inverse does not exist that breaks down the algorithm. Moreover, in some cases, such as the complex functional \mathcal{L} as in this study, the exact formulation of Hessian matrix is usually impossible to obtain. The Newton direction as presented previously may not be the descent direction that means the update function value at iteration i is larger than the function value at $i - 1$ because the Hessian is not positive definite at this point. Moreover, this method is attracted by the minimum as much as the attraction of maximum, therefore it just tries to find the solution of equation $\nabla\mathcal{L} = 0$. To overcome these drawback of this method, there are several modified versions to convert the original Newton one into a true gradient descent method by manipulating on the Newton direction such as replacing the Newton direction by the steepest descent one whenever it is not defined (singular Hessian matrix) or making diagonal modification on the non-positive definite Hessian matrix (modified Cholesky factorization and trust region method whose details can be found in [Ber99]). These variations of Newton method can not guarantee the fast convergence at early iteration as the pure method by provide an more reliable algorithm with asymptotic rate of convergence. But they always require both the computation of both gradient and Hessian matrix which demand a lot of computational effort for a large and complex optimization problem. This obstacle can be dealt with by the approach proposed in the Quasi-Newton methods.

The Quasi-Newton algorithm is basically the same algorithm of gradient descent 1 with the same approaches to find the step size but with a different philosophy to choose the searching direction as presented under the following form:

$$\pi^{i+1} = \pi^i + \eta^i u^i \tag{II.67}$$

where the Quasi-Newton direction u^i holds

$$u^i = -\mathcal{D}^i \nabla \mathcal{L}(\pi^i) \quad (\text{II.68})$$

The term \mathcal{D}^i , a position definite matrix, represents an approximation of Hessian matrix of Newton direction. It is adjusted to adapt to the situation of each iteration step based on the iteration increment π_{inc}^i and gradient increment $\nabla \mathcal{L}_{inc}^i$ of two successive iterates. Their definitions are

$$\begin{aligned} \pi_{inc}^i &= \pi^{i+1} - \pi^i \\ \nabla \mathcal{L}_{inc}^i &= \nabla \mathcal{L}^{i+1} - \nabla \mathcal{L}^i \end{aligned} \quad (\text{II.69})$$

which allows to get the approximate equality

$$\nabla \mathcal{L}_{inc}^i \approx \mathcal{H}(\mathcal{L}^{i+1}) \pi_{inc}^i \quad (\text{II.70})$$

Extending this equation to the case of n iteration increments $\pi_{inc}^0, \dots, \pi_{inc}^{n-1}$ corresponding to n gradient increments $\nabla \mathcal{L}_{inc}^0, \dots, \nabla \mathcal{L}_{inc}^{n-1}$ allows to get the approximation of the inverse Hessian matrix at iteration n as

$$\mathcal{H}^{-1}(\mathcal{L}^n) \approx [\pi_{inc}^0, \dots, \pi_{inc}^{n-1}] [\nabla \mathcal{L}_{inc}^0, \dots, \nabla \mathcal{L}_{inc}^{n-1}]^{-1} \quad (\text{II.71})$$

This approximation is exact in the case of quadratic cost function. In other cases, more sophisticated approaches are needed to progressively tend to the inverse Hessian. Some classes of Quasi-Newton method use the equation

$$\mathcal{D}^{i+1} = \mathcal{D}^i + \frac{\pi_{inc}^i \pi_{inc}^{i T}}{\pi_{inc}^{i T} \nabla \mathcal{L}_{inc}^i} - \frac{\mathcal{D}^i \mathcal{L}_{inc}^i \mathcal{L}_{inc}^{i T} \mathcal{D}^i}{\mathcal{L}_{inc}^{i T} \mathcal{D}^i \mathcal{L}_{inc}^i} + \kappa \zeta^i \nu^i \nu^{i T} \quad (\text{II.72})$$

to update the approximation matrix \mathcal{D}^{i+1} from its preceding value and vector of iteration and gradient increments. The first guessed value of Quasi-Newton direction \mathcal{D}^0 must be a positive definite matrix. In equation (II.72), the intermediate vectors ζ^i, ν^i are defined as

$$\begin{aligned} \zeta^i &= \mathcal{L}_{inc}^{i T} \mathcal{D}^i \mathcal{L}_{inc}^i \\ \nu^i &= \frac{\pi_{inc}^i}{\pi_{inc}^{i T} \nabla \mathcal{L}_{inc}^i} - \frac{\mathcal{D}^i \mathcal{L}_{inc}^i}{\zeta^i} \end{aligned} \quad (\text{II.73})$$

The value of real scalar coefficient κ is chosen in the interval $[0, 1]$ and varies between the different methods. The first Quasi-Newton method, Davidon-Fletcher-Powell method, uses the null value of κ while the well-known Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm chooses $\kappa = 1$. Thank to their choice of update direction the Quasi-Newton methods reaches a very fast convergence velocity near the local minimum and this property does not depend on the choice of \mathcal{D}^0 . It is also not sensitive to the inaccuracy of the line search method that allows to manipulate some relaxation of line search to increase the

algorithm speed without having affect on the optimization accuracy. Beside these strong points, they also have some disadvantages including the high computational requirement (function, gradient evaluation, operation to calculate the matrix \mathcal{D}^i) per iteration and the high memory demand to store the calculated vector of iteration and gradient increments also the matrix of \mathcal{D}^i especially whenever the size of optimization is large.

The steepest descent method previously described in 5.b has been replaced by the BFGS method whose basic ideas are presented in this subsection. The implementation of BFGS in nonlinear and non convex optimization like the one in this thesis is pretty complicated. Moreover, because the intention of this work is to propose an estimation methodology for state and parameter, the BFGS algorithm is therefore used via optimization tool called *fmincon* in *Matlab*. The weak form of gradient of cost functional with respect to the initial state and parameters $\nabla\mathcal{L}$ is provided to this tool as well. Some linear constraints are also used to prevent the estimated variables for both state and parameters to become negative and be out of the limits describing the physical meaning. One can notice that the *fmincon* tool has itself a finite difference method to calculate the gradient of estimated variable, which means that it is not obligatory to provide it via adjoint method in some simple estimation problems. In the numerical examples that will be presented in the next chapter, due to the estimation size and the complexity of considered nonlinear hyperbolic models, the optimization problem converges with long estimation time or does not converge without feeding gradients.

6 Conclusion

In this chapter, an introduction of the methodology of adjoint based optimal estimation method for a class of (switched) hyperbolic systems whose dynamics are described by a PDE coupled with an ODE has been proposed. The initial condition of PDE along with some parameters in both ODE and PDE are supposed to be unknown and needing to be identified. The system dynamics can also included switching characteristics in the source term. A smooth activation approach is then employed to add the smoothness to the system equation. The optimal estimation is considered as an optimization problem by defining a cost function which is basically made of the least square gaps between measurement and simulated values. By optimizing this function, the found solutions are the desired state and parameters. The presented concave nonlinear constrained optimization is transformed to the unconstrained one by using the Lagrangian multiplier technique. The calculus of variations is a typical tool to solve the new optimization problem of augmented cost functional. By applying the first variation calculation on the functional and using the weak optimality condition, the gradient of objective functional with respect to the initial state and parameters can be obtained depending on the intermediate adjoint models.

Three discretization schemes, the midpoint method for solving the ODE, implicit Preissmann and Lax-Wendroff schemes (both forward and backward simulation) for PDE are presented after a short general presentation of numerical methods for differential equation. The continuous gradients are also discretized according to the discretization scheme to get the gradients at each discrete point of spatial domain. These gradients are the inputs of the optimization algorithms, and play the role as an important factor to compute the direction for optimum seeking. The details direction calculation with the algorithms are then discussed to give more insight about the steepest descent and Quasi-Newton methods.

To summarize, this chapter presents one of the original points of the first contribution of the thesis, the adjoint-based estimation method for state and parameters in a pretty general case of hyperbolic system in infinite dimensional way, and to the best of our knowledge, there are very few researches realized with such a spirit.

At this point, all the necessary tools to solve the estimation problem, have been discussed. Their applications in some concrete issues will be clarified in the next chapter.

III | Estimation application

Based on the estimation methodology developed in the previous chapter, the present one proposes its application to estimation problems in two types of physical examples, illustrative of hyperbolic systems: overland flow and traffic flow. The chapter is organized from cases with synthetic data to cases with real data, and from simpler estimation problems to more complex ones, following progresses and publications realized along the thesis. The chapter starts with the traffic flow case, on the basis of classical Lighthill–Whitham–Richards model, for which first (constant) parameters, and then state and (constant) parameters, can be estimated. The model is subsequently modified by addition of a relief route at some point along the main road, resulting in a switched hyperbolic system for which the estimation method is applied. The second part of the chapter deals with the overland flow example, starting with a smooth (non switched) model derived from classical Saint-Venant equations, and for which only initial conditions are first estimated. Then the Manning roughness coefficient is estimated when considered as a distributed one, with the help of Radial Basic Function Network approximation. The estimation problem is subsequently extended to the case of both state and distributed coefficient estimation for the same model. Here the parameters needing to be estimated only appear in the flow function and not in the source term. An extension to identification of parameters also in the source term is then proposed (here infiltration), turning the model into a switched one (due to the so-called Green-Ampt model for infiltration).

The next part of the chapter presents estimation results obtained with real measurements, here limited to the case of overland flow for which data of water discharges could be obtained from LTHER lab for the Tondi Kiboro catchment (Sahel). The first application is to estimate a constant Manning coefficient of the overland flow, with the infiltration being modeled by Green-Ampt formula. This application is then extended to the case of simultaneous estimation of Manning coefficient and infiltration parameters, when considering here the empirical Horton model.

The robustness of the proposed adjoint-based estimation method with different configuration of real values of initial state and parameters and different initialization value of optimization algorithm (the beginning points where the algorithm starts) is tested in the

next section. Moreover, at the end of this section, a small section is presented to give a comparison about the effectiveness of optimization algorithm (given by *fmincon* tool of *Matlab*) in two cases with adjoint-based gradient and without adjoint gradient.

Finally, the aim of the last section is to provide an example of state and distributed parameter estimation with the noisy measurements. A potential solution to deal with the presence of noise in measurement is also provided.

Notice that since the calculation details have been provided in chapter II for the general case, they are omitted here, except for some specific situations.

1 Traffic flow

1.a Overview of free traffic flow

The traffic phenomena, from the very beginning of traffic flow modeling in the 1920s up to now, have been described and analyzed mainly via 3 modeling approaches, which are classified according to the scale of observation: microscopic scale, mesoscopic scale and macroscopic scale. These scales can be visually compared with the views of a map at a high altitude above the ground where the traffic flow takes place on. The macroscopic scale corresponds to the case when an observer stays at one thousand meters above the region of highway. The possible observable quantities are the average velocity, density of a group of vehicles which propagates like a wave. When decreasing the altitude to a hundred meters, or "zoom" in more details, the observer can see each individual vehicle but instead of exact measurement of its position or velocity, the probability of its appearance on a position and at a moment is a more reasonable quality to measure. Finally, the microscopic scale is to see each vehicle at a low altitude which allows to characterize its dynamics easily. More precisely, the microscopic model of traffic flow analysis takes in account all the individual vehicles circulating on highway as a particle of the whole traffic system and the dynamics of each one is individually described by an ordinary differential equation. Some of its typical sub-model families are car-following family including continuous time models (Wiedemann model, intelligent driver model, Gipps model) and integer variable models (cellular automaton family including Biham–Middleton–Levine model, Nagel–Schreckenberg model). The so-called mesoscopic scale, or kinetic scale, observes the flow of vehicle under the statistical point of view in order to find the probability distribution of appearance of a vehicle in space and time. As a result, the integro-differential equations are used to describe the dynamics of the traffic system in this case. The idea of the third analysis approach is to use some physical variables to represent the characteristics of a set of vehicles on the highway, such as average velocity and density. The relations between these variables are

thus illustrated by partial differential equations. One of its typical model is the Lighthill-Whitham-Richards (LWR) model [Pap98]. In some large and complicated traffic systems, these three approaches can be combined together to get hybrid models.

Thanks to its advantages, the LWR model plays a central role in many studies associated with traffic dynamics and control. Some of its strong points can be cited here are its simplicity for understanding and its low calculation demand, as compared to the other two models, having much fewer parameters. Moreover, the formation and the propagation of vehicle flow is well reproduced in this kinetic wave model. It is constructed, firstly by Lighthill and Whitham in 1955 [LW55] and by Richards in 1956 [Ric56], at a macroscopic level and in a deterministic way, used to describe the dynamics of density of a vehicle's set through time and space. By denoting the density of vehicle per space unit by $\rho(x, t)$ and the average velocity $v(x, t)$ with $(x, t) \in \mathbb{R} \times \mathbb{R}^+$, and considering a part of highway limited by two positions x_1 and x_2 in a time interval $[t_1, t_2]$, the number of vehicles remaining on the interval $[x_1, x_2]$, as in Fig.III.1, of highway is given by

$$\frac{d}{dt} \int_{x_1}^{x_2} \rho(x, t) dx = \rho(x_1, t)v(x_1, t) - \rho(x_2, t)v(x_2, t) \quad (\text{III.1})$$

By taking the time integral in $[t_1, t_2]$ of equation (III.1), the following equality holds:

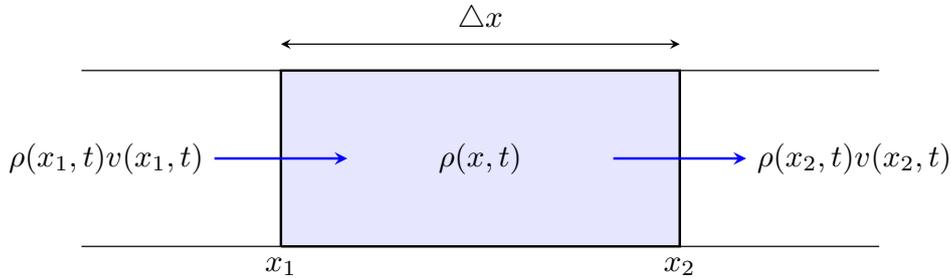


Figure III.1: Interval $[x_1, x_2]$ of highway.

$$\int_{x_1}^{x_2} \int_{t_1}^{t_2} \frac{\partial \rho(x, t)}{\partial t} dt dx = - \int_{x_1}^{x_2} \int_{t_1}^{t_2} \frac{\partial [\rho(x, t)v(x, t)]}{\partial x} dt dx \quad (\text{III.2})$$

and after removing double integration, the equation of traffic flow is as follows:

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial [\rho(x, t)v(x, t)]}{\partial x} = 0 \quad (\text{III.3})$$

where $\rho(x, t)$ is the density of vehicles (*vehicle/km*); $v(x, t)$ is the average velocity of vehicles (*km/hour*). This free traffic flow is thus described by a one dimensional, time-varying and non-linear hyperbolic PDE. The basic assumption of this model is that velocity depends on vehicle density. The next step is thus to choose the relation between the velocity and density of a given group of vehicles at position x . Among some deterministic models like Greenshields model, Greenberg model, Underwood model etc., the first and simplest one, Greenshields model [GCM+35] is exploited to characterize this relation. Its idea is depicted in Fig.III.2 where the straight line describes the linear speed-density liaison. This

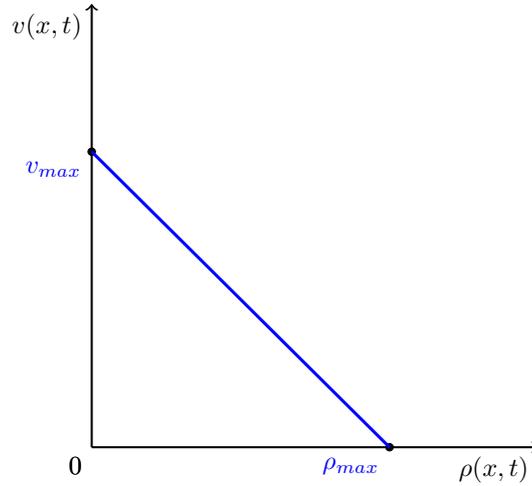


Figure III.2: Relation density-velocity of Greenshields model.

means that when density becomes zero, free highway, the vehicles can get the maximal velocity $v = v_{max}$. In contrast, the vehicles can not move $v = 0$ if the density gets its maximal value $\rho = \rho_{max}$, in the tailback moment.

$$v(x, t) = v_{max} \left[1 - \frac{\rho(x, t)}{\rho_{max}} \right] \quad (\text{III.4})$$

The final equation of traffic flow can be established by inserting equation (III.4) into equation (III.3) and given by

$$\begin{cases} \frac{\partial \rho(x, t)}{\partial t} + v_{max} \left[1 - \frac{2\rho(x, t)}{\rho_{max}} \right] \frac{\partial \rho(x, t)}{\partial x} = 0 \\ \rho(x, 0) = \rho_0^i(x) \quad \text{and} \quad \rho(0, t) = \rho_0^b(t) \end{cases} \quad (\text{III.5})$$

where $\rho_0^i(x)$ and $\rho_0^b(t)$ are respectively the initial and boundary conditions of vehicles density $\rho(x, t)$. Being affected by the characteristics of vehicles and highway, like vehicles type and geometric and constructive characteristics of the highway such as the road width

or road's surface properties, the two parameters v_{max} and ρ_{max} are normally empirical parameters. For a given short enough highway length, the values of these parameters usually remain constant.

1.b Estimation problems using synthetic measurements

1.b.1 Estimation of parameters

Estimation problem statement

The optimal estimation problem of two parameters v_{max} and ρ_{max} of equation (III.5) is considered in this part with assumption that the initial and boundary conditions $\rho_0^i(x)$ and $\rho_0^b(t)$ are perfectly known. The optimal values of parameters are estimated by using the estimation methodology presented in chapter II. The PDE are solved via the nonlinear implicit Preissmann scheme of subsection 4.c. The parameters to be estimated are then obtained through an optimization process being constructed by steepest descent method and backtracking line search method described in subsection 5.b of chapter II. This estimation problem has been published in [NGB15b]. The calculation of variations is not necessary to be repeated, except the gradient of the cost functional \mathcal{L} with respect to the parameter v_{max} and ρ_{max} and the adjoint system because they have a important role in the optimization algorithm. By comparing system equation (III.5) with the general equation of switched hyperbolic system (II.7), one can notice that this is just an continuous model without the switched source term. The associated coupled ODE does not exist and the same with the sub functions $\{g_1(x, t, p, y), g_2(x, t, p, y)\}$ and $\{h_1(y, p), h_2(y, p)\}$. The flow function f is just a function of vehicle density $\rho(x, t)$ and does not depend on x anymore because of the constant values of parameters v_{max} and ρ_{max} . Its expression is

$$f(\rho(x, t)) = v_{max}\rho(x, t) - \frac{v_{max}}{\rho_{max}}\rho^2(x, t)$$

By correlating this equation with the general form of flow function in equation (II.2), the corresponding vector of parameter α is $\alpha = [v_{max} \quad v_{max}/\rho_{max}]$ denoted $[\alpha_1 \quad \alpha_2]$. But the parameter needing to be estimated in this application is ρ_{max} , some transformations are thus needed to get the gradient for ρ_{max} from gradients for α_2 or v_{max}/ρ_{max} . Moreover, the formulation of α_2 contains v_{max} , the variation of v_{max} thus appears in variation of α_2 . The first variation of α_2 is written in the relation with variation of ρ_{max} and v_{max} as

$$\delta\alpha_2 = -\frac{v_{max}}{\rho_{max}^2}\delta\rho_{max} + \frac{1}{\rho_{max}}\delta v_{max}$$

By inserting this relation into the first variation of \mathcal{L} in the direction of v_{max}/ρ_{max} or α_1 , one obtains

$$\begin{aligned} \delta\mathcal{L}_{\alpha_2} = & -\frac{v_{max}}{\rho_{max}^2} \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_2} \right] \Big|_0^L \delta\rho_{max} dt + \frac{v_{max}}{\rho_{max}^2} \int_0^T \int_0^L \frac{\partial f}{\partial \alpha_2} \frac{\partial \lambda}{\partial x} \delta\rho_{max} dx dt + \epsilon_2 [\rho_{max} - \rho_{Fmax}] \delta\rho_{max} \\ & + \frac{1}{\rho_{max}} \int_0^T \left[\lambda \frac{\partial f}{\partial \alpha_2} \right] \Big|_0^L \delta v_{max} dt - \frac{1}{\rho_{max}} \int_0^T \int_0^L \frac{\partial f}{\partial \alpha_2} \frac{\partial \lambda}{\partial x} \delta v_{max} dx dt \end{aligned} \quad (\text{III.6})$$

with the partial derivatives $\frac{\partial f}{\partial \alpha_2} = -\rho^2(x, t)$ and $\frac{\partial f}{\partial \alpha_1} = \rho(x, t)$. The first order optimality condition, the relation of Gâteaux derivative and inner product allow to have the form of gradient for ρ_{max} as follows:

$$\nabla\mathcal{L}_{\rho_{max}} = \frac{v_{max}}{\rho_{max}^2} \int_0^T \left[\lambda \rho^2(x, t) \right] \Big|_0^L dt - \frac{v_{max}}{\rho_{max}^2} \int_0^T \int_0^L \rho^2(x, t) \frac{\partial \lambda}{\partial x} dx dt + \epsilon_2 [\rho_{max} - \rho_{Fmax}] \quad (\text{III.7})$$

The gradient for v_{max} is obtained easily by inserting the derivative $\frac{\partial f}{\partial \alpha_1} = \rho$ in the general form of gradient for α in equation (II.38) and adding the term relating δv_{max} in variation of $\delta \alpha_2$ in equation (III.6). It then satisfies

$$\begin{aligned} \nabla\mathcal{L}_{v_{max}} = & \int_0^T \left[\lambda \rho(x, t) \right] \Big|_0^L dt - \int_0^T \int_0^L \rho(x, t) \frac{\partial \lambda}{\partial x} dx dt + \epsilon_2 [v_{max} - v_{Fmax}] \\ & - \frac{1}{\rho_{max}} \int_0^T \left[\lambda \rho^2(x, t) \right] \Big|_0^L dt + \frac{1}{\rho_{max}} \int_0^T \int_0^L \rho^2(x, t) \frac{\partial \lambda}{\partial x} dx dt \end{aligned} \quad (\text{III.8})$$

All the partial derivatives of sub-functions $\{g_1(x, t, p, y), g_2(x, t, p, y)\}$ and $\{h_1(y, p), h_2(y, p)\}$ with respect to ρ do not exist, the dynamic of adjoint equation of variable $\lambda(x, t)$ in equation (II.28) in this case is presented as

$$-\frac{\partial \lambda}{\partial t} - \left[v_{max} - 2 \frac{v_{max}}{\rho_{max}} \rho(x, t) \right] \frac{\partial \lambda}{\partial x} + \sum_{j=1}^N \delta_A(x - x_j) \times \left[\int_0^L \delta_A(x - x_j) \rho(x, t) dx - \rho_j^{meas}(x_j, t) \right] = 0 \quad (\text{III.9})$$

Estimation result

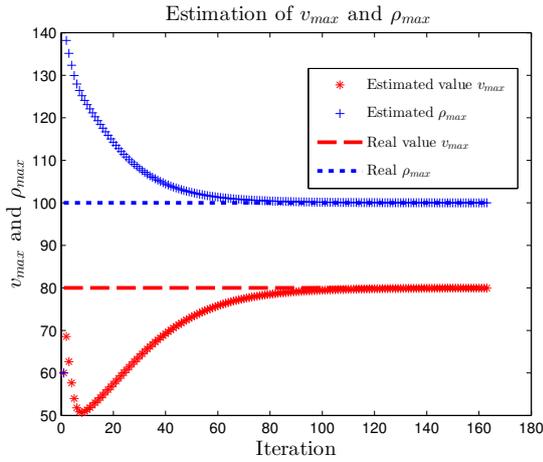
The estimation methodology is applied on a highway with the geometric and numerical data given in Table III.1.

Parameter	Notion	Value	Unit
Maximal speed	v_{Rmax}	80	<i>km/hour</i>
Maximal density	ρ_{Rmax}	100	<i>vehicle/km</i>
Highway length	L	50	<i>km</i>
Observation horizon	T	4	<i>hour</i>
Time step	Δt	5	<i>minute</i>
Space step	Δx	1	<i>km</i>
Calibration coefficient	ϵ_2	$[10^{-3} \ 10^{-3}]$	<i>Unitless</i>
Guess value of ρ_{max}	ρ_{Fmax}	70	<i>vehicle/km</i>
Guess value of v_{max}	v_{Fmax}	110	<i>km/hour</i>
Cost function tolerance	ξ_3	10^{-2}	<i>Unitless</i>
Gradient tolerance	ξ_2	10^{-2}	<i>Unitless</i>
Number of observation value	M	10	<i>Unitless</i>

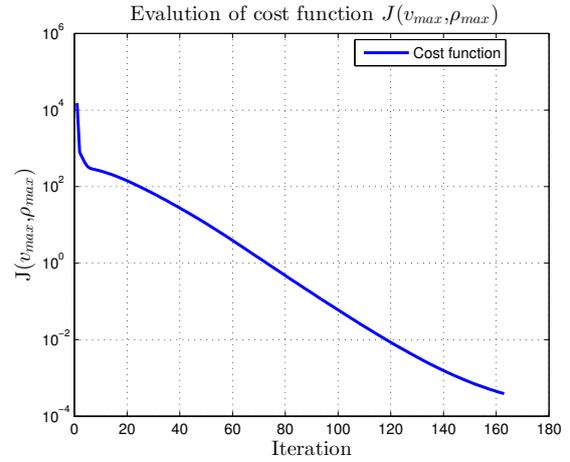
Table III.1: Parameters of highway and numerical values of optimization parameters used for estimating of v_{max} and ρ_{max}

At 10 randomly selected observation positions, 10 observation values of flow density $\rho(x, t)$ are obtained from the exact simulation of system equation. The initial values used for v_{max} and ρ_{max} in the optimization algorithm are 60 (*km/hour*) and 60 (*vehicle/km*), respectively. The steepest descent method along with line search in algorithm 1 will be stopped if the 2-norm of gradients is less than gradient tolerance ξ_2 or the distance between two consecutive values of cost function $\mathcal{J}(v_{max}, \rho_{max})$ is less than tolerance ξ_3 .

Figure III.3a depicts the convergence of the estimations of v_{max} and ρ_{max} from initial values to the real ones. After 163 iterations, the final estimated values of v_{max} and ρ_{max} are 79.96 and 100.01, respectively, which are pretty close to the expected values. The relative errors are pretty small 0.05% and 0.01%.

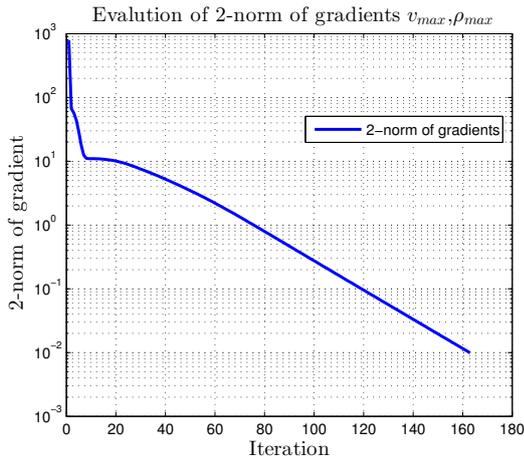


(a) Estimation of v_{max} and ρ_{max} with 10 measurements.

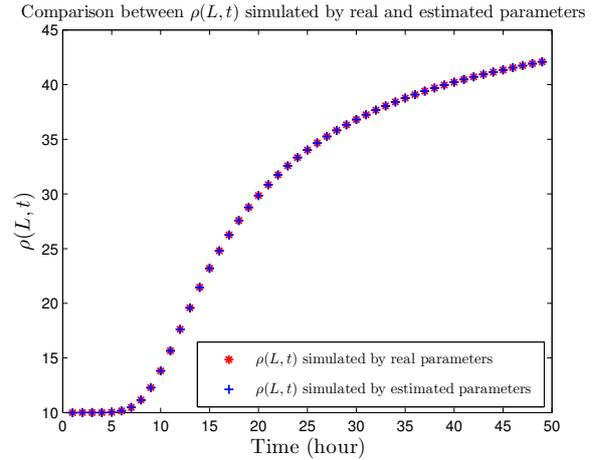


(b) Cost function $\mathcal{J}(v_{max}, \rho_{max})$ in logarithmic scale.

Figure III.3: Parameters estimation results in LWR traffic flow (I).



(a) 2-norm of gradient in logarithmic scale.

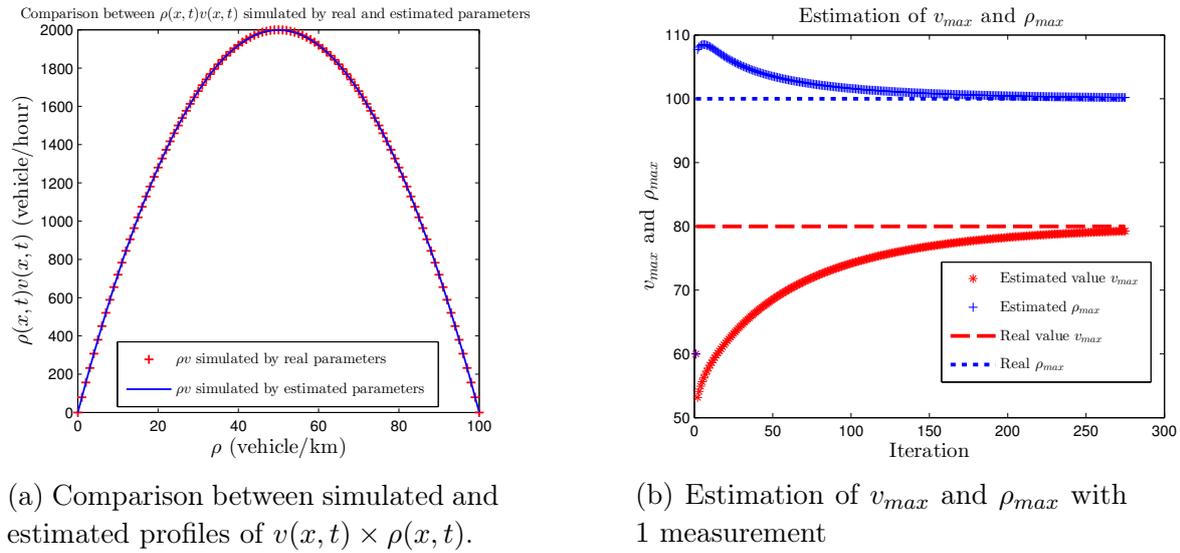


(b) Comparison between two case of flow density at the end of highway.

Figure III.4: Parameters estimation results in LWR traffic flow (II).

Moreover, the value of cost function $\mathcal{J}(v_{max}, \rho_{max})$ indeed decreases to zero, as seen in Figure III.3b, and reaches 0.00039 at the end of optimization process. The reached minimum is not exactly zero possibly due to numerical errors. As seen in Figure III.4a, the 2-norm of the gradient of \mathcal{L} with respect to v_{max} and ρ_{max} decreases also to 0.00991. After the estimation step, it is worth to re-simulate the system with estimated parameters in order to compare the simulated flow density at the end of highway with the real one. The density

of vehicle simulated with estimated optimal values v_{max} and ρ_{max} is really close to the one simulated by real values of parameters that presents the previously mentioned accuracy of parameters estimation, see Figure III.4b. Based on the estimated values of parameters, the velocity \times density curve which represents the characteristic of the considered road can be seen. Figure III.5a presents two curves $v(x, t) \times \rho(x, t)$ (the number of vehicles per hour), in two cases: using real parameters and using estimated parameters. The similarity between the two curves is consistent with the accuracy of parameters estimation as seen before. It is interesting to see how the estimation algorithm react to the decreasing number



(a) Comparison between simulated and estimated profiles of $v(x, t) \times \rho(x, t)$.

(b) Estimation of v_{max} and ρ_{max} with 1 measurement

Figure III.5: Parameters estimation results in LWR traffic flow (III).

of observation points or sensor numbers. This is also a test to examine the sensitivity of optimization problem to measurements. There is only one measurement available, in the worst case. Let us consider the estimation issue when the only observed value of flow density is obtained at 25 km the middle of highway. After using the same algorithm and same starting point, the optimal obtained values of v_{max} and ρ_{max} are $v_{max} = 79.268$ and $\rho_{max} = 100.199$ which are very close to the results when using 10 sensors. This means that for parameter estimation, 10 sensors provide redundant information. However, the number of iterations to get the optimal solution is 275 which is much larger than in the first case. The estimation of parameters in this case is described in Figure III.5b. As a conclusion here, after 163 iterations for the case of 10 measurements and 275 iterations for the case of only 1 measurement, the optimal solutions of the estimation problem are found, which is very promising. One may notice that gradient descent is relatively slow close to the minimum and the optimization time is pretty long, but this can be improved by changing the numerical method.

1.b.2 Estimation of state and parameter

Estimation problem statement

The initial state of vehicle density $\rho_0^i(x)$ which was supposed to be perfectly known in the precedent application example, will be estimated in this example at the same time as parameter v_{max} and quotient v_{max}/ρ_{max} . The value of maximal vehicles density ρ_{max} can be easily recovered from estimation of v_{max} and v_{max}/ρ_{max} . This is also an illustrative example in our publication [NGB16b]. The ODE coupled with the PDE does not exist in this case, and neither its adjoint equation $\gamma(t)$. The adjoint equation of the first Lagrangian $\lambda(x, t)$ has the same form as in the previous case, equation (III.9). The gradients for parameters $\alpha = [\alpha_1 \quad \alpha_2]$ are given by the following equations:

$$\begin{aligned}\nabla \mathcal{L}_{\alpha_1} &= \int_0^T \left[\lambda \rho(x, t) \right] \Big|_0^L dt - \int_0^T \int_0^L \rho(x, t) \frac{\partial \lambda}{\partial x} dx dt + \epsilon_2 [\alpha_1 - \alpha_{1F}] \\ \nabla \mathcal{L}_{\alpha_2} &= - \int_0^T \left[\lambda \rho^2(x, t) \right] \Big|_0^L dt + \int_0^T \int_0^L \rho^2(x, t) \frac{\partial \lambda}{\partial x} dx dt + \epsilon_2 [\alpha_2 - \alpha_{2F}]\end{aligned}\tag{III.10}$$

And the gradient for initial condition $\rho_0^i(x)$ is:

$$\nabla \mathcal{L}_{\rho_0^i(x)} = -\lambda(x, 0) + \epsilon_1 [u_0^i(x) - u_{0F}^i(x)]\tag{III.11}$$

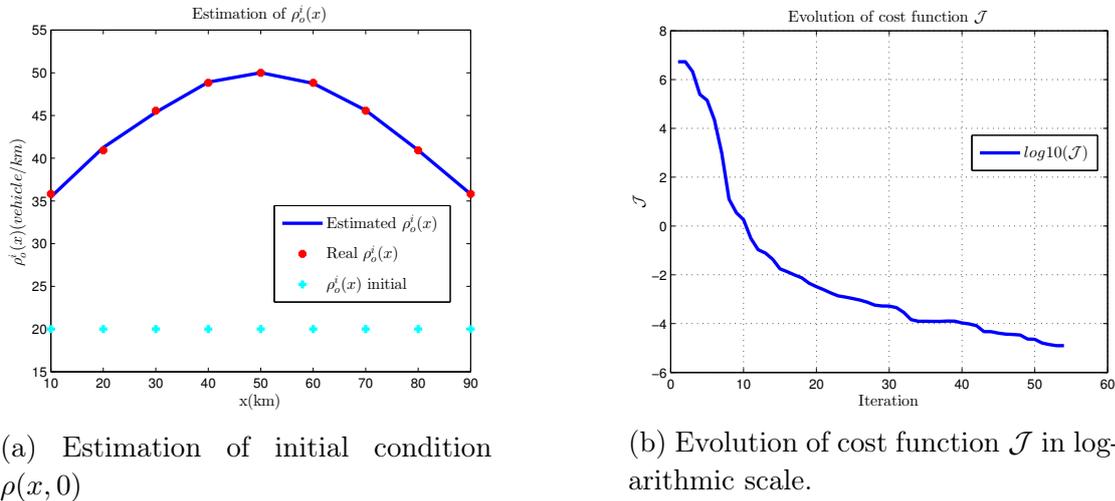
The numerical schemes used for discretizing the differential equation are forward Lax-Wendroff scheme for the LWR model and the backward Lax-Wendroff scheme for the PDE of adjoint equation. The algorithm of optimization process is based on the interior point method of *fmincon* with the gradient provided by equations (III.10) and (III.11).

Estimation result

The estimation process for identification the optimal values of initial condition of vehicles density and two parameters α_1 and α_2 is realized on a straight traffic road with the detailed characteristics presented in Table III.2.

Parameter	Value	Unit
Highway length L	100	km
Observation horizon T	3	$hour$
Space step Δx	10	km
Time step Δt	0.01	$hour$
Calibration coefficients ϵ_1	1×10^{-7}	<i>Unitless</i>
Calibration coefficients ϵ_2	$[1 \times 10^{-7} \ 1 \times 10^{-7}]$	<i>Unitless</i>
Observation number N	1	<i>Unitless</i>

Table III.2: Parameters of highway and numerical value of optimization parameter used for estimating of $\alpha_1 = v_{max}$ and $\alpha_2 = v_{max}/\rho_{max}$ and initial condition $\rho_0^i(x)$.



(a) Estimation of initial condition $\rho(x, 0)$

(b) Evolution of cost function \mathcal{J} in logarithmic scale.

Figure III.6: State and parameters estimation in LWR traffic flow (I).

The real values of maximal vehicles velocity α_1 is $150 (km/hour)$ and the parameter α_2 is $0.5 (km/hour \times km/vehicle)$ which need to be estimated. With spatial discretization step $\Delta x = 10 (km)$, the 100 km highway length is divided into 10 sections. In this example, the boundary condition at $x = 0$, $\rho_0^b(t)$ is a sum of four sinusoidal signals in time with different frequencies and phases. The first values of optimization algorithm are $\rho_{init}(x_j) = 20 (vehicle/km) \ \forall x_j$, $\alpha_1 = v_{max} = 100 (km/hour)$, $\alpha_2 = v_{max}/\rho_{max} = 0.25 (km/hour \times km/vehicle)$. The first guessed values of initial condition $\rho_{F0}^i(x_j)$ and two parameters are equal to zero. The optimization tool *fmincon* is stopped whenever the cost function tolerance or tolerance of estimation variable is smaller than the predefined termination tolerance values. There is only 1 observation value at the end of the considered domain or the discrete position 10, from an exact simulation of direct system (III.3) with

real value of initial state and parameters.

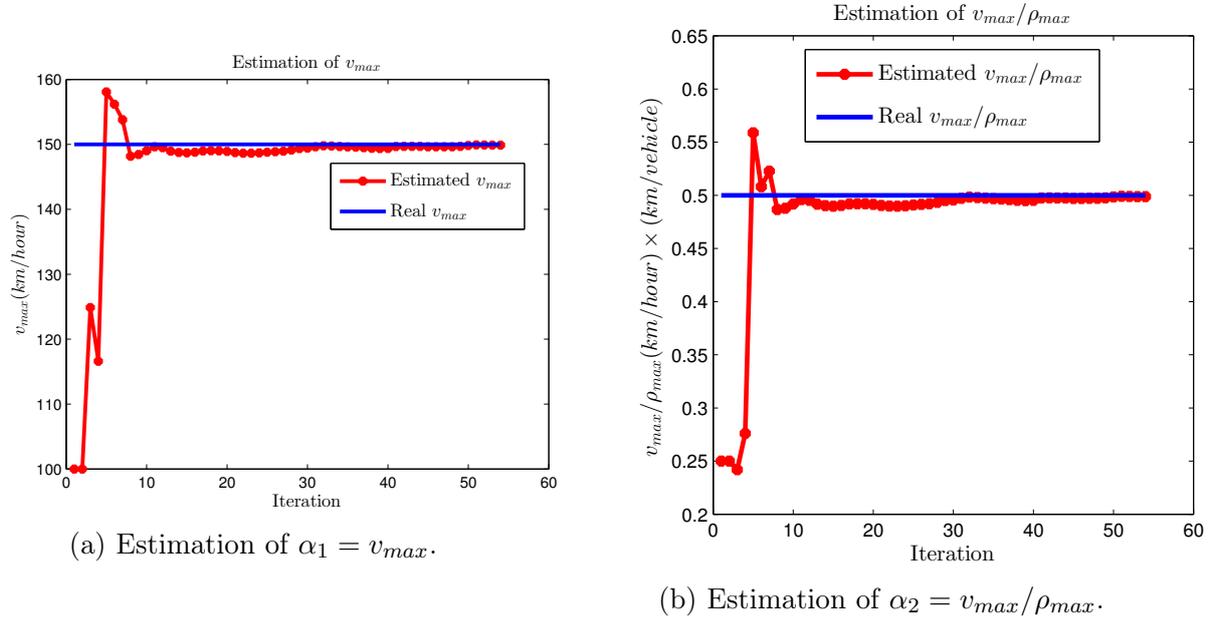


Figure III.7: State and parameters estimation in LWR traffic flow (II).

Figure III.6a compares estimated optimal values of $\rho_0^i(x_j)$ with initial values and the real ones. From the initiations, its estimated values converge to the respective real state with some small bias values at some positions in the considered spatial section with relative error 0.0812%. The found local minimal value of cost function \mathcal{J} is 0.0074, after 54 iterations, slightly larger than zero. The estimated value of parameters α_1 and α_2 are 149.8782 and 0.4988 respectively with relative errors 0.08% and 0.24%. Their convergences are described in Figure III.7a and Figure III.7b. These biases of estimation are possibly caused by the numerical simulation errors on direct and adjoint system simulation and optimization algorithm's convergence limitation. The optimal value of maximal vehicles density is found by simple division $\rho_{max} = \alpha_1/\alpha_2$.

1.b.3 Estimation of state and parameter in switched system

Estimation problem statement

The idea of this estimation example is to go further to estimate the parameter in the source term of the equation of traffic flow. Here the model is modified by considering a relief route on the highway at position \bar{x} . This situation is very common in a real traffic system, to avoid traffic jam at rush hours for instance. The relief phenomenon happens when the density of vehicles becomes larger than a defined threshold called β , the gate

at relief position is opened and a percentage of vehicles turn into this relief route. The threshold value is appropriately chosen to adapt the capacity of both main highway and relief route. To represent this phenomenon a source term can be added to the right side of equation (III.5) as follows:

$$\begin{cases} \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial f[\rho(x, t)]}{\partial x} = 0 & \text{if } \rho(\bar{x}, t) \leq \beta \\ \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial f[\rho(x, t)]}{\partial x} = -\delta(x - \bar{x})\zeta\rho(x, t) & \text{if } \rho(\bar{x}, t) > \beta \end{cases} \quad (\text{III.12})$$

where \bar{x} is the relief position and belongs to the spatial domain of highway; ζ is the relief percentage and its value is in the section $[0, 1]$; $f[\rho(x, t)]$ is the traffic flow and equal to $v_{max}[\rho(x, t) - \rho^2(x, t)/\rho_{max}]$. Along with the initial state $\rho_0^i(x)$ and the parameters v_{max} , v_{max}/ρ_{max} , two specifications β and ζ of relief route also need to be identified.

Equation (III.12) becomes a switched hyperbolic system, and has the same form as compared with the general model of equation (II.5) without the associated ODE. In this example, the parameter vector $p = [\zeta \ \beta]^T$ is denoted by $[p_1 \ p_2]^T$ and $\alpha = [v_{max} \ v_{max}/\rho_{max}]^T$ is denoted by $[\alpha_1 \ \alpha_2]^T$. By comparing with the general case presented in section 2, one can get: the switching condition $\xi = \rho(\bar{x}, t) - \beta$ depending on the state of system; the sub-functions $g_1 = 0$ and $g_2 = -\delta(x - \bar{x})\zeta\rho(x, t)$. The functions $\{h_1 \ h_2\}$ do not exist due to the absence of the ODE. The necessary partial derivatives, in equation (II.28) must be calculated, to get the dynamics of adjoint variable λ : $\frac{\partial g_1}{\partial \rho(x, t)} = 0$, $\frac{\partial g_2}{\partial \rho(x, t)} = -\delta(x - \bar{x})\zeta$, $\frac{\partial f}{\partial \rho(x, t)} = \alpha_1 - 2\alpha_2\rho(x, t)$, $\frac{\partial \varphi_a}{\partial \rho(x, t)} = \frac{\partial \varphi_a}{\partial \rho(\bar{x}, t)} \frac{\partial \rho(\bar{x}, t)}{\partial \rho(x, t)} = \varphi_a^2 \delta_A(\bar{x}) e^{-K_{ac}\xi}$. The second adjoint equation of $\gamma(t)$ does not exist while the corresponding adjoint equation of variable $\lambda(x, t)$ is described by:

$$\begin{aligned} -\frac{\partial \lambda}{\partial t} - [\alpha_1 - 2\alpha_2\rho(x, t)] \frac{\partial \lambda}{\partial x} - \lambda \left[-\zeta\varphi_a\delta(x - \bar{x}) + \delta(x - \bar{x})\zeta\varphi_a^2\delta_A(\bar{x})e^{-K_{ac}\xi} \right] \\ + \sum_{j=1}^N \delta_A(x - x_j) \left[\int_0^L \delta_A(x - x_j)\rho(x, t)dx - \rho_j^{meas}(x_j, t) \right] = 0 \end{aligned} \quad (\text{III.13})$$

By replacing the partial derivatives of the sub-functions g_1 , g_2 and activation function φ_a with respect to each element of p , the partial derivative of traffic flow f with respect to α in equations (II.35), (II.37), the gradient of cost functional with respect to parameters p and α can be easily formulated and are presented in equation (III.14) below:

$$\begin{aligned} \frac{\partial g_1}{\partial p_1} = 0 \ ; \quad \frac{\partial g_1}{\partial p_2} = 0 \ ; \quad \frac{\partial g_2}{\partial p_1} = -\delta(x - \bar{x})\rho(x, t) \ ; \quad \frac{\partial g_2}{\partial p_2} = 0 \\ \frac{\partial \varphi_a}{\partial p_1} = 0 \ ; \quad \frac{\partial \varphi_a}{\partial p_2} = K_{ac}\varphi_a^2 e^{-K_{ac}\xi} \ ; \quad \frac{\partial f}{\partial \alpha_1} = \rho(x, t) \ ; \quad \frac{\partial f}{\partial \alpha_2} = -\rho^2(x, t) \end{aligned} \quad (\text{III.14})$$

The gradients of the Lagrangian functional \mathcal{L} with respect to the constant parameters follows the general gradient in equation (II.38), and are thus as follows:

$$\nabla \mathcal{L}_{\alpha_1} = \int_0^T \left[\lambda \rho(x, t) \right] \Big|_0^L dt - \int_0^T \int_0^L \rho(x, t) \frac{\partial \lambda}{\partial x} dx dt + \varepsilon_2 [\alpha_1(x) - \alpha_{1F}(x)] dx \quad (\text{III.15})$$

$$\nabla \mathcal{L}_{\alpha_2} = - \int_0^T \left[\lambda \rho^2(x, t) \right] \Big|_0^L dt + \int_0^T \int_0^L \rho^2(x, t) \frac{\partial \lambda}{\partial x} dx dt + \varepsilon_2 [\alpha_2(x) - \alpha_{2F}(x)] dx \quad (\text{III.16})$$

$$\nabla \mathcal{L}_{p_1} = - \int_0^T \int_0^L \lambda \varphi_a \delta(x - \bar{x}) \rho(x, t) dx dt + \varepsilon_3 [p_1 - p_{1F}] \quad (\text{III.17})$$

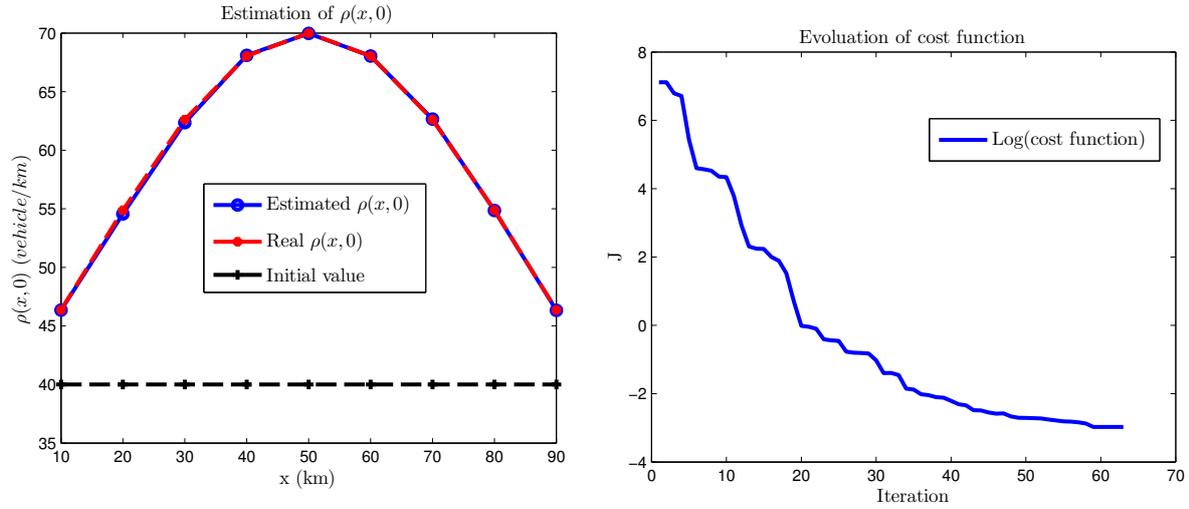
$$\nabla \mathcal{L}_{p_2} = -\zeta K_{ac} \int_0^T \int_0^L \lambda \varphi_a^2 e^{-K_{ac}\xi} \delta(x - \bar{x}) \rho(x, t) dx dt + \varepsilon_3 [p_2 - p_{2F}] \quad (\text{III.18})$$

In addition, the gradient for the initial condition of the vehicle density is not changed, the same as equation (III.11).

Estimation result

Parameter	Value	Unit
Highway length L	100	<i>km</i>
Observation horizon T	3	<i>hour</i>
Space step Δx	10	<i>km</i>
Time step Δt	0.005	<i>hour</i>
Calibration coefficients ϵ_1	$\underbrace{[1 \ 1 \dots 1]}_{\text{size 9}} \times 10^{-9}$	<i>Unitless</i>
Calibration coefficients ϵ_2	$[1 \ 1] \times 10^{-9}$	<i>Unitless</i>
Calibration coefficients ϵ_3	$[5 \times 10^{-4} \ 5 \times 10^{-1}]$	<i>Unitless</i>
First guessed value of v_{max} , α_{1F}	200	<i>km/hour</i>
First guessed value of v_{max}/ρ_{max} , α_{2F}	0.8	<i>km²/(hour × vehicle)</i>
First guessed value of leaving percentage ζ , p_{1F}	0.8	<i>Unitless</i>
First guessed value of vehicle threshold β , p_{2F}	60	<i>vehicle/km</i>
First guessed value of initial condition $\rho_0^i(x)$, $\rho_{0F}^i(x)$	$0 \ \forall x \in [0, L]$	<i>vehicle/km</i>
K_{ac}	50	<i>Unitless</i>
Observation number N	4	<i>Unitless</i>

Table III.3: Parameters of highway and numerical values of optimization parameters used for estimating of $\alpha_1 = v_{max}$ and $\alpha_2 = v_{max}/\rho_{max}$, initial condition $\rho_0^i(x)$ and relief characteristics $p = [\zeta \ \beta]$.



(a) Estimation of initial state of system.

(b) The evolution of cost function \mathcal{J} .

Figure III.8: State and parameters estimation results in switched LWR traffic flow (I).

The considered parameters are synthesized in table III.3, the 100 (km) highway is divided into 10 sections with 10-km space step Δx . The real initial state is in the form of a Gaussian function as in Figure III.8a. The boundary condition, at position $x = 0$ of system is a sum of some sinusoidal signals with different frequencies and phases and there is no interest to present it here. The maximal velocity possible on highway v_{max} or $\alpha_1 = 150$ (km/hour); the real value of second parameter $\alpha_2 = v_{max}/\rho_{max}$ is chosen to be equal to 0.7, the maximal density of vehicles circulating is thus approximately 214.2857 (vehicles/km). The relief route is located at position $\bar{x} = 40$ (km), which corresponds to discrete position 5 of the numerically simulated model. The percentage of leaving vehicles is equal to $\zeta = 0.5$, while the vehicle density threshold $\beta = 50$ (vehicles/km). In this simulation example, 4 measurements are taken from the exact simulation at positions $\{40, 50, 60, 90\}$ with no disturbance, to define the cost function \mathcal{J} . The term "exact simulation", in this case, means the simulation realized with the original equation of system, described in equation (III.12), and with the real values of initial state and parameters. The simulation for estimation process is based on the smoothed one using the activation function. Comparing the estimated values with the simulated ones then also gives a validation of the approximation model.

The cost function \mathcal{J} , from the initial value 1.2308×10^3 , decreases to the minimal one equal to 0.0511 as in Fig.III.8b after 63 iterations. Figure III.8a illustrates the real, initial and final estimated values of the initial state of system. The initial values of parameters α_1 , α_2 , p_1 and p_2 , which are equal to 180 (km/hour), 0.3, 0.4% and 40 (vehicles/km), converge to the final ones 149.9609 (km/hour), 0.4995, 0.6854% and 50.1782 (vehicles/km), respectively. Their convergences are depicted in Fig.III.10a, Fig.III.10b, Fig.III.9a and

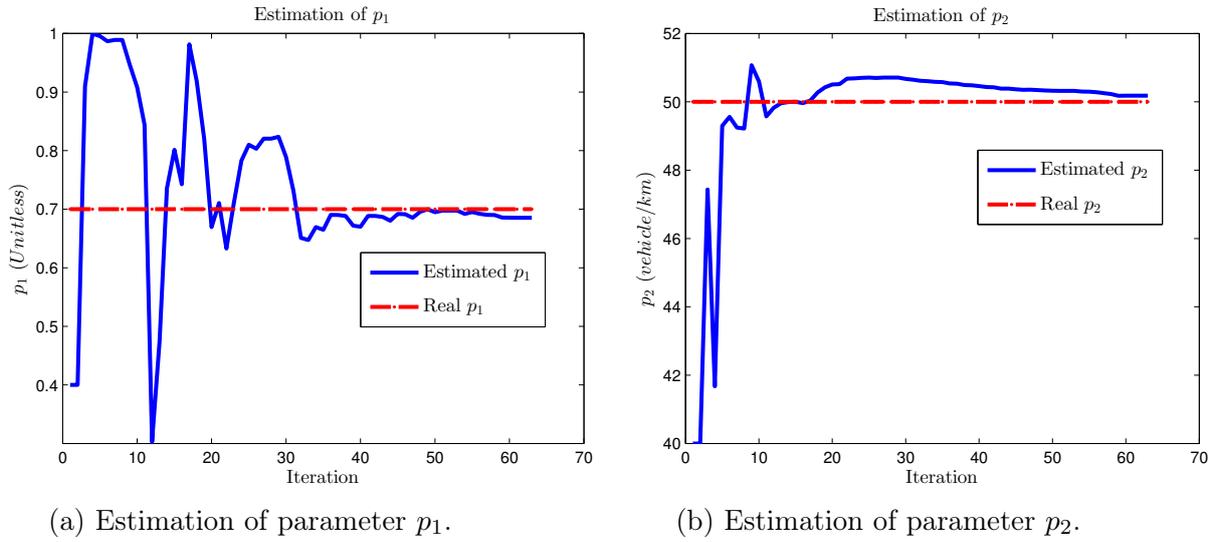


Figure III.9: State and parameters estimation results in switched LWR traffic flow (II).

Fig.III.9b). The corresponding relative errors between the real and estimated values of these parameters are pretty small and equal to 0.0261%, 0.0972%, 2.0877% and 0.3563%.

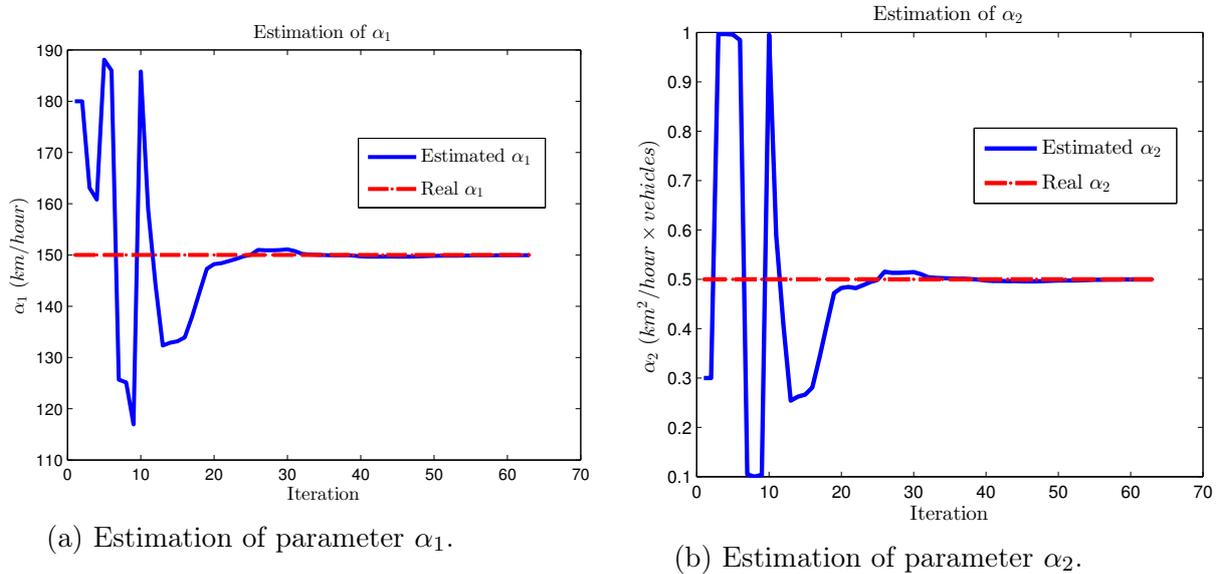


Figure III.10: State and parameters estimation results in switched LWR traffic flow (III).

One can observe that the estimated values are pretty close to the real ones with an average relative error of all estimation at discrete positions equal to 0.1487%. The activation function approximation is possibly one of the main reasons that causes the estimation errors. The cost function calculated with the real values of state and parameter is already

0.0309, slightly smaller than the found optimal one 0.0511. Moreover, the estimation result, especially the estimation of p_2 is greatly affected by the factor K_{ac} . According to the gradient formulation of p_2 in equation (III.18), a large value of K_{ac} makes this gradient close to zero while a small value allows to have a good estimation of p_2 but give a bad effect and causes larger estimation error on other parameters. The value of K_{ac} is finally chosen equal to 50 after some manual tests. This value allows to keep the balance between the estimation of all variables.

2 Overland flow

2.a Overview of overland flow

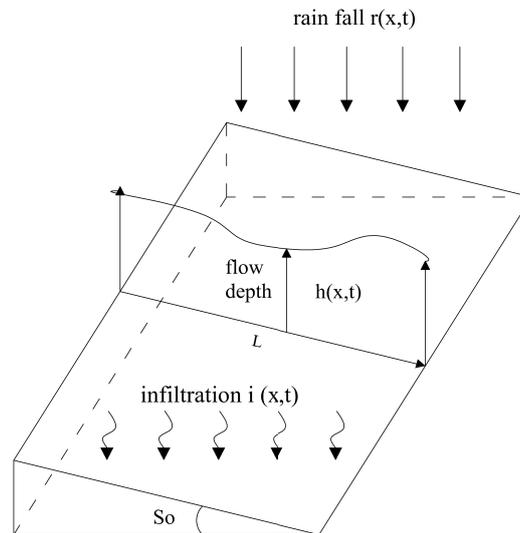


Figure III.11: Illustration of overland flow model.

Being one of the major components in water cycle and the main agent of soil erosion, the so-called overland flow or surface runoff is the hydrological concept used to describe the movement of water over soil surface and towards some hydraulic channels such as streams, small water channel or rivers under gravitational force (see Figure III.11 for an illustration). This type of flow is different from the channel flows which start at the end of overland flow and are concentrated in a channel. Both overland flow and channel flow are considered as surface flow. They play an essential role in various hydrological topics, especially flood prediction and surveillance. In nature, there are many types of overland flows which are classified according to the generating physical mechanisms: rainfall, snow and glacier melting or ground water. The first type of overland is much more common and

happens during and after the rainfall process. There are mainly two factors concerning an overland flow: fluid flow on soil surface and the infiltrated water in the ground which makes the soil to become saturated. The infiltration physically begins at the same moment as rainfall event with the infiltration rate depending on the initial state of soil, and ends when the soil reaches its saturation threshold. One of the common ways to mathematically model overland flow is to determine the flow depth and flow rate at a point on watercourse from the continuity equation of the hydrodynamic model or Saint-Venant model for flow dynamics and some common infiltration model, such as the physical Green-Ampt model or the empirical Horton model for the infiltration process (see [Sco92] and [Par+97]). The well-known Saint-Venant model belongs to a class of distributed flow routing, or sometime referred as unsteady flow routing, where the flow rate, water level and water discharge are functions of time and space instead of a function of time only as in the case of lumped flow routing (see [May10] for more details). The relationships between water flow's height and rainfall, infiltration rates are as follows:

$$\begin{cases} \frac{\partial h}{\partial t} + \frac{\partial uh}{\partial x} = r - i \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + g \frac{\partial h}{\partial x} = g(S_0 - S_f) + \frac{u}{h}(r - i) \end{cases} \quad (\text{III.19})$$

where h is the water flow depth, (m); u is the water flow velocity, (m/s); x is the space variable, (m); t is the time, (s); r is the rainfall intensity, (m/s); i is the infiltration intensity, (m/s); g is the acceleration of gravity, (m/s^2); S_0 is the bed slope, (m/m); S_f is the friction slope, (m/m).

The first and second equations of model (III.19) are called respectively the continuity and momentum ones. The mathematical description of overland flow now can be simplified by omitting the momentum equation. The resulting model is the kinematic-wave equation where the motion of water is the only subject being taken into account with the influence of force and mass. Notice that in the momentum equation, from left to right, the first term is the local acceleration describing the change of velocity overtime, the second one is the convective acceleration describing the change of velocity along spatial length, the third one is called pressure force term equal to gravitational acceleration times spatial change of water depth, the fourth term includes gravity force and friction force term and the last one is the momentum caused by excess rainfall. In the kinematic-wave model, all the local and convective accelerations, and pressure force are neglected. An assumption of kinematic wave is also made to keep the balance between gravity and friction forces, $S_0 = S_f$. Moreover, the bottom slope of overland flow is small, $\sin(S_0) \simeq S_0$, which allows to neglect the so-called scour's effect and deposit. The dynamic of water flow is then only described by the continuity equation.

The relation of the flow rate velocity and the flow depth is expressed by the Manning

equation (III.20):

$$u = \frac{R^{2/3} S_f^{1/2}}{n} \quad \text{with} \quad R = \frac{hw}{2h + w} \quad (\text{III.20})$$

where R is the hydraulic radius, (m); n is the Manning roughness coefficient, ($s/m^{1/3}$); w is the elemental flow width, (m). In an overland flow, the water depth h is usually much smaller than the elemental flow width ($h \ll w$) so $R \simeq h$ and the Manning equation then becomes:

$$u = \frac{S_0^{1/2}}{n} h^{2/3} \quad (\text{III.21})$$

By denoting the flow of water as f , which is a function of water depth $h(x, t)$ and spatial variable x due to the distributed Manning coefficient $n(x)$ in some cases, the equation of overland flow is finally:

$$\begin{cases} \frac{\partial h(x, t)}{\partial t} + \frac{\partial f(h(x, t), x)}{\partial x} = r - i \\ h(x, 0) = h_0^i(x) \\ h(0, t) = h_0^b(t) \end{cases} \quad (\text{III.22})$$

where $h_0^i(x)$ and $h_0^b(t)$ are respectively the initial condition and boundary conditions.

According to the empirical or physical points of view, there are many ways to characterize the infiltration rate, such as Horton model or Green-Ampt model. Due to its simplicity and good physical meaning preservation, the classical Green-Ampt model is used in the example using synthetic measurement and the Horton one is employed in the case of real data. The details of Horton model are discussed in the subsection where it is used. From the physical point of view, under the effect of rainfall and depending on the initial state of soil and rainfall intensity, the soil surface becomes ponded with infiltrated water. There are thus two distinct stages of an infiltration process under a rainfall event: a stage without surface ponding and a stage when ground is ponded. In the first case, before the ponding moment, all the rainfall infiltrates into the soil. As a result the infiltration rate is equal to rainfall rate and less than the infiltration capacity. With the movement of water into the soil, the soil surface becomes ponded and after a moment, so-called "ponding time", the infiltration process is independent of rainfall time distribution. The Green-Ampt model follows this idea. But this model was given initially (by Green and Ampt in 1911, see [GA11]) for the ponded surface. It has been modified by Mein and Larson [1973] and Swartzendruber [1974] to determine the time when the surface ponding starts. One can consider that this is the first physically and most widely used infiltration model for steady rainfall. The derivation of this model from the Darcy's law and the principle of mass conservation in homogeneous soil condition can be seen in [ML73]. An infiltration rate according to Green-Ampt model is finally divided into two periods: before and after "ponding time" denoted by t_{pon} , as

follows:

$$\begin{cases} i(t) = K_i \left(\frac{\Psi \eta (1 - \theta)}{I(t)} + 1 \right) & \text{if } t > t_{pon} \\ i(t) = r & \text{if } t \leq t_{pon} \end{cases} \quad (\text{III.23})$$

where K_i is the effective hydraulic conductivity of the soil, (m/s); Ψ is the soil suction at wetting front, (m); η is the soil porosity, (%); θ is the relative initial soil moisture, (*unitless*); $I(t)$ is the cumulative infiltration m which is the accumulated depth of water infiltrating during the time period t , defined by equation

$$I(t) = \int_0^t i(\tau) d\tau$$

The ponding time t_{pon} depends on rainfall rate and soil characteristics as in equation (III.24):

$$t_{pon} = \frac{\Psi K_i \eta (1 - \theta)}{r(r - K_i)} \quad (\text{III.24})$$

One can notice that the equation used to describe the infiltration phenomenon, after ponding time, in Green-Ampt model, is implicit in time. No formal algebraic formulation can be determined to calculate the infiltration rate. Instead, an iteration method (e.g. Newton-type iteration) can be used to get the cumulative infiltration at time $t + \delta t$ from the one at time t through the equation below by

$$I_{t+\Delta t} = I_t + \Psi \Delta \theta \ln \left[\frac{I_{t+\Delta t} + \Psi \Delta \theta}{I_t + \Psi \Delta \theta} \right] + K_i \Delta t \quad (\text{III.25})$$

2.b Estimation problems using synthetic measurements

The synthetic measurements are the values obtained from the exact simulation of system without noise at some fixed observation positions. This simulated values of system dynamics are realized by using the true initial state and exact values of parameter. These measurements are used as inputs of our optimization process, to define the cost function \mathcal{J} .

2.b.1 Estimation of initial condition

Estimation problem statement

The problem of estimation for initial condition of the flow per width unit is presented in this example. This is also the content of our publication [NGB14]. The dynamics of

overland flow used here is of the flow per unit width $f(x, t)$. Equation (III.22) must be written for variable $f(x, t)$ as

$$\begin{cases} \frac{\partial f^{3/5}}{\partial t} + K_{hf} \frac{\partial f}{\partial x} = K_{hf}(r - i) \\ f(x, 0) = f_0^i(x) \\ f(0, t) = f_0^b(t) \end{cases} \quad (\text{III.26})$$

where the coefficient $K_{hf} = \left[\frac{S_0^{1/2}}{n} \right]^{3/5}$; $f_0^i(x)$ is the initial condition needing to be identified and $f_0^b(t)$ is the upstream boundary condition which is fixed. The cost function \mathcal{J} is defined in the same form as the general one (II.4) but using both the measurement of flow depth and flow per unit width:

$$\begin{aligned} \mathcal{J} = & \frac{1}{2} \sum_{j=1}^M \int_0^T \left\{ \int_0^L \frac{1}{K_{hf}} \delta_A(x - x_j) f^{3/5}(x, t) dx - h_j^{meas}(x_j, t) \right\}^2 dt \\ & + \frac{1}{2} \sum_{k=1}^N \int_0^T \left\{ \int_0^L \delta_A(x - x_k) f(x, t) dx - f_k^{meas}(x_k, t) \right\}^2 dt + \varepsilon_1 \frac{1}{2} \int_0^L \|f_0^i(x) - f_{F_0}^i(x)\|^2 dx \end{aligned} \quad (\text{III.27})$$

where M , N are respectively the number of measurements of flow dept and flow per unit width. The procedure to solve the estimation problem is almost the same as in the presentation of the estimation methodology (section 3.c in chapter II) except that the variation of water flow depth must be expressed as a function of the variation of $f(x, t)$ as

$$\delta h(x, t) = \frac{3}{5K_{hf}} f^{-2/5}(x, t) \delta f(x, t)$$

The source term $r - i$ does not relate to the initial condition function $f_0^i(x)$, and the system equation is a continuous one without the switch function g_1 , g_2 and there is no associated ODE. Then all the partial derivatives of g_1 , g_2 , h_1 , h_2 and activation function with respect to $f(x, t)$ do not exist. The corresponding adjoint equation of $\lambda(x, t)$ is thus

$$\begin{aligned} -\frac{3}{5} f^{-2/5} \frac{\partial \lambda}{\partial t} - K_{hf} \frac{\partial \lambda}{\partial x} + \sum_{j=1}^M \left[\frac{3}{5K_{hf}} \delta_A(x - x_j) f^{-2/5}(x, t) \left\{ \int_0^L \frac{1}{K_{hf}} \delta_A(x - x_j) f^{3/5}(x, t) dx \right. \right. \\ \left. \left. - h_j^{meas}(x_j, t) \right\} \right] + \sum_{k=1}^N \left[\delta_A(x - x_k) \left(\int_0^L \delta_A(x - x_k) f(x, t) dx - f_k^{meas}(x_k, t) \right) \right] = 0 \end{aligned} \quad (\text{III.28})$$

And the gradient of \mathcal{L} for $f_0^i(x)$ has the form

$$\nabla \mathcal{L}_{f_0^i(x)} = -\frac{3}{5} f^{-2/5}(x, 0) \lambda(x, 0) + \varepsilon_1 [f_0^i(x) - f_{F_0}^i(x)]$$

Parameters	Value	Unit
Length of plan	1000	<i>m</i>
Simulation time	100	<i>hour</i>
Space step Δx	50	<i>m</i>
Time step Δt	2	<i>hour</i>
Bed slope S_0	0.04	<i>m/m</i>
Manning coefficient n	0.025	<i>Unitless</i>
Effective hydraulic conductivity K_i	0.145	<i>cm/s</i>
Soil suction at wetting front Ψ	21.85	<i>cm</i>
Soil porosity η	0.434	<i>%</i>
Relative initial soil moisture θ	0.463	<i>Unitless</i>
Weighting factor ϵ_1	10^{-4}	<i>Unitless</i>
Coefficient α ; β of Preissmann	0.68; 0.5	<i>Unitless</i>
Observation number M, N	5; 5	<i>Unitless</i>

Table III.4: Parameters of overland flow used in example of estimation the initial condition of flow per unit width $f(x, t)$.

Estimation result

A numerical simulation of overland flow in this case takes place on an hill-slope under a rainfall event. All the geometric and numerical simulation are summarized in Table III.4.

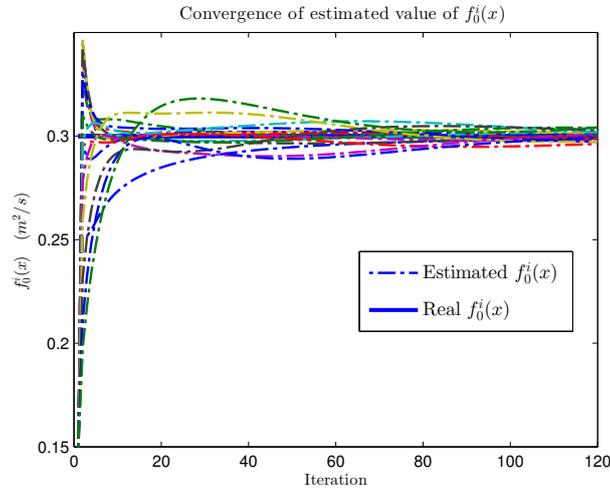


Figure III.12: Estimation of $f_0^i(x)$

The infiltration process is simulated by using a Newton-Raphson method. The flow's depth and flow per unit width observed values are got from the first simulation of system (III.26) with a chosen initial state $f_0^i(x)$ which needs to be estimated. The forward and backward Preissmann schemes and the steepest descent method are used as the numerical tool for the optimization problem. In order to reduce time of optimization process without

any loss of generality, the author considered a case when the initial state of flow per unit width is identical for all x . The real value of condition initial is $f_0^i(x_j) = 0.3 \text{ (m}^2/\text{s)}$, $\forall x_i \in$

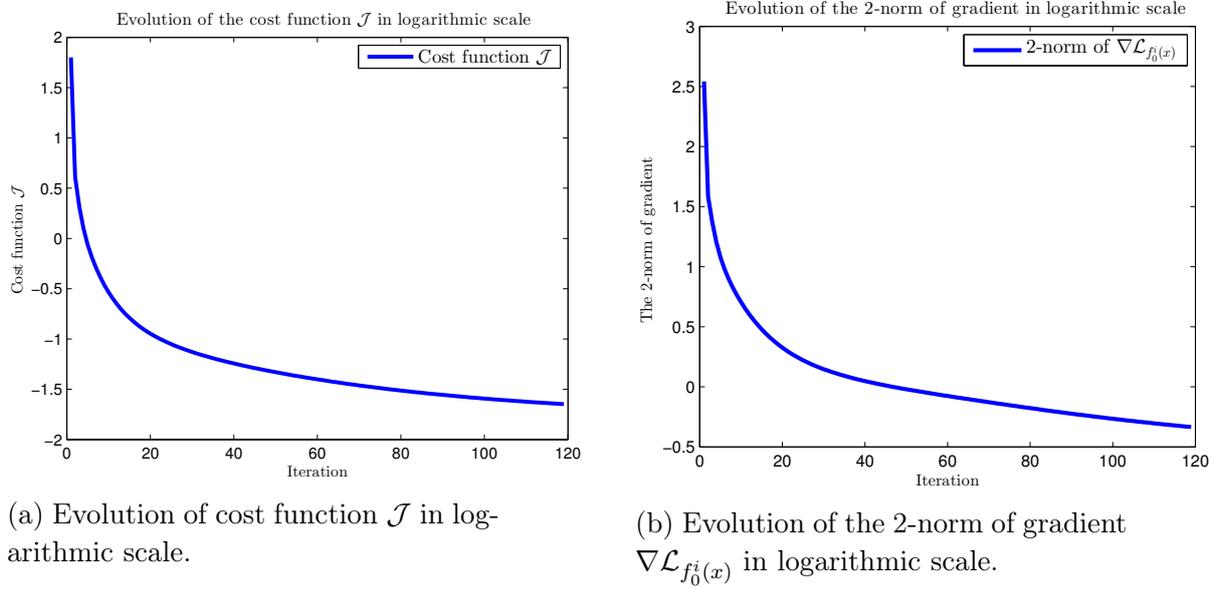


Figure III.13: Estimation results of state in overland flow (II).

$]0, L[$ and the optimization is started with $f_{0Init}^i(x) = 0.15$, $\forall x_i \in]0, L[$, and $f_{F0}^i(x_j) = 0.2$. It can be seen how the estimation of $f_0^i(x_j)$ indeed converges in Figure III.12, with an average relative error between the estimated values of $f_0^i(x_j)$ and the desired one of 0.6%, which is pretty small. Figure III.13a depicts the convergence of the value of the cost functional to the minimal one which is here equal to 0.022 that is almost 0. The convergence rate is indeed slow due to the characteristic of steepest descent method which is used here. One can finally see in Figure III.13b that the norm of gradient indeed decreases to zero. The optimization process is stopped when its value is smaller than the chosen gradient tolerance $\xi = 10^{-3}$.

2.b.2 Estimation of distributed Manning coefficient

Estimation problem statement

In the present estimation problem, the Manning roughness coefficient $n(x)$ of continuity equation is assumed to be spatially distributed in space and unknown. A new parameter called $K_{fh}(n, S_0)$ is introduced and is a simple function of the roughness coefficient $n(x)$ and the bed slope S_0 . This example uses the equation of water flow depth $h(x, t)$ to describe

the dynamics of overland flow as follow:

$$\begin{cases} \frac{\partial h}{\partial t} + \frac{\partial [h^{5/3} K_{fh}]}{\partial x} = r - i & \text{with } K_{fh}(n, S_0) = \frac{S_0^{1/2}}{n(x)} \\ h(x, 0) = h_0^i(x) & \text{and } h(0, t) = h_0^b(t) \end{cases} \quad (\text{III.29})$$

The estimation of $n(x)$ transform to the estimation of $K_{fh}(n, S_0)$. Moreover, an assumption is made where the function K_{fh} can be approximated by a Radial Basis Function Network (RBFN) which is a type of artificial neural networks containing three layers called input, hidden and output layer, depict in Figure III.14 and described in equation (III.30).

$$K_{fh}(n, S_0) = \sum_{k=1}^P w_k \phi(\|x - x_{ck}\|) \quad (\text{III.30})$$

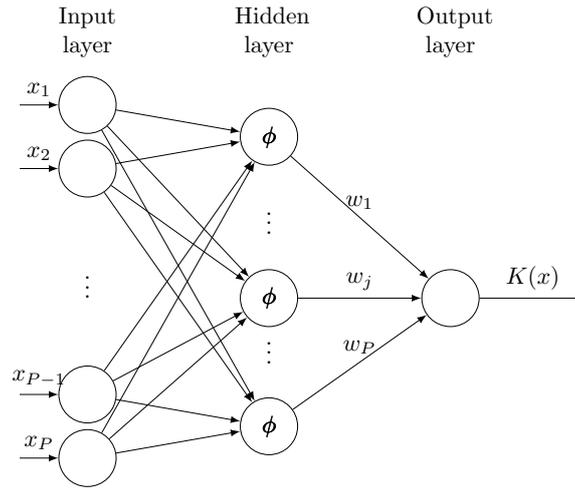


Figure III.14: Overview of Radial Basic Function Network to approximate $K_{fh}(n, S_0)$

where P is the number of neurons in the hidden layer; x is the spatial variable and also input data of the network; x_{ck} is the predefined center vector of each neuron of the network, belonging to the same set as x ; $\phi(\|x - x_i\|)$ is a predefined activation function in hidden layer, which can be Gaussian, multiquadric, inverse quadric, polyharmonic spline function; w_k denotes the weighting factors. The Euclidean or Mahalanobis distance can be used to define the operation $\|\cdot\|$ and the activation function is described by the Gaussian function

$$\phi(\|x - x_i\|) = e^{-\frac{\|x - x_{ck}\|^2}{\sigma_K^2}}$$

where $\|\cdot\|$ is the Euclidean distance. The problem of how to choose optimal value of center vector and parameters of activation function is not taken into account. From the knowledge of $K(x)$ and other parameters, the values of w_k are found by just a simple linear optimization. With the assumption that we know all information of RBFN except the weighting factors w_k and the constant bed slope S_0 throughout the plan, the estimation of Manning roughness coefficient or the estimation of $K_{fh}(n, S_0)$ can be reduced to estimation of w_k . This means that one can get the value of Manning coefficient through the estimated values of w_k . The reason why the estimation problem of $n(x)$ or $K_{fh}(n, S_0)$ is transformed to estimation problem of w_k is the fact that the dimension of w_k or size of parameter that needs to be estimated, is usually much smaller than that of $n(x)$. A reduction of complexity of optimization problem is realized. On this basis, the point here is to look for an optimal estimation of the weighting factors w_k of RBFN by means of adjoint based estimation methodology in chapter II. The cost function \mathcal{J} is defined with the same esprit as in the previous case, estimating initial condition of the flow per unit width in equation (III.27), except the calibration term of $f_i^0(x)$ being replaced by the one of w_k , $\frac{1}{2}\varepsilon_1 \sum_{k=1}^P \|w_k(0) - w_{kF}(0)\|^2$. Due to the change of system equation, the adjoint equation in this example becomes:

$$\begin{aligned}
& -\frac{\partial \lambda}{\partial t} - \frac{5K_{fh}}{3}h^{2/3}\frac{\partial \lambda}{\partial x} + \sum_{j=1}^M \delta_A(x - x_j) \left[\int_0^L \delta_A(x - x_j)h \, dx - h_j^{meas}(x_j, t) \right] \\
& + \sum_{k=1}^N \left[\frac{5}{3} \sum_{k=1}^P w_k \phi(\|x - x_{ck}\|) \delta_A(x - x_k) h^{2/3} \left\{ \int_0^L \delta_A(x - x_k) \sum_{k=1}^P w_k \phi(\|x - x_{ck}\|) h^{5/3} \, dx \right. \right. \\
& \quad \left. \left. - f_k^{meas}(x_k, t) \right\} \right] = 0
\end{aligned} \tag{III.31}$$

By applying the general calculation of first variation to this special case, the gradient of the cost functional \mathcal{L} with respect to the weighting factors w_k can be deduced by considering the $K_{fh}(n, S_0)$ as parameter $\alpha(x)$ of general case and by replacing the first variation $\delta\alpha(x)$ by $\frac{\partial K_{fh}}{\partial w_k} \delta w_k = \phi(\|x - x_{ck}\|) \delta w_k$ and adding the variation δw_k in the first variation of \mathcal{J} .

This gradient is thus

$$\begin{aligned} \nabla \mathcal{L}_{w_k} = & \sum_{k=1}^N \int_0^T \int_0^L \delta_A(x - x_k) h^{5/3} \phi(\|x - x_{ck}\|) \left[\int_0^L \delta_A(x - x_k) \sum_{k=1}^P w_k \phi(\|x - x_{ck}\|) h^{5/3} dx \right. \\ & \left. - f_k^{meas}(x_k, t) \right] dx dt - \int_0^T \int_0^L h^{5/3} \frac{\partial \lambda}{\partial x} \frac{\partial K_{fh}}{\partial w_k} dx dt + \int_0^T \lambda \left[h^{5/3} \frac{\partial K_{fh}}{\partial w_k} \delta w_k \right] \Big|_{x=0}^{x=L} dt + \varepsilon_1(w_k - w_{kF}) \end{aligned} \quad (\text{III.32})$$

The optimization process is solved by using steepest descent method. The four points implicit schemes are the discretization tool to solve the PDE of overland flow and adjoint equation. This result is presented in author's publication [NGB15a].

Estimation result

The illustrative example to estimate w_k is chosen with the same condition of rainfall

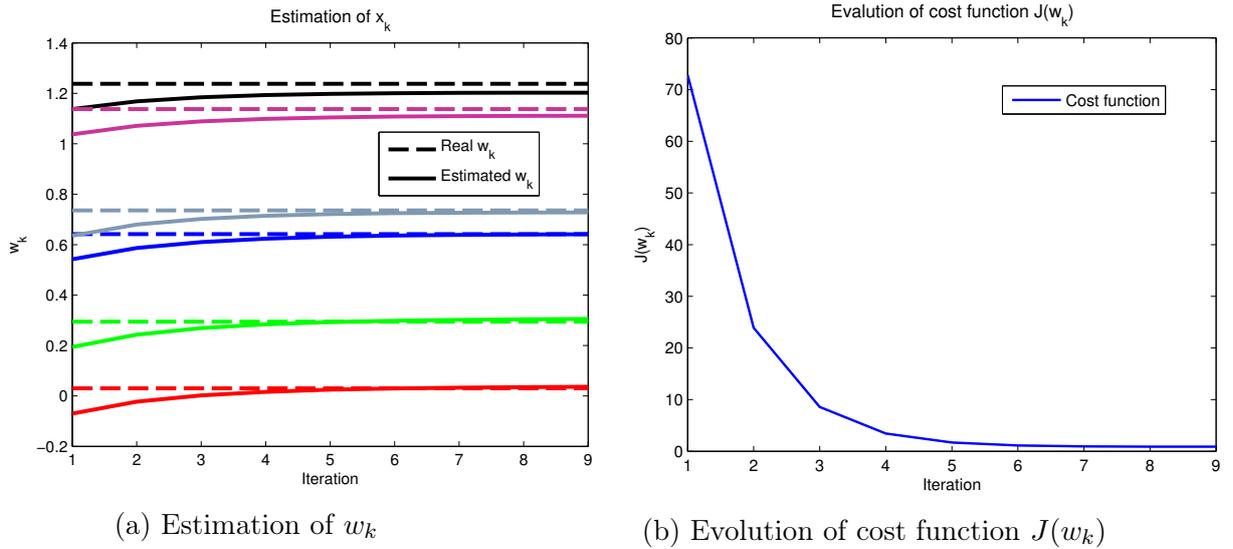
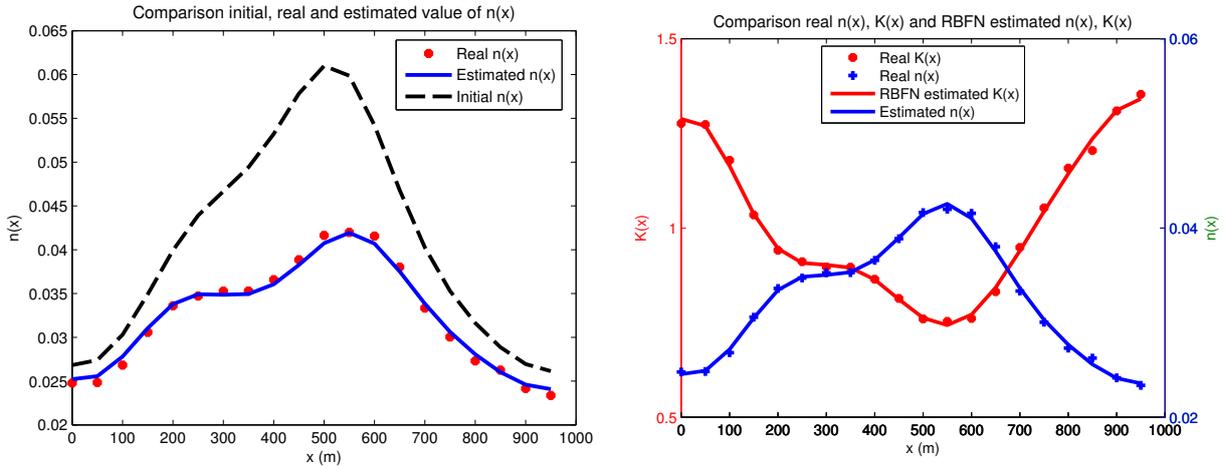


Figure III.15: Estimation results of distributed Manning coefficient in overland flow (I).

and infiltration as an identical hill slope. The detail of all parameters used in this example can be seen in Table III.4. From the real value of $n(x)$, K_{fh} is calculated via formula $K_{fh}(n, S_0) = \frac{S_0^{1/2}}{n(x)}$. By solving the RBFN formula with predefined parameters, the weighting factors w_k , called real weighting factors of RBFN, can be then obtained. 5 values of flow depth and 5 values of flow per unit width are taken from the simulation of system with real value of Manning coefficient, as observation values. The positions of observations are

chosen randomly in interval $[0, L]$. The RBFN parameters used to approximate $K(x)$ are: 6 hidden functions, $P = 6$; $x_{ck} = [0 \ 320 \ 400 \ 500 \ 750 \ 1000]$ and $\sigma_K = 200$. After a simple inversion, the real values of w_k is $w_k = [1.2964 \ 0.6027 \ 0.0906 \ 0.2671 \ 0.7707 \ 1.1499]$. Due to the error of RBFN approximation of $K(x)$, the corresponding cost function value is 0.1664 larger than zero. The initial value of w_k for steepest descent is $w_k \text{ initial} = [1.2064 \ 0.5127 \ 0.0006 \ 0.1771 \ 0.6807 \ 1.0599]$ not far from the real one. One can see how



(a) Comparison real and estimated value of $n(x)$

(b) Comparison real value and RBFN estimated value of $n(x)$, $K(x)$

Figure III.16: Estimation results of distributed Manning coefficient in overland flow (II).

the estimation w_k indeed converges in Figure III.15a with some biases on estimation values of w_1 and w_2 . The errors on w_k are possibly caused by convergence limitation of steepest descent method. As depicting in Figure III.15b, the value of cost function \mathcal{J} decreases from 15670 to the minimum 0.8964 which is a bit larger than the theoretical one 0.1664.

After getting the optimal estimated w_k , one can get the estimated K_{fh} and also the estimated values of $n(x)$. Figure III.16b presents the comparisons between the real K_{fh} , and $n(x)$ and estimated ones. One can observe more clearly, in Figure III.16a, the convergence from initial value of $n(x)$ to the optimal value. In the same figure, the estimated $n(x)$ is pretty close to the desired one and the average relative error between them is small 1.6818%. By using the RBFN, the estimation of 19 discrete values of $n(x)$ is successfully transformed and solved by an equivalent problem but the size of optimization problem reduced to 5 values of w_k .

2.b.3 Estimation of initial state and distributed Manning roughness

Estimation problem statement

In this example, the author tries to increase the complexity of estimation problem by considering to estimate directly (without RBFN approximation) the Manning roughness and also the initial condition of flow depth $h_0^i(x)$ at the same time and using the same measurement. This problem is one among two illustrative examples in author's publication [NGB16b] (the second example is the state and parameter estimation in traffic flow which has been previously presented). The Manning coefficient will be estimated via the estimation of parameter $\alpha(x) = S_0^{1/2}/n(x)$. The system equation is still the PDE of water depth $h(x, t)$ as in equation (III.22). The cost function \mathcal{J} has a simpler form, defined by the errors between simulations and some lumped observation values of only water depth $h(x, t)$, as equation below:

$$\begin{aligned} \mathcal{J} = & \frac{1}{2} \sum_{j=1}^N \int_0^T \left\{ \int_0^L \delta_A(x - x_j) h \, dx - h_j^{meas}(x_j, t) \right\}^2 dt + \frac{1}{2} \varepsilon_1 \int_0^L \|h_0^i(x) - h_{0F}^i(x)\|^2 dx \\ & + \frac{1}{2} \varepsilon_2 \int_0^L \|\alpha(x) - \alpha_F(x)\|^2 dx \end{aligned} \quad (\text{III.33})$$

By comparing this to the general form of flow in equation (II.1), one has the vector of only one parameter $\alpha(x)$ and the corresponding function of system variable $\varphi(h) = h^{5/3}$. The switching sub-functions g_1 , g_2 , h_1 and h_2 do not exist, neither their partial derivatives. By inserting the derivative $\frac{\partial f(u, x)}{\partial h} = \frac{5}{3} \alpha(x) h^{2/3}$ into equation (II.28), the adjoint system of $\lambda(x, t)$ can be constructed. Replacing the derivative $\frac{\partial f(h, x)}{\partial \alpha(x)} = h^{5/3}$ into equations (II.38), the gradient of cost functional \mathcal{L} with respect to α for the interior spatial section is obtained.

Estimation result

This example is assumed to be taking place on the same hills as in the two earlier ones with characteristics described in Table. III.4. The real value of the initial condition of flow height $h_0^i(x) \forall x \in [0 \dots L]$ is depicted in Figure III.17a. The real value of distributed parameter $\alpha(x) \forall x \in [0 \dots L]$ (being calculated via equation $\alpha(x) = S_0^{1/2}/n(x)$ with chosen values of $n(x)$) is sketched in Figure III.17b. From the exact simulation of the system, 2 observation values of water depth, 1 at the begin and 1 at the end, are taken out of 10 discretized sections. The interior point algorithm of fmincon is initialized by $0.35(m) \forall x \in [0 \dots L]$ for the initial state and $7 \forall x \in [0 \dots L]$ for $\alpha(x)$. The estimation of $h_0^i(x)$ indeed converges to the

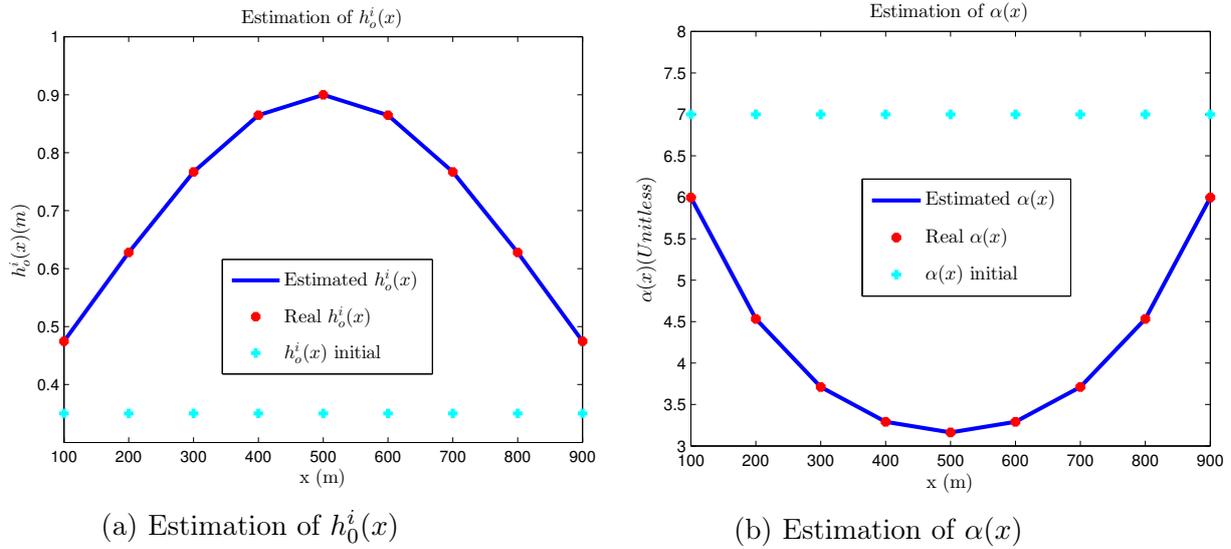


Figure III.17: Estimation results of distributed Manning coefficient and initial condition in overland flow (III).

desired values in Figure III.17a and parameter $\alpha(x)$ also converges in Figure III.17b after 192 iterations with the average relative errors of estimated values of $h_0^i(x)$ and $\alpha(x)$ equal to $3.4240 \times 10^{-8}\%$ and $2.4165 \times 10^{-8}\%$ respectively. One can see that the estimation is pretty accurate with such small errors. Along with the convergence of state end parameter, the cost function \mathcal{J} reaches its minimum at 1.1478×10^{-18} . After the termination of optimization algorithm, the Manning coefficient $n(x)$, recovered via optimized parameter $\alpha(x)$, has also a small relative error ($3.4240 \times 10^{-8}\%$).

2.b.4 Estimation of initial state, distributed Manning roughness and infiltration parameters

Estimation problem statement

With the same spirit as in the state and parameter estimation for switched traffic flow, the object of this example is to consider the estimation of the parameters of the infiltration process at the same time with state and water flow parameter. But unlike the case of switched traffic flow, the original equation of overland flow with Green-Ampt infiltration is already a switched PDE coupled with the ODE of cumulative infiltration because of the ponding time. On the basic the presentation of system equation in subsection 2.a, the

overland flow system can be rewritten as equation (III.34) below with switching time t_{pon} :

$$\left\{ \begin{array}{ll} \frac{\partial h(x,t)}{\partial t} + \frac{\partial f(h(x,t),x)}{\partial x} = 0 & \text{if } t \leq t_{pon} \\ \frac{\partial h(x,t)}{\partial t} + \frac{\partial f(h(x,t),x)}{\partial x} = r - K_i \left(\frac{\Psi\eta(1-\theta)}{I(t)} + 1 \right) & \text{if } t > t_{pon} \\ \frac{dI(t)}{dt} = rt & \text{if } t \leq t_{pon} \\ \frac{dI(t)}{dt} = K_i \left(\frac{\Psi\eta(1-\theta)}{I(t)} + 1 \right) & \text{if } t > t_{pon} \end{array} \right. \quad (\text{III.34})$$

These equations have exactly the same form in the general one (II.5). One can notice that the cumulative infiltration $I(t)$ plays the role of ODE variable $y(t)$. The unknown parameters of infiltration are $p_1 = K_i$, $p_2 = K_i\Psi\eta(1-\theta)$ and the vector of parameters $p = [p_1 \ p_2]$. There is only one element of vector α defined as $\alpha_1 = S_0^{1/2}/n(x)$, like in the previous example. Moreover, the switching condition in this example is a function of time, and defined by $\xi = t - t_{pon}$; the sub-switching functions are $\{g_1 \ g_2\} = \{0 \ r - p_2/I(t) - p_1\}$ and $\{h_1 \ h_2\} = \{rt \ p_2/I(t) + p_1\}$. The adjoint models of variables λ and γ in equations (II.28) and (II.29) can be rewritten by replacing the following partial derivatives:

$$\begin{array}{llll} \frac{\partial f}{\partial h(x,t)} = \frac{5}{3}\alpha_1 h^{2/3}(x,t) & ; & \frac{\partial g_1}{\partial h(x,t)} = 0 & ; & \frac{\partial g_2}{\partial h(x,t)} = 0 & ; & \frac{\partial g_1}{\partial I(t)} = 0 \\ \frac{\partial g_2}{\partial I(t)} = \frac{p_2}{I^2(t)} & ; & \frac{\partial h_1}{\partial I(t)} = 0 & ; & \frac{\partial h_2}{\partial I(t)} = -\frac{p_2}{I^2(t)} & & \end{array} \quad (\text{III.35})$$

The corresponding adjoint systems of $\lambda(x,t)$ and $\gamma(t)$ are subsequently presented as follows:

$$-\frac{\partial \lambda}{\partial t} - \frac{5}{3}\alpha_1 h^{2/3}(x,t) \frac{\partial \lambda}{\partial x} + \sum_{j=1}^N \delta_A(x - x_j) \times \left[\int_0^L \delta_A(x - x_j) h(x,t) dx - h(x_j,t)_j^{meas} \right] = 0 \quad (\text{III.36})$$

$$-\frac{\partial \gamma}{\partial t} + \gamma \varphi_a \frac{p_2}{I(t)} - \int_0^L \lambda \varphi_a \frac{p_2}{I^2(t)} dx = 0 \quad (\text{III.37})$$

The gradient of objective functional with respect to the parameters p and α is also found by the same method, replacing the derivatives in equations (II.35) and (II.59), respectively,

by the following derivatives:

$$\frac{\partial \varphi_a}{\partial p_1} = -\frac{rK_{ac}p_2}{r^2(r-p_1)^2}\varphi_a^2 e^{-K_{ac}\xi}; \quad \frac{\partial \varphi_a}{\partial p_2} = -\frac{K_{ac}}{r(r-p_1)}\varphi_a^2 e^{-K_{ac}\xi}; \quad \frac{\partial h_1}{\partial p_2} = 0; \quad \frac{\partial h_2}{\partial p_1} = 1 \quad (\text{III.38})$$

$$\frac{\partial f}{\partial \alpha_1} = h^{5/3}(x, t); \quad \frac{\partial g_2}{\partial p_1} = -1; \quad \frac{\partial g_2}{\partial p_2} = -\frac{1}{I(t)}; \quad \frac{\partial h_1}{\partial p_1} = 0 \quad (\text{III.39})$$

$$\frac{\partial h_2}{\partial p_2} = \frac{1}{I(t)}; \quad \frac{\partial g_1}{\partial p_1} = 0; \quad \frac{\partial g_1}{\partial p_2} = 0 \quad (\text{III.40})$$

Estimation result

Parameters	Value	Unit
Overland flow length L	300	m
Simulation time	100	s
Space step Δx	30	m
Time step Δt	0.01	s
Bed slope S_0	0.1	m/m
K_i	0.05	m/s
Ψ	1.4	cm
η	0.8340	%
θ	0.0863	Unitless
r	0.07	m/s
K_{ac}	1×10^{-7}	Unitless
Calibration coefficients ϵ_1	$\underbrace{[1 \ 1 \dots 1]}_{\text{size 9}} \times 10^{-9}$	Unitless
Calibration coefficients ϵ_2	$\underbrace{[1 \ 1 \dots 1]}_{\text{size 9}} \times 10^{-9}$	Unitless
Calibration coefficients ϵ_3	$[1 \ 1] \times 10^{-9}$	Unitless
First guessed value of $n(x)$, α_{1F}	$0 \ \forall x \in [0, L]$	$s/m^{1/3}$
First guessed value of K_i , p_{1F}	0.1	m/s
First guessed value of $K_i\Psi\eta(1-\theta)$, p_{2F}	0.05	Unitless
First guessed value of initial condition $h_0^i(x)$, $h_{0F}^i(x)$	$0 \ \forall x \in [0, L]$	m
Observation number N	5	Unitless

Table III.5: Parameters of overland flow and numerical value of optimization parameter used for estimating of $\alpha_1 = S_0^{1/2}/n(x)$, initial condition $h_0^i(x)$ and infiltration characteristics $p = [K_i \ K_i\Psi\eta(1-\theta)]$.

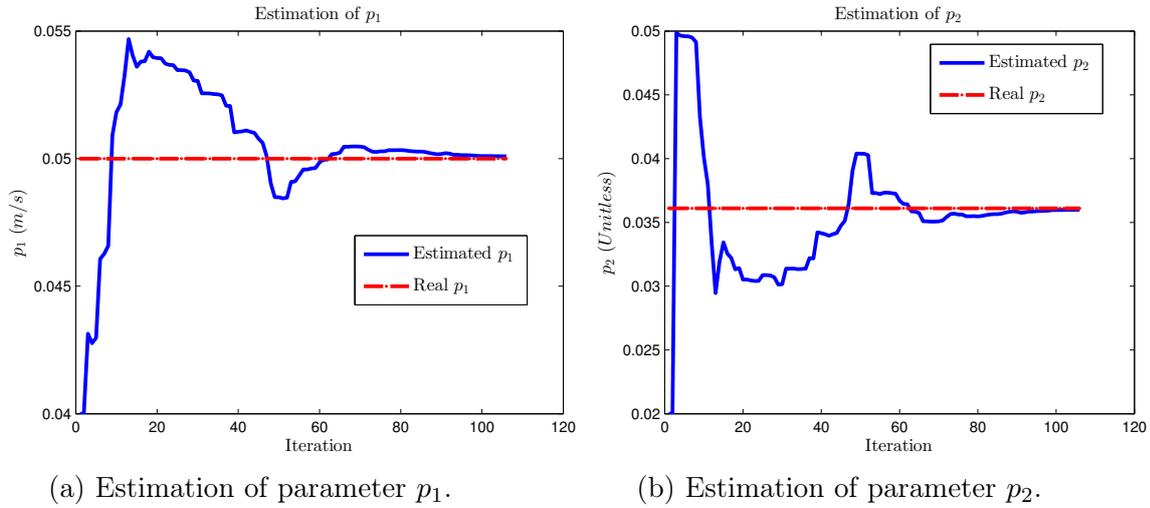


Figure III.18: State and parameter estimation result in switched overland flow (II).

The numerical simulation model of an overland flow with the characteristic presented in Table III.5 is considered. There are 9 discrete values of the Manning coefficient and initial state will be estimated along with two parameters p_1 and p_2 of the infiltration process. For the measurements, 5 observed values of water flow height at discrete positions $\{2, 3, 6, 9, 10\}$, among 11 discrete positions of 300-m long hill slope, are used for estimation.

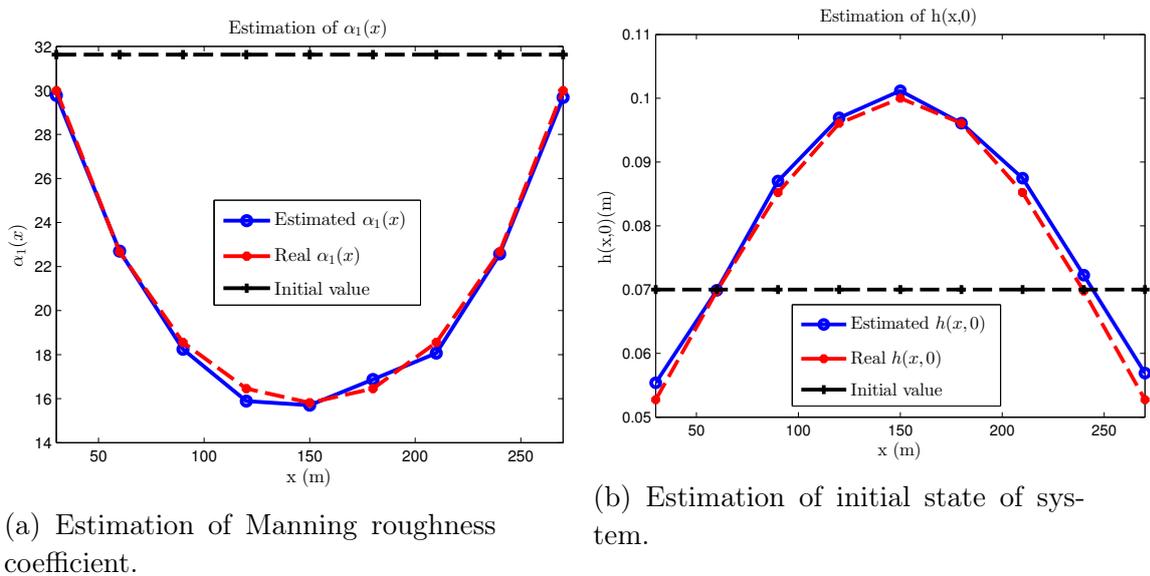


Figure III.19: State and parameter estimation result in switched overland flow (I).

Figure III.19a and Figure III.19b illustrate the estimation results for the state and the parameters. Initial values are shown in dashed line with plus sign. The convergence to the real values can be seen in the same figures, under the blue continuous line with circle. The average relative between the estimated and real one are 1.4703% and 2.6109% for the Manning coefficient and the state respectively. One can notice some slight different accuracies depending on the position. Two figures Fig.III.18a and Fig.III.18b provide the insight of the convergence to the desired ones 0.0501 and 0.0360 of infiltration parameters p_1 and p_2 from the initial values 0.04 and 0.02. The estimation relative errors are 0.1859% and 0.3182% respectively. in parallel with these convergences, the evolution of cost function \mathcal{J} reaches it minimal value 1.8248×10^{-4} , seen in Fig.III.20.

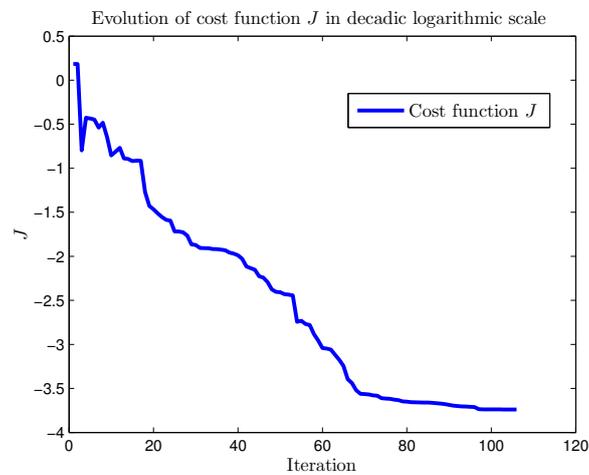


Figure III.20: State and parameter estimation result in switched overland flow (III): the evolution of cost function \mathcal{J} .

From all those examples, it results that the proposed methodology can be applied pretty efficiently, in estimation problem which can be quite complex. In the next section, an application with real data is further given.

2.c Estimation problems using real measurement of Tondi Kiboro catchment

2.c.1 Overview of Tondi Kiboro catchment and real measurements

This subsection is intended for introducing the parameter estimation problem for the so-call Tondi-Kiboro catchment. Being one of the small and typical catchments of the Sahelian area, the Tondi-Kiboro has been affected by the global climate change leading to hydro-

logical change which is one of the major importance environmental challenges nowadays (further information can be found in [Mou+11], [Sou08]). In the last four decades for example, strong rainfall decrements together with a significant change in the land use have been noted in this catchment. Moreover, the water cycle in such locations are strongly modified by this phenomenon. The increment of flow channel infiltration and runoff factor also causes topsoil debasement and vegetation clearing.

Some studies in hydrological domains have investigated this area of Sahelian region such as two researches of Descroix et al. to find out the effect of land use change on the runoff evolution [Des+11] or the possibility to define this area of Sahelian region as a deep infiltration one [Des+12]. There is no dynamical model of water flow in those references, and some parameters used in these works, including the Manning roughness coefficient of the soil, characterizing the friction of water flow on the soil surface and the factors of infiltration process are, in general not easily, nor accurately, known. These factors indeed are usually considered as empirical parameters, and their values depend on the soil surface characteristics, water topology, vegetation and crop. They are usually obtained from some empirical calculations applied to measured data or from a lookup table, with small adjustment to adapt to the specific context if necessary. On the basis of the presented adjoint-based estimation methodology of the chapter II, an alternative approach for Manning coefficient and infiltration parameters estimation is proposed. The essential assumption in this method is that the water flow on Tondi Kiboro catchment is described by the continuity equation of Saint-Venant model as equation (III.19). This hypothesis can be verified by comparing the simulated water discharge (system simulation with the estimated values of parameters) with the measurement after the estimation process.

A good estimation of these parameters allows to have a simple numerical model to represent the real water flow of Tondi Kiboro basin and promote further hydrology research to study other effects of rain fall on the water cycle, soil erosion and fallow period. The inverse problems investigated on this basin are the subject of two publications [Ngu+15], dealing with Manning roughness estimation and [Ngu+16] for Manning coefficient and parameters of Horton infiltration model.

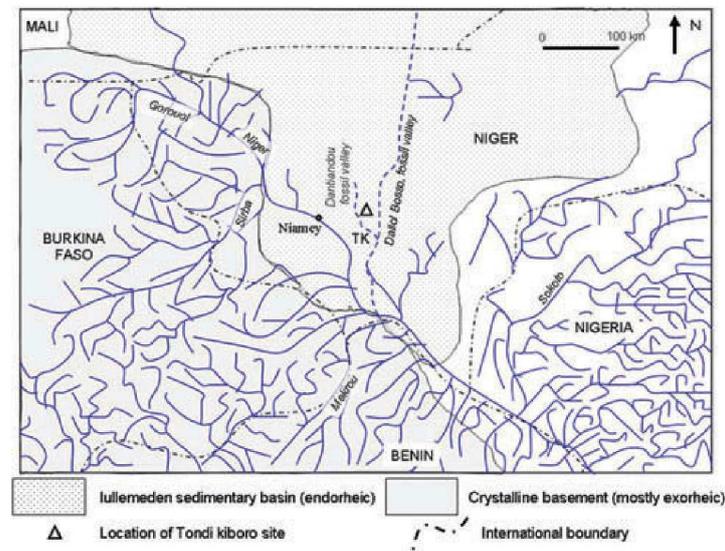


Figure III.21: Location of Tondi Kiboro catchment.

Location of Tondi-Kiboro catchment

Banizoumbou Tondi Kiboro catchment is a set of basins, located near Tondi Kiboro village, 70 km east of the Niamey region, at the Fakara plateau southwest Niger (Figure III.21). Its bed slope S_0 is about 3.5%. This catchment consists of three small basins and two of them are nested as illustrated in Figure III.22 with surfaces of 46800 m^2 and 63720 m^2 for the upper and lower basins respectively and the total ground surface of this area (including upstream and downstream stations) is 110520 m^2 . Only the up basin of the two nested ones (with a length of 600 m approximately, between the upper-stream end and downstream end and measurements stations) is taken into account.

The top soil of this area is made of 10% of silt and clay, 90% of sands. These components play an important role on not only vegetation and crop growth but also the infiltration and overland processes. They also contribute significantly to fix the value of Manning coefficient. Moreover, due to the small size of the considered Tondi Kiboro basin, the rain fall rate and infiltration rate can be considered to be uniform all over the area. Manning coefficient n is also constant because the characteristics of soil do not vary spatially.

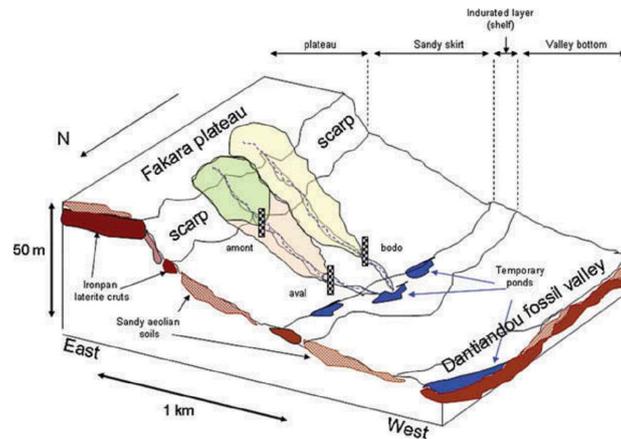


Figure III.22: Overview of Tondi Kiboro basins.

Real measurements

The only measurement on flow dynamics available is the water discharge during all rainfall events at the end of the basin, equivalent to the position L in the simulator. This observed value of water discharge is provided by stream flow sensor equipped on this basin, with locations presented in Figure III.23, in the measurement campaign from 2004 to 2012. Each observed water discharge is associated to the corresponding measurements of rainfall rate provided by 1 day raingauge (see its location in Figure III.23).

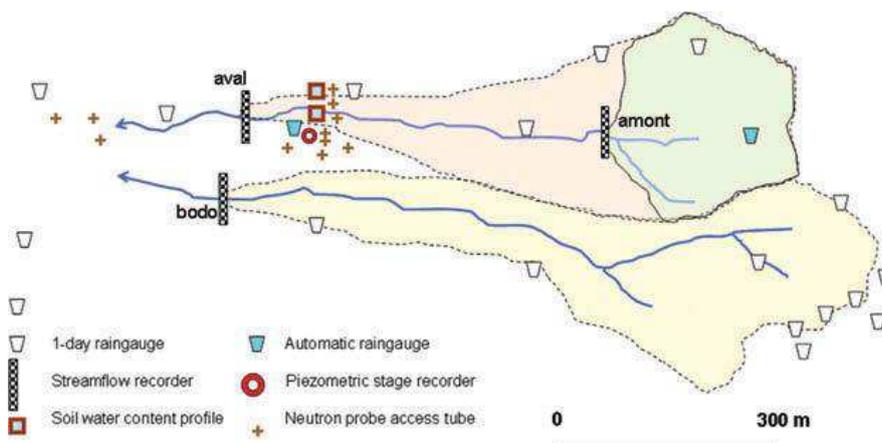


Figure III.23: Observation stations on Tondi Kiboro catchment.

Parameters	Value	Unit
Length of basin	600	<i>m</i>
Simulation time	3960	<i>second</i>
Space step Δx	60	<i>m</i>
Time step Δt	60	<i>second</i>
Bed slope S_0	3.5%	<i>m/m</i>
Observation number N	1	<i>Unitless</i>
Surface flow width at position L, W	78	<i>m</i>

Table III.6: Geometric characteristic of Tondi Kiboro catchment.

All the geometric characteristics of this basin and numerical simulations factors are summarized in Table. III.6.

2.c.2 Estimation of Manning coefficient with Green-Ampt model for unsteady rain

Estimation statement

The purpose of this part is to estimate the Manning coefficient n in equation (III.20). The infiltration model used to calculate the infiltration rate is the Green-Ampt modified model for unsteady rainfall process. This model is different from the one used for steady rain which is already presented in previous sections. Let us recall the dynamical model of an overland flow:

$$\begin{cases} \frac{\partial h(x, t)}{\partial t} + \frac{\partial f(h(x, t), x)}{\partial x} = r - i \\ h(x, 0) = h_0^i(x) \quad h(0, t) = h_0^b(t) \end{cases} \quad (\text{III.41})$$

The initial state function $h_0^i(x)$ and boundary function $h_0^b(t)$ in this case are supposed to be equal to zero at the begin of rainfall process. This assumption is reasonable because this region is generally dry all the time between two rain events. The measured rainfall in this real case is not a steady one, its intensity varies as a function of time $r(t)$ and the same with infiltration rate $i(t)$. The original Green-Ampt model was developed to deal with the case of constant or steady rainfall and is not suitable for unsteady one. To overcome this difficulty, S.T. Chu [Chu78] proposed modifications to adapt to the new complex situation. The idea is that there is almost one ponding time in a steady rain. The infiltration process begins with an unponded surface, transforms to a stage with surface ponding and keeps lasting to the end of rainfall event. The equation of this original model allows to calculate the ponding time beforehand which only depends on the rainfall intensity and some soil's parameters. But in the case of unsteady rain, there are more than one ponding moments because the rainfall rate exceeds the infiltration rate at several periods and the infiltration process can shift back to original stage from the current one.

Parameters	Value	Unit
Effective hydraulic conductivity K_i	0.0479	cm/s
Average suction at the wetting front Ψ	28.5	cm
Saturated water content	0.1	cm^3/cm^3
Initial water content at start of the rainfall event	0.01	cm
Soil surface storage	10	cm

Table III.7: Infiltration parameter of modified Green-Ampt model on Tondi Kiboro catchment.

The details of the modified model in unsteady rain are not presented here. They can be found in [Chu78] and [SK82] with entire formula and a calculation validation. Based on these formula, the infiltration rate is calculated offline (following the equations and Fortran code developed in *WinGAmpt* tool, Agricultural and Biological Engineering Department University of Florida, [MCEP09]) with rainfall data and some soil parameters presented in Table III.7. Only measurements on water discharge are available and the water discharge $Q(x, t)$ at position and time t depends also the unknown Manning coefficient

$$Q(x, t) = Wf(h, x) = W \frac{S_0^{1/2}}{n} h^{5/3}$$

with W the surface flow width at position x . Because the values of bed slope S_0 and water flow width W are perfectly known, the estimation of Manning coefficient can be related to the estimation of variable $\alpha_n = \frac{S_0^{1/2}}{n}$. Notice that the flow sensor equipped on Tondi Kiboro catchment is at the end of watercourse or corresponding to the point near the last position of numerical model (the length of the numerical model is 660 m , longer than the real flow with one spatial step). The cost function must be redefined, for this special case of one sensor at numerical position $x_{meas} = L$, as follow:

$$\mathcal{J} = \frac{1}{2} \int_0^T \left\{ \int_0^L \delta_A(x - x_{meas}) Wf(h, x) dx - Q^{meas}(x_{meas}, t) \right\}^2 dt \quad (III.42)$$

Comparing with the general case, the adjoint equation of $\lambda(x, t)$ becomes:

$$-\frac{\partial \lambda}{\partial t} - \frac{5}{3} \alpha_n h^{2/3} \frac{\partial \lambda}{\partial x} + \frac{5}{3} W \alpha_n h^{2/3} \delta_A(x - x_{meas}) \left[\int_0^L \delta_A(x - x_{meas}) Wf dx - Q^{meas}(x_{meas}, t) \right] = 0 \quad (III.43)$$

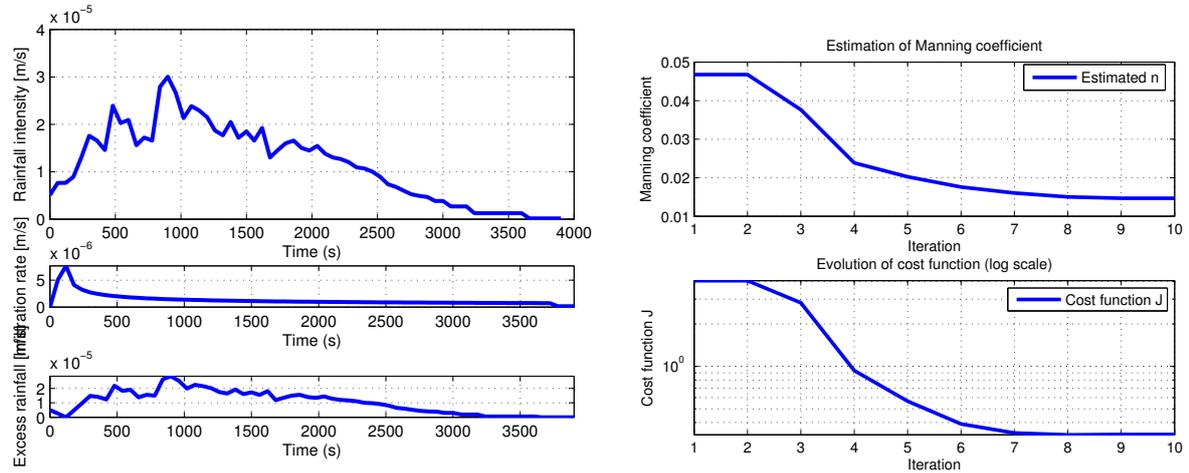
And the equation of gradients for parameter α_n is

$$\begin{aligned} \nabla \mathcal{L}_{\alpha_n} = & \int_0^T \left[\lambda h^{5/3} \right]_0^L dt - \int_0^T \int_0^L h^{5/3} \frac{\partial \lambda}{\partial x} dx dt + \int_0^T \int_0^L \delta_A(x - x_{meas}) W h^{5/3} \\ & \times \left[\int_0^L \delta_A(x - x_{meas}) W \alpha h^{5/3} dx - Q^{meas}(x_{meas}, t) \right] dx dt \end{aligned} \quad (\text{III.44})$$

The numerical tools used in this example are Lax-Wendroff scheme and the interior point algorithm of *fmincon*.

Estimation result

Among the used data (provided by LTHE), two sets are chosen and used: one for optimization and one for verifying the consistency of the estimated model. Each data set includes one measurement of water discharge $Q^{meas}(x_{meas}, t) \forall t \in [0, T]$ and one rainfall rate value $r(t) \forall t \in [0, T]$. The rainfall data set employed in the optimization process is taken on 2012/06/21 from 08:12 to 09:17 universal time. The infiltration rate and excess rain fall are calculated offline. One can observe in Figure III.24a the values of rainfall, infiltration rate and excess rainfall which are calculated offline and independent of the optimization algorithm.



(a) Rainfall rate, soil moisture, infiltration rate, excess rainfall rate on 2012/06/21

(b) Estimation of Manning coefficient and cost function \mathcal{J} .

Figure III.24: Manning coefficient estimation results using real data of Tondi-Kiboro catchment (I).

The estimation of Manning coefficient (via estimation of parameter α_n) converges to the value of 0.0151 (see convergence of n in Figure III.24b). This estimated value of n is pretty reasonable and suitable with the mentioned characteristics of soil on Tondi Kiboro basin. The found minimal cost function value is $\mathcal{J}^* = 0.3288$ after 10 iterations. One can see the cost function evolution in Figure III.24b). After the optimization procedure, the value of estimated value of n is verified by simulating again the system equation with this value of n and compare the discharge at the end of watercourse with the measured data. The small biases between them shown in Figure III.25 allow to conclude that the model simulates pretty well the characteristics of overland flow on Tondi Kiboro catchment. In

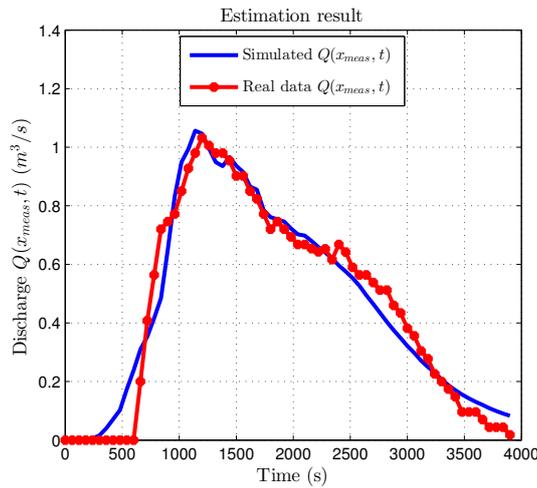
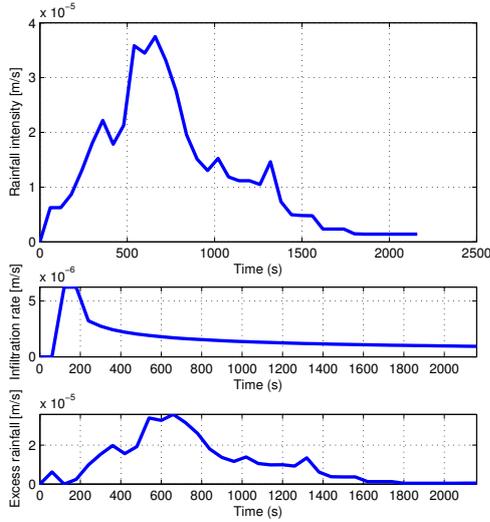
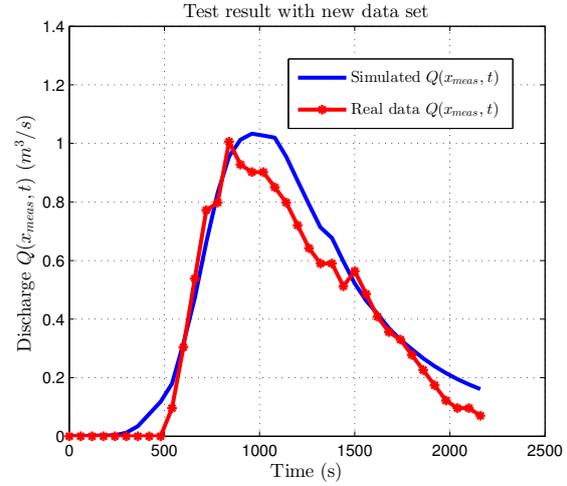


Figure III.25: Manning coefficient estimation results using real data of Tondi-Kiboro catchment (II): The comparison between simulated discharge and measured discharge.

order to verify the consistency of the numerical model and effectiveness of the optimization process, the second data set measured on 2012/08/03 from 20:33 to 21:09 universal time is thus used to simulate the system. Figure III.26a describes this rainfall rate of in this data and the corresponding infiltration rate.



(a) Rainfall rate, soil moisture, infiltration rate, excess rainfall rate on 2012/08/03



(b) Comparison between simulated discharge and measured discharge with another data set

Figure III.26: Manning coefficient estimation results using real data of Tondi-Kiboro catchment (III).

The simulated water discharge at position x_{meas} follows the measured value of this variable with an acceptable error as shown in Figure III.26b. These biases are possibly caused by the error on sensor and in measurement procedure, and by the fact that only the continuity equation is taken in consideration and the momentum one (which describes a part of the system dynamics) is neglected. This validation allows to consider that equation (III.22) with mentioned parameter and estimated Manning coefficient as a validated numerical model of overland flow taking place on Tondi Kiboro basin.

2.c.3 Estimation of Manning coefficient and infiltration parameter with Horton model

Estimation statement

After the estimation of Manning coefficient, the question is how can one can estimate also the infiltration process only from the rainfall infiltration and the water discharge. The information about the Green-Ampt infiltration in the preceding case presented in Table III.7 is assumed to be unknown. But the complexity (discontinuities due to multi-ponding time) of this model under variable rainfall condition creates a lot of difficulty for the formulation of adjoint-gradient. The empirical Horton model is adopted instead of the physical Green-Ampt model to overcome this obstacle. Being a purely empirical model, the

Horton model is adopted to describe the dynamics of infiltration process after the rainfall intensity exceeds the infiltration capacity of soil. With the meaning of a fitting equation to data, Horton model does not depend on rainfall or soil properties and is described by a function of time rather than the rainfall rate. The infiltration capacity tends to decrease from the initial infiltration rate in an exponential manner, as:

$$i(t) = i_c + (i_0 - i_c)e^{-kt} \quad (\text{III.45})$$

where $i(t)$ is the infiltration rate at time t , (m/s); k is the constant representing the exponential rate, ($1/s$); i_c is the equilibrium infiltration rate, (m/s); i_0 is the initial infiltration capacity, (m/s). As a consequence, this model only makes sense if $i(t)$ is lower than $r(t)$. This model can be compared to the Green-Ampt model after the ponding time. The soil surface of Tondi Kiboro catchment consists of 90% sand, which means that it has a very large infiltration capacity. As a result, the ponding time is usually very small which allows the use of the Horton model.

The model of overland flow at Tondi Kiboro is adjusted by turning 4 parameters, including Manning coefficient n , equilibrium infiltration rate i_c , initial infiltration rate i_0 and exponential rate k . They are also the parameters which will be identified. The cost function \mathcal{J} in this application case is defined the same as the previous one in equation (III.42), using only one observed value of water discharge at the end of considered domain. The system equation of overland flow can be seen as an non-switched hyperbolic system which corresponds to equation (II.1) or equation (II.5) without switching function g_2 and the associated ODE. Basically, the calculations to find the adjoint system and the weak form of gradient for estimated parameters can be directly obtained from the general case in section 3. And the PDE of adjoint variable $\lambda(x, t)$, the gradient for parameter $\alpha = S_0^{1/2}/n$ are the same with the case of Green-Ampt infiltration. On the other hand, the gradients for the parameters are below

$$\begin{aligned} \nabla \mathcal{L}_k &= \int_0^T \int_0^L \lambda \frac{\partial g}{\partial k} dxdt \text{ where } \frac{\partial g}{\partial k} = -t(i_0 - i_c)e^{-kt} \\ \nabla \mathcal{L}_{i_0} &= \int_0^T \int_0^L \lambda \frac{\partial g}{\partial i_0} dxdt \text{ where } \frac{\partial g}{\partial i_0} = e^{-kt} \\ \nabla \mathcal{L}_{i_c} &= \int_0^T \int_0^L \lambda \frac{\partial g}{\partial i_c} dxdt \text{ where } \frac{\partial g}{\partial i_c} = 1 - e^{-kt} \end{aligned} \quad (\text{III.46})$$

Estimation result

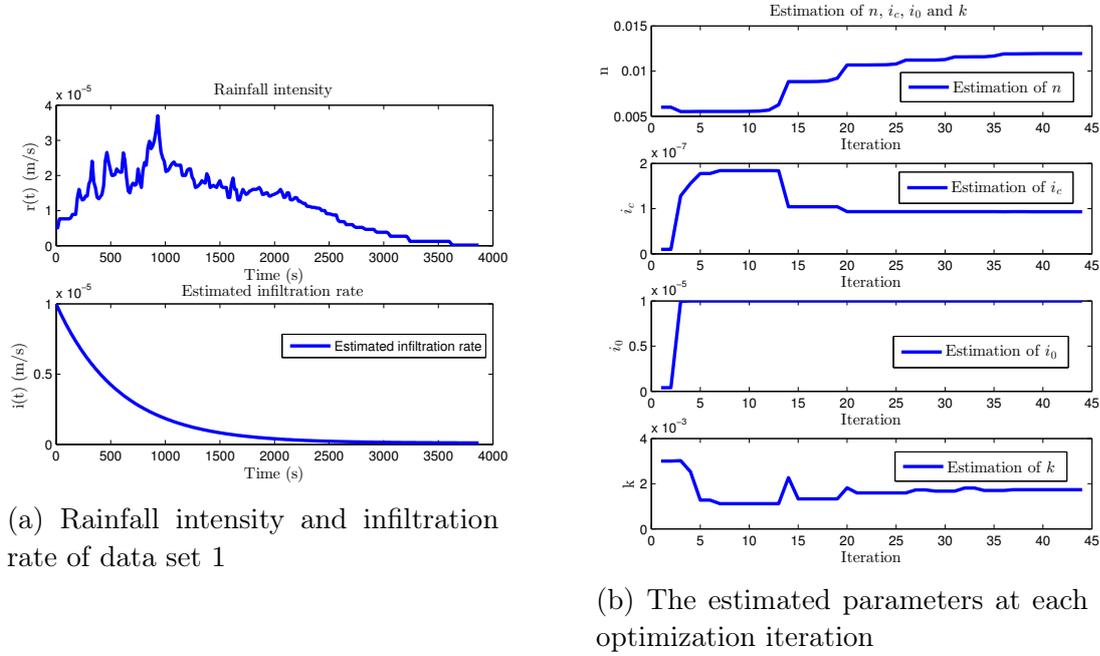


Figure III.27: Manning coefficient and Horton infiltration parameters estimation results using real data of Tondi-Kiboro catchment (I).

The geometric characteristics of Tondi-Kiboro in this estimation problem are the same as in Table III.6. The data used hereby are the same as in the preceding example with the single one observation value of water discharge, provided by stream flow sensor at position $x_{meas} = L$. In fact the same two data sets are chosen. The parameter estimation is realized on the first data set taken on 2012/06/21 from 08:12 to 09:17 universal time whose rainfall intensity is depicted in the high part of Figure III.27a. After 44 iterations as illustrated in Figure III.27b, and from the initial guessed value of parameters $[n^0; i_c^0; i_0^0; k^0] = [0.006; 1 \times 10^{-8}; 4 \times 10^{-7}; 3 \times 10^{-3}]$, the parameters converge to the optimal values $[0.0119; 9.3 \times 10^{-8}; 9.995 \times 10^{-6}; 0.0017]$. The estimated friction coefficient in this example is slightly different from the one obtained in the earlier one (0.0151). But both of these two values are consistent with the characteristics of the basin soil. The gap between them is possibly due to the different infiltration models. The cost function \mathcal{J} , shown in Figure III.28a decreases from 13.0643 to minimum value of 0.9833.

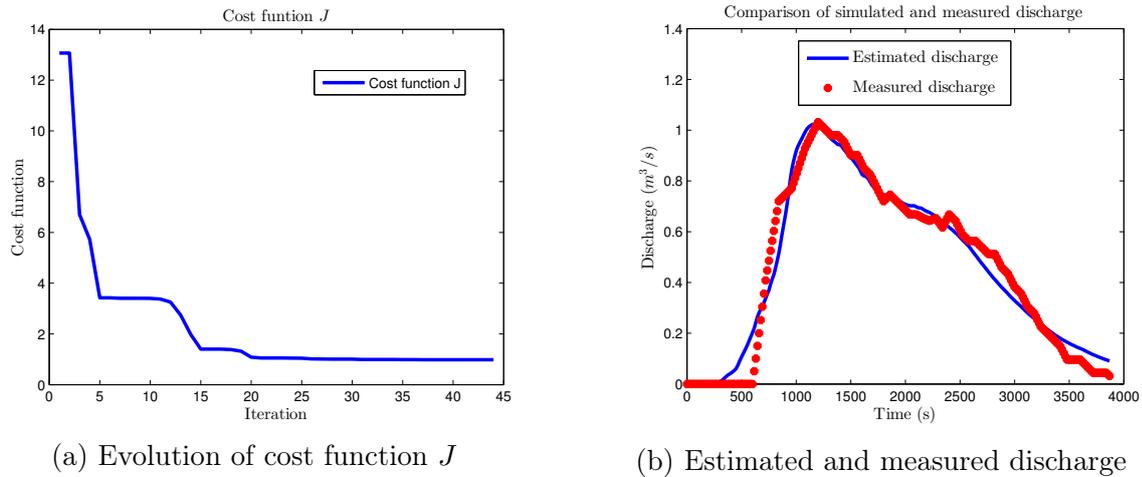


Figure III.28: Manning coefficient and Horton infiltration parameters estimation results using real data of Tondi-Kiboro catchment (II).

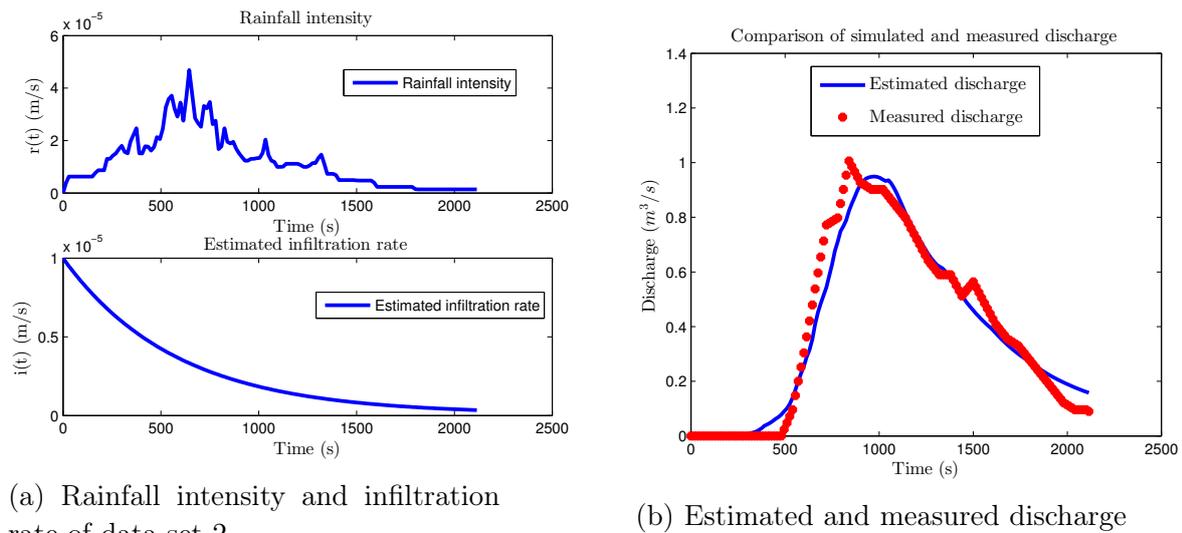


Figure III.29: Manning coefficient and Horton infiltration parameters estimation results using real data of Tondi-Kiboro catchment (III).

It is interesting to compare the water discharge measurement with the simulated one obtained by simulating system equation with the estimated variables. Figure III.28b describes this comparison. One can see that there are no big differences between them, confirming the correctness of estimation result. Moreover, the obtained result is very similar to the one already obtained in example 2.c.2. The roles of the second rainfall data set (taken on 2012/08/03 from 20:33 to 21:09 universal time) is to verify the consistency of estimated

values. The rainfall and infiltration rate (simulated with estimated parameters) are given in Figure III.29a. As observed in Figure III.29b, the simulated discharged fits with the measurements provided by sensors with some discrepancies which can be sensor error and the changes of soil surface and cultivated plant over time.

3 Estimation results with different configurations

Almost all the presented estimation examples have been carried out in the case where the initial state and/or distributed parameter are under the form of Gaussian functions. But the accuracy of estimation results and convergence of optimization algorithm possibly depend on the form of initial state and parameter of the system, as well as initial conditions fed to optimization process. It is thus necessary to test the proposed estimation approach with different configurations. Coming back to the case of state and distributed parameter estimation of non-switched overland flow presented in subsection 2.b.3, this numerical example will be re-carried out on different situations: 3 forms of initial state and distributed parameter, and for each form, there are 6 different tests with 6 initial conditions for optimization algorithm where 3 initial conditions are chosen manually and the 3 remaining are generated randomly. The geometric and numerical parameters for simulation is the same as example 2.b.3 (see Table III.4).

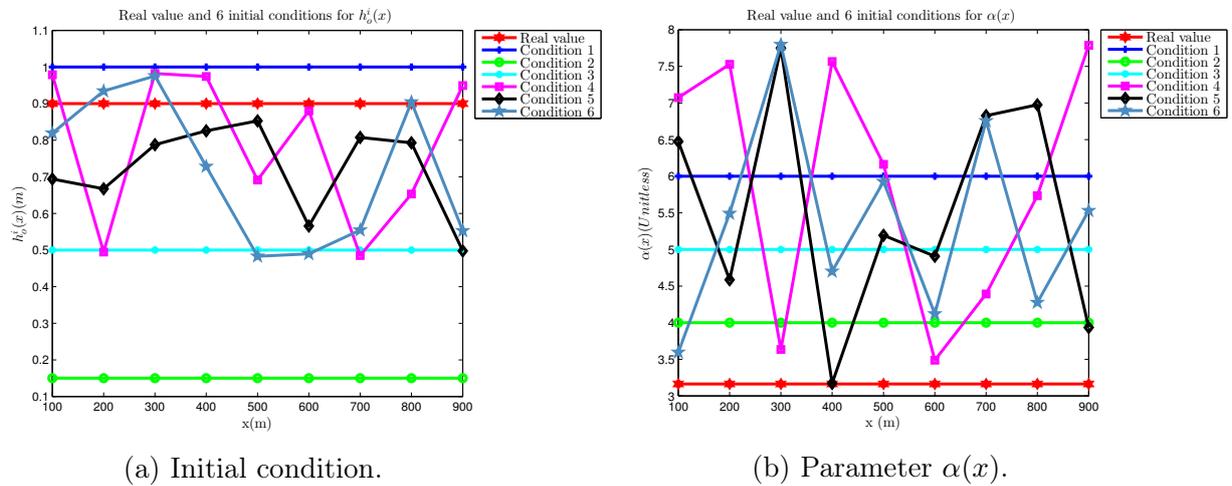


Figure III.30: Real values and 6 initial conditions of initial state and parameter $\alpha(x)$ in the case of constant initial state and parameter.

Constant initial state and parameter

In this case, the real values of initial state and parameter are chosen as constant distri-

butions (i.e. identical all over the spatial domain). Figure III.30 depicts all the 6 pairs of initial condition and parameter. The optimization process to estimate the chosen state and parameter starts firstly with 3 manually selected initial values and then with 3 random ones. Table III.8 summarizes the corresponding relative estimation error for the 6 cases. In all the different situations of starting condition, the optimization process is successful to recover the real states and parameter with very small relative error.

	Estimation relative errors and minimal \mathcal{J} for each initial condition of optimization algorithm					
	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
for $\alpha(x)$	5.09×10^{-8}	5.92×10^{-8}	5.92×10^{-8}	5.07×10^{-8}	8.50×10^{-8}	3.08×10^{-9}
for $h_o^i(x)$	3.09×10^{-8}	4.25×10^{-8}	3.79×10^{-8}	3.43×10^{-8}	5.30×10^{-8}	4.84×10^{-9}
\mathcal{J}	3.31×10^{-18}	3.39×10^{-18}	1.82×10^{-18}	3.74×10^{-18}	2.83×10^{-18}	2.37×10^{-20}

Table III.8: Estimation errors with 6 different initial conditions in the case constant initial state and parameter.

Gaussian initial state and parameter

The second case is to test the estimation accuracy when the real values of initial state and parameter follow a Gaussian distribution like in the example which has been presented in subsection 2.b.3. Similarly to the previous case, the error between real values and estimated values of both state and parameter are very small. Figure III.31 and Table III.9 represent respectively all the initial conditions and summary of estimation error.

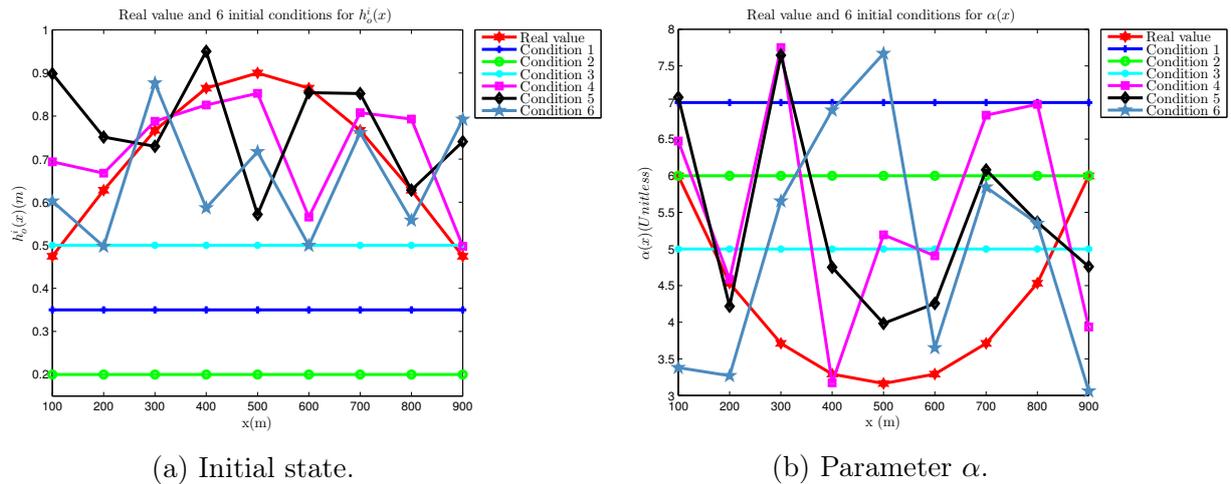


Figure III.31: Real values and 6 initial conditions of initial state and parameter $\alpha(x)$ in the case of Gaussian initial state and parameter.

	Estimation relative errors and minimal \mathcal{J} for each initial condition of optimization algorithm					
	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
for $\alpha(x)$	3.42×10^{-8}	2.58×10^{-8}	1.94×10^{-8}	1.26×10^{-8}	1.79×10^{-8}	1.38×10^{-7}
for $h_o^i(x)$	2.42×10^{-8}	1.82×10^{-8}	1.36×10^{-8}	9.80×10^{-8}	1.60×10^{-8}	1.04×10^{-7}
\mathcal{J}	1.15×10^{-18}	6.39×10^{-19}	2.09×10^{-19}	9.26×10^{-18}	4.60×10^{-15}	1.17×10^{-17}

Table III.9: Estimation errors with 6 different initial conditions in the case Gaussian initial state and parameter.

Random initial state and parameter

This last case is devoted to examine the results of adjoint based estimation approach in the situation where both real initial state and parameter are generated randomly in a reasonable range section (in order to preserve the physical meaning of variable). Figure III.32a and Figure III.32b show the comparison between estimated values, real values and initial conditions (of optimization algorithm) for initial state and parameter $\alpha(x)$. These figures correspond to the case of condition 6 in Table III.10.

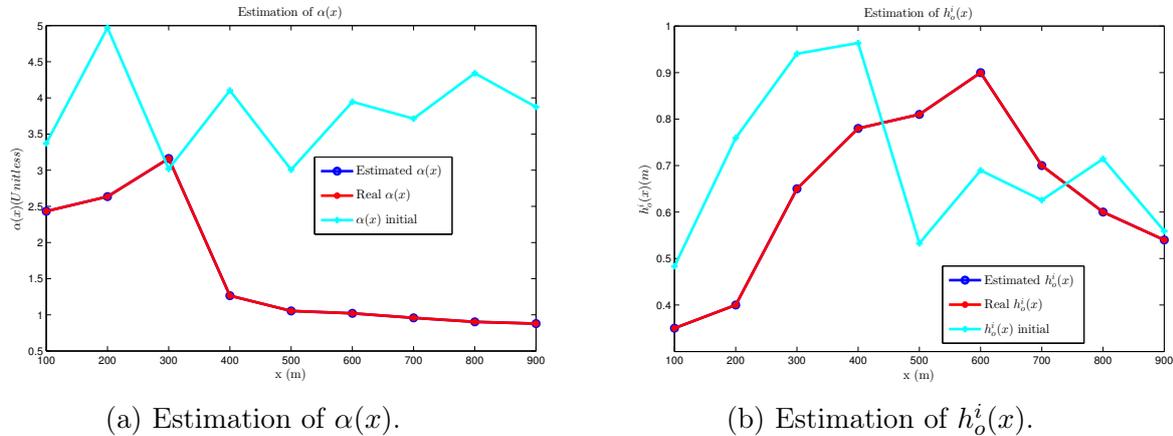


Figure III.32: Estimation result of state and parameter estimation with initial condition 5 in Table III.10.

The relative errors for this case are also small. Figure III.33 and Table III.10 summarize all the initial conditions, estimation error and final value of cost function \mathcal{J} of 6 optimization situations in this case.

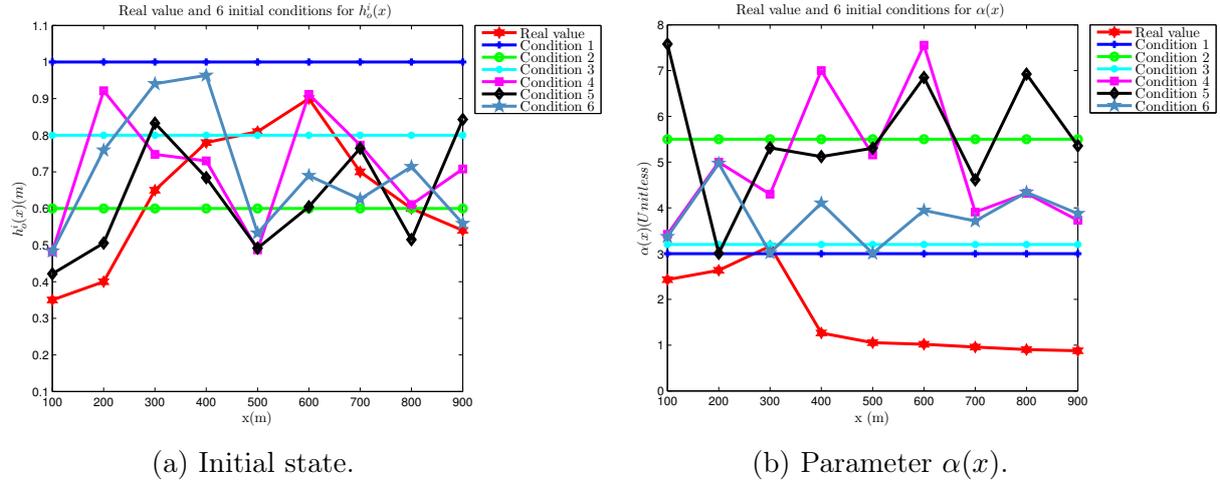


Figure III.33: Real values and 6 initial conditions of initial state and parameter α in the case of random initial state and parameter.

	Estimation relative errors and minimal \mathcal{J} for each initial condition of optimization algorithm					
	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
for $\alpha(x)$	1.04×10^{-5}	1.32×10^{-7}	2.27×10^{-4}	2.44×10^{-6}	9.04×10^{-8}	3.13×10^{-7}
for $h_o^i(x)$	1.82×10^{-5}	1.20×10^{-7}	2.26×10^{-4}	4.39×10^{-6}	1.43×10^{-7}	7.74×10^{-7}
\mathcal{J}	7.88×10^{-13}	1.27×10^{-16}	3.96×10^{-11}	4.31×10^{-14}	9.98×10^{-15}	2.04×10^{-15}

Table III.10: Estimation errors with 6 different initial conditions in the case random initial state and parameter.

The convergence of optimization algorithm and the accuracy of estimation result (with very small average relative error) in various estimation situations previously presented (3 forms of initial state and parameter and 6 initial conditions of optimization algorithm for each form) of overland flow allows to validate the robustness of proposed approach with changes in both real values of initial state and parameter and starting values of algorithm. The same results are expected in a more general case of 1-D hyperbolic system.

4 Optimization algorithm with adjoint-gradient and without adjoint-gradient

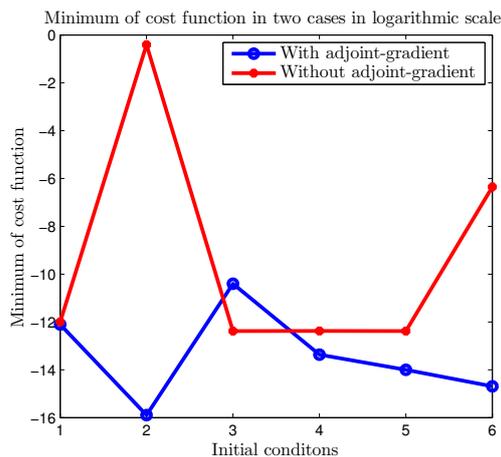
As previously mentioned in subsection 5.c of chapter II, it is worth to test the optimization algorithm without using the calculated adjoint gradients vector $\nabla \mathcal{L}$. This means that

4. Optimization algorithm with adjoint-gradient and without adjoint-gradient 93

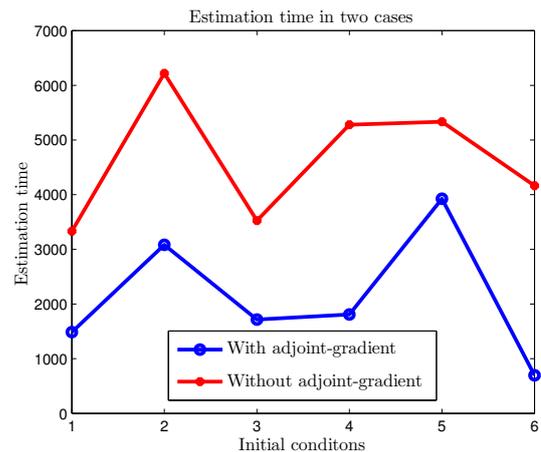
the function *fmincon* of *Matlab* will calculate by itself the gradient (by finite difference) instead of using the supplied adjoint gradient. This test is carried out in the case of random initial state and parameters (with the 6 initial conditions of Table III.10). The estimated results of two cases (with and without adjoint-gradient) will be compared in three aspects: the average value of average estimation relative errors (the mean value of the average relative errors of initial state and the average relative errors parameter α), minimum of cost function and estimation time. The first aspect of both two cases is summarized in the following table.

	Average estimation relative errors					
	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
with adjoint-gradient	1.43×10^{-5}	1.26×10^{-7}	2.27×10^{-4}	3.37×10^{-6}	1.17×10^{-7}	5.44×10^{-7}
without adjoint-gradient	3.48×10^{-5}	43.74	2.08×10^{-5}	2.06×10^{-5}	1.86×10^{-5}	0.01

Table III.11: Comparison of average relative error in 6 different initial conditions between two cases: with adjoint-gradient and without adjoint-gradient.



(a) Minimum of cost function in logarithmic scale.



(b) Estimation time.

Figure III.34: Comparison of minimum of cost function and estimation time in two cases: with adjoint-gradient and without adjoint-gradient.

In the 6 configurations of initial condition of optimization algorithm, one can observe that the adjoint-gradient allows the optimization algorithm to reach more accurate estimation results in most of cases (case 1, 2, 4, 5, 6). More specially, in case 2, the average relative error while using optimization algorithm without adjoint gradient is very large. The *fmincon* tool in this case seems not to converge. It is necessary to take a look at the minimum of cost function and estimation time. Figure III.34a and Figure III.34b respectively show that the *fmincon* tool fed with adjoint based gradient almost always gives

smaller minimal value of cost function \mathcal{J} (excepting initial condition 3) with shorter estimation time. These points allow to conclude about the advantages of adjoint-gradient for *fmincon* tool compared to its self calculated gradient by finite difference.

5 Estimation with noisy measurements

It is necessary to investigate the quality of the result of the proposed estimation approach in the case of noisy measurements, that is the robustness of the adjoint method with respect to the ill-posedness of the inverse problem caused by uncertainty of observation. Let us reconsider the estimation problem which has been presented in example 2.b.3 and the initial configuration 6 (in Figure. III.33) for parameter $\alpha(x)$ and $h_0^i(x)$. To represent the noise in measurements, an additive white Gaussian noise is added to each observed value $h^{meas}(x_j, t)$ with specified signal to noise ratio (SNR). The smaller SNR is, the more noisier measurements are. By denoting the added noise at observation position x_j as $\mathcal{N}(x_j, t)$, the SNR (in decibel, dB) at this position is given by :

$$SNR_{dB}^j = 10 \log_{10} \left[\frac{|h^{meas}(x_j, t)|^2}{|\mathcal{N}(x_j, t)|^2} \right] \quad (\text{III.47})$$

The SNR used in this example varies from 15 dB to 60 dB. There are 2 sensors used to estimate $\alpha(x)$ and $h_0^i(x)$. Figure III.35 represents the comparison between noise-free measurement and noisy measurement at sensor location 10 in two cases, $SNR_{dB} = 15$ dB and $SNR_{dB} = 30$ dB.

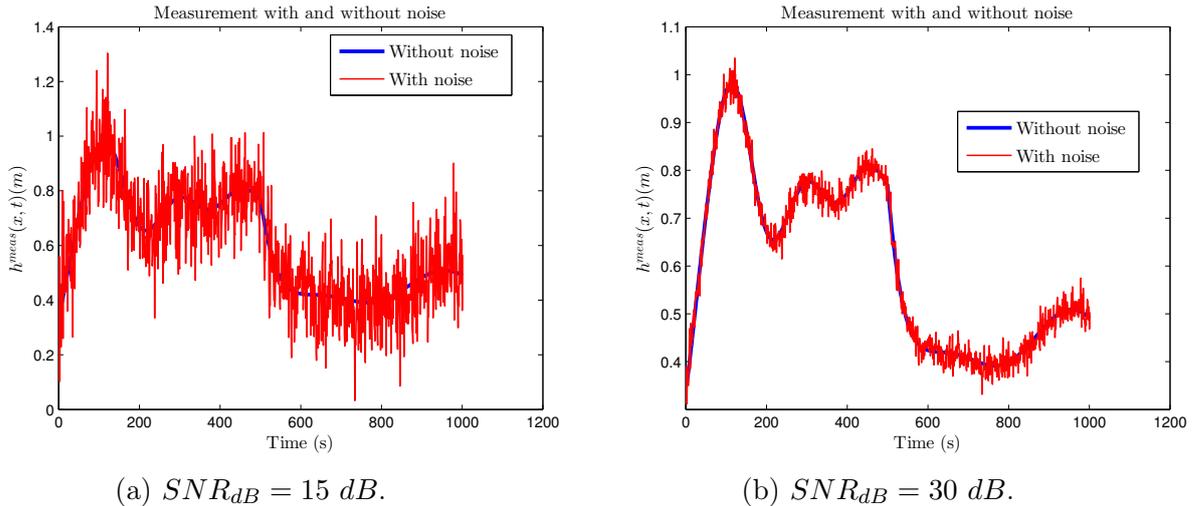


Figure III.35: Comparison between noise-free measurement and noisy measurement at sensor location 10

In the case of noise-free, the average relative error of $\alpha(x)$ and $h_0^i(x)$ is $5.44 \times 10^{-7}\%$ (presented in Table III.11). The estimation process is carried out with various values of the SNR_{dB} and the average relative error of the estimated state and parameters are synthesized in the following table.

SNR_{dB}	Erreur (%)
15	13.97
25	6.15
30	3.09
40	1.23
60	0.11

Table III.12: Average relative error with different values of the SNR_{dB} .

From the above table, one can observe that for the value of SNR_{dB} being larger than or equal to 30 dB, the estimation process gives a reasonable accuracy. The estimation of parameter $\alpha(x)$ in two cases $SNR_{dB} = 15$ dB and $SNR_{dB} = 30$ dB, for instance, is presented in Figure III.36.

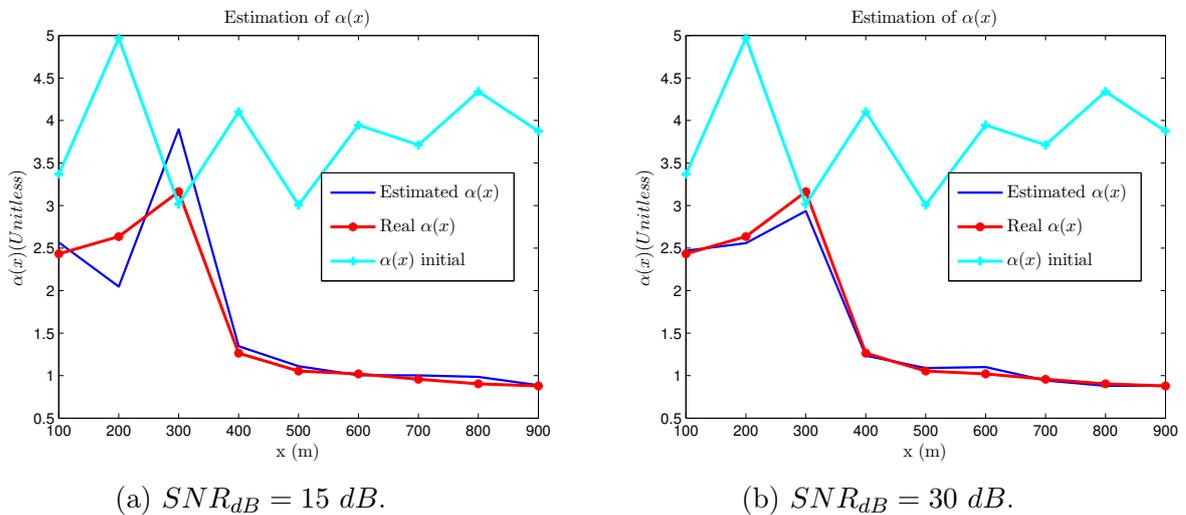


Figure III.36: Estimation of $\alpha(x)$ in two cases $SNR_{dB} = 15$ dB and $SNR_{dB} = 30$ dB.

By comparison with the estimation of $\alpha(x)$ in the case noise-free, illustrated in Figure III.32a, the estimation errors are mainly at the begin of the spatial domain. These results show that the adjoint based estimation method is sensitive to measurement noise. However,

with a less noisy measurement ($SNR_{dB} \geq 30$ dB), an reasonable estimation accuracy can be achieved.

In order to minimize the effect of noise, the idea is to weight the cost function \mathcal{J} with the inverse of the variance of the noise at each observation position x_j , denoted by $\mathcal{R}(x_j)$. The cost function \mathcal{J} in equation II.4 is rewritten as follow:

$$\mathcal{J} = \frac{1}{2} \sum_{j=1}^S \int_0^T \left[\int_0^L \delta_A(x-x_j)u \, dx - u_j^{meas}(x_j, t) \right] \mathcal{R}^{-1}(x_j) \left[\int_0^L \delta_A(x-x_j)u \, dx - u_j^{meas}(x_j, t) \right] dt \quad (\text{III.48})$$

where the initial guessed value terms are omitted. This is equivalent to what is done in the Kalman filter. The validation of this approach is a perspective for future works.

6 Conclusion

In this chapter, the application of adjoint-based estimation methodology has been investigated through various scenarios in two illustrative examples of hyperbolic system: overland flow, described by the continuity equation of Saint-Venant model, and traffic flow, characterized by the LWR model. The scenarios include cases of constant parameter estimation, distributed parameter estimation, initial state estimation, or simultaneous state and parameter estimation. They also include smooth or switched models. In all cases, the proposed methodology gives pretty good results. In addition, the estimation methodology applicability is also validated by successfully solving two parameters estimations of the soil of overland flow on the Tondi-Kiboro catchment using only one measurement of water discharge at the end of flow, and real data.

The tests of the adjoint-based estimation method with various configurations of initial state and parameter and different algorithm initial values show some robustness of the approach. In addition, the adjoint gradient is more effective than the finite difference gradient for *fmincon* tool.

The sensor location or the position of measurements in these examples are sometimes randomly chosen, sometimes manually chosen. For cases of parameter estimation, it seems that the position of sensor does not affect a lot the convergence and accuracy of estimation algorithm. However, after some simulations in the presented examples to determine the sensitivity of estimation accuracy with respect to the number and the location of measurements, a conclusion can be drawn that for estimating the initial state, the sensor must be placed at the end of spatial domain to achieve observability. But only one sensor at the boundary may not be enough the get a good estimation result. Moreover, the more measurements available, the more accurate results can be obtained. But the accuracy improvement is not noticeable while being compared to the cost paid for sensor number

increment. The question is how to optimally locate the sensors to get as much useful information as possible. The next two chapters of the second part are devoted to answer this question.

Part B

Optimal sensor locations for state estimation

IV | Optimal sensor location methodology

The location of sensor used in the previous chapter dedicated to some optimal estimation applications is selected manually. This means that the best locations of sensors, in the sense of giving the smallest estimation errors, are chosen after several trials. The manual method can be performed on the system for a small number of discretized segments or when the number of possible locations is small. But in more complex situations, an investigation of optimal positions of sensors is thus needed to ensure at least observability and also the best possible performance of the estimation process. Moreover, the number of sensor is limited due to the economical and technical reasons. For a given sensor resource with these constraints, the sensor placement will find the best possible locations to recover the expected system information from measurements. A-state-of-the-art about sensor placement techniques will be proposed in section 1. Some background about observability for both linear and nonlinear system and the sensor placement criteria based on observability Gramian will be presented in the section of preliminaries. For sensor placement of infinite dimensional nonlinear systems, such as the overland flow described in section 2 of chapter III, a focus will be made on the approach based on the Fisher information matrix. The main drawback of the Fisher information matrix based approach will be discussed. To overcome this drawback a new sensor optimal location criteria based on the adjoint PDE presented in subsection 3.c of chapter II will be proposed.

1 State-of-the-art

The problem of sensor locations for infinite-dimensional systems has been firstly addressed in 1972 in the work of Bensoussan [Ben72]. Further approaches have been studied in various works in the last four decades. Most of the proposed approaches can be classified into four mains categories: some optimization methods based on the error covariance matrix of

the Kalman filter, some methods based on the notion of Fisher information matrix, and some approaches based on the observability Gramian or using the risk of failure of sensors combined with the measurement cost.

Firstly introduced by Kalman et al. [Kal60] in 1960, the Kalman filter has been used widely in signal processing, navigation and system and control. With its extensions and generalizations such as the extended Kalman filter or the unscented Kalman filter, this algorithm allows to combine the measurements from sensors for example, the control inputs and the system dynamics to predict the system state. The Kalman filter belongs to the class of asymptotical state observers. A. Bensoussan [Ben72] proved the existence of optimal sensor placement problems for a filtering problem of the Riccati equation in infinite dimensional spaces. The location of sensors was considered as a control variable of the optimal control problem of the filter covariance equation. Athans et al. [Ath72] obtained an analogous result to the work of Bensoussan by using Pontryagin's maximum principle. Following the same idea and with more detailed developments, the work of Curtain et al. [CI78] has to be mentioned in the case of linear distributed systems in the framework of the semi-group theory. Kumar et al. [KS78] developed a new approach that minimizes a measure linked to the upper bound of the filter covariance while Omatu et al. [OKS78] minimized the trace of the optimal error covariance matrix of the optimal filter to prove the existence and uniqueness of a optimal sensor location for the operator Riccati equation of partial differential equations. More recently, Khan et al. [KMS15] showed that the sensor must be placed at the position of disturbance by applying the Kalman filter for state estimation in 1-D linear dispersive wave equation. The positions of measurements are found by minimizing the trace of the solution of the Riccati equation associated with the optimal filter.

The Fisher information is a statistical notion firstly introduced by statistician R. Fisher to measure the amount of information contained in the observation of a model, which is supposed to be observable, in terms of state or parameters. In the state and parameter estimation, the measurements taken at some positions of the spatial domain are the main inputs for the estimation algorithm. The Fisher information matrix provides to be an intuitive way to describe the amount of knowledge about the states and parameters to be estimated. As a result, one can use some scalar functions of the Fisher information or Fisher information matrix (in the case of multiple variables) to determine the optimal positions of sensors. Two approaches based on the normalized Fisher information matrix for locating sensors in parameter estimation process are proposed by T.D. Fadale et al. [FNE95]. The D-optimality criterion, i.e the determinant of Fisher information matrix, has been investigated by Uciński et al. [Uci99]; [Uci00]; [UP07]; [Tri+08] to find out the trajectories of mobile sensor used to estimate parameters of distributed processes. An extension of this work to the case of mobile sensor networks is presented in [Uci12]. Autrique et al. [Aut+00]; [ALR02] used the Fisher information matrix to determine sensor location in thermal processes. Nahor et al. [Nah+03] where the modified E-criterion, minimizing

the ratio between the largest and smallest eigenvalues of the Fisher information matrix, is used to optimize the temperature sensor position in a hot wire probe in food process. Alaña [Ala10] also used the determinant of Fisher information matrix to study the role of measurement locations in the parameter estimation of a tabular reactor. By the same approach S. Martinez et al. [MB06] studied the maximization of D-optimality criterion to find the set of positions of mobile sensor network. After that the information from sensors at chosen locations is used as input for target tracking realized by extended Kalman filter. The observability gramian can be employed as an optimality criterion for optimal sensor placement. Originally developed for linear dynamical systems case, this method can be extended to the case of nonlinear systems. In the case of linear systems, Müller et al. [MW72] discussed the optimal sensor location through the minimization of some scalar functional, such as the determinant, the trace or the maximal eigenvalue of the observability matrix. The problem of sensor positioning in a tabular reactor, modeled by an infinite dimensional model is treated in the work of Waldraff et al. [Wal+98] and van den Berg et al. [Ber+00]. The authors considered the minimal singular value of an observability gramian matrix as a criterion for optimal sensor location. Georges firstly developed a placement criterion based on the minimum eigenvalue of the transient observability gramian in [Geo95] for finite-dimensional systems and then used this idea to find the optimal sensor network configuration for air pollution monitoring. Due to the complexity of observability gramian in the case of nonlinear and complex systems, such as in chemical reactions in fixed-bed bio-reactors, Singh et al. [SH05] replaced it by the empirical one to determine the optimal sensor location.

2 Preliminaries

2.a Observability

2.a.1 Linear time-invariant systems

It is necessary to briefly provide some background on observability because this property is useful for developing some sensor positioning strategies. Basically, observability is required to be able to reconstruct the values of the state variables from a given output or measurement \mathbf{y} obtained on a given period of time. If we consider some linear time-invariant system of the form

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} = C\mathbf{x} + D\mathbf{u} \end{cases} \quad (\text{IV.1})$$

For asymptotically stable linear systems, observability can be related to the notion of output energy function generated by any initial state x :

$$\mathbf{E}_o = \int_0^{\infty} \|y(t)\|^2 dt = \int_0^{\infty} x^T e^{A^T t} C^T C e^{At} x dt = x^T \mathbf{W}_{linear} x \quad (\text{IV.2})$$

where

$$\mathbf{W}_{linear} = \int_0^{\infty} e^{A^T t} C^T C e^{At} dt \quad (\text{IV.3})$$

where state vector $\mathbf{x} \in \mathbb{R}^n$; output $\mathbf{y} \in \mathbb{R}^m$ and input vector $u \in \mathbb{R}^v$. The matrix \mathbf{W}_{linear} is called the observability gramian for linear system. If it has full rank, the system is then observable. And in the opposite case, some of the states can not be recovered from information of output \mathbf{y} and the system is simply detectable.

For an observation horizon $[0, \mathcal{T}]$, the observability gramian is a function of \mathcal{T} defined by

$$\mathbf{W}_{linear}(\mathcal{T}) = \int_0^{\mathcal{T}} e^{A^T t} C^T C e^{At} dt$$

For linear systems, it can be interpreted as the solution of the differential Lyapunov equation:

$$\begin{cases} \dot{\mathbf{W}}_{linear}(\mathcal{T}) = A^T \mathbf{W}(\mathcal{T}) + \mathbf{W}(\mathcal{T}) A + C^T C \\ \mathbf{W}_{linear}(0) = 0 \end{cases} \quad (\text{IV.4})$$

If the matrix A is stable and the observation time reaches infinity $\mathcal{T} \rightarrow \infty$, \mathbf{W}_{linear} can be obtained as the non negative definite symmetric matrix solution of the Lyapunov equation:

$$A^T \mathbf{W}_{linear} + \mathbf{W}_{linear} A + C^T C = 0 \quad (\text{IV.5})$$

Observability can also be assessed by using the observability matrix \mathbf{M}_{obs} (which does not require the stability assumption).

$$\mathbf{M}_{obs} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

2.a.2 Link between the Fisher information matrix and the observability gramian

The Fisher information matrix in this linear system provides an alternative way to derive $\mathbf{W}_{linear}(\mathcal{T})$. By denoting the initial state $\mathbf{x}(0)$ of the linear system in equation (IV.1) by \mathbf{x}_0 ,

the sensitivity of the output \mathbf{y} with respect to \mathbf{x}_0 can be given by the following formulation of Fisher information matrix:

$$\mathcal{F}(\mathcal{T}) = \int_0^{\mathcal{T}} \frac{\partial \mathbf{y}(t)}{\partial \mathbf{x}_0} \frac{\partial \mathbf{y}(t)^T}{\partial \mathbf{x}_0} dt \quad (\text{IV.6})$$

Taking the partial derivative of equation (IV.1) with respect to initial condition \mathbf{x}_0 and changing the derivative order, one can obtain

$$\begin{cases} \frac{d}{dt} \left(\frac{\partial \mathbf{x}^T(t)}{\partial \mathbf{x}_0} \right) = A \frac{\partial \mathbf{x}^T(t)}{\partial \mathbf{x}_0} \\ \frac{\partial \mathbf{y}^T(t)}{\partial \mathbf{x}_0} = C \frac{\partial \mathbf{x}^T(t)}{\partial \mathbf{x}_0} \end{cases} \quad (\text{IV.7})$$

with initial condition $\frac{\partial \mathbf{x}^T(0)}{\partial \mathbf{x}_0} = I_d$ where I_d is the identity matrix. The partial derivative of output with respect to initial condition is then

$$\frac{\partial \mathbf{y}^T(t)}{\partial \mathbf{x}_0} = C e^{At} \quad (\text{IV.8})$$

As a result, the Fisher information matrix in (IV.6) becomes

$$\mathcal{F}(\mathcal{T}) = \int_0^{\mathcal{T}} e^{A^T t} C^T C e^{At} dt = \mathbf{W}(\mathcal{T}) \quad (\text{IV.9})$$

We can see that the Fisher Information Matrix coincides with the observability gramian for linear systems. This gives an interpretation of the observability gramian in terms of the measure of the output sensitivity with respect to the initial conditions.

For infinite dimensional linear systems, some extension of the notion of observability gramian has been proposed in the semi-group theory framework by Curtain et al. [CI78]. The observability gramian becomes an infinite-dimensional operator, solution of an operator Lyapunov equation which generalizes Lyapunov equation. In practice, the numerical solution of this equation requires complicated finite-dimensional reduction techniques, which will not be explored in this thesis.

2.a.3 Nonlinear systems

However observability matrix \mathbf{M}_{obs} and observability gramian \mathbf{W}_{linear} in equation (IV.3) cannot assess the observability of a finite-dimensional nonlinear system. Some alternative conditions have been derived in various researches in last decades (see [Bes07] for an

overview on observer tools for nonlinear systems). In practice, these conditions appear to be complicated to develop some systematic approaches. A local approach consists in linearizing the nonlinear system around an operating point and in considering the observability at this point by using the rank test applied to the observability matrix of the linearized system. The observability gramian of the linearized system can also be used to assess local observability. However, observability of the linearized system does not allow to assess the observability of the nonlinear system. In order to avoid the computation of the observability gramian through the numerical solution of large Lyapunov equations, a new concept developed by S. Lall et al. [LMG02], empirical observability Gramian has been developed. Let us consider a nonlinear system of the following form:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) \end{cases} \quad (\text{IV.10})$$

where the notations $\mathbf{f}(\cdot)$ and $\mathbf{g}(\cdot)$ are state and output functions; x, y, u play the same role as the linear system. Some intermediate matrices need to be defined to get the definition of the so-called observability covariance matrix, which is strongly related to the observability gramian, as mentioned later:

$$\begin{aligned} \mathbf{T}^n &= \{\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_r\} \quad \text{with } \mathbf{T}_i \in \mathbb{R}^{n \times n}; \quad \mathbf{T}_i^T \mathbf{T}_i = \mathbf{I} \quad \text{and } i = 1, 2, \dots, r \\ \mathbf{M} &= \{c_1, c_2, \dots, c_s\} \quad \text{with } c_i \in \mathbb{R}; \quad c_i > 0 \quad \text{and } i = 1, 2, \dots, s \\ \mathbf{E}^n &= \{e_1, e_2, \dots, e_n\} \quad \text{with } \mathbf{E}^n \text{ is the standard vector in } \mathbb{R}^n \end{aligned} \quad (\text{IV.11})$$

where n is the system state size; r is the perturbation direction matrices number and for each perturbation direction s is the number of different sizes. By denoting the initial condition by

$$x(0) = c_m \mathbf{T}_j e_k + x_{ss} \quad (\text{IV.12})$$

with x_{ss} being the steady state of system, the system output corresponding to initial state $x(0)$ is denoted by $y^{ilm}(t)$. A new intermediate matrix $\Psi^{lm}(t) \in \mathbb{R}^{n \times n}$ is defined, as the norm of the difference between the current output at position i for perturbation l and perturbation number m and output at steady state of system $y_{ss}(t)$ by equation below at each element position $\{i, j\}$.

$$\Psi_{ij}^{lm}(t) = (y^{ilm}(t) - y_{ss})^T (y^{ilm}(t) - y_{ss}) \quad (\text{IV.13})$$

The expression of the the observability covariance matrix is thus given by

$$\mathbf{W}_{nonlinear} = \sum_{i=1}^r \sum_{j=1}^s \frac{1}{r^2 s^2 c_j^2} \int_0^{\infty} \mathbf{T}_i \Psi^{ij}(t) \mathbf{T}_i^T dt \quad (\text{IV.14})$$

If the initial state of the nonlinear system $x(0)$ given in equation (IV.12) is in the domain of attraction of the steady state x_{ss} or the system (IV.10) is asymptotically stable, the

observability covariance matrix in equation (IV.14) reduces to the case of an empirical observability Gramian. Moreover, by applying the technique of observability covariance matrix on the case of linear and stable system described in equation (IV.1) as shown in [LMG02]. For the perturbed initial condition (IV.12), the output of linear system (IV.1) at steady state is

$$y^{ilm}(t) = Ce^{At}c_m\mathbf{T}_l e_i + Cx_{ss} \quad (IV.15)$$

with the stable state of output $y_{ss}^{ilm} = Cx_{ss}$

As a result, the element at position $\{i, j\}$ of matrix Ψ^{lm} is

$$\Psi_{ij}^{lm}(t) = (Ce^{At}c_m\mathbf{T}_l e_i)^T (Ce^{At}c_m\mathbf{T}_l e_j) = c_m^2 e_i^T \mathbf{T}_l^T e^{A^T t} C^T C e^{At} \mathbf{T}_l e_j \quad (IV.16)$$

The matrix Ψ^{lm} is finally written as

$$\Psi^{lm}(t) = c_m^2 \mathbf{T}_l^T e^{A^T t} C^T C e^{At} \mathbf{T}_l \quad (IV.17)$$

And the corresponding observability covariance matrix holds

$$\mathbf{W}_{nonlinear} = \sum_{i=1}^r \sum_{j=1}^s \frac{1}{r^2 s^2 c_j^2} \int_0^{\infty} c_m^2 \mathbf{T}_l^T e^{A^T t} C^T C e^{At} \mathbf{T}_l dt = \int_0^{\infty} e^{A^T t} C^T C e^{At} dt \quad (IV.18)$$

and is identical to the observability matrix \mathbf{W}_{linear} . The choice of matrices \mathbf{T}^n and \mathbf{M} play an important role for calculating the observability covariance matrix. The choice of matrix \mathbf{T} must satisfy the condition in its definition. Its value is usually chosen as the negative identity matrix $-\mathbf{I}$ if the negative perturbations are used, and vice-versa as the positive identity matrix \mathbf{I} for positive perturbations. For the value of matrix \mathbf{M} , the authors in [SH05] state that the largest perturbation size must be selected to allow the system to rest in the attraction domain of the considered operating point. Moreover, the work of S. Lall et al [LMG02] shows that in the case of linear system, any value of \mathbf{M} give the same value of observability covariance matrix, as shown in equation (IV.18). The local observability of nonlinear system is guaranteed if and only if the rank of the observability covariance matrix or empirical observability gramian is full. The application of this approach to infinite-dimensional systems, such as the balance laws studied in this thesis, is an interesting perspective in future works.

2.b Criteria for optimal sensor locations

As previously presented, the rank of the observability matrix just answers the question of whether or not the considered system is completely observable. It does give enough information to investigate the optimal sensor positions or which locations gives the most

useful information for the state estimation (that is the maximal sensitivity with respect to unknown parameters to be estimated). In other words, it is necessary to define some criteria to measure the degree of observability. By duality with the controllability problem, Muller et al. [MW72] proposed three candidates for physically meaningful measures: the maximum eigenvalue, the trace and the determinant of the inverse controllability matrix. Because of the duality between observability and controllability, one can apply the same approaches to find the potential locations of sensors. The details of these criteria applied on the linear observability Gramian in equation (IV.3) will be presented below.

2.b.1 Smallest eigenvalue of the observability Gramian

By denoting the smallest eigenvalue of the inverse linear observability Gramian by notation κ_1 . The formulation of this criterion is given by:

$$\kappa_1 = \lambda_{min}(\mathbf{W}_{linear}) \quad (IV.19)$$

The notation λ_{min} denotes the minimal eigenvalue of \mathbf{W}_{linear} . If the matrix is full rank or its determinant is non null, the maximum eigenvalue of the inverse observability Gramian can be used to replace this criterion. The smallest eigenvalue λ_{min} represents the least observable direction of system. In the same way, higher degree of observability is represented by larger value of the smallest eigenvalue. Each sensor location configuration gives different values of κ_1 . By maximizing κ_1 on the sensor position space, one can determine the optimal locations giving the largest κ_1 . This criterion is also called as the E-optimality criterion.

2.b.2 Trace of the inverse observability Gramian

With n being the states number of system, the trace criterion or sensitivity criterion, κ_2 is

$$\kappa_2 = \frac{n}{trace(\mathbf{W}_{linear}^{-1})} \quad (IV.20)$$

Another possibility to define this criterion with the same meaning is to use the trace of \mathbf{W}_{linear}^{-1} , the so-called A-optimality criterion. A large value of κ_2 , corresponding to a small trace of the observability matrix, gives the higher observability degree. As a result, the optimal sensor location can be transformed into a maximization problem, $\max(\kappa_2)$. But this criterion may become problematic when system is not observable, since the inverse observability matrix \mathbf{W}_{linear}^{-1} is then not defined. D. Ucinski [Uci00] proposed a similar criterion but with the original observability matrix to overcome this issue

$$\kappa_2 = trace(\mathbf{W}_{linear}) \quad (IV.21)$$

This is denoted as the sensitivity criterion.

2.b.3 Determinant of observability Gramian

This D-optimality criterion is defined apart from the determinant of the linear observability Gramian as:

$$\kappa_3 = [\det(\mathbf{W}_{linear})]^{1/n} \quad (\text{IV.22})$$

If κ_3 is equal to zero, the system is not completely observable. And similarly, the improvement of the observability degree corresponds to a larger value of the determinant.

Finally F. van Den Berg et al. [Ber+00] suggested two other similar criteria, the maximal singular value, in equation (IV.23) and the trace, in equation (IV.24) of the observability matrix.

$$\kappa_4 = \lambda_{max}(\mathbf{W}_{linear}) \quad (\text{IV.23})$$

$$\kappa_5 = \sum_{i=1}^n \lambda_i(\mathbf{W}_{linear}) \quad (\text{IV.24})$$

where λ_i is the i^{th} eigenvalue of \mathbf{W}_{linear} . One can remark that by definition, the observability Gramian is a symmetric positive (semi)-definite matrix, so the two terminologies singular values and eigenvalues are equivalent.

3 Two approaches for the optimal location of sensors

Now the Fisher information matrix and the so-called new adjoint-based approaches will be presented in more details in the next section of the thesis. After that, some comparisons between their characteristics will be also performed. Both of these approaches will be developed on a general non switched hyperbolic system similar to the one presented in chapter II. To avoid a far backward reference, the system equation and the optimal estimation method will be briefly recalled. The dynamic of system is

$$\begin{cases} \frac{\partial u(x, t)}{\partial t} + \frac{\partial f(u(x, t), x)}{\partial x} & = g(x, t) \\ u(x, 0) = u_0^i(x) \quad \text{and} \quad u(0, t) & = u_0^b(t) \end{cases} \quad (\text{IV.25})$$

with the flow function $f(u(x, t), x) = \sum_{i=1}^N \alpha_i(x) \varphi_i(u)$ and the unknown initial state $u_0^i(x)$.

The least-square cost function \mathcal{J} is then

$$\mathcal{J} = \frac{1}{2} \sum_{j=1}^N \int_0^T \left\{ \int_0^L \delta_A(x - x_j) u \, dx - u_j^{meas}(x_j, t) \right\}^2 dt \quad (\text{IV.26})$$

with the assumption that there is N available sensors and without the calibration terms related to the first guessed values.

3.a A Fisher information matrix approach

In equation (IV.26) of cost function \mathcal{J} , the measurement $u_j^{meas}(x_j, t)$ depends on the position x_j of the j^{th} sensor. This means that the result, of course, depends on these positions. The Fisher information matrix provides a way to formulate the sensitivity of $u_j^{meas}(x_j, t)$ with respect to the estimated state and/or parameters. The formulation of this approach is developed for both the initial state $u_0^i(x)$ and some parameters $\alpha_i(x)$. Because of the independence between these variables, the obtained formulation can be used separately without any constraints. In order to formulate the Fisher information matrix, let us denote the unknown values, which are the discretized variables obtained at M discretization points on spatial domain $[0, L]$: $\theta = \underbrace{[\alpha_i(x_k) \ u_0^i(x_k)]^T}_{\text{size } (N+1) \times (M-2)}$ with $k \in [2 \dots M-1]$. By supposing that

the number of sensors is S , denoting the set of discrete positions of sensors by \mathcal{X} , and x^j denoting variable in this set corresponding to the j^{th} sensor, the Fisher information matrix is

$$\mathcal{F}(x^j) = \frac{1}{ST} \sum_{j=1}^S \int_0^T \nu(x^j, t)^T \nu(x^j, t) dt \quad (\text{IV.27})$$

where the term

$$\nu(x^j, t) = \left[\frac{\partial u(x^j, t)}{\partial \theta} \right] \quad (\text{IV.28})$$

is the partial derivative of system variable $u(x, t)$ at sensor position x^j with respect to variable θ . It is also called the *sensitivity vector*. The reason behind this choice is that the inverse of the Fisher information matrix in (IV.27) is a good approximation of the covariance error matrix, $cov(\hat{\theta})$ of the estimated value of $\hat{\theta}$ with a large time horizon [WP97]. It is calculated by taking the partial derivative of the two sides of equation (IV.25) with respect to θ , with the notation $F(u, x, \theta)$ representing the derivative $\frac{\partial f(u(x, t), x)}{\partial x}$. The sensitivity is obtained as the solution of

$$\partial_t \left[\frac{\partial u}{\partial \theta} \right] = -\frac{\partial F}{\partial \theta} - \frac{\partial F}{\partial u} \frac{\partial u}{\partial \theta} + \frac{\partial g}{\partial \theta} + \frac{\partial g}{\partial u} \frac{\partial u}{\partial \theta} \quad (\text{IV.29})$$

The last two components of equation (IV.29) are equal to zero because they are not function of θ . Equation (IV.30) finally becomes

$$\partial_t \left[\frac{\partial u}{\partial \theta} \right] = -\frac{\partial F}{\partial \theta} - \frac{\partial F}{\partial u} \frac{\partial u}{\partial \theta} \quad (\text{IV.30})$$

By changing the partial derivative orders (do the derivative of all terms with respect to θ before to x and to t), the derivatives $\frac{\partial F}{\partial \theta}$ and $\frac{\partial F}{\partial u}$, in equation (IV.30) are as bellows:

$$\begin{aligned} \frac{\partial F}{\partial \theta} &= \frac{\partial}{\partial x} \left\{ \sum_{i=1}^N \left[\frac{\partial \alpha_i(x)}{\partial \theta} \varphi_i(u) + \alpha_i(x) \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial u}{\partial \theta} \right] \right\} \\ &= \sum_{i=1}^N \left[\frac{\partial \alpha_i(x)}{\partial \theta} \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial \alpha_i(x)}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \times \frac{\partial u}{\partial \theta} + \alpha_i(x) \left[\frac{\partial}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial u}{\partial \theta} + \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial}{\partial \theta} \frac{\partial u}{\partial x} \right] \right] \\ \frac{\partial F}{\partial u} &= \frac{\partial}{\partial x} \sum_{i=1}^N \alpha_i(x) \frac{\partial \varphi_i(u)}{\partial u} = \sum_{i=1}^N \left[\frac{\partial \alpha_i(x)}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} + \alpha_i(x) \frac{\partial}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial u}{\partial \theta} \right] \end{aligned} \quad (\text{IV.31})$$

where

$$\frac{\partial \alpha_i(x)}{\partial \theta} = [0 \dots \underbrace{1}_{\text{position } i} \dots 0]$$

By inserting these terms into equation (IV.30), one can obtain the equations describing the dynamics of sensitivity $\frac{\partial u}{\partial \theta}$, as defined below:

$$\begin{aligned} \frac{\partial}{\partial t} \left[\frac{\partial u}{\partial \theta} \right] + \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial \theta} \right] \sum_{i=1}^N \alpha_i(x) \frac{\partial \varphi_i(u)}{\partial u} + \frac{\partial u}{\partial \theta} \sum_{i=1}^N \left[2 \frac{\partial \alpha_i(x)}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} + \alpha_i(x) \frac{\partial}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \right] \\ + \left[\frac{\partial u}{\partial \theta} \right]^2 \sum_{i=1}^N \alpha_i(x) \frac{\partial}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} + \sum_{i=1}^N \frac{\partial \alpha_i(x)}{\partial \theta} \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial u}{\partial x} = 0 \end{aligned} \quad (\text{IV.32})$$

More specifically, equation (IV.32) can be rewritten for each parameter $\alpha_i(x_k)$ as:

$$\begin{aligned} \frac{\partial}{\partial t} \left[\frac{\partial u}{\partial \alpha_i(x_k)} \right] + \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial \alpha_i(x_k)} \right] \sum_{i=1}^N \alpha_i(x) \frac{\partial \varphi_i(u)}{\partial u} + \frac{\partial u}{\partial \alpha_i(x_k)} \sum_{i=1}^N \left[2 \frac{\partial \alpha_i(x)}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \right. \\ \left. + \alpha_i(x) \frac{\partial}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \right] + \left[\frac{\partial u}{\partial \alpha_i(x_k)} \right]^2 \sum_{i=1}^N \alpha_i(x) \frac{\partial}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} + \sum_{i=1}^N \frac{\partial \alpha_i(x)}{\partial \alpha_i(x_k)} \frac{\partial \varphi_i(u)}{\partial u} \frac{\partial u}{\partial x} = 0 \end{aligned} \quad (\text{IV.33})$$

The initial condition for this PDE is $\frac{\partial u}{\partial \alpha_i(x_k)}(x, 0) = 0 \forall x \in [0 \dots L]$ and the boundary condition $\frac{\partial u}{\partial \alpha_i(x_k)}(0, t) = 0 \forall t \in [0 \dots T]$. Similarly, sensitivity equation (IV.32) with

respect to the initial condition $u_0^i(x_k)$ is given by

$$\begin{aligned} \frac{\partial}{\partial t} \left[\frac{\partial u}{\partial u_0^i(x_k)} \right] + \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial u_0^i(x_k)} \right] \sum_{i=1}^N \alpha_i(x) \frac{\partial \varphi_i(u)}{\partial u} + \frac{\partial u}{\partial u_0^i(x_k)} \sum_{i=1}^N \left[2 \frac{\partial \alpha_i(x)}{\partial x} \frac{\partial \varphi_i(u)}{\partial u} \right. \\ \left. + \alpha_i(x) \frac{\partial^2 \varphi_i(u)}{\partial x^2} \right] + \left[\frac{\partial u}{\partial u_0^i(x_k)} \right]^2 \sum_{i=1}^N \alpha_i(x) \frac{\partial^2 \varphi_i(u)}{\partial x^2} = 0 \end{aligned} \quad (\text{IV.34})$$

with following initial and boundary conditions

$$\begin{aligned} \frac{\partial u}{\partial u_0^i(x_k)}(x, 0) &= [0 \dots \underbrace{1}_{\text{position } k} \dots 0] \\ \frac{\partial u}{\partial u_0^i(x_k)}(0, t) &= 0 \quad \forall t \in]0, T] \\ \frac{\partial u}{\partial u_0^i(x_k)}(0, 0) &= 0 \end{aligned} \quad (\text{IV.35})$$

Equations (IV.33) and (IV.34) are two nonlinear PDEs. In order to determine $\frac{\partial u}{\partial u_0^i(x_k)}$ and $\frac{\partial u}{\partial \alpha_i(x_k)}$, these two equations must be discretized and solved via the use of numerical schemes which have been presented in section 4 of chapter II. Similarly to the case of the direct system discretization, their spatial domains $[0 \dots L]$ are divided into P sections with step Δx and time domain $[0 \dots T]$ into Q sections with step Δt . For the first equation (IV.33), the solution (at each observation position x^j) is a vector of $N \times P$ elements (because there are totally N parameter α) and each element is a vector size $1 \times Q$.

$$\frac{\partial u(x^j, t)}{\partial \alpha(x)} = \left[\begin{array}{c} \frac{\partial u(x^j, t)}{\partial \alpha_1(x_1)} \\ \dots \\ \frac{\partial u(x^j, t)}{\partial \alpha_i(x_k)} \\ \dots \\ \frac{\partial u(x^j, t)}{\partial \alpha_N(x_P)} \end{array} \right] \left. \vphantom{\frac{\partial u(x^j, t)}{\partial \alpha(x)}} \right\} N \times P \quad (\text{IV.36})$$

For the second equation, the solution $\frac{\partial u}{\partial u_0^i}$ is a vector of only P elements and each element is a matrix of size $1 \times Q$.

$$\frac{\partial u(x^j, t)}{\partial u_0^i(x)} = \left. \begin{array}{c} \frac{\partial u(x^j, t)}{\partial u_0^i(x_1)} \\ \dots \\ \frac{\partial u(x^j, t)}{\partial u_0^i(x_k)} \\ \dots \\ \frac{\partial u(x^j, t)}{\partial u_0^i(x_P)} \end{array} \right\} P \quad (\text{IV.37})$$

The size of sensitivity vector $\frac{\partial u(x^j, t)}{\partial \theta}$ is then $(N + 1)P$. The Fisher matrix defined by expression (IV.27) is consequently a $(N + 1)P \times (N + 1)P$ symmetric matrix. Each sensor position x_j generates a Fisher matrix. All the matrices given by all the sensors have to be added and normalized by dividing by the number of sensor S and observation horizon T . The resulting Fisher information matrix reflects the sensitivity of the sensor measurements with respect to the parameters to be estimated through sensitivity vectors $\nu(x^j, t)$. But the direct use of this matrix does not seem to be a possible way to directly derive sensor locations and a metric (a measure of sensitivity) has to be introduced. Some scalar functions denoted κ_F (with the same meaning with κ in subsection 2.b) can play this role. Several choices for such a function have been also presented in 2.b. The optimal positions of sensors based on the Fisher information matrix are the solutions of the maximization problem with function κ_F , as follows:

$$\max_{x^j \in \mathcal{X}} \{ \kappa_F(\mathcal{F}(x^j)) \} \quad (\text{IV.38})$$

The main drawback of this Fisher matrix approach applied to PDEs is the need to compute the solutions of a very large number of sensitivity PDEs, whose number grows exponentially with the number of parameters to be estimated.

3.b An adjoint based approach

The previously-defined Fisher information matrix approach along with the other recalled approaches, such as the approach based on the error covariance matrix of the Kalman or extended Kalman filter, approaches using the observability Gramian, are very demanding in terms of computational complexity when large-scale systems are considered. This is especially the case if we consider complex systems governed by PDEs. Moreover, the sensor placement process uses different criterions and is completely separated from the optimal estimation process. This needs additional calculation efforts. The idea is now to

connect the two processes. A new approach based the adjoint PDE is proposed for the initial state estimation. However this approach can be extended without restriction to parameter estimation problems. From the previous analysis and derivation in section 2 of the chapter devoted to the estimation method methodology, it appears that the gradients in (II.36) represent the sensitivity of the objective functional with respect to the initial state of system. This sensitivity function should not be equal to zero to ensure observability of the initial state. The basic idea is thus to use the absolute value of the gradients as a measure of the degree of observability. Let us recall the form of gradients of the Lagrangian objective functional with respect to state $u_0^i(x_j)$ of system (IV.25) and cost function (IV.26).

$$\nabla \mathcal{L}_{u_0^i(x_j)} = -\lambda(x_j, 0) \quad (\text{IV.39})$$

Since this gradient is the ingredient of the steepest descend method used to solve the optimal estimation problem, there is now a direct connection between the degree of observability and the performance of the iterative method based on this gradient. The criterion used for solving the optimal sensor placement problem is now defined as follows:

$$\kappa_g(x_j) = \min\{\text{abs}(-\lambda(x_j, 0))\} \quad (\text{IV.40})$$

and $\{\text{abs}(-\lambda(x_j, 0))\}$ is a vector of $M - 2$ element corresponding to $M - 2$ discrete values of initial state to estimate. The optimal sensor positions will be consequently the solutions of the maximization of κ_g :

$$\max_{x^j \in \mathcal{X}} \{\kappa_g(x_j)\} \quad (\text{IV.41})$$

Since this approach directly connects the adjoint-based gradients used for the optimal estimation to the criterion used for sensor placement, the computation is less demanding is reduced because there is no need to make additional computation. Moreover, it helps to measure the impact of the sensor positions on the performance of the estimation algorithm. To the best of our knowledge, this approach has not been studied before.

4 Conclusion

Along with the state of the art on optimal sensor placement methods and some background on observability, two sensor placement approaches for optimal sensor locations have been developed in the present chapter. The Fisher information matrix method relies on the maximization of some scalar criteria, such as the determinant, the minimal eigenvalue and the trace of a Fisher information matrix of sensitivity functions measuring the degree of observability. The main drawbacks of this method is the computational complexity, which grows exponentially with the system size, especially in the case of infinite-dimensional systems. This drawback has motivated the proposal of a new approach based on the adjoint

state of the optimal estimation problem as a measure of both the degree of observability and the performance of the iterative method used to solve the optimal estimation problem at the same time. The adjoint-based method has been applied only to the case of state estimation. An extension to both state and parameters estimation should be investigated. The application of both the Fisher information matrix approach (for single overland flow and traffic flow) and the adjoint-based approach (for overland flow network and traffic flow) will be studied in the next chapter. The adjoint-based sensor placement methodology developed in this chapter is presented as a part of publication [\[NGB16a\]](#).

V | Optimal sensor location application

The objectives of this chapter are to illustrate the effectiveness of the two approaches developed in chapter IV for solving the optimal sensor placement problem. Three case studies: a traffic flow, an overland flow, and an overland flow network, will be successively investigated. The optimal placement of one sensor on the traffic flow will be performed by using the new adjoint-based criterion proposed in this thesis. The impact of the choice of the sensor location on the estimation of the initial state will be then investigated. In a second part of this chapter, the optimal placement of one sensor on a single overland flow will be assessed by using the Fisher information matrix criterion and the adjoint-based criterion. Again, the impact of the choice of the sensor location on the state estimation process will be studied. Finally a more complex case study consisting in a network of three water flows will be also investigated. The optimal placement of multiple sensors will be considered in this section together with a analysis of the impact on the state estimation process.

1 Sensor placement in a traffic flow

1.a Using the adjoint-based criterion

In this section, the effectiveness of the adjoint-based location technique of one single sensor is investigated on a traffic flow described by the LWR model already presented. The parameters of the studied highway are the same as the ones defined in Table III.2. Figure V.1 shows that the sensor must be placed at the discrete position 10. This result is consistent with the case of an overland flow.

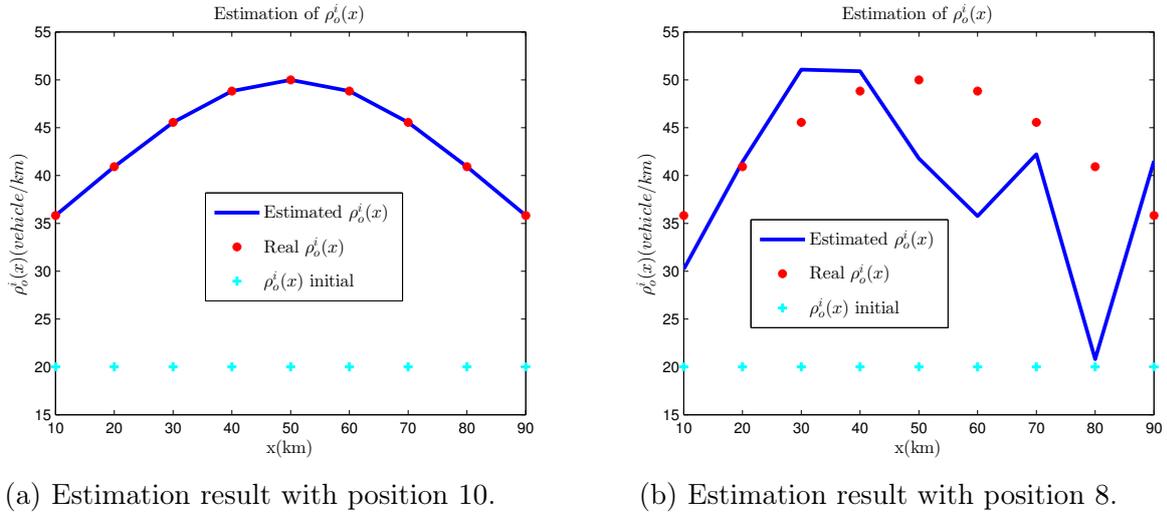


Figure V.2: State estimation in traffic flow with sensor position 10 and position 8.

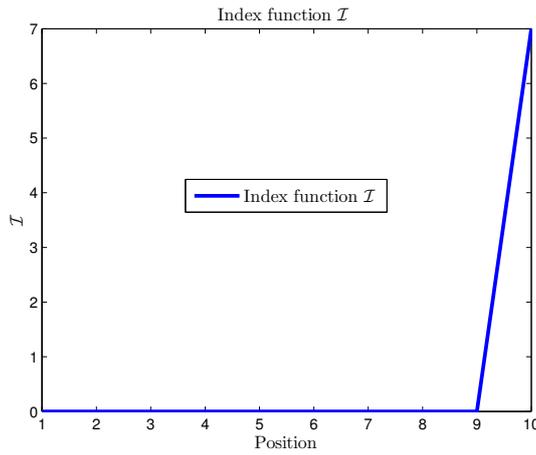


Figure V.1: One-sensor placement using the adjoint-based criterion.

1.b State estimation using optimal and non optimal locations of a single sensor

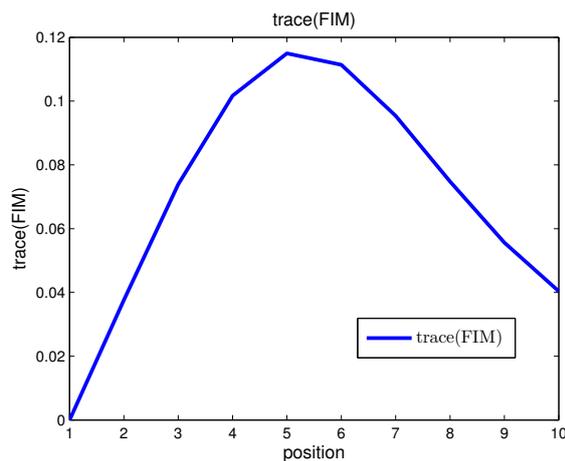
On the basis of the previous sensor placement result, the estimation algorithm is carried out with two scenarios: using the optimal position 10 and using the non optimal one, i.e. position 8. Figure V.2a and Figure V.2b show that the optimal location of the sensor at position 10 allows a very accurate estimation of the vehicle initial state (the average relative errors are $8 \times 10^{-8}\%$, and 16.5% respectively for position 10 and 8).

2 One-sensor placement in overland flows

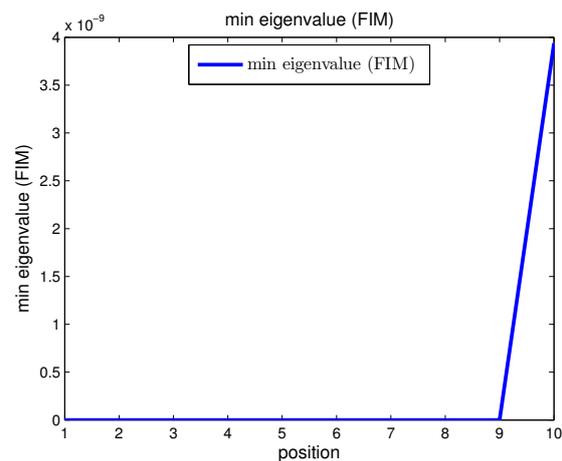
2.a One-sensor placement in a single overland flow

2.a.1 Using the Fisher Information Matrix criterion

The criterion for sensor location which has been formulated in section 3.a of chapter IV will be applied to determine the appropriate position of one sensor in order to estimate the initial state of an overland flow. All the criteria including the D-optimality criterion (determinant of Fisher information matrix), E-optimality criterion (minimal eigenvalue of Fisher information matrix) or the sensitivity criterion (trace of Fisher information matrix) are used and their characteristics will be analyzed. The parameters of the overland flow are the same as the ones presented in Table III.5. This case takes in account only 1 sensor which can be placed on a discrete positions set (integer number from 1 to 10 with 11 discretized points) The computation of the three placement criteria are presented in Figures V.4, V.3a and V.3b where the y-axis gives the value of the criterion, while the x-axis represents the discrete positions of the sensor.



(a) E-optimality criterion.



(b) Sensitivity criterion.

Figure V.3: Sensor placement used Fisher information matrix criterion in an overland flow: trace-based criterion and E-optimality criterion.

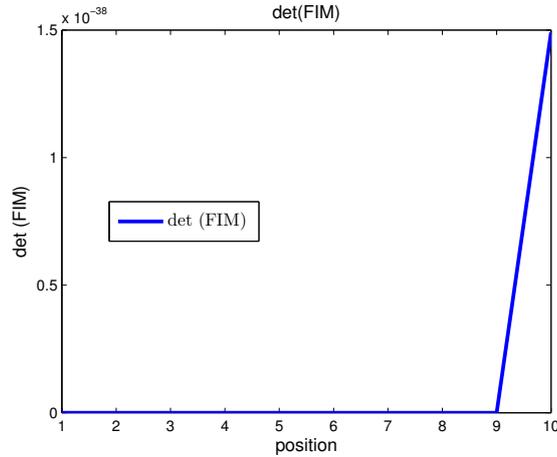


Figure V.4: Sensor placement used Fisher information matrix criterion in an overland flow: D-optimality criterion.

One can see that the optimal location according to the trace-based sensitivity criterion is given by the highest peak (0.17) in Figure V.3a, which corresponds to position 6. This result is not relevant because if the sensors are placed at this position, the sensitivity of the measurement with respect to the values of initial state after position 5 are all equal to zero. This means that the trace-based criterion is not suitable, since it cannot guarantee the observability of all the distributed state. On the contrary, E-optimality criterion shows that the optimal position is the last discrete position 10. This position is known as being the one which ensures observability. The same result is found in the case of the D-optimality criterion, as shown in Figure V.4. Note that, due to the ill-conditioned computation, the largest determinant of Fisher information matrix, given by measurement at position 10, is here very small 1.5×10^{-38} , but still larger than the other ones. This position corresponds to the optimal sensor position given using both D and E-optimality criterion.

2.a.2 Using the adjoint-based criterion

On the same system, the adjoint-based criterion is now used to find the optimal location of a single sensor. The evolution of this criterion over the position set is depicted in Figure V.5 where one can easily observe that position 10 gives the maximum value of index function \mathcal{I} . This result is similar to the one obtained with the Fisher information matrix criterion.

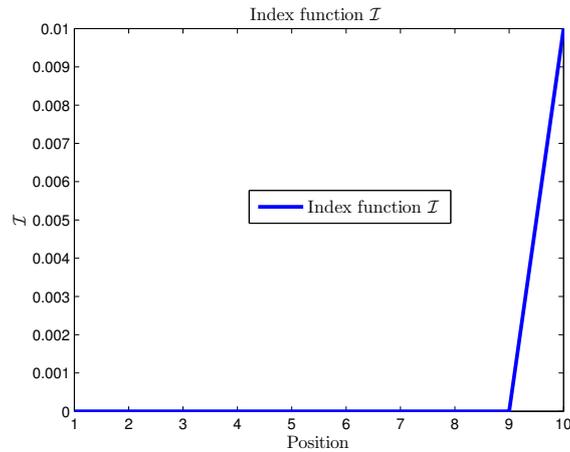


Figure V.5: One-sensor placement ubased on the adjoint-based criterion.

2.a.3 State estimation using optimal and non optimal locations

Each of the two proposed methods shows that the optimal location is the position 10 or at $900 - m$, that is the end of the interior section of spatial domain. It is now important to use the measurement at this position to define the cost function \mathcal{J} and to assess the performance of the estimation algorithm. In addition, the measurement at two positions 5 and 6 (i.e. at $400 - m$ and $500 - m$) will be evaluated also to illustrate the algorithm behavior when the condition ensuring full state observability over the domain are not fulfilled. Figure V.6a presents the estimation result of the initial state of the system with measurements performed at position 10. The estimated values are pretty close to the real one with a small average relative error of 0.41%. It appears that the relative errors corresponding to measurements performed at positions 5 and 6 are 20.75% and 29.75%, which are much larger. Not surprisingly, the only good estimated value of the state components, see in Figure V.6b and Figure V.7 are obtained at positions 5 and 6.

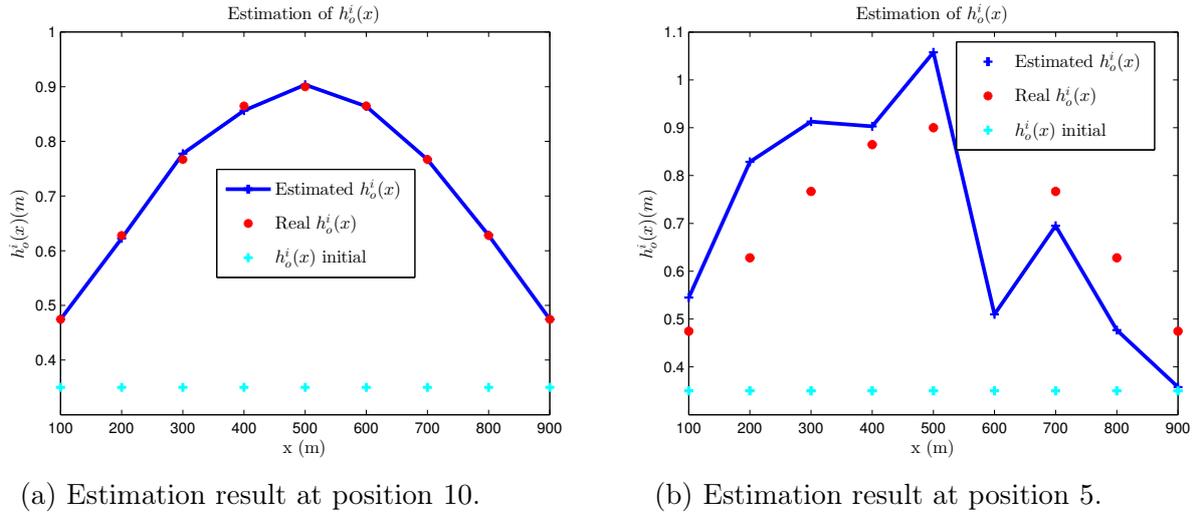


Figure V.6: State estimation in an overland flow with a single sensor at position 10 and position 5.

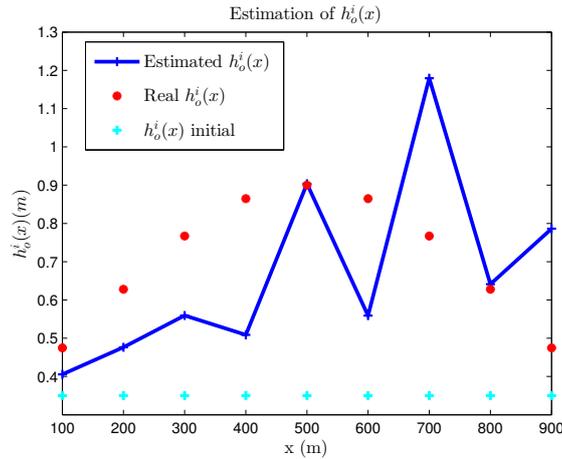


Figure V.7: State estimation in an overland with a single sensor at position 6.

2.b Sensor placement in an overland flow network

The adjoint-based criteria is now applied to a more complex system: a network of three water flow pools, with two upstream flow pools whose waters are pouring into a third pool (i.e, a Y topology). The problem is not to locate multiple sensors all over the spatial domains of the three flows and to investigate the optimal position combinations in order to get the best estimate of all the initial states of the three flows.

2.b.1 Dynamics of overland flow network

As depicted in Fig. V.8, this network consists in three small overland flows defined on three spatial domains called $\Omega_1 = [a_1, a_3]$, $\Omega_2 = [a_2, a_3]$ and $\Omega_3 = [a_3, a_4]$ with spatial variables $(x_1, x_2, x_3) \in \Omega_1 \times \Omega_2 \times \Omega_3$, respectively denoted by Σ_1 , Σ_2 and Σ_3 . The downstream end of the two upstream flow pools Σ_1 and Σ_2 , denoted as a_3 , is the interconnection point where the water runs into the downstream flow pool Σ_3 .

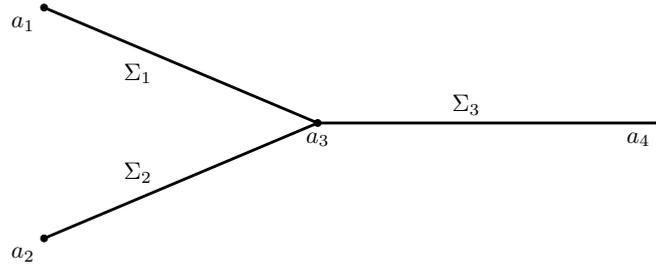


Figure V.8: Overland flow network configuration

Each of the three flows are governed by the water flow depth $h(x, t)$ PDE, already defined in equation (III.22) of chapter III. It is more convenient for reading if the system dynamics are represented hereafter by equation (V.1).

$$\frac{\partial h_j(x_j, t)}{\partial t} + \frac{\partial f_j(h_j(x_j, t))}{\partial x_j} = r_j(t) - i_j(t) \quad (\text{V.1})$$

$$j = \{1, 2, 3\} \quad \text{for } \Sigma_1, \Sigma_2 \text{ and } \Sigma_3$$

with the same notations as the ones in equation (III.22). The subscript j is used to refer to the j^{th} system. Some Dirichlet conditions are used to define the upstream end boundary conditions of system 1 and 2, together with some initial state conditions, as follows

$$(\Sigma_1) \quad \begin{cases} h_1(x_1, 0) = h_{01}^i(x_1) \\ h_1(a_1, t) = h_{01}^b(t) \end{cases} \quad (\Sigma_2) \quad \begin{cases} h_2(x_2, 0) = h_{02}^i(x_2) \\ h_2(a_2, t) = h_{02}^b(t) \end{cases}$$

The initial state of system Σ_3 is defined similarly, $h_3(x_3, 0) = h_{03}^i(x_3)$. At the interconnection point between the three pools at position a_3 , the mass conservation principle is used to define the confluence condition, which is specified by equation (V.2)

$$f_3(a_3, t) = f_1(a_3, t) + f_2(a_3, t) \quad (\text{V.2})$$

This condition can be expressed in terms of water flow depth $h_j(x_j, t)$ as

$$h_3(a_3, t) = \left\{ [\alpha_1(a_3)h_1^{5/3}(a_3, t) + \alpha_2(a_3)h_2^{5/3}(a_3, t)] / \alpha_3(a_3) \right\}^{3/5} \quad (\text{V.3})$$

This means that the boundary condition of Σ_3 at a_3 is specified by the water depths of the two upstream flows. The information about initial states of upstream flows is transmitted to the downstream system through this confluence condition, as the derivation of the adjoint equations of the network will show it in the next section.

2.b.2 Generalized adjoint-based criterion for the water flow network

Derivation of the adjoint equations

Because the general criterion of gradient defined in the previous chapter is defined for a single onverland flow only, a more general derivation of the adjoint equations of the network is needed. The initial conditions functions $h_{01}^i(x_1)$, $h_{02}^i(x_2)$ and $h_{03}^i(x_3)$ of all the systems are supposed to be unknown and need to be estimated by solving also an inverse problem based on a new cost function \mathcal{J} . Without any loss of generality, it is assumed that multiple sensors can be located along all the three pools. As a consequence, \mathcal{J} can be defined by equation (V.4).

$$\begin{aligned} \mathcal{J} = & \frac{1}{2} \sum_{j_1=0}^{M_1} \int_0^T \left\{ \int_{a_1}^{a_3} \delta_A(x_1 - x_{1j_1}) h_1 dx_1 - h_1^{meas}(x_{1j_1}, t) \right\}^2 dt + \frac{1}{2} \sum_{j_2=0}^{M_2} \int_0^T \left\{ \int_{a_2}^{a_3} \delta_A(x_2 - x_{2j_2}) h_2 dx_2 \right. \\ & \left. - h_2^{meas}(x_{2j_2}, t) \right\}^2 dt + \frac{1}{2} \sum_{j_3=0}^{M_3} \int_0^T \left\{ \int_{a_3}^{a_4} \delta_A(x_3 - x_{3j_3}) h_3 dx_3 - h_3^{meas}(x_{3j_3}, t) \right\}^2 dt \end{aligned} \quad (\text{V.4})$$

The sensor positions are denoted by x_{1j_1} , x_{2j_2} and x_{3j_3} , respectively for system 1, system 2 and system 3. Again the goal is to derive an adjoint-based criterion suitable to derive the optimal locations of the sensors in order to get the best performance of the inverse problem algorithm, in a similar way to the one presented in the previous chapter. In this case, the variables to be estimated are the set of the three initial states of the network. The derivation of the adjoint system will be now briefly presented. The objective functional \mathcal{L} , defined on the function space $L_2([a_1, a_3], [0, T]) \times L_2([a_2, a_3], [0, T]) \times L_2([a_3, a_4], [0, T]) \times L_2([a_1, a_3]) \times L_2([a_2, a_3]) \times L_2([a_3, a_4])$ into \mathbb{R} with three Lagrangian multipliers $\lambda_1(x_1, t)$,

$\lambda_2(x_2, t)$ and $\lambda_3(x_2, t)$ is given by the following equation.

$$\begin{aligned} \mathcal{L}(h_j(x, t), h_{0j}^i(x_j)) = \mathcal{J} + \int_0^T \int_{a_1}^{a_3} \lambda_1(x_1, t) \Sigma_1 dx_1 dt + \int_0^T \int_{a_2}^{a_3} \lambda_2(x_2, t) \Sigma_2 dx_2 dt \\ + \int_0^T \int_{a_3}^{a_4} \lambda_3(x_3, t) \Sigma_3 dx_3 dt \quad \text{with } j=\{1, 2, 3\} \end{aligned} \quad (\text{V.5})$$

The inner product \langle, \rangle is here defined as

$$\begin{aligned} \langle \nabla \mathcal{L}, \pi \rangle = \int_0^T \int_{a_1}^{a_3} \nabla \mathcal{L}_{h_1(x_1, t)} \delta h_1(x_1, t) dx_1 dt + \int_0^T \int_{a_2}^{a_3} \nabla \mathcal{L}_{h_2(x_2, t)} \delta h_2(x_2, t) dx_2 dt \\ + \int_0^T \int_{a_3}^{a_4} \nabla \mathcal{L}_{h_3(x_3, t)} \delta h_3(x_3, t) dx_3 dt + \int_{a_1}^{a_3} \nabla \mathcal{L}_{h_{01}^i(x_1)} \delta h_{01}^i(x_1) dx_1 + \int_{a_2}^{a_3} \nabla \mathcal{L}_{h_{02}^i(x_2)} \delta h_{02}^i(x_2) dx_2 \\ + \int_{a_3}^{a_4} \nabla \mathcal{L}_{h_{03}^i(x_3)} \delta h_{03}^i(x_3) dx_3 \end{aligned} \quad (\text{V.6})$$

where $\nabla \mathcal{L}$ is a vector of 6 elements $\nabla \mathcal{L} = [\nabla \mathcal{L}_{h_1(x_1, t)} \quad \nabla \mathcal{L}_{h_2(x_2, t)} \quad \nabla \mathcal{L}_{h_3(x_3, t)} \quad \nabla \mathcal{L}_{h_{01}^i(x_1)}$
 $\mathcal{L}_{h_{02}^i(x_2)} \quad \mathcal{L}_{h_{03}^i(x_3)}]$ with the variation vector in all directions $\pi = [\delta h_1(x_1, t) \quad h_2(x_2, t)$
 $h_3(x_3, t) \quad h_{01}^i(x_1) \quad h_{02}^i(x_2) \quad h_{03}^i(x_3)]$. The first variation of \mathcal{L} is equal to the variation of two parts, \mathcal{J} and the sum of three remaining terms called \mathcal{K} in direction π . The variation of cost function \mathcal{J} is given as,

$$\begin{aligned} \delta \mathcal{J} = \sum_{j_1=1}^{M_1} \int_0^T \int_{a_1}^{a_3} \delta_A(x_1 - x_{1j_1}) \left[\int_{a_1}^{a_3} \delta_A(x_1 - x_{1j_1}) h_1 dx_1 - h_1^{meas}(x_{1j_1}, t) \right] \delta h_1 dx_1 dt \\ + \sum_{j_2=1}^{M_2} \int_0^T \int_{a_2}^{a_3} \delta_A(x_2 - x_{2j_2}) \left[\int_{a_2}^{a_3} \delta_A(x_2 - x_{2j_2}) h_2 dx_2 - h_2^{meas}(x_{2j_2}, t) \right] \delta h_2 dx_2 dt \quad (\text{V.7}) \\ + \sum_{j_3=1}^{M_2} \int_0^T \int_{a_3}^{a_4} \delta_A(x_3 - x_{3j_3}) \left[\int_{a_3}^{a_4} \delta_A(x_3 - x_{3j_3}) h_3 dx_3 - h_3^{meas}(x_{3j_3}, t) \right] \delta h_3 dx_3 dt \end{aligned}$$

end also the first variation of \mathcal{K} is obtained by using an integration by part technique:

$$\begin{aligned}
\delta\mathcal{K} = & \int_{a_1}^{a_3} \left[\lambda_1 \delta h_1 \right] \Big|_0^T dx_1 - \int_0^T \int_{a_1}^{a_3} \frac{\partial \lambda_1}{\partial t} \delta h_1 dx_1 dt + \int_0^T \left[\lambda_1 \frac{\partial f_1}{\partial h_1} \right] \delta h_1 \Big|_{a_1}^{a_3} dt - \int_0^T \int_{a_1}^{a_3} \frac{\partial f_1}{\partial h_1} \frac{\partial \lambda_1}{\partial x_1} \delta h_1 dx_1 dt \\
& + \int_{a_2}^{a_3} \left[\lambda_2 \delta h_2 \right] \Big|_0^T dx_2 - \int_0^T \int_{a_2}^{a_3} \frac{\partial \lambda_2}{\partial t} \delta h_2 dx_2 dt + \int_0^T \left[\lambda_2 \frac{\partial f_2}{\partial h_2} \right] \delta h_2 \Big|_{a_2}^{a_3} dt - \int_0^T \int_{a_2}^{a_3} \frac{\partial f_2}{\partial h_2} \frac{\partial \lambda_2}{\partial x_2} \delta h_2 dx_2 dt \\
& + \int_{a_3}^{a_4} \left[\lambda_3 \delta h_3 \right] \Big|_0^T dx_3 - \int_0^T \int_{a_3}^{a_4} \frac{\partial \lambda_3}{\partial t} \delta h_3 dx_3 dt + \int_0^T \left[\lambda_3 \frac{\partial f_3}{\partial h_3} \right] \delta h_3 \Big|_{a_3}^{a_4} dt - \int_0^T \int_{a_3}^{a_4} \frac{\partial f_3}{\partial h_3} \frac{\partial \lambda_3}{\partial x_3} \delta h_3 dx_3 dt
\end{aligned} \tag{V.8}$$

The three adjoint systems are then derived as the weak form of the gradient of \mathcal{L} with respect to $h_1(x, t)$, $h_2(x, t)$ and $h_3(x, t)$. They are obtained by collecting the terms under the double integral multiplied by variations δh_1 , δh_2 and δh_3 . The three adjoint PDEs are given by equations (V.9), (V.10) and (V.11).

$$-\frac{\partial \lambda_1}{\partial t} - \frac{\partial f_1}{\partial h_1} \frac{\partial \lambda_1}{\partial x_1} + \sum_{j_1=0}^{M_1} \delta_A(x_1 - x_{1j_1}) \left[\int_{a_1}^{a_3} \delta_A(x_1 - x_{1j_1}) h_1(x_1, t) dx_1 - h_1^{meas}(x_{1j_1}, t) \right] = 0 \tag{V.9}$$

$$-\frac{\partial \lambda_2}{\partial t} - \frac{\partial f_2}{\partial h_2} \frac{\partial \lambda_2}{\partial x_2} + \sum_{j_2=0}^{M_2} \delta_A(x_2 - x_{2j_2}) \left[\int_{a_2}^{a_3} \delta_A(x_2 - x_{2j_2}) h_2(x_2, t) dx_2 - h_2^{meas}(x_{2j_2}, t) \right] = 0 \tag{V.10}$$

$$-\frac{\partial \lambda_3}{\partial t} - \frac{\partial f_3}{\partial h_3} \frac{\partial \lambda_3}{\partial x_3} + \sum_{j_3=0}^{M_3} \delta_A(x_3 - x_{3j_3}) \left[\int_{a_3}^{a_4} \delta_A(x_3 - x_{3j_3}) h_3(x_3, t) dx_3 - h_3^{meas}(x_{3j_3}, t) \right] = 0 \tag{V.11}$$

The weak gradients of \mathcal{L} with respect to the initial states are the terms multiplied by $\delta h_j(x_j, 0)$ where $j \in \{1, 2, 3\}$. The optimization algorithm will ensure that these gradients will vanish in order to satisfy the first order necessary conditions for optimality.

$$\begin{aligned}
\nabla \mathcal{L}_{h_{01}^i(x_1)} &= -\lambda_1(x_1, 0) & \nabla \mathcal{L}_{h_{02}^i(x_2)} &= -\lambda_2(x_2, 0) \\
\nabla \mathcal{L}_{h_{03}^i(x_3)} &= -\lambda_3(x_3, 0)
\end{aligned} \tag{V.12}$$

After the derivation of the gradients and the adjoint equations, the remaining terms of (V.7) and (V.8) are gathered together and set to zero to derive the initial state and boundary

conditions for the adjoint PDEs.

$$\begin{aligned}
& - \int_{a_1}^{a_3} \lambda_1(x_1, T) \delta h_1(x_1, T) dx_1 - \int_{a_2}^{a_3} \lambda_2(x_2, T) \delta h_2(x_2, T) dx_2 - \int_{a_3}^{a_4} \lambda_1(x_3, T) \delta h_3(x_3, T) dx_3 \\
& + \int_0^T \left[\lambda_1 \frac{\partial f_1}{\partial h_1} \right] \delta h_1 \Big|_{a_1}^{a_3} dt + \int_0^T \left[\lambda_2 \frac{\partial f_2}{\partial h_2} \right] \delta h_2 \Big|_{a_2}^{a_3} dt + \int_0^T \left[\lambda_3 \frac{\partial f_3}{\partial h_3} \right] \delta h_3 \Big|_{a_3}^{a_4} dt = 0
\end{aligned} \tag{V.13}$$

Since the boundary conditions of the three systems are fixed, their variations are equal to zero, i.e., $\delta h_1(a_1, t) = 0$, $\delta h_2(a_2, t) = 0$. To satisfy the above equality for all variation directions $\delta h_1(x_1, T)$, $\delta h_2(x_2, T)$ and $\delta h_3(a_4, t)$, the following equalities must hold: $\lambda_1(x_1, T) = 0$, $\lambda_2(x_2, T) = 0$, $\lambda_3(x_3, T) = 0$ and $\lambda_3(a_4, t) = 0$ which are also the initial conditions for the adjoint systems. Finally, the remaining time integrals are gathered as follows:

$$\begin{aligned}
& \int_0^T \left[\lambda_1(a_3, t) \frac{\partial f_1}{\partial h_1}(a_3, t) \delta h_1(a_3, t) + \lambda_2(a_3, t) \frac{\partial f_2}{\partial h_2}(a_3, t) \delta h_2(a_3, t) \right. \\
& \qquad \qquad \qquad \left. - \lambda_3(a_3, t) \frac{\partial f_3}{\partial h_3}(a_3, t) \delta h_3(a_3, t) \right] dt = 0
\end{aligned} \tag{V.14}$$

By considering the first variation of interconnection condition (V.2) at the interconnection point a_3 , one can obtain

$$\delta f_3(a_3, t) = \delta f_1(a_3, t) + \delta f_2(a_3, t) \tag{V.15}$$

By removing the time integral of equation (V.14) and putting equation (V.15) into (V.14) the following equalities hold:

$$\begin{aligned}
\lambda_1(a_3, t) &= \lambda_3(a_3, t) \\
\lambda_2(a_3, t) &= \lambda_3(a_3, t)
\end{aligned} \tag{V.16}$$

These boundary conditions of the three adjoint systems are therefore induced by the confluence condition (V.2) of the direct PDEs.

Derivation of the generalized sensor placement criterion

As in the approach proposed in section 3.b of the previous chapter, the gradients defined by equation (V.12) are the ingredients for the generalized sensor placement criterion. Following the idea developed in 3.b, the absolute minimal value of gradients is used as an

index function. However in this case, there are three initial states to be estimated. This means that the selected locations will guarantee the observability of all states over the spatial domains of all three systems. By denoting the combination of discretized observation positions selected as ξ and the set of all the position locations on the three domains Ω_1, Ω_2 and Ω_3 as $\chi = \{x_{i1}, x_{j2}, x_{k3}\}$ where $1 < i \leq N_1 - 1$ with $\Delta x_1 = (a_3 - a_1)/N_1$, $1 < j \leq N_2 - 1$ with $\Delta x_2 = (a_3 - a_2)/N_2$ and $0 < k \leq N_3 - 1$ with $\Delta x_3 = (a_4 - a_3)/N_3$, the discretized versions of the gradients for initial states in equation (V.12) are given by

$$\begin{aligned} \nabla \mathcal{L}_{h_{01}^i(x_{i1})} &= -\lambda_1(x_{i1}, 0) & \nabla \mathcal{L}_{h_{02}^i(x_{j2})} &= -\lambda_2(x_{j2}, 0) \\ \nabla \mathcal{L}_{h_{03}^i(x_{k3})} &= -\lambda_3(x_{k3}, 0) \end{aligned} \quad (\text{V.17})$$

The set $\{x_{i1}, x_{j2}, x_{k3}\}$ is defined as the set of the discretized positions of the sensor location domain $\{x_{1j_1}, x_{2j_2}, x_{3j_3}\}$ in equation (V.4). As a result, the index function is formulated as follows:

$$\mathcal{I}(\xi) = \min \left[\text{abs} \begin{pmatrix} -\lambda_1(x_{1i}) \\ -\lambda_2(x_{2j}) \\ -\lambda_3(x_{3k}) \end{pmatrix} \right] \quad (\text{V.18})$$

The optimal location problem consists in maximizing function $\mathcal{I}(\xi)$

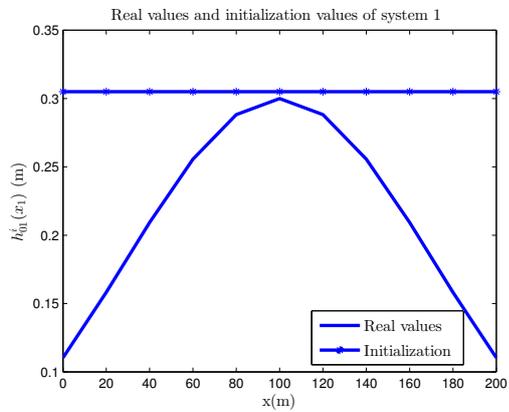
$$\max_{\xi \in \chi} \{ \mathcal{I}(\xi) \} \quad (\text{V.19})$$

Sensor placement result

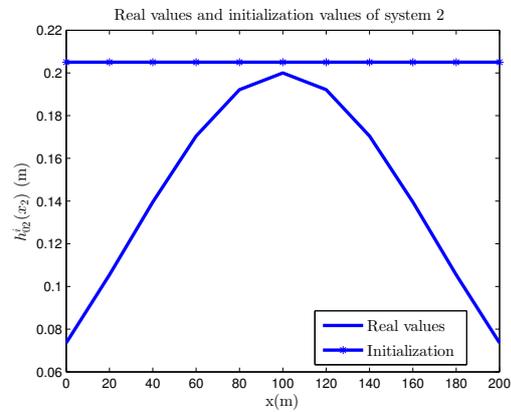
Defining the number of sensors to be used is also an important issue beside the issue of defining the optimal sensor locations. In this thesis the maximum number of sensors will be fixed to three and the impact of the number of sensors on the performance of the inverse problem computation will not be evaluated. Three cases will be considered in this section: state estimation with 1, 2 or 3 sensors. Since the number of sensors is limited to three, the computation of the optimal sensor location problem is still tractable by using a simple enumeration of all the possible combinations. For a larger number of sensors or a with larger number of discrete spatial positions, the number of combinations can be very large. As a result, some adequate integer programming algorithms should be used to overcome the combinatorial burden issue. In this example, the parameters used for the numerical simulations of the three overland flows are summarized in Table V.1. The real value of the initial states and the initialization parameters of algorithm for 3 systems are depicted in Figures V.9a, V.9b and V.10.

Parameters	Values			Unit
	System 1	System 2	System 3	
Length of flow	200	200	300	<i>m</i>
Simulation time	1	1	1	<i>hour</i>
Space step Δx	20	20	30	<i>m</i>
Time step Δt	0.1	0.1	0.1	<i>hour</i>
Bed slope S_0	1%	1%	1%	<i>m/m</i>
Manning coefficient n	0.002	0.002	0.0025	<i>hour/m^{1/3}</i>

Table V.1: Parameters used for the numerical simulations of the three overland flows.



(a) Real initial state and initialization of system 1.



(b) Real initial state and initialization of system 2.

Figure V.9: Real initial state and initialization of systems 1 and 2.

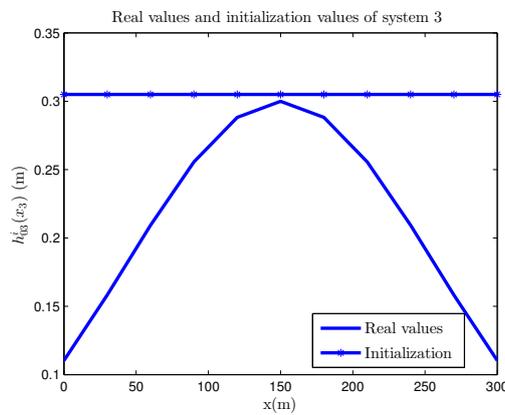


Figure V.10: Real initial state and initialization of system 3.

The lengths and the step sizes of systems are different but each system is always divided into 10 sections with 11 discrete positions. The 10 interior positions from 1 to 10 define the set of possible locations. It is important to notice that the final positions of system 1 and system 2 at the interconnection point are physically the same, and identical to the first position of system 3. There are consequently 28 discretized positions all along the flow pools, which define the accessible set. The optimal combination(s) is/are computed using an enumeration procedure based on all the possible combinations of sensor locations over the domain $\Omega_1 \times \Omega_2 \times \Omega_3$ for a given number of sensors. The application of this procedure is now considered for the three cases of sensor number.

One-sensor case

Because there is only one sensor available, there are only 28 combinations. The value of the placement criterion defined by (V.19) for all the possible discrete position combinations is presented in Fig. V.11, where the highest value of index function is 0.4047×10^{-13} . This means that a single sensor must be located at combination number 28, which corresponds to the discrete position 10 of system 3, on spatial domain Ω_3 . Moreover, one can observe that, all the other positions give an index function value equal to zero. This means that if one places a single sensor at these locations, the values of all three initial states cannot be all recovered. This result is interesting since the maximum of the minimal gradient ensures the observability for all the discrete states of the three systems. It appears that position 10 on system 3 is the only one which receives the information propagated from the upper-streams and its preceding points. The sensor position 7 (on system 3) for instance, allows to estimate all the information (the discrete states) propagated through systems 1, 2 from their upper-stream ends, and the discrete states from 1 to 7 of system 3.

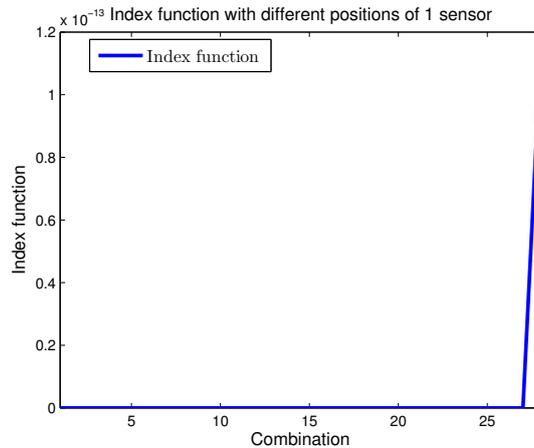


Figure V.11: Values of the adjoint-based criterion for different locations of a single sensor

Two-sensors case

With two sensors, the number of possible combinations is now 378. Figure V.12 shows that the maximal value of the criterion is 0.5790×10^{-3} , which is obtained at 3 combinations 350, 357, 363 represented by the highest peaks in this figure. Its corresponding discrete positions are $\{2, 10\}$, $\{3, 10\}$ and $\{4, 10\}$ all located on $\Omega_3 \times \Omega_3$. There are some sub-optimal locations located on $\Omega_3 \times \Omega_3$ such as $\{5, 10\}$ (index function 0.4935×10^{-3}), $\{6, 10\}$ (index function 0.1967×10^{-3}) and $\{7, 10\}$ (index function 0.0089×10^{-3}) which are also some potential sub-optimal locations for the sensors. One can remark that all of them share the presence of position 10 of system 3. This result can be interpreted as the fact that the system 3 catches all the information from the two previous ones and if a sensor is placed near the conjunction point, the obtained information will not be heavily influenced by the initial state of system 3 and the upper-stream end information of system 3 reflecting the information propagated through both systems 1 and 2 will be available through this sensor. The result is also consistent with the result of the one-sensor case.

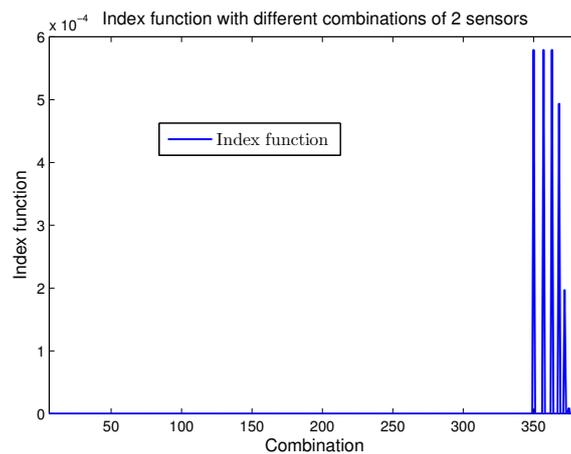


Figure V.12: Index function with different location of 2 sensors

Three-sensors case

In this case, there are 3276 position combinations for 3 sensors. Figure V.13 shows that there is a group of 6 optimal combinations which give the highest sensitivity to estimate states (all equal to 0.0011). This group is depicted in the last blue column at the end of figure V.13. They are $\{2, 9, 10\}$, $\{3, 9, 10\}$, $\{4, 9, 10\}$ (all of them are on $\Omega_3 \times \Omega_3 \times \Omega_3$) and $\{2, 8, 10\}$, $\{3, 8, 10\}$ and $\{4, 8, 10\}$ (also on $\Omega_3 \times \Omega_3 \times \Omega_3$). Some of other suboptimal sensor configurations are $\{2, 7, 10\}$, $\{3, 7, 10\}$ and $\{4, 7, 10\}$ (all of them is on $\Omega_3 \times \Omega_3 \times \Omega_3$ sharing the same index function value of 0.0010).

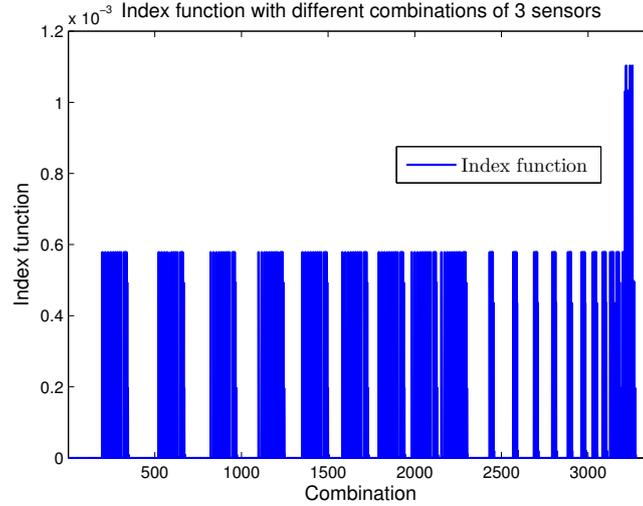


Figure V.13: Adjoint-based criterion w.r.t. locations of 3 sensors

From this analysis, it appears that at least one sensor must be placed at the downstream end of system 3 to ensure the observability of all the system states. The locations of the three sensors on Ω_3 always provide a larger index function than the location of one of the sensors at the downstream end and the two others on Ω_1 and/or Ω_2 . The optimal locations of the other sensors depend on the physical characteristics of the network. However the initial values fed to the inverse problem algorithm may have some impact on the sensor placement result. A different configuration of algorithm initialization is used to investigate the independence between the chosen sensor locations with the initial value of the estimation algorithm (in this case, the values of the initial state of the three systems are not changed). After the calculation, this initialization almost gives the same sensor positions. The choice of optimal sensor locations provided by adjoint-gradient approach seems to be not really sensitive to the choice of initialization. In practice, averaging of the adjoint-based criterion over a set of initialization scenarios would be a solution to minimize the potential effect of initialization. However this aspect has not been done yet.

In what follows, the focus is made on the evaluation of the inverse problem performance with respect to the locations of the sensors on the basis of a fixed initial value profile. All the 8 sensor combinations which have been derived, are presented with the corresponding sensor positions and criterion values.

2.b.3 State estimation using optimal locations

The relative errors of the state estimation using 8 sensor configurations are presented in Table V.2. With case 1, when only one sensor is used and located at position 10 on Ω_3 , we can achieve a pretty good estimation result for initial condition of system 3 (average

relative error equal to 2.4%) but some pretty large estimation errors always remain for the upstream systems. With case 2, when the sensor is located at position $\{7\}$ also on Ω_7 , one can observe better results concerning the estimation of systems 1 and 2 while the estimation of system 3 gets worse (with a larger estimation relative error). In fact, as already discussed, the sensor position at 7 on Ω_3 does not have any state information after it (point $\{8, 9, 10\}$ of Ω_3) and then the three corresponding discrete states are not observable, since the adjoint gradient at these points are equal to zero. The estimation algorithm cannot recover the real values of initial state at these positions. This is also the reason why the index function at position 7 is zero. However, the shorter distance from position 7 to the interconnection point a_3 reinforce the observability of both systems 1 and 2, since this location provides more information from systems 1 and 2 and, as a result, the relative errors of the upstream systems 1 and 2 in this case are smaller than in case 1.

Case	Sensor number	Combination	Index function value	Estimation relative error		
				System 1	System 2	System 3
1	1	$\{10\}$ on Ω_3	0.4047×10^{-13}	19.27%	19.22 %	2.40 %
2	1	$\{7\}$ on Ω_3	0	10.27 %	10.15 %	6.77 %
3	2	$\{2, 10\}$ on $\Omega_3 \times \Omega_3$	0.5790×10^{-3}	3.88%	5.91 %	0.12 %
4	2	$\{3, 10\}$ on $\Omega_3 \times \Omega_3$	0.5790×10^{-3}	4.76 %	7.91%	0.014 %
5	2	$\{6, 10\}$ on $\Omega_3 \times \Omega_3$	0.1967×10^{-3}	10.56 %	11.33%	0.75 %
6	3	$\{2, 9, 10\}$ on $\Omega_3 \times \Omega_3 \times \Omega_3$	0.0011	3.46 %	6.15 %	0.12 %
7	3	$\{2, 8, 10\}$ on $\Omega_3 \times \Omega_3 \times \Omega_3$	0.0011	4.05%	5.62 %	0.09 %
8	3	$\{3, 7, 10\}$ on $\Omega_3 \times \Omega_3 \times \Omega_3$	0.0010	3.65 %	6.10 %	0.11 %

Table V.2: Estimation errors with different sensor configurations

The number of sensors is then increased to 2, as in cases 3, 4 and 5 in Table V.2, the estimation accuracy also increases for all the systems. The estimation errors of two optimal combination $\{2, 10\}$ and $\{3, 10\}$ are smaller than the one of case 5. This result appears to be correlated with the result of the sensor placement. The combination giving more sensitivity (higher criterion value), i.e cases 3, 4, allows to get smaller estimation errors than for case 5. More over two cases 3, 4 almost give a close estimation error.

Three last cases of 3 sensors are then considered. The obtained results appear to be not very different when compared between them. However, the estimation errors between these 3 cases are really close to case 2 where 2 sensors are located at 2, 10 on $\Omega_3 \times \Omega_3$. This leads to the conclusion that sensor positions $\{7, 8, 9\}$ on Ω_3 (which are present in cases 6, 7, 8) seem not to bring significant additional and useful information for state estimation, especially for the estimation of the initial state of the two upstream systems. Finally it appears that the proposed adjoint-based criterion provides an effective way to assess and predict the performance of the state estimation algorithm.

3 Conclusion

The two optimal sensor placement methodologies proposed in this chapter (based on the Fisher Information Matrix and the adjoint-based criteria) have been validated on an overland flow and an overland flow network. The optimal sensor locations are determined by solving a maximization problem via an enumerative approach if the number of combinations remains small. Then cost function \mathcal{J} of the inverse problem problems described in chapter II is modified to include The performance analysis of the state estimation performed in this chapter proves that the optimal sensor location results provided by the proposed optimal sensor location methodologies are the most appropriate, in the sense that they provide the greatest sensitivity of the cost functional with respect to the unknown initial states to estimate. Even though the adjoint-based approach has not been applied to the traffic flow example, one can guess that similar results could be obtained, since the problem is similar to the overland flow one. It is important to mention that the optimal sensor location configurations depends on the initial values used by the optimization algorithm. In order to make the proposed approach more robust with respect to the initial values, the idea is to introduce randomly-chosen initialization scenarios and to consider an average criterion computed with these scenarios. The remaining errors present in the estimation results, (see in Table V.2), may be due to mismatches between the numerical methods used for sensor placement (based on a simple gradient descend approach) and the state estimation algorithm (based on a quasi-Newton approach).

But beside the sensor position, the determination of the number of sensors, which also impacts the estimation process has not been extensively investigated. Moreover, if the sensor location optimization is performed with a large number of sensors, the number of combinations will increase dramatically. An enumerative procedure is no more possible in this case and some integer programming algorithms should be employed.

All these different issues will be investigated in some future works.

The sensor placement example on the overland flow network which has been presented in this chapter is a part of publication [NGB16a].

VI | Conclusions and perspectives

The main goal of this chapter is to synthesize the works that have been presented in this thesis and to give some perspectives for future works to go deeper in some of the topics.

1 Conclusions

The contributions of this thesis can be divided into two main parts: some contributions to both state and parameter estimation of nonlinear balance equations (hyperbolic partial differential equations) in the first part and some contributions to optimal sensor placement for the same partial differential equations in second one. Each of them is developed in the same way: the proposed approaches (for estimation and for sensor positioning) are first presented in a methodology chapter and then their validation is considered with examples in an application chapter.

For the first part of the thesis, the main contribution can be summarized as follows: An adjoint based estimation method is developed for a class of 1-D switched hyperbolic systems. The calculus of variations plays a central role for the solution of the estimation problems. The adjoint method used for state and distributed parameter estimation is formulated with a very general manner, which can be easily reduced to simpler estimation cases: Estimation of the initial state only or of some constant/distributed parameters only, and for non-switched hyperbolic equations. It is important to point out that, to the best of my knowledge, it is the first time that an estimation methodology is proposed for switched hyperbolic equations, thanks to the use of continuous activation/switching functions, which allows to apply smooth calculus of variations. This method allows, to simultaneously estimate the initial state, the distributed system parameters, and also the parameters in the switching source terms, in an infinite dimensional framework. In this thesis, a direct nonlinear approach is used: The calculus of variations is applied directly to

the original PDE and ODE governing the system without linearization and reduction of the PDE in finite dimensions. The necessary conditions for optimality lead to a set of direct and adjoint PDEs which are then solved by using time-space discretization techniques. Some numerical approaches and gradient-based optimization are also investigated. The proposed estimation method is validated with various application scenarios: Estimation of constant parameters, state and distributed parameters in non-switched/switched traffic flows, of the initial state only, of distributed coefficients only, of both state and coefficients in overland flow using synthetic measurements; parameters estimation using real data on Tondi-Kiboro catchment. Moreover, from an application point of view, Tondi-Kiboro estimation application offers the opportunity to derive a realistic model of water flow on Tondi-Kiboro catchment which promotes also further researches dealing with hydrological changes including soil erosion, components of water cycle, fallow period changes in this region. A new switched traffic flow model, based on the LWR equation which can handle relief route has also been proposed in this thesis.

For the second part of the thesis which is dedicated to the issue of optimally locating sensors, a focus is made on the extension of an existing optimal sensor placement technique, the so-called Fisher information matrix technique which is here applied on a class of non-switched hyperbolic system. To overcome the main disadvantage of Fisher information matrix approach related to the need of solving a large set of sensitivity nonlinear partial differential equations, a new method has then been proposed: this approach relies on the use of the adjoint state provided by the necessary conditions for optimality of the optimal estimation problem. It ensures a strong connection between observability (in the sense of guaranteeing non zero sensitivity of the estimation cost functional with respect to the state or parameters to estimate) and performance of the gradient-based iterative method used to numerically solve the optimal estimation problem. While the Fisher information matrix approach is validated only a single traffic flow and overland flow, the validation of adjoint-gradient criterion is realized not only on a single flow but also on a overland flow network.

2 Perspectives

In the view of the obtained results, several perspectives/extensions can be deduced:

- The validation of the adjoint-based estimation method has been performed with real data only for the case of overland flow, a real date validation or some laboratory experiments are expected to evaluate the applicability of the presented method to real problems in traffic flow.

-
- Results of optimal sensor placement in chapter V, especially for the case of the overland flow network, depend on the choice of initial conditions (the initial values used in the iterative optimization method, not the ones used to generate the synthetic measurements). Each initial condition configuration can possibly give a different sensor position. A deeper robustness analysis of the adjoint based approach will be part of future works. An extension to the case of optimal sensor location for both state *and parameter* estimation is also an interesting topic to investigate.
 - Inclusion of additional constraints in the optimal sensor location methodology such as communication constraints should be also investigated in the future.
 - The generalization of both the optimal estimation method and the optimal sensor location one to more complex network of balance laws is possible and should be next investigated.
 - Some more general problems in infinite dimensional hyperbolic system such as output-feedback stabilization or optimal control with output constraints could help to widen the applicability of the presented estimation method. The parameters considered in this thesis are time-invariant which can restrict the generality of the method. A time varying parameter version will be then a possible part of future developments.

Bibliography

- [AHB08] Saurabh Amin, Falk M Hante, and Alexandre M Bayen. “Stability analysis of linear hyperbolic systems with switching parameters and boundary conditions.” In: *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE. 2008, pp. 2081–2086 (cit. on p. 11).
- [Ala10] JE Alaña. “Optimal measurement locations for parameter estimation of non linear distributed parameter systems.” In: *Brazilian Journal of Chemical Engineering* 27.4 (2010), pp. 627–642 (cit. on p. 103).
- [Ali94] Oleg M Alifanov. “Iterative Regularization of Inverse Problems.” In: *Inverse Heat Transfer Problems*. Springer, 1994, pp. 227–328 (cit. on p. 13).
- [ALR02] L Autrique, JP Leyris, and N Ramdani. “Optimal Sensor Location for the Identification of Moving Heat Sources.” In: *IFAC Proceedings Volumes* 35.1 (2002), pp. 187–192 (cit. on p. 102).
- [Arm66] Larry Armijo. “Minimization of functions having Lipschitz continuous first partial derivatives.” In: *Pacific Journal of mathematics* 16.1 (1966), pp. 1–3 (cit. on p. 40).
- [ARR05] L Autrique, N Ramdani, and S Rodier. “Mobile source estimation with an iterative regularization method.” In: *5th International Conference on Inverse Problems in Engineering: Theory and Practice, Cambridge, UK*. 2005, pp. 11–15 (cit. on p. 12).
- [Ath72] Michael Athans. “On the determination of optimal costly measurement strategies for linear stochastic systems.” In: *Automatica* 8.4 (1972), pp. 397–412 (cit. on p. 102).
- [Aut+00] Laurent Autrique et al. “Optimal sensor strategy for parametric identification of a thermal system.” In: *Proceedings of IFAC-SYSID 2000, Santa Barbara* (2000) (cit. on p. 102).

- [BB90] A Bagchi and P ten Brummelhuis. “Simultaneous ML estimation of state and parameters for hyperbolic systems with noisy boundary conditions.” In: *Decision and Control, 1990., Proceedings of the 29th IEEE Conference on.* IEEE. 1990, pp. 222–224 (cit. on p. 10).
- [BBP07] Yolanda Bolea, Joaquín Blesa, and Vicenç Puig. “LPV modelling and identification of an open canal for control.” In: *Control Conference (ECC), 2007 European.* IEEE. 2007, pp. 3584–3591 (cit. on p. 10).
- [Ben72] Alain Bensoussan. “Optimization of sensors’ location in a distributed filtering problem.” In: *Stability of stochastic dynamical systems.* Springer, 1972, pp. 62–84 (cit. on pp. 101, 102).
- [Ber+00] FWJ Van den Berg et al. “Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria.” In: *Chemical Engineering Science* 55.4 (2000), pp. 827–837 (cit. on pp. 103, 109).
- [Ber99] Dimitri P Bertsekas. *Nonlinear programming.* Athena scientific Belmont, 1999 (cit. on p. 41).
- [Bes07] Gildas Besançon. *Nonlinear observers and applications.* Vol. 363. Springer, 2007 (cit. on p. 105).
- [BRT06] Alexandre M Bayen, Robin L Raffard, and Claire J Tomlin. “Adjoint-based control of a new Eulerian network model of air traffic flow.” In: *Control Systems Technology, IEEE Transactions on* 14.5 (2006), pp. 804–818 (cit. on p. 12).
- [BY72] Leonard Becker and William W-G Yeh. “Identification of parameters in unsteady open channel flows.” In: *Water Resources Research* 8.4 (1972), pp. 956–965 (cit. on p. 10).
- [Cac81a] Dan G Cacuci. “Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach.” In: *Journal of Mathematical Physics* 22.12 (1981), pp. 2794–2802 (cit. on p. 11).
- [Cac81b] Dan G Cacuci. “Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses.” In: *Journal of Mathematical Physics* 22.12 (1981), pp. 2803–2812 (cit. on p. 11).
- [Cal+13] Victor M Calo et al. “Gradient-based estimation of Manning’s friction coefficient from noisy data.” In: *Journal of Computational and Applied Mathematics* 238 (2013), pp. 1–13 (cit. on p. 12).
- [Cas+09] William Castaings et al. “Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods.” In: *Hydrology and Earth System Sciences Discussions* 13.4 (2009), pp. 503–517 (cit. on p. 12).

- [CFL67] Richard Courant, Kurt Friedrichs, and Hans Lewy. “On the partial difference equations of mathematical physics.” In: *IBM journal* 11.2 (1967), pp. 215–234 (cit. on p. 33).
- [Cha04] Hubert Chanson. *Environmental hydraulics for open channel flows*. Butterworth-Heinemann, 2004 (cit. on p. 31).
- [Chu78] Shu Tung Chu. “Infiltration during an unsteady rain.” In: *Water Resources Research* 14.3 (1978), pp. 461–466 (cit. on pp. 81, 82).
- [CHV80] Jan Andrzej Cunge, Forrest M Holly, and Adri Verwey. *Practical aspects of computational river hydraulics*. Pitman publishing, 1980 (cit. on p. 30).
- [CI78] Ruth F Curtain and Akira Ichikawa. *Optimal location of sensors for filtering for distributed systems*. Springer, 1978 (cit. on pp. 102, 105).
- [Des+11] L Descroix et al. “Runoff evolution according to land use change in a small Sahelian catchment.” In: *Hydrology and Earth System Sciences Discussions* 8.1 (2011), pp. 1569–1607 (cit. on p. 78).
- [Des+12] Luc Descroix et al. “Experimental evidence of deep infiltration under sandy flats and gullies in the Sahel.” In: *Journal of Hydrology* 424 (2012), pp. 1–15 (cit. on p. 78).
- [DGB04] Jean-François Dulhoste, Didier Georges, and Gildas Besançon. “Nonlinear control of open-channel water flow based on collocation control model.” In: *Journal of Hydraulic Engineering* 130.3 (2004), pp. 254–266 (cit. on p. 12).
- [DJW04] Yan Ding, Yafei Jia, and Sam SY Wang. “Identification of Manning’s roughness coefficients in shallow water flows.” In: *Journal of Hydraulic Engineering* 130.6 (2004), pp. 501–510 (cit. on p. 10).
- [DL91] SK Das and RW Lardner. “On the estimation of parameters of hydraulic models by assimilation of periodic tidal data.” In: *Journal of Geophysical Research: Oceans* 96.C8 (1991), pp. 15187–15196 (cit. on p. 12).
- [DW05] Yan Ding and Sam SY Wang. “Identification of Manning’s roughness coefficients in channel network using adjoint analysis.” In: *International Journal of Computational Fluid Dynamics* 19.1 (2005), pp. 3–13 (cit. on p. 10).
- [DW06] Yan Ding and Sam SY Wang. “Optimal control of open-channel flow using adjoint sensitivity analysis.” In: *Journal of Hydraulic Engineering* 132.11 (2006), pp. 1215–1228 (cit. on p. 12).
- [DW12a] Yan Ding and Sam SY Wang. “Optimal control of flood diversion in watershed using nonlinear optimization.” In: *Advances in Water Resources* 44 (2012), pp. 30–48 (cit. on p. 12).

- [DW12b] Yan Ding and Sam SY Wang. “Optimal control of flood water with sediment transport in alluvial channel.” In: *Separation and purification technology* 84 (2012), pp. 85–94 (cit. on p. 12).
- [EL88] Richard E Ewing and Tao Lin. “A direct method for parameter estimation in a hyperbolic partial differential equation.” In: *Decision and Control, 1988., Proceedings of the 27th IEEE Conference on.* IEEE. 1988, pp. 1662–1667 (cit. on p. 10).
- [EWD03] M Egerstedt, Y Wardi, and F Delmotte. “Optimal control of switching times in switched dynamical systems.” In: *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on.* Vol. 3. IEEE. 2003, pp. 2138–2143 (cit. on p. 11).
- [Fer04] James Ferguson. “A Brief Survey of the History of the Calculus of Variations and its Applications.” In: *arXiv preprint math/0402357* (2004) (cit. on pp. 11, 17).
- [FNE95] TD Fadale, AV Nenarokomov, and Ashley F Emery. “Two approaches to optimal sensor locations.” In: *Journal of Heat Transfer* 117.2 (1995), pp. 373–379 (cit. on p. 102).
- [GA11] W Heber Green and GA Ampt. “Studies on Soil Physics.” In: *The Journal of Agricultural Science* 4.01 (1911), pp. 1–24 (cit. on p. 63).
- [GC06] I Yu Gejadze and GJM Copeland. “Open boundary control for Navier-Stokes equations including the free surface: adjoint sensitivity analysis.” In: *Computers and Mathematics with Applications* 52.8-9 (2006), pp. 1243–1268 (cit. on p. 12).
- [GCM+35] BD Greenshields, Ws Channing, Hh Miller, et al. “A study of traffic capacity.” In: *Highway research board proceedings.* Vol. 1935. National Research Council (USA), Highway Research Board. 1935 (cit. on p. 48).
- [Geo08] Didier Georges. “Infinite-dimensional nonlinear predictive controller design for open-channel hydraulic systems.” In: *Workshop on Irrigation Channels and Related Problems.* Vol. 4. 2008, pp. 267–285 (cit. on p. 12).
- [Geo09] Didier Georges. “Infinite-dimensional nonlinear predictive control design for open-channel hydraulic systems.” In: *Networks and Heterogeneous Media* 4.2 (June 2009), pp. 267–285 (cit. on p. 10).
- [Geo95] Didier Georges. “The use of observability and controllability gramians or functions for optimal sensor and actuator location in finite-dimensional systems.” In: *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on.* Vol. 4. IEEE. 1995, pp. 3319–3324 (cit. on p. 103).

- [GM16] Igor Gejadze and Pierre-Olivier Malaterre. “Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model.” In: *International Journal for Numerical Methods in Fluids* (2016) (cit. on p. 12).
- [Gol12] Herman Heine Goldstine. *A History of the Calculus of Variations from the 17th through the 19th Century*. Vol. 5. Springer Science & Business Media, 2012 (cit. on p. 17).
- [Had02] Jacques Hadamard. “Sur les problèmes aux dérivées partielles et leur signification physique.” In: *Princeton university bulletin* 13.49-52 (1902), p. 28 (cit. on p. 12).
- [HI14] Agus Hasan and Lars Imsland. “Moving horizon estimation in managed pressure drilling using distributed models.” In: *2014 IEEE Conference on Control Applications (CCA)*. IEEE. 2014, pp. 605–610 (cit. on p. 11).
- [HK+10] Bernd Hofmann, Stefan Kindermann, et al. “On the degree of ill-posedness for linear problems with noncompact operators.” In: *Methods and Applications of Analysis* 17.4 (2010), pp. 445–462 (cit. on p. 13).
- [Hof94] Bernd Hofmann. “On the degree of ill-posedness for nonlinear problems.” In: *Journal of Inverse and Ill-Posed Problems* 2.1 (1994), pp. 61–76 (cit. on p. 13).
- [HW05] Bernd Hofmann and Lothar von Wolfersdorf. “Some results and a conjecture on the degree of ill-posedness for integration operators with weights.” In: *Inverse Problems* 21.2 (2005), p. 427 (cit. on p. 13).
- [JDWK05] Denis Jacquet, Carlos Canudas De Wit, and Damien Koenig. “Traffic control and monitoring with a macroscopic model in the presence of strong congestion waves.” In: *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC’05. 44th IEEE Conference on*. IEEE. 2005, pp. 2164–2169 (cit. on p. 12).
- [JOB91] Y Jarny, MN Ozisik, and JP Bardon. “A general optimization method using adjoint equation for solving multidimensional inverse heat conduction.” In: *International journal of heat and mass transfer* 34.11 (1991), pp. 2911–2919 (cit. on p. 13).
- [Kal60] Rudolph Emil Kalman. “A new approach to linear filtering and prediction problems.” In: *Journal of basic Engineering* 82.1 (1960), pp. 35–45 (cit. on p. 102).
- [KMS15] Tawsif Khan, Kirsten Morris, and Marek Stastna. “Computation of the optimal sensor location for the estimation of an 1-D linear dispersive wave equation.” In: *2015 American Control Conference (ACC)*. IEEE. 2015, pp. 5270–5275 (cit. on p. 102).

- [KS78] Sudarshan Kumar and John H Seinfeld. “Optimal location of measurements for distributed parameter estimation.” In: *Automatic Control, IEEE Transactions on* 23.4 (1978), pp. 690–698 (cit. on p. 102).
- [LA90] Daniel A Lashof and Dilip R Ahuja. “Relative contributions of greenhouse gas emissions to global warming.” In: (1990) (cit. on p. 2).
- [LCC14] Yanning Li, Edward Canepa, and Christian Claudel. “Optimal traffic control in highway transportation networks using linear programming.” In: *Control Conference (ECC), 2014 European*. IEEE. 2014, pp. 2880–2887 (cit. on p. 10).
- [LFB92] Fubo Liu, Jan Feyen, and Jean Berlamont. “Computation method for regulating unsteady flow in open channels.” In: *Journal of irrigation and drainage engineering* 118.5 (1992), pp. 674–689 (cit. on p. 30).
- [LG99a] X Litrico and D Georges. “Robust continuous-time and discrete-time flow control of a dam–river system.(I) Modelling.” In: *Applied mathematical modelling* 23.11 (1999), pp. 809–827 (cit. on p. 10).
- [LG99b] Xavier Litrico and Didier Georges. “Robust continuous-time and discrete-time flow control of a dam–river system.(II) Controller design.” In: *Applied Mathematical Modelling* 23.11 (1999), pp. 829–846 (cit. on p. 10).
- [LMG02] Sanjay Lall, Jerrold E Marsden, and Sonja Glavaški. “A subspace approach to balanced truncation for model reduction of nonlinear control systems.” In: *International journal of robust and nonlinear control* 12.6 (2002), pp. 519–535 (cit. on pp. 106, 107).
- [Lon08] Han Longxi. “Parameter estimation in channel network flow simulation.” In: *Water Science and Engineering* 1.1 (2008), pp. 10–17 (cit. on p. 10).
- [LW55] Michael J Lighthill and Gerald Beresford Whitham. “On kinematic waves. II. A theory of traffic flow on long crowded roads.” In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 229.1178 (1955), pp. 317–345 (cit. on p. 47).
- [LW60] Peter Lax and Burton Wendroff. “Systems of conservation laws.” In: *Communications on Pure and Applied mathematics* 13.2 (1960), pp. 217–237 (cit. on p. 31).
- [Mal+09] Pierre-Olivier Malaterre et al. “Data assimilation for real-time estimation of hydraulic states and unmeasured perturbations in a 1D hydrodynamic model. Application to water management problems and comparison of Kalman filter and sequential Monte Carlo approaches.” In: *3rd Mamern International Conference on Approximation Methods and Numerical Modelling in Environment and Natural Resources*. Pau, France: CNRS, June 2009, pp. 629–634 (cit. on p. 10).

- [Mat+10] PA Matson et al. *Advancing the Science of Climate Change*. The National Academy of Sciences, 2010 (cit. on p. 1).
- [May10] Larry W Mays. *Water resources engineering*. John Wiley & Sons, 2010 (cit. on p. 62).
- [MB06] Sonia MartíNez and Francesco Bullo. “Optimal sensor placement and motion coordination for target tracking.” In: *Automatica* 42.4 (2006), pp. 661–668 (cit. on p. 103).
- [MCEP09] Rafael Muñoz-Carpena and John E. Parsons. *WinGAmpt document*. 2009. URL: <http://abe.ufl.edu/carpena/files/pdf/software/WinGAmpt.pdf> (cit. on p. 82).
- [ML73] Russell G Mein and Curtis L Larson. “Modeling infiltration during a steady rain.” In: *Water resources research* 9.2 (1973), pp. 384–394 (cit. on p. 63).
- [Mou+11] Ibrahim Bouzou Moussa et al. “Les changements d’usage des sols et leurs conséquences hydrogéomorphologiques sur un bassin-versant endoréique sahélien.” In: *Science et changements planétaires/Sécheresse* 22.1 (2011), pp. 13–24 (cit. on p. 78).
- [MR97] Pierre-Olivier Malaterre and J Rodellar. “Multivariable predictive control of irrigation canals. Design and evaluation on a 2-pool model.” In: *International Workshop on Regulation of Irrigation Canals, Marroco*. 1997, pp. 230–238 (cit. on p. 10).
- [MW72] PC Müller and HI Weber. “Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems.” In: *Automatica* 8.3 (1972), pp. 237–246 (cit. on pp. 103, 108).
- [MW90] James J MacKenzie and Michael P Walsh. “Driving forces: Motor vehicle trends and their implications for global warming, energy strategies, and transportation planning.” In: (1990) (cit. on p. 2).
- [Nah+03] HB Nahor et al. “Optimization of the temperature sensor position in a hot wire probe set up for estimation of the thermal properties of foods using optimal experimental design.” In: *Journal of food engineering* 57.1 (2003), pp. 103–110 (cit. on p. 102).
- [Nav98] IM Navon. “Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography.” In: *Dynamics of Atmospheres and Oceans* 27.1 (1998), pp. 55–79 (cit. on p. 12).
- [NGB14] Van Tri Nguyen, Didier Georges, and Gildas Besançon. “Optimal state estimation in an overland flow model using the adjoint method.” In: *Control Applications (CCA), 2014 IEEE Conference on*. IEEE. Antibes, France, 2014, pp. 2034–2039 (cit. on pp. 38, 64).

- [NGB15a] Van Tri Nguyen, D. Georges, and G. Besançon. “Adjoint-method-based estimation of Manning roughness coefficient in an overland flow model.” In: *American Control Conference (ACC), 2015*. Chicago, United States, July 2015, pp. 1977–1982 (cit. on pp. 38, 70).
- [NGB15b] Van Tri Nguyen, Didier Georges, and Gildas Besançon. “Traffic flow parameter estimation based on an adjoint method.” In: *Control Conference (ECC), 2015 European*. IEEE. Linz, Austria, July 2015, pp. 1333–1338 (cit. on pp. 38, 49).
- [NGB16a] Van Tri Nguyen, Didier Georges, and Gildas Besançon. “Optimal sensor location for overland flow network monitoring.” In: *3rd International Conference on Control and Fault-Tolerant Systems, 2016*. IEEE. Barcelona, Spain, Sept. 2016 (cit. on pp. 115, 134).
- [NGB16b] Van Tri Nguyen, Didier Georges, and Gildas Besançon. “State and parameter estimation in 1-D hyperbolic PDEs based on an adjoint method.” In: *Automatica* 67 (2016), pp. 185–191 (cit. on pp. 20, 54, 72).
- [Ngu+15] Van Tri Nguyen et al. “Application of adjoint method for estimating Manning-Strickler coefficient in Tondi Kiboro catchment.” In: *Control Applications (CCA), 2015 IEEE Conference on*. Sydney, Australia, Sept. 2015, pp. 551–556 (cit. on p. 78).
- [Ngu+16] Van Tri Nguyen et al. “Parameter estimation of a real hydrological system using an adjoint method.” In: *Adaptation and Learning in Control and Signal Processing (ALCOSP), 2016 IFAC International Workshop on*. Eindhoven, The Netherlands, July 2016 (cit. on p. 78).
- [OKS78] Sigeru Omatu, Satoru Koide, and Takasi Soeda. “Optimal sensor location problem for a linear distributed parameter system.” In: *Automatic Control, IEEE Transactions on* 23.4 (1978), pp. 665–673 (cit. on p. 102).
- [Pac+14] Rajendra K Pachauri et al. *Climate change 2014: synthesis Report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*. IPCC, 2014 (cit. on p. 1).
- [PAG08] L Perez, L Autrique, and M Gillet. “Implementation of a conjugate gradient algorithm for thermal diffusivity identification in a moving boundaries system.” In: *Journal of Physics: Conference Series*. Vol. 135. 1. IOP Publishing. 2008, p. 012082 (cit. on p. 12).
- [Pap98] Markos Papageorgiou. “Some remarks on macroscopic traffic flow modelling.” In: *Transportation Research Part A: Policy and Practice* 32.5 (1998), pp. 323–329 (cit. on p. 47).
- [Par+97] Anthony J Parsons et al. “Distributed dynamic modelling of interrill overland flow.” In: *Hydrological Processes* 11.14 (1997), pp. 1833–1859 (cit. on p. 62).

- [PGB12] Thang V Pham, Didier Georges, and Gildas Besançon. “Receding horizon boundary control of nonlinear conservation laws with shock avoidance.” In: *Automatica* 48.9 (2012), pp. 2244–2251 (cit. on p. 10).
- [PGW14] Christophe Prieur, Antoine Girard, and Emmanuel Witrant. “Stability of switched linear hyperbolic systems by Lyapunov techniques.” In: *IEEE Transactions on Automatic Control* 59.8 (2014), pp. 2196–2202 (cit. on p. 11).
- [Pie+02] Roger A Pielke et al. “The influence of land-use change and landscape dynamics on the climate system: relevance to climate-change policy beyond the radiative effect of greenhouse gases.” In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 360.1797 (2002), pp. 1705–1719 (cit. on p. 1).
- [PK97a] Michael Piasecki and Nikolaos D Katopodes. “Control of contaminant releases in rivers. I: Adjoint sensitivity analysis.” In: *Journal of hydraulic engineering* 123.6 (1997), pp. 486–492 (cit. on p. 11).
- [PK97b] Michael Piasecki and Nikolaos D Katopodes. “Control of contaminant releases in rivers. II: Optimal design.” In: *Journal of Hydraulic Engineering* 123.6 (1997), pp. 493–503 (cit. on p. 11).
- [Pre61] A Preissmann. “Propagation des intumescences dans les canaux et rivières.” In: *First Congress of the French Association for Computation AFCAL* (1961), pp. 690–698 (cit. on p. 29).
- [Pui+05] Vicenç Puig et al. “Identification and control of an open-flow canal using LPV models.” In: *Proceedings of the 44th IEEE conference on decision and control*. IEEE. 2005, pp. 1893–1898 (cit. on p. 10).
- [RDK07] Michael Rinehart, Munther Dahleh, and Ilya Kolmanovsky. “Optimal control of switched homogeneous systems.” In: *2007 American Control Conference*. IEEE. 2007, pp. 1377–1382 (cit. on p. 11).
- [Rei+15] Jack Reilly et al. “Adjoint-Based Optimization on a Network of Discretized Scalar Conservation Laws with Applications to Coordinated Ramp Metering.” In: *Journal of Optimization Theory and Applications* 167.2 (2015), pp. 733–760 (cit. on p. 12).
- [Ric56] P.I Richards. “Shock waves on the highway.” In: *Operations Research* 4.1 (1956), pp. 42–51 (cit. on p. 47).
- [RT05] Robin L Raffard and Claire J Tomlin. “Second order adjoint-based optimization of ordinary and partial differential equations with application to air traffic flow.” In: *American Control Conference, 2005. Proceedings of the 2005*. IEEE. 2005, pp. 798–803 (cit. on p. 12).

- [Sco92] H Scoging. “Modelling overland-flow hydrology for dynamic hydraulics.” In: *in Parsons, A. J. and Abrahams, A. D. (Eds), Overland Flow Hydraulics and Erosion Mechanics* (1992), pp. 89–103 (cit. on p. 62).
- [SH05] Abhay K Singh and Juergen Hahn. “Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems.” In: *Industrial & engineering chemistry research* 44.15 (2005), pp. 5645–5659 (cit. on pp. 103, 107).
- [SK11] Mohamed T Shamaa and Hmida M Karkuri. “Implicit numerical scheme for regulating unsteady flow in open channel.” In: *Fifteenth International Water Technology Conference-IWTC-15*. Alexandria, Egypt, 2011 (cit. on p. 30).
- [SK82] RW Skaggs and R Khaleel. “Infiltration.” In: *Hydrologic modeling of small watersheds. ASAE Monogr* 5 (1982), pp. 121–166 (cit. on p. 82).
- [SK99a] Brett F Sanders and Nikolaos D Katopodes. “Active flood hazard mitigation. I: Bidirectional wave control.” In: *Journal of hydraulic engineering* 125.10 (1999), pp. 1057–1070 (cit. on p. 12).
- [SK99b] Brett F Sanders and Nikolaos D Katopodes. “Control of canal flow by adjoint sensitivity method.” In: *Journal of irrigation and drainage engineering* 125.5 (1999), pp. 287–297 (cit. on p. 12).
- [SM03] Endre Süli and David F Mayers. *An introduction to numerical analysis*. Cambridge university press, 2003 (cit. on p. 27).
- [SMB10] Jacques Sau, Pierre-Olivier Malaterre, and Jean-Pierre Baume. “Sequential Monte Carlo hydraulic state estimation of an irrigation canal.” In: *Comptes Rendus Mécanique* 338.4 (2010), pp. 212–219 (cit. on p. 10).
- [Sou08] YERO K Souley. *L’évolution de l’occupation des sols à l’échelle des bassins versants de Wankama et Tondi Kiboro: Conséquences sur les débits*. 2008 (cit. on p. 78).
- [SW05] Xubin Sun and Hong Wang. “Dynamic system control of gas jet flames distribution by predictive control method.” In: *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005*. IEEE. 2005, pp. 1458–1463 (cit. on p. 10).
- [Tri+08] Christophe Tricaud et al. “D-optimal trajectory design of heterogeneous mobile sensors for parameter estimation of distributed systems.” In: *American Control Conference, 2008*. IEEE. 2008, pp. 663–668 (cit. on p. 102).
- [Uci00] Dariusz Ucinski. “Optimal sensor location for parameter estimation of distributed processes.” In: *International Journal of Control* 73.13 (2000), pp. 1235–1248 (cit. on pp. 102, 108).

- [Uci12] Dariusz Uciński. “Sensor network scheduling for identification of spatially distributed processes.” In: *International Journal of Applied Mathematics and Computer Science* 22.1 (2012), pp. 25–40 (cit. on p. 102).
- [Uci99] Dariusz Uciński. *Measurement optimization for parameter estimation in distributed systems*. Technical University Press Zielona Góra, 1999 (cit. on p. 102).
- [UP07] Dariusz Uciński and Maciej Patan. “D-optimal design of a monitoring network for parameter estimation of distributed systems.” In: *Journal of Global Optimization* 39.2 (2007), pp. 291–322 (cit. on p. 102).
- [Wal+98] Walter Waldraff et al. “On the use of observability measures for sensor location in tubular reactor.” In: *Journal of Process Control* 8.5 (1998), pp. 497–505 (cit. on p. 103).
- [WB08] Daniel B Work and Alexandre M Bayen. “Convex formulations of air traffic flow optimization problems.” In: *Proceedings of the IEEE* 96.12 (2008), pp. 2096–2112 (cit. on p. 12).
- [WP97] Eric Walter and Luc Pronzato. *Identification of parametric models from experimental data*. Springer Verlag, 1997 (cit. on p. 110).
- [XA02] Xuping Xu and PJ Antsaklis. “Optimal control of switched autonomous systems.” In: *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*. Vol. 4. IEEE. 2002, pp. 4401–4406 (cit. on p. 11).
- [YI15] Keisuke Yoshida and Tadaharu Ishikawa. “Flood hydrograph estimation using an adjoint shallow-water model.” In: *Journal of Hydro-environment Research* 9.3 (2015), pp. 429–440 (cit. on p. 12).
- [YJC14] Hao Yang, Bin Jiang, and Vincent Cocquempot. “A survey of results and perspectives on stabilization of switched nonlinear systems with unstable modes.” In: *Nonlinear Analysis: Hybrid Systems* 13 (2014), pp. 45–60 (cit. on p. 11).
- [YM14] Keisuke Yoshida and Shiro Maeno. “Inverse estimation of distributed roughness coefficients in vegetated flooded rivers.” In: *Journal of Hydraulic Research* 52.6 (2014), pp. 811–823 (cit. on p. 12).
- [Yu99] Wenhuan Yu. “A quasi-Newton method for estimating the parameter in a nonlinear hyperbolic system.” In: *Journal of mathematical analysis and applications* 231.2 (1999), pp. 397–424 (cit. on p. 10).
- [ZA15] Feng Zhu and Panos J Antsaklis. “Optimal control of hybrid switched systems: A brief survey.” In: *Discrete Event Dynamic Systems* 25.3 (2015), pp. 345–364 (cit. on p. 11).
- [ZC11] Xiao Zhang and Ximing Cai. “Climate change impacts on global agricultural land availability.” In: *Environmental Research Letters* 6.1 (2011), p. 014014 (cit. on p. 1).

Approche à base d'adjoint pour estimation et placement de capteurs dans les systèmes hyperboliques 1D avec application à l'hydrologie et au trafic routier

Résumé — Ce travail de thèse propose une approche générique pour l'estimation de l'état/ des paramètres et pour le placement de capteurs de systèmes hyperboliques non linéaires en dimension infinie. Le travail est donc divisé en deux parties principales : une partie consacrée à l'estimation optimale et une partie dédiée au placement optimal de capteurs. La méthode d'estimation optimale utilise une approche par calcul des variations et utilise la méthode des multiplicateurs de Lagrange. Ces multiplicateurs jouent un rôle important en donnant accès aux sensibilités des mesures par rapport aux variables qui doivent être estimées. Ces sensibilités, décrites par les équations adjointes, sont aussi à l'origine d'une nouvelle approche, dite méthode de l'adjoint, pour le placement optimal de capteurs. Divers exemples, construits sur la base de simulations mais également de données réelles et pour différents scénarios, sont aussi étudiés afin d'illustrer l'efficacité des approches développées. Ces exemples concernent les écoulements à surface libre (en hydrologie des bassins versants) et le trafic routier représentés par des équations aux dérivées partielles hyperboliques non linéaires.

Mots clés : Estimation optimale d'état et de paramètres, Placement optimal de capteurs, Calcul des variations, Systèmes hyperboliques non linéaires; Écoulements de surface, Trafic routier

Adjoint-based approach for estimation and sensor location on 1D hyperbolic systems with applications in hydrology and traffic

Abstract — The thesis proposes a general framework for both state/parameters estimation and sensor placement in nonlinear infinite dimensional hyperbolic systems. The work is therefore divided into two main parts: a first part devoted to the optimal estimation and a second one to optimal sensor location. The estimation method is based on the calculus of variations and the use of Lagrange multipliers. The Lagrange multipliers play an important role in giving access to the sensitivities of the measurements with respect to the variables to be estimated. These sensitivities, described by the adjoint equations, are also the key idea of a new approach, so-called the adjoint-based approach, for the optimal sensor placement. Various examples, either based on some simulations with synthetic measurements or real data sets and for different scenarios, are also studied to illustrate the effectiveness of the developed approaches. These examples concern the overland flow systems and the traffic flow, which are both governed by nonlinear hyperbolic partial differential equations.

Keywords: Optimal state and parameter estimation, Optimal sensors placement, Calculus of variations, nonlinear hyperbolic systems, Overland flows, Traffic flow
