Modeling historical floods of the Saar

Sleimane Hariri

▶ **To cite this version:**

Sleimane Hariri. Modeling historical floods of the Saar. Fluid mechanics [physics.class-ph]. Université de Strasbourg, 2021. English. NNT : 2021STRAD028 . tel-03850417

HAL Id: tel-03850417
https://theses.hal.science/tel-03850417v1

Submitted on 13 Nov 2022

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.

UNIVERSITÉ DE STRASBOURG

École doctorale 269 Mathématiques, Sciences de l'Information et de l'Ingénieur

Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie

THÈSE présentée par :

Sleimane Hariri

soutenance prévue le : **10 novembre 2021**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/Spécialité : Mécanique des fluides

Modélisation des inondations historiques de la Sarre

THESE dirigée par :

CHARPENTIER Isabelle	Directrice de recherche, CNRS

RAPPORTEURS :

ANQUETIN Sandrine	Directrice de recherche, CNRS
VIALLE Stéphane	Professeur, CentraleSupélec
GUSTEDT Jens	Directeur de recherche, INRIA
JEANNOT Benjamin	Docteur, Ministère de la Transition écologique
PONS Marie-Noelle	Directrice de recherche, CNRS

Contents

List of Figures

List of Tables

Listings

Introduction

Ces dernières années, les cycles hydrologiques sont perturbés par le dérèglement climatique qui se traduit par des événements naturels extrêmes plus fréquents (La Jeunesse, 2016), (Bergholt and Lujala, 2012), (Huang et al., 2017). Selon la base de données des événements d'urgence (EM-DAT), les catastrophes naturelles liées au climat sont passées de 90 événements entre 1970 et 1989 à plus de 300 événements après 2000. Parmi ces catastrophes naturelles, les inondations sont considérées comme les plus fréquentes affectant les populations au niveau socio-économique et causant plus de décès humains que tout autre phénomène naturel (Berz, 2000).

L'étude des événements de crue nécessite une connaissance du comportement hydrologique de la zone d'étude. Le développement des modèles numériques a permis une meilleure compréhension des processus hydrologiques. Ces modèles expriment les différents processus hydrologiques et la relation entre eux sous forme d'équations et sont largement utilisés dans la gestion des ressources en eau, la prévision des inondations et l'évaluation des risques (Fan et al., 2016). A partir de 1980, l'utilisation des équations aux dérivées partielles (EDP) non linéaires et de paramétrisations physiques ont permis de représenter les processus hydrologiques complexes d'une manière de plus en plus complète (Rodriguez et al., 2008). Cette complexité requiert une résolution itérative d'autant plus coûteuse en temps de calcul et en mémoire que le bassin versant modélisé est grand.

L'objectif de cette thèse est de modéliser les événements de crues extrêmes dans les grands bassins versants. Le défi dans ces modélisations est de considérer toute la surface du bassin versant comme domaine de calcul, ce qui suppose de prendre en compte des données de précipitation. Ceci permet un diagnostic de l'événement modélisé sur toute la surface du bassin versant concerné et apporte une plus-value à la simulation. Ces simulations nécessitent de très longs temps de calcul et des outils puissants. Par opposition, la majorité des études de crues ne considèrent que le lit majeur et utilisent des données de débit pour modéliser l'événement pluvial (David and Schmalz, 2021). La modélisation des événements de crue en considérant les données de précipitation comme données d'entrée sur le bassin de la Sarre ($1800\,km^2$) et en utilisant un maillage de 25 m, (qui est bien adapté à la résolution de 25 m de notre MNT), impliquant environ 2 800 000 éléments. HEC-RAS, qui est l'un des modèles hydrauliques les plus utilisés et les plus connus, a échoué à exécuter cette simulation en raison d'un dépassement de la capacité de mémoire disponible. Cette thèse propose des méthodes de modélisations spatiales et décrit des outils logiciels conçus pour permettre la réalisation de ce type de simulation sur des grand bassins versants et améliorer la performance de ces simulations.

Comme illustration, le cas d'étude choisi est le district hydrographique Moselle-Sarre-Nieds qui a connu plusieurs crues d'importance significative de décembre 1919 jusqu'à nos jours. Le projet PRIM'Eau auquel je suis associé s'intéresse à la modélisation du bassin de la Sarre pour les aspects hydrologiques très particuliers, notamment la mise en place d'un dispositif de plusieurs barrages et étangs pour initier des crues défensives pendant la seconde guerre mondiale. Plusieurs personnes impliquées dans le projet PRIM'Eau ont travaillé sur d'autres thématiques relatives au bassin du Moderbach (sous-bassin de la Sarre) telles que la qualité de l'eau, le socio-écosystème et la compréhension par le public des cartes de risques.

La simulation de grands domaines a toujours été conditionnée par les performances de calcul. Ces 20 dernières années, dans de nombreux domaines en sciences et en ingénierie, l'utilisation du calcul à haute performance s'est considérablement accrue en raison du développement de nouvelles architectures et de la diminution de leurs coûts, facilitant l'accès à ces super-calculateurs (Kumar and Duffy, 2015). Les méthodes de décomposition de domaines et de calcul parallèle ont alors permis de résoudre des problèmes plus grands avec des temps de calculs raisonnables, lorsqu'elles ont été combinées dans des algorithmes efficaces. Dans ce contexte, cette thèse présente une méthode de partition des bassins versants pour distribuer la charge de calcul de façon équilibrée. L'implémentation de la méthode de partition dans la modélisation hydrologique est présentée pour deux modèles utilisant des plateformes de calcul différentes, que ce soit en parallèle ou en séquentiel. Les résultats hydrologiques des deux modèles sont comparés entre eux et avec des données observées. Les performances de calcul des deux stratégies utilisées dans ce travail de recherche sont comparées pour évaluer leur efficacité. Ce travail développe une nouvelle méthode hybride pour corriger les dépressions réelles ou trompeuses des modèles numériques de terrain.

Plus précisément, ***le chapitre 1*** rappelle brièvement quelques notions de base de la modélisation hydrologique. Les différentes composantes du cycle de l'eau, les interactions entre elles, la notion de bassin versant et ses caractéristiques sont décrites. Ensuite, le bassin versant de la Sarre est présenté et les données disponibles extraites des différentes stations de mesure sont analysées. Une étude bibliographique des crues historiques et l'identification de leurs typologies ont permis de mieux comprendre le comportement hydrologique du bassin. ***Le chapitre 2*** se concentre sur les outils de modélisation de ces événements historiques dans les grands bassins comme celui de la Sarre. Les différents types de modèles hydrologiques sont présentés, justifiant notre choix d'un modèle physique distribué pour aboutir à l'objectif de l'étude. Le chapitre aborde la prise en compte de la dimension spatiale dans la modélisation hydrologique, notamment pour les modèles distribués, et discute l'intérêt de l'implantation des Systèmes d'Information Géographique comme support des modèles hydrologiques.

Ces éléments de contexte posés, ***le chapitre 3*** aborde et compare les méthodes de partition spatiale des bassins versants qui sont développées à l'aide d'outils SIG tels que ArcGIS et Topotoolbox (Schwanghart and Kuhn, 2010). Les méthodes les plus utilisées dans la littérature sont présentées et appliquées pour partitionner le bassin versant du Moderbach afin de déterminer leurs limites. Ce chapitre présente une nouvelle méthode de partition qui permet de distribuer la charge de calcul d'une manière équilibrée et de prendre en compte les particularités du terrain comme les structures hydrauliques ou les stations de mesure. ***Le chapitre 4*** présente l'intérêt de l'implémentation des méthodes de partition de bassin dans la modélisation hydrologique. Deux décompositions différentes du bassin du Moderbach, chacune considérant un nombre bien défini de sous-bassins, sont utilisées pour exécuter des simulations 2D sur leurs sous-bassins en parallèle avec le modèle 2D (NIHM) Jeannot (2018) supporté par un outil parallèle (ORWL) (Clauss and Gustedt, 2010). La comparaison du temps de calcul requis pour la simulation de chaque décomposition permet de déterminer le gain en temps de calcul associé au nombre de sous-bassins d'une décomposition. Les résultats hydrologiques de ce modèle sont validés en les comparant avec HEC-RAS (Brunner, 2016), qui utilise quant à lui un solveur basé sur la décomposition du système linéaire et non pas du bassin versant. Ce chapitre présente également les méthodes classiques de correction des modèles numériques de terrain et détermine leurs limites sur le bassin de Moderbach. Une nouvelle méthode hybride est présentée pour surmonter les

limites des méthodes classiques et convient bien à notre cas d'étude.

Le chapitre 5 répond à l'objectif de ce projet de recherche visant à modéliser les événements historiques et extrêmes dans la région de la Sarre. Trois événements extrêmes sont présentés : l'inondation de la vallée de Moderbach due aux ouvrages de génie civil mis en œuvre pour cette raison en 1939, l'événement exceptionnel de décembre 2011-janvier 2012 montrant trois crues majeures en moins de deux mois dans le bassin de l'Albe et la crue extrême de 2010 dans le bassin de la Sarre. Afin de réduire le long temps de calcul nécessaire pour réaliser ces modélisations, ce chapitre présente une nouvelle stratégie pour le calcul des modèles 2D dans HEC-RAS en s'inspirant des méthodes proposées dans les chapitres 3 et 4. L'implémentation des méthodes de partition de bassin dans HEC-RAS permet de distribuer la charge de calcul par décomposition du bassin versant et tire profit de la décomposition du système linéaire proposée par HEC-RAS dans chaque sous-bassin. L'algorithme qui organise la simulation dans HEC-RAS est automatisé par un script VBA en utilisant l'outil HECRASController et un couplage de modèles 2D/1D. Enfin, cette stratégie est appliquée pour la simulation des événements historiques et extrêmes sur les bassins du Moderbach, de l'Albe et de la Sarre. Les résultats des simulations sont validés par une comparaison avec les cartes de crues et les débits observés.

Chapter 1

Physical environment of a watershed

Contents

French Abstract

Les inondations sont considérées comme la catastrophe naturelle la plus fréquente, susceptible d'affecter une grande partie de la population mondiale. Entre 1998 et 2002, l'Europe a subi plus de 100 inondations sévères, dont celles du Danube et de l'Elbe. Récemment, les fortes inondations du 14 et 15 juillet 2021 ont causé la mort d'au moins 220 personnes, notamment en Allemagne et en Belgique. Les conséquences des inondations expliquent les efforts entrepris pour analyser, comprendre et cartographier cet aléa et les enjeux (biens et personnes), afin de mettre en place des programmes de prévention sur les territoires concernés. La France s'est ainsi dotée de programmes de prévention (PPRI : Plans de Prévention des Risques d'inondation, PAPI : Programmes d'Actions de Prévention des Inondations, etc.), qui visent à gérer et réduire le risque d'inondation.

Le nord-est de la France, en particulier le district hydrographique Moselle-Sarre-Nieds, a connu des crues d'importance significative, notamment en 1919-1920, 1947, 1983, 1995, 1997, 2006 et 2010. Dans le cadre du projet PRIM'Eau, nous nous sommes intéressés à la modélisation des inondations historiques extrêmes survenues dans le bassin de la Sarre. Caractérisé par la dominance de sols argileux, le bassin de la Sarre a de plus été équipé comme une ligne d'inondations défensives dans le dispositif militaire de la ligne Maginot construit en anticipant une guerre avec le IIIème Reich.

Après un bref rappel des processus du cycle de l'eau (Section 1.2) et de la notion de bassin versant (Section 1.3), le bassin de la Sarre est présenté au regard des données disponibles et utiles pour une modélisation hydrologiques : topographie, sols, météorologie et descriptions d'inondations historiques extrêmes (Section 1.4).

1.1 Introduction

During the period from 1998 to 2002, Europe suffered from more than 100 severe floods, including the Danube and Elbe rivers in 2002. The floods caused 700 deaths and at least €25 billion in economic losses (EPRI, 2011). Recently, the heavy rainfall of 14 and 15 July 2021 caused the death of at least 220 people, especially in Germany and Belgium. These serious implications arising from floods, explain the efforts undertaken to analyse and understand this phenomenon in order to reduce the risk and the consequences. Most of the European countries have already set prevention programs and mitigation plans. These are intended to reduce the consequences of floods through the control of the location of goods and people and the characterization of the vulnerability of the existing situation. For example, France has efficient prevention programs (PPRI: Flood Risk Prevention Plan , PAPI: Flood Prevention Action Programmes, etc.), which aims to manage and reduce flooding risks.

In France, floods represent more than half of the natural risks faced by humans (Berz, 2000). One out of three French communes is affected by floods (including 300 large urban areas). For 160 000 km of rivers, an area of 22 000 km^2 is recognized as particularly flood-prone: two million local residents are concerned. The damage caused by floods represents about 80% of the cost of damage attributable to natural hazards, an average of €600 million annual costs calculated over the period 1980-2000 (Gresillon, 2004).

North-eastern France, in particular in the Moselle-Saar-Nieds river basin district, has experienced several floods of significant importance, in particular in December 1919 - January 1920, December 1947, April and May 1983, January 1995, February 1997, October 2006, December 2010 and July 2021. The flows measured on the Saar at the Sarreinsming station during the 1970, 1983 and 1993 floods were in the order of 500 m^3/s (Flavien and Laurent, 2014).

In this chapter, an overview of the basic concepts of hydrological modeling is presented. The first part addresses the water cycle, in particular the processes involved in the generation of flow, and the watershed which represents the scale at which hydrological models are mainly applied. The second part introduces the physical and anthropic characteristics of the case study area. We provide a detailed presentation of the Saar watershed, the geographical and geological context, the hydro-climatic data used, and a summary of the historical extreme flood events that have occurred.

1.2 Water cycle

The water cycle, also known as the hydrologic cycle or the hydrological cycle, describes the movement of water masses in the earth system between the atmosphere, the land surface, the watercourses, the subsurface, and the ocean (Ficchi, 2017). Water constantly travels from one reservoir to another, driven in an endless cycle powered by the sun and the geothermal fluxes (Gilli et al., 2008). Water exchange is handled through the physical processes of precipitation, evaporation, transpiration, runoff, infiltration, percolation, and open-channel flow.

As a result of solar radiation, the water evaporated from the ground, oceans and other water surfaces rise into the atmosphere. Due to the rise of a humid air mass, the evaporated water condenses into drops constituting clouds. Then, the evaporated water that has been stored in the clouds returns to the oceans and lands in the form of liquid or snowy precipitation. When precipitation reaches the ground, a part of it can be intercepted by vegetation. The water reaching then the soil surface can follow a more or less complex path depending on the nature of the soil. In permeable soils, part of the precipitation infiltrates into the subsurface. There may be percolation to deeper zones to assist recharge of groundwater reserves. Infiltration is almost negligible when the soil is impermeable (clays). Precipitation that is not intercepted by vegetation or infiltrated in the soil produces runoff, which forms rivers and streams. Evaporation from soil, streams, and plant transpiration completes the cycle (Musy and Laglaine, 1992).

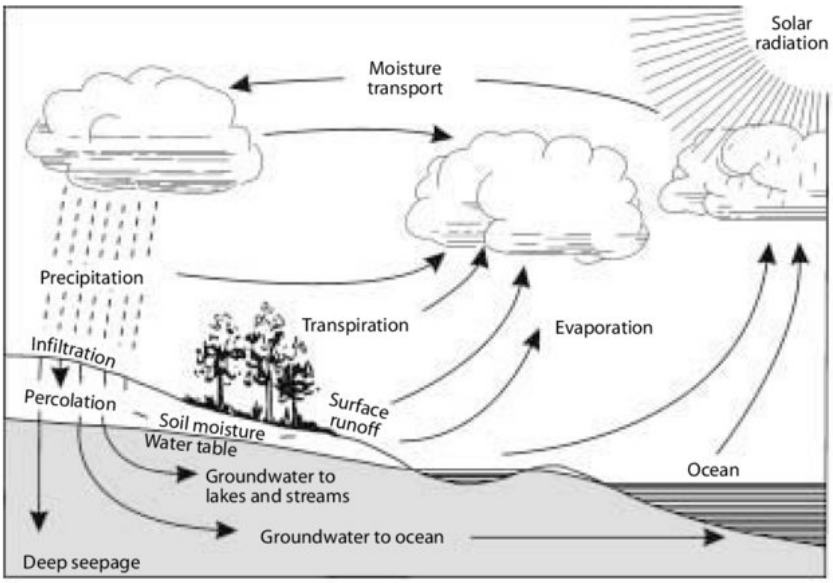


Figure 1.1: The water cycle (McVicar et al., 2002).

1.2.1 Water cycle processes

The water cycle is the result of a series of diverse and inter-related processes. In the following, we briefly present the concept of these processes.

- **Precipitation:** One of the major factors acting on the streamflow behavior is the precipitation amount falling in the watershed area, both in liquid form (rain) and in solid form (snow). They are caused by a change in temperature or pressure. Precipitation is expressed in intensity $\mathrm{mm/h}$ or in water depth mm.

- **Evaporation:** Evaporation refers to the transition of the water from liquid to vapor phase to reach the atmosphere. The evaporation of the oceans, enabling large-scale evaporation, is the main pathway for water to enter the atmosphere. Most of the water evaporated from the oceans ends up in the oceans as precipitation. Unlike precipitation or runoff, it is currently very difficult to directly measure the flow of water vapor in the atmosphere.

- **Transpiration:** Transpiration refers to the water evaporated through the vegetation. The roots of plants absorb water from the surrounding soil to escape into the atmosphere through the plant leaves. Plant transpiration is an invisible process. The amount of water that plants release varies greatly depending on many parameters such as temperature, wind motion and the type of plants.

- **Interception:** Most of the precipitation reaches the ground, but there is always a part that never reaches the ground. This is known as interception. The water is directly captured by the vegetation. The amount of water that can be intercepted varies considerably and depends on the type and surface of the vegetation cover.

- **Water storage:** Water stored in the form of ice, snow and glaciers is part of the water cycle. The National Snow and Ice Data Center states that if these glaciers, covering 10-11% of the Earth's surface, were to melt today, the seas would rise by about 70 m. At the surface level, water can be retained in soil depressions and hollows, ponds, lakes, reservoirs (man-made) during and after a rainfall.

- **Infiltration and Percolation:** Infiltration and percolation refer to the movement of water penetrating the soil, driven by gravity and the effects of pressure. The infiltration process describes the water that remains in the shallow soil layer, while percolation corresponds to deep infiltration allowing to recharge the groundwater reservoirs. The infiltration rate is given by the volume of water that infiltrates per unit of time ($\mathrm{mm/h}$ or $\mathrm{m^3/s}$).

- **Runoff:** Runoff is defined as the movement of water over the surface running down slopes into rivers, streams and lakes. Runoff is generally expressed as a volume/surface/time ratio ($\mathrm{m^3/m^2/t}$).

1.2.2 Water balance

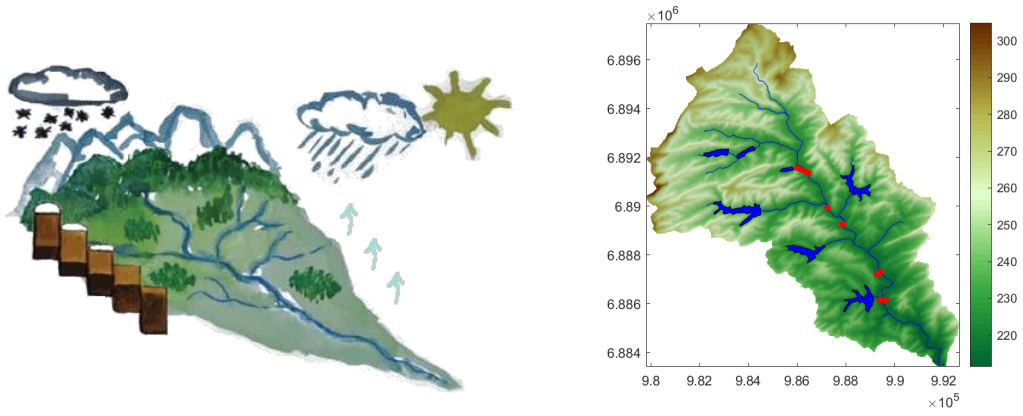
The water exchanges between the different components of the water cycle are irregular which can produce extremes events as floods and droughts. A hydrological system's water balance is developed by taking into account the different physical processes that can act on the state of the water (Henine, 2005). The mathematical model of the water balance can also be written as follow:

$$P - R - G - E - T = \Delta S, \tag{1.1}$$

where P is Precipitation, R is Runoff surface flow, G is Ground water flow, E is Evaporation, T is Transpiration and ΔS is Storage variation.

1.3 Drainage basin - Watershed

A watershed, also called a drainage basin, river basin, or catchment, is a closed surface drained by a river and its tributaries (Fig. 1.2). It is characterized by an outlet, which is assumed to be the only possible escape for runoff generated in the watershed area. A watershed represents the geographic environment for analyzing the various processes of the water cycle (Musy and Laglaine, 1992). The surface area of a watershed can vary from a few square kilometers (a small mountain stream) to a few million of square kilometers such as the Amazon basin, in the north of South America, the largest in the world with a surface area of over 7 million km^2.



(a) Artificial drainage basin (Seibert and Vis, 2012) (b) Moderbach basin

Figure 1.2: Examples of drainage basins.

For the sake of a more detailed description, it may be worthwhile to point out the difference between a topographic and a real drainage basin, see Fig.1.3. The topographic basin corresponds to a hydrologically closed area, where no external flow can enter the basin and any water flow originating within the basin can only flow through its outlet. It is delineated only on a topographical basis. When permeable soil overlies impermeable bedrock, interactions between the surface and subsoil can occur. The real basin corresponds to the drained area, considering both surface and subsoil, contributing to the water flow in the basin. Therefore, the basin boundary based on topography may no longer correspond to the real basin boundary.

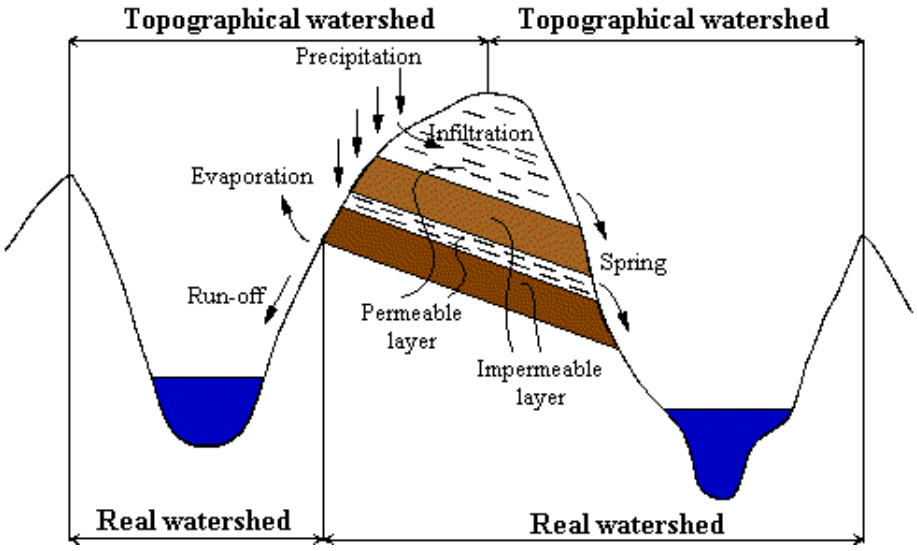


Figure 1.3: Difference between real and topographic watersheds (Roche, 1963).

1.3.1 Watershed characteristics

Several physiographic characteristics can act on the hydrological behavior of a watershed. Two watersheds of the same area may have different behavior if they are different in terms of shape, elevation, slope, vegetation, and soil type.

- **Shape:** The shape of a watershed has a major effect on its hydrological behavior (Fig.1.4). A stretched watershed has a long time of concentration (Tc) which helps to reduce the intensity of the peak flow once it reaches the outlet. A fan-shaped watershed has a shorter time of concentration and therefore the flow reaches the outlet rapidly with a high intensity.
- **Slope:** The slope is an important parameter in the watershed topography. It acts on the time needed for runoff to reach the outlet (concentration-time), the velocity of the flow, and the intensity of the peak discharge. Steep slopes enhance the runoff velocity while a gentle slope helps to retain water and favors infiltration.

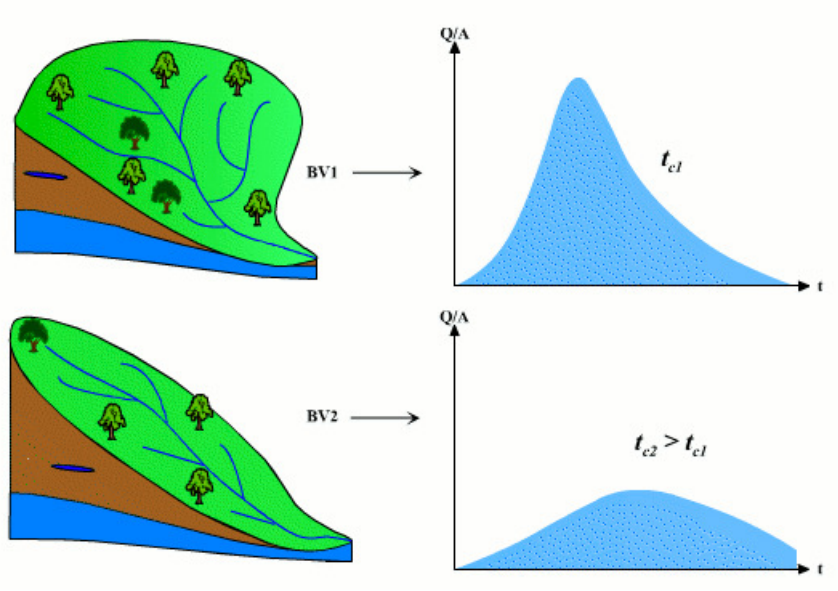


Figure 1.4: The effect of the watershed shape on its hydrological behavior (Q is the discharge and A is the watershed area) (Andre, 2001).

- **Land use land cover and vegetation:** Urban areas have a low permeability ratio. The infiltration process is negligible in these areas which enhances the runoff. Therefore, urbanizations play a major role in the vulnerability to flood risks. Vegetation cover plays a major role in reducing runoff. Depending on its type and density, it allows reducing runoff through interception by retaining water into the soil.
- **Soil type:** Some of the hydrological cycle processes depend on the soil type. Sandy soils, for example, have a high permeability rate, which favors the infiltration process and reduces runoff. Other soil types, such as clay soils, have very low permeability rates. In this case, infiltration is almost negligible, resulting in increased runoff and reduced concentration time.
- **Water bodies:** The placement of water bodies or storage areas, such as lakes and ponds, helps to reduce the intensity of a flood by temporarily storing a certain volume of water.
- **The hydrographic network:** The hydrographic network is the ensemble of streams and waterways in the watershed area. It is one of the major factors acting on the hydrological behavior of the watershed.

1.3.2 Hydrological behavior

The hydrological behavior of a watershed is characterized by its velocity (Vmax), its intensity (Qmax), and its rising time (Tm) (temps de montée) which is the time between the start of the precipitation event and the peak flow (Qmax). This should not be confused with the concentration-time (Tc), which is the longest time needed for a drop in the watershed to reach the outlet (Musy and Laglaine, 1992).

The analysis of the hydrological behavior of a watershed is based on its response when subjected to a stimulus. In this case, the stimulus is the rainfall and the response of the watershed corresponds to the outflow. This can be represented by a stage graph showing the precipitation data and a hydrograph showing the variation of the discharge at the outlet of the watershed.

The hydrological behavior of the Albe watershed (405 km^2) during the December 2011 - January 2012 flood event is illustrated in Fig.1.5. It shows the evolution of the discharge at the outlet regarding the precipitation data. The precipitations data are provided by the Kappelkinger station, located in the basin and available on the site of Meteo-France.The observed discharge is provided by the gauging station l'Albe a Sarralbe (Station A9192060) and available on the Banque Hydro website.

This event will be further investigated and modeled as a part of the extreme events presented in Chapter 5.

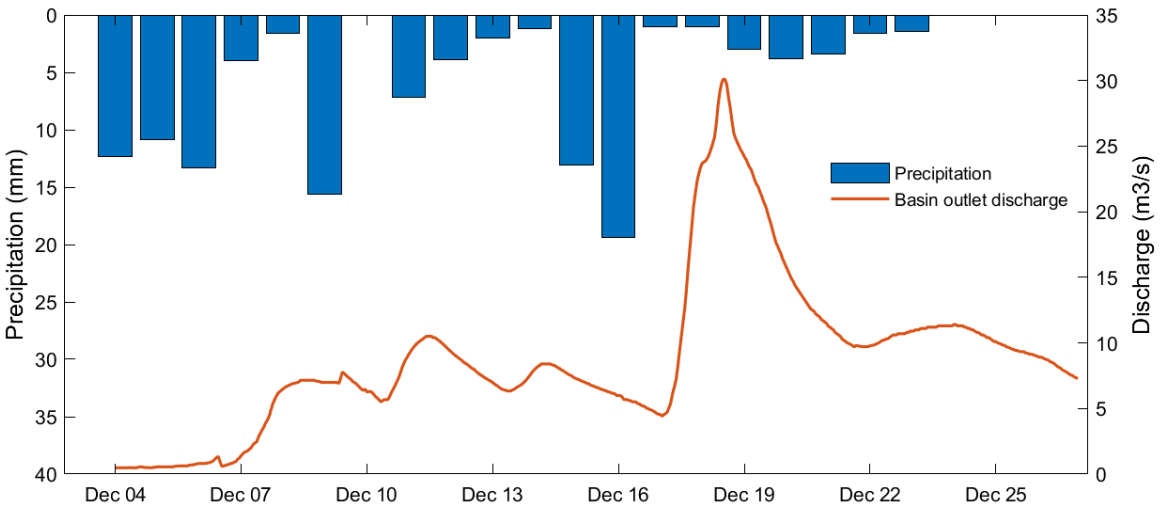


Figure 1.5: Hydrological behavior of the Albe basin during the December 2011 - January 2012 event.

1.4 Case Study

The Saar, a 246 km long river in north east of France/south west of Germany, is an important tributary of the Moselle river (Fig. 1.6). The Saar river takes its principal resources from the foot of the Vosges mountains at the Donon, it flows for about 126 km in the French region, then it flows along the border between France and Germany for about 11 km before it penetrates entirely on the German territory for 117 km.

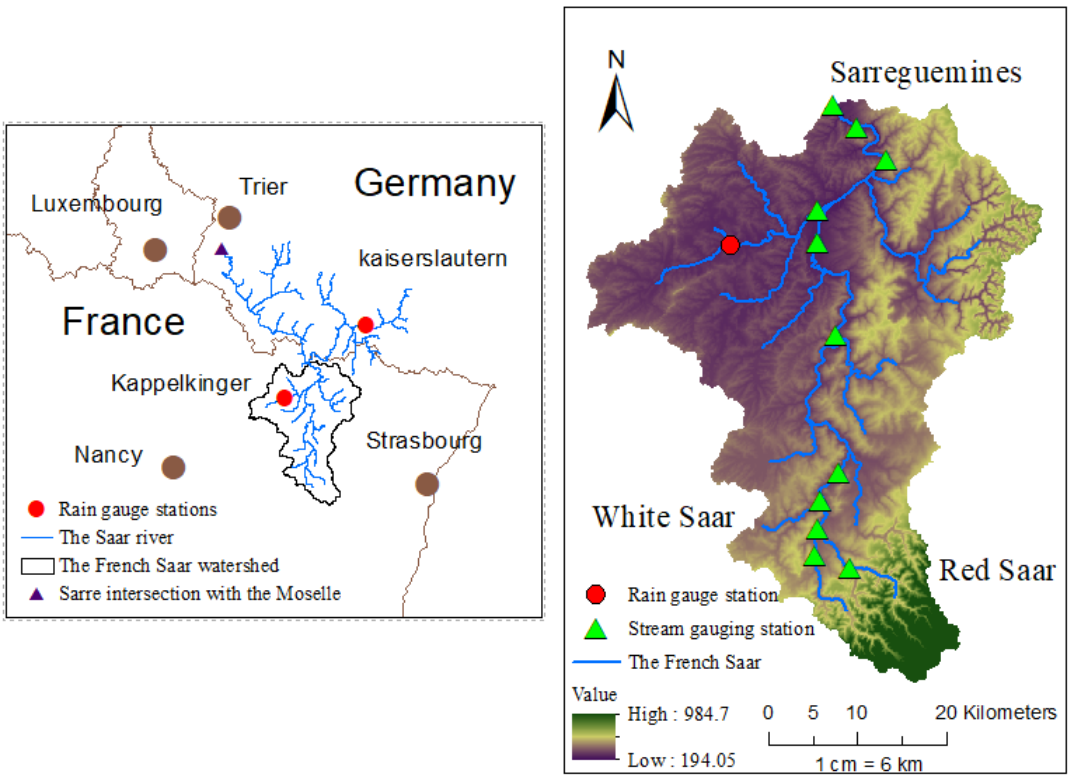


Figure 1.6: Location and topography of the Saar watershed.

The Saar river starts from the union of its two main resources near Hermelange: the red Saar and the white Saar. According to the SANDRE (National Water Data and Repository Administration Service) an official French organization, the white Saar is the upper course of the Saar, and the red Saar is its tributary. The White Saar starts near the Grandfontaine region in the Bas-Rhin department while the Red Saar takes its resources for Abreschviller in the Moselle department.

In this research work, we are interested in the French part of the Saar River which drains a basin of 1800 km^2 (Fig.1.6). The outlet of this basin is located in the city of Sarreguemines, on the French-German border. This area is classified as a territory of significant flood risk (TRI) according to the Preliminary Flood Risk Assessment (EPRI) of the Rhine District, approved on 22 December 2011 (SLGRI, 2017). The French part of the river has been equipped with eleven gauging stations: on the White Saar at Laneuveville-les-Lorquin, on the Red Saar at Vasperviller, on the Saar at Sarrebourg, Imling, Hermelange, Keskastel, Diedendorf, Sarralbe, Sarreinsming, Wittring and Sarreguemines [centre].

The digital elevation models used to illustrate the geographical information of the Saar watershed, are downloaded from the IGN. The IGN offers freely available DEM tiles of 1000×1000 elements with a resolution of $25m \times 25m$. Covering the entire case study area required to merge several DEM tiles.

In the continuity of this work on modeling flood events in the Saar, see Chapter 5, we focus on two particular sub-watersheds of the Saar: the Albe and the Moderbach due to the very specific historical circumstances where these watersheds were involved in the implementation of a defensive water line during the regional conflicts between France and Germany in the 20th century. The Albe stream that drains a watershed of 405 km^2 is one of the most important tributaries of the Saar in the French part. It meets the Saar at Sarralbe. The Moderbach stream that drains a watershed of 89 km^2 flows into the Albe near to Sarralbe.

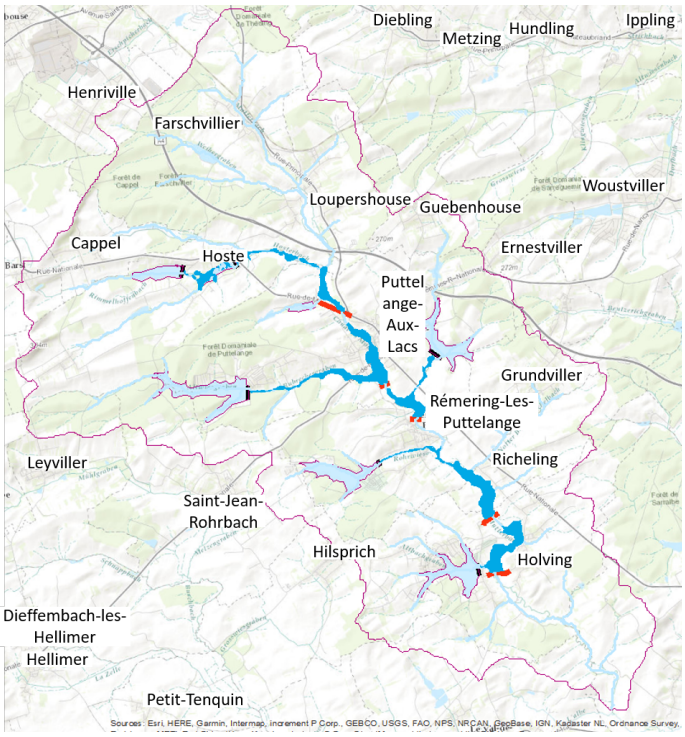


Figure 1.7: Inundation of the Moderbach watershed by accumulation of water upstream of the 5 dams (red obstacles).

Historically, this region had a number of small ponds supplied by temporary watercourses. After the First World War, the French government set up a defensive flooding system. In the east of Sarralbe, the large dams of Herbitzeim and Wittring were supplied by the ponds of Stock and Mittersheim in order to flood the major river bed of 1 to 2 km wide. In the west, the marl soil of the Moderbach watershed was used to build six ponds and five dams to block the main stream, in order to initiate flooding and stop the enemy's attack (Fig. 1.7). This event is detailed and modeled in Chapter 5. The ponds remained under military administration until 1974. In 2004, restoration and renaturation actions were coordinated by the municipalities gathered in a syndicate. Today, the ponds are used for fishing, tourism and nautical activities (Viviani et al., 2020).

1.4.1 Geological context

The flood risk prevention plan for the Saar river has classified the basin into three geological zones (PPRI, 1998):

- the upper basin: composed of Vosges sandstone, the land is permeable and the hydrographic network is embedded;
- the right bank of the Saar river starting from Sarrebourg: composed of limestone banks, the permeability is variable in a general way and the rivers are deepened;
- the left bank of the Saar starting from Sarrebourg (Fig. 1.8): the terrains are made up of marl and clay of various shades. The marls are very often covered with fine clayey silts.

Clay soils dominate the Saar region and especially the Albe and Modebrach watersheds. This is the reason why this area was useful as a water defense line to store surface water and flood the valley. This area can be considered as highly impermeable and the infiltration effect may be negligible (Fig. 1.8). This assumption is considered for modeling the extreme flood events in the Saar watershed, see Chapter 5.

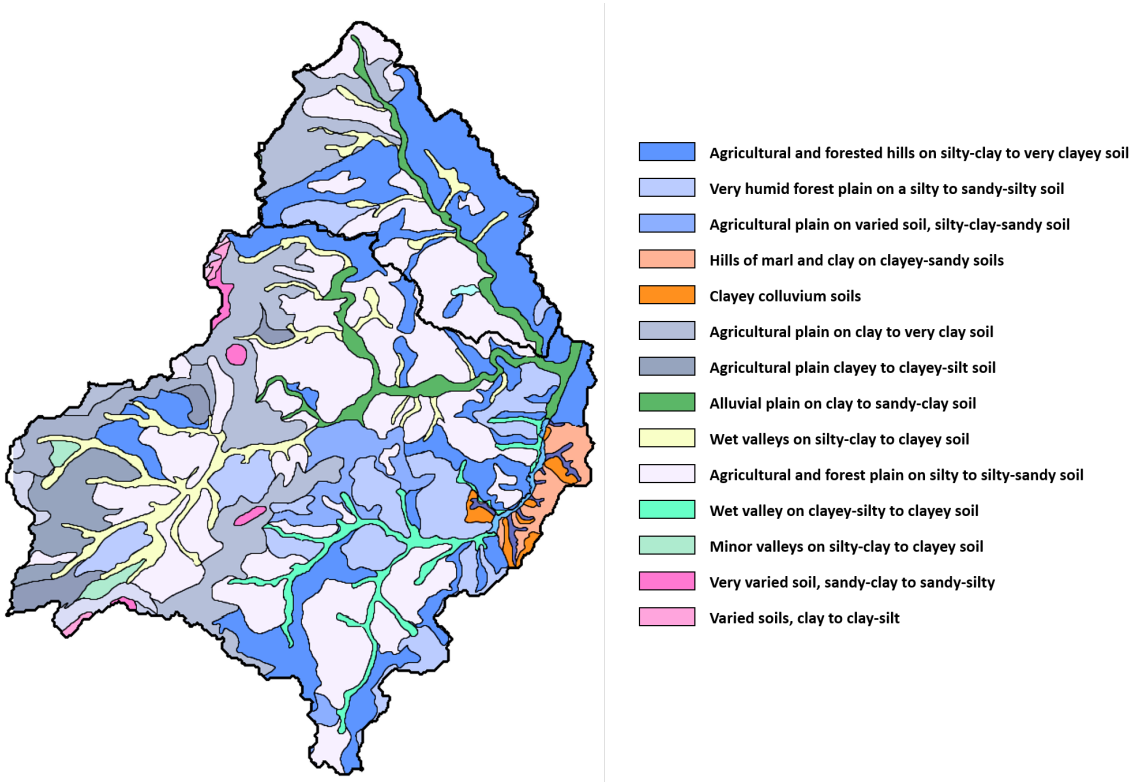


Figure 1.8: Soil map of the Moderbach and Albe basins (RRP, 2011).

The geological environment strongly conditions the land use of the basin. Associated with the clayey nature of the soils, the major bed of the Saar and its tributaries are occupied by green areas. About 70% of the Moderbach basin is covered by agricultural land and 20% corresponds to forests. The rest consists of water bodies and discontinuous urban areas (Viviani et al., 2020).

1.4.2 Hydrological context

In the Vosges, the Saar river benefits from intense precipitation events. The water level in the river experienced high amplitudes depending on the seasons.

The period of low water level is marked from the end of the spring until the beginning of autumn, see Fig. 1.9. The average discharge of the Saar calculated over a reference period of 56 years (1964 - 2021) is 17.5 m^3/s. It records much greater variations during the winter period between December and March reaching 37 m^3/s. These data are provided by the station of the Saar at Sarreguemines [center] (A9311050) from the Banque Hydro.

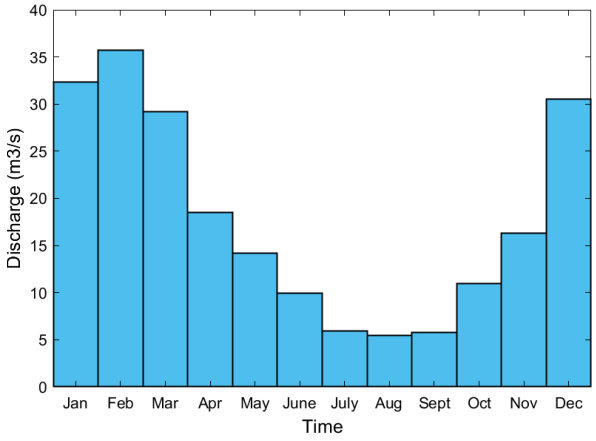


Figure 1.9: Monthly discharge at the Saar outlet near Sarreguemines for the period of 1964-2021 (Banque Hydro).

The catchment area can be characterized by areas with homogeneous hydrological and hydraulic behaviors, see Fig.1.10.

- from Hermelange to upstream of Sarrebourg, the Saar valley is relatively tight, the width of the major bed is 400 m and the land use is mainly rural. The flow rate of the 100-year flood is 105 m^3/s at the entrance of Sarrebourg;
- from Sarrebourg to Sarre-Union, the valley remains tight but is more urbanized. The width of the major bed is 300 to 400 m. The flow rate of the 100-year flood is 415 m^3/s at the entrance of the Saar-Union;
- from downstream of Sarre-Union to Herbitzheim, urbanized residential areas are vulnerable especially in Sarre-Union and Herbitzheim where the river bed is widening to reach 2 km. The flow rate of the 100-year flood is 603 m^3/s at Sarralbe;
- from downstream Herbitzheim to the German-French border, the valley is deeply incised and the river bed narrows considerably to an average width of 100 m. The Blies flows into the Saar on this section at Sarreguemines. The flow rate of the 100-year flood is 650 m^3/s upstream of the Blies confluence.

Figure 1.10: The hydrographic network of the Saar watershed (PPRI, 1998).

1.4.3 Climatological data

The region of the Saar basin is characterized by an oceanic climate with strong continental influences. The data presented in this section are provided by the station of Kaiserslautern, located at 60 km from Sarreguemines, the city classified as a territory at high risk of flooding (TRI) in the Saar basin. These data presented below are available on the ECAD website (European Climate Assessment & Datasets) (Klein Tank et al., 2002).

From the data analysis, it is possible to distinguish two distinct seasons (Fig. 1.11):

- a cold season from November to March with minimum precipitation in February;
- a hot season from April to October with high temperature.

The values below can be retained from the analysis of precipitation and temperature data extracted from the Kaiserslautern station for the period 1901 - 2018 (Tab 1.1 and 1.2) :

- The mean annual temperature is 9.1 ° C;
- The mean annual precipitation is 607.3 mm.

The analysis of the cumulative monthly precipitation for the period 1901-2018 is presented in Fig. 1.12. The boxes indicate the 25th, 50th, and 75th percentiles. The red line inside the box represents the median value. The ends of the whiskers indicate the lowest and highest data within 1.5 times the interquartile range of the lower and upper quartile. Outliers (red crosses) are data that exceed the boundary of the whiskers. Similar to Fig. 1.11, the data analysis shows a slightly wet seasonal behavior starting in May with an increase of the median until August. The month of December shows the highest number of outliers, indicating a high frequency of heavy precipitation during this period.

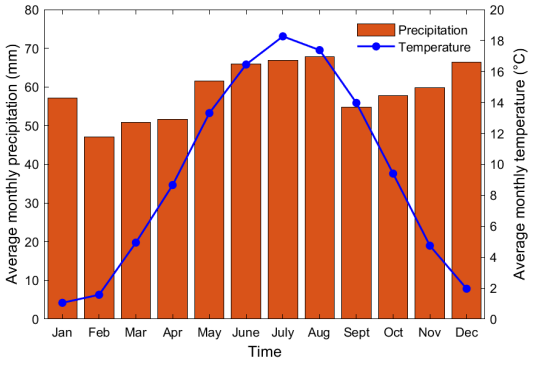


Figure 1.11: Ombrothermic diagram for the Kaiserslautern station (1901-2018).

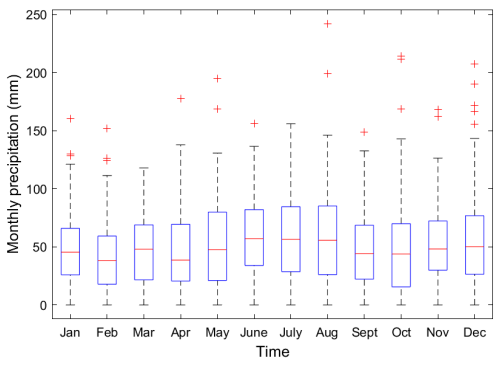


Figure 1.12: Boxplot of cumulated monthly rainfall from Kaiserslautern station (1901-2018) .

Monthely average temperature (°C) 1901-2018												
	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
Tmax	5.5	7.3	8.7	13.2	16.6	19.6	23.2	21.1	17.3	12.8	8.5	7.1
Tmin	-8.4	-9.6	0.4	5.1	9.2	12.4	14.6	14.5	9.5	5.6	0.5	-3.2
Taverage	1.0	1.5	4.9	8.6	13.3	16.4	18.2	17.3	13.9	9.4	4.7	1.9

Table 1.1: Monthly average temperature (°C) 1901-2018 at Kaiserslautern station.

Average Monthely precipitation (mm) 1901-2018											
Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
57.1	47.5	50.8	51.6	61.5	65.8	66.8	67.7	54.7	57.7	59.7	66.4

Table 1.2: Monthly average precipitation (mm) 1901-2018 at Kaiserslautern station.

Even though the station of Kaiserslautern (285 m) is located at 60 km from the Saar, we assume that the climatology their is not very different from the one in the Saar basin. As a proof of this assumption, we compare the precipitation and temperature data of the station of Kaiserslautern to those of the station of Kappelkinger (250 m) which is located in the Saar basin. The data of Kappelkinger are available from Meteo-France only for the period of 2008 to 2018.

In the following, the precipitation data from both stations are analyzed and illustrated. The amount of cumulated annual precipitation is compared in Fig.1.13. The annual correlation coefficient, calculated considering the daily precipitation data for each year from the two stations, varies between 0.56 and 0.81. In addition, the comparison of the number of dry days recorded in both stations show a strong agreement (Fig.1.14).

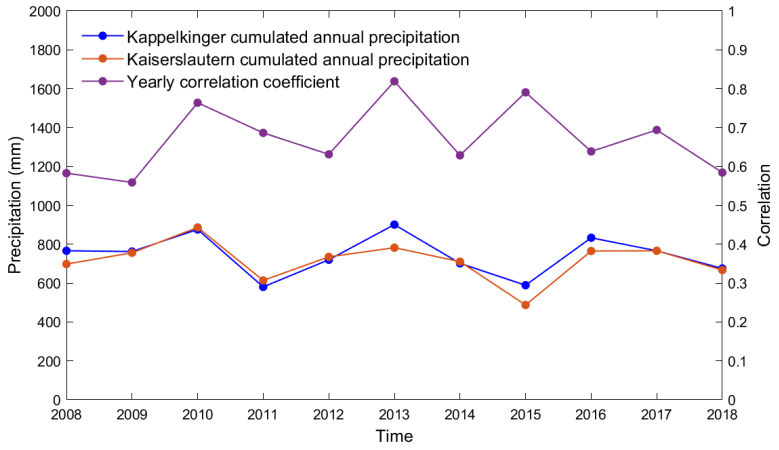


Figure 1.13: Cumulated annual precipitation recorded in Kappelkinger and Kaiserslautern stations (2008-2018) and annual correlation coefficient calculated using daily precipitation data.

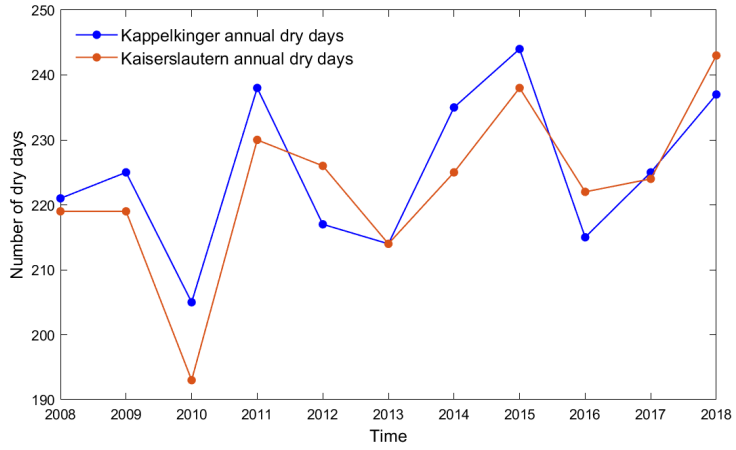


Figure 1.14: Number of dry days recorded in Kappelkinger and Kaiserslautern stations (2008-2018).

The average annual temperature in both stations shows a strong correspondence. The annual correlation coefficient, computed considering the daily temperature for each year from the two stations, shows a tight relation between the two data sets for a value of 0.98 (Fig.1.15). The number of days with a negative temperature were also similar in both stations (Fig.1.16).

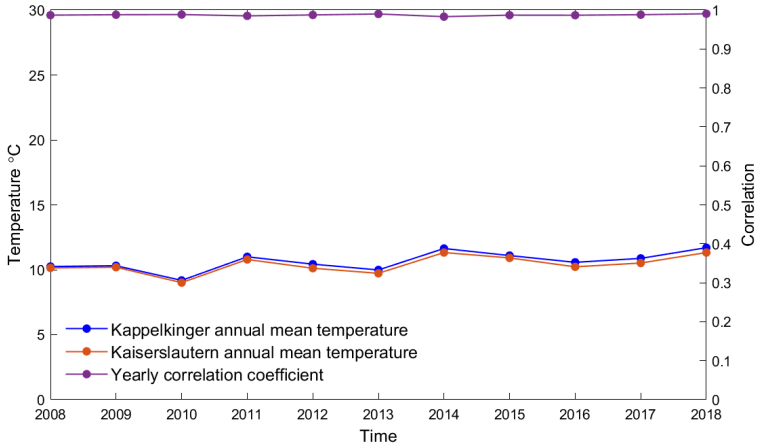


Figure 1.15: Mean annual temperature recorded in Kappelkinger and Kaiserslautern stations (2008-2018) and annual correlation coefficient calculated using daily temperature data.

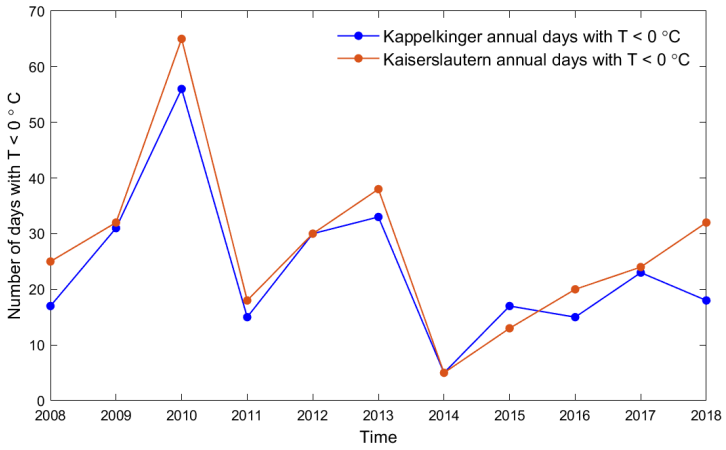


Figure 1.16: Number of days with negative temperature recorded in Kappelkinger and Kaiserslautern stations (2008-2018).

The Standardized Precipitation Evapotranspiration Index (SPEI) is a parameter that has been found useful for the assessment of drought severity (Vicente-Serrano et al., 2010). The SPEI overcomes the SPI index, which is computed using only precipitation data, by considering in addition to the precipitation the effect of evapotranspiration based on temperature data. Even though droughts are mainly controlled by the temporal variability of precipitation, the temperature has also an important impact on drought assessment.

The following results were computed using the SPEI package in R using the Hargraves method (Hargreaves and Samani, 1985) requiring precipitation, max and min temperature to compute the Potential Evapotranspiration (PET). It should be mention that the PET can be computed using other methods as the Thornthwaite method (Thornthwaite, 1948) requiring only mean temperature or Penman-Monteith method (Allen et al., 1998) requiring several datasets as wind speed, surface humidity, and solar radiation. However, these additional data may have large uncertainties.

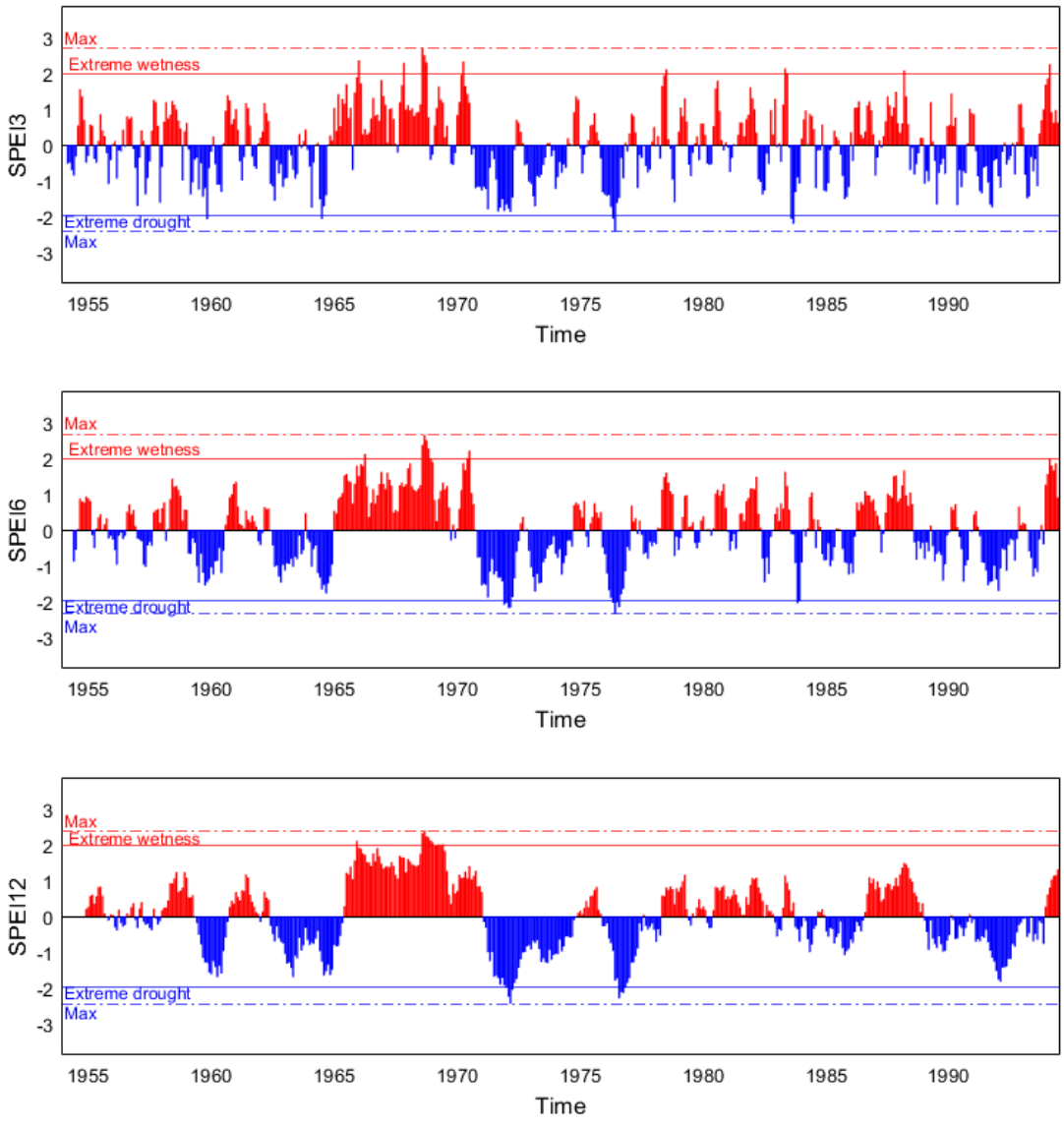


Figure 1.17: SPEI computed on 3, 6 and 12-month scale with the Kaiserslautern data.

The data used for the computation of the SPEI in the Saar area were monthly precipitation, minimum and maximum temperature from the Kaiserslauterm station for the period of January 1954 - December 1994. The evolution of the SPEI at different time scales computed for 3, 6 and 12-month time scale are presented in Fig. 1.17. The different time scale allows to distinguish different drought types based on the accumulated conditions for a certain time scale (Beguería et al., 2014). For the 3 and 6-month time scales, the drought series shows a high frequency in a short duration. These time scales allow determining agriculture droughts and events based on soil water content, river discharge, and storage areas (reservoirs). However, the computation on a long time scale, which is mainly affected by the groundwater storage variation, shows a lower frequency on a longer period. The SPEI computed on a 12-month time scale allows determining the longest wet periods and most severe droughts which occurred mainly in the period of 1959-1965, 1971-1975, 1976-1978, and 1990-1994. These data illustrate the major impact on the hydrological regime.

1.4.4 Historical extreme events

The EPRI (preliminary flood risk assessment) (EPRI, 2011) has identified the most important historical floods in the Saar basin according to several criteria taking into account the intensity, the return period, the extent of the floods and also considered the typology of the floods.

The last and most devastating floods in the Saar watershed occurred in December 1919, December 1947, December 1983, February 1997, and more recently to a lesser extent in October 2006 (EPRI, 2011).

- **Flood event of December 1919 - January 1920**

  The flood of December 1919 - January 1920 was the result of heavy precipitation on the north of the Lorraine and the Saar watershed and most importantly a sudden snow melt. This event was classified as “a winter flood due to a rapid snow melt”. The flow generated by the snow melt and the volume of the snow cover were the main causes of occurrence of this event.

  Precipitation records has reached 162.5 mm at Sarreguemines during November 1919 while the average was 59 mm. During December heavy rain continues, it snows on the Vosges and the temperature drops and becomes lower than normal. The temperature suddenly rises on the 18 of December, while rainfall continues which causes a very rapid snow melt.

  Most of the Moselle and Saar valleys were flooded. The recorded impacts were at least one death and all roads, railway houses were flooded.

- **Flood event of December 1947 - January 1948**

  The flood event at the end of December 1947 is also “a winter flood due to a rapid snow melt”.

  Precipitation during December reached almost 3 times more than the normal amount. Records during 48 h exceeded the normal value for one month and in five days the fifth or sixth of what falls in a year, and important snowfall on the Vosges was noted. An

increase in temperature associated with torrential rains was the origin of this exceptional hydrological event. The return period of this event is 30 years.

The damages and losses recorded in the archived press documents in the State Services are exceptional. Three dead in the Vosges and two people disappeared on the Saar. At Sarralbe, the water level reaches the first floor of the buildings. Phone call totally interrupted isolating Sarralbe from the rest of the department. In Sarreguemines, a district is entirely underwater and 300 people are evacuated.

- **Flood event of May 1983**

  The May 1983 flood was caused by a heavy rainfall event that occurred primarily between May 22 and 26. This flood was classified as a “Spring flood”. This type of flood is mainly due to the warm and moist air arriving from the Mediterranean basin providing a massive rain fronts that moves very slowly. These floods are the most violent causing remarkable rainfall in terms of intensity and quantity.

  The low pressure from the Mediterranean brought a warm front from the south-east on the 23rd of May. This depression moved up to Germany on the 24th, which brought the fronts back from the northeast. On May 25 and 26, the situation did not change, which blocked the rain over the Moselle basin (and Alsace). Rain fell continuously from May 23 until May 27. The total rainfall in May 1983 was about three times the normal amount. The soils were prepared by the heavy rains of April, but the vegetation had probably pumped a good part of the water from the soil. The return period of the May 1983 flood on the Saar is of the order of 30 years in Wittring and 50 years in Sarreguemines.

  The total cost of the damage in Saarbrucken was estimated at 24 million euros (damage to private individuals as well as agricultural, industrial and commercial activities) to which can be added 1.5 million euros of damage to public facilities and infrastructure.

- **Flood event of December 1997**

  The flood of February 1997 is the result of the rainy episode of February 24 to 26. This flood is classified as "winter flood by west current". These floods are due to heavy precipitation that causes the saturation of the soil and the increase of the runoff coefficient.

  Precipitation levels were high and reached 150% of the normal amount recorded in general in February. On February 24, rainfall totals exceeded 65 mm. In addition, the soil was already saturated due to the rainfall event of February 11 to 19. The return period of the February 1997 flood on the Saar is about 20 years in Sarralbe, 50 years in Wittring and 20 years in Sarreguemines.

  The flooding of the Saar and the Blies rivers caused considerable damage in the old town districts of Sarreguemines. Traffic is cut off and families were evacuated. Along the entire Saar Valley, from Zetting to Wittring, about 50 houses were flooded. In Sarrebourg, neighborhoods were invaded by water flows from the hills and flooded dozens of basements. The scenario is repeated in Réding at the entrance to Sarrebourg, where about thirty houses were damaged, forcing their occupants to leave the area.

- **Flood event of October 2006**

  Normally, rainfall is evenly distributed over the watershed. However, the distribution of floods may not follow the same concept. This is mostly related to the geology and soil type of the catchment area. For example, the Vosges Mountains play a key role in the evolution of floods. The geological characteristics of the Vosges watershed (granitic, clayey) provide a runoff coefficient higher than the one in the rest of the watershed.

  The inundation in October 2006 was the result of heavy rainfall from October 3 to 6. The entire Saar basin experienced high water levels. However, it was the upstream part that experienced a strong impact. The return period of the flood peak in Sarrebourg is more than 20 years. The flooding has decreased downstream of Sarralbe with a return period of 10 years in Wittring.

  The damages were very serious. About fifty businesses were damaged. The Vosges fire department carried out more than two thousand interventions. Thousands of EDF subscribers were deprived of electricity. Two hundred and fifty employees of eight companies were technically unemployed.

- **Flood event of December 2010**

  The December 2010 flooding was due to a severe rainfall event that occurred from December 20 to 24. This precipitation fell on saturated soils due to an earlier heavy precipitation episode between December 6 and 9. This event will be modeled in Chapter 5. The results will be compared to an observed flood map (Fig. 1.18) produced by the SERTIT (Service régional de traitement d'images et de télédétection, Strasbourg) from SAR Cosmo-Skymed images (5 m resolution)

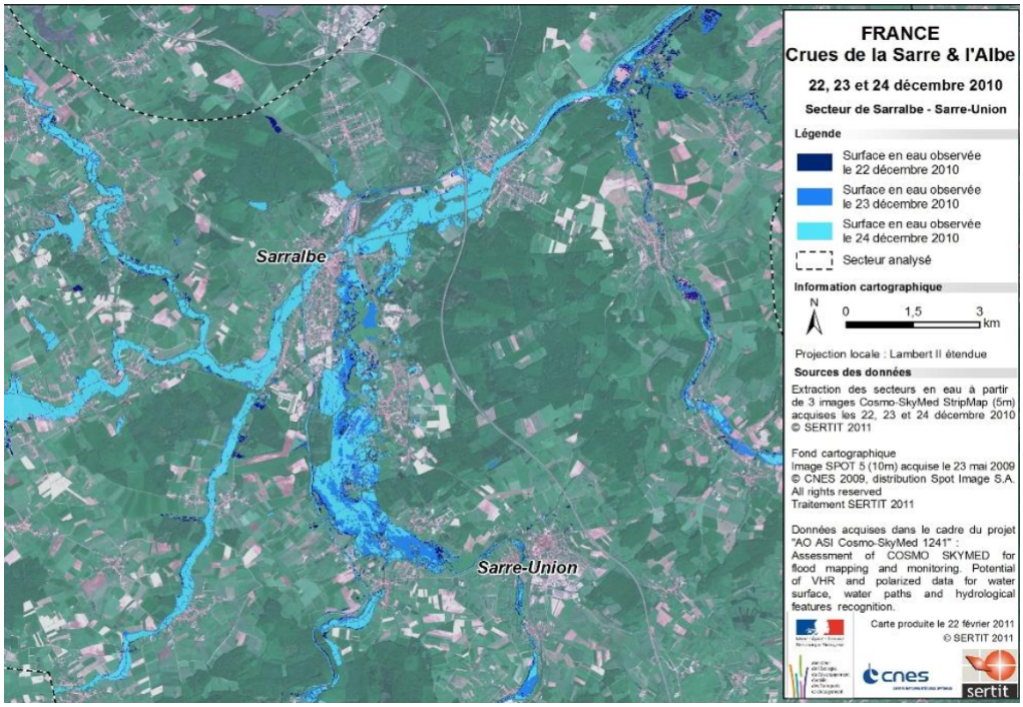


Figure 1.18: Flood map of December 2010 (SERTIT-ICUBE)

1.4.5 Flood potential impacts

The EPRI (Preliminary Flood Risk Assessment) identified the most important historical floods in the Moselle-Sarre-Nieds hydrographic district and their impacts. The potential impacts of future floods were also evaluated (SLGRI, 2017). Thirty-eight communities in the Saar area were classified as "areas at high risk of flooding" by the PPRI (Fig. 1.19).

The stakes of flooding affect the people, goods, activities, facilities and heritage. They result from urbanization and the establishment of human activities in flood-prone areas. Floods can damage the economic activity of the territory, with direct impacts on the industry sector. In the Saar area, the employments are mainly located in the largest agglomerations which are particularly vulnerable to flooding. The dominant sectors of activity in this area are mainly energy, automotive industry, chemicals and plastics, metallurgy, food processing, wood and paper industry.

Floods can cause damage to transport networks, resulting in international economic consequences. A major transport axis crosses the Saar basin to connect Paris to Strasbourg via the high-speed rail line and the A4 freeway. Cross-border employment is also highly developed in the region with the proximity of Germany, Luxembourg and Belgium.

Floods can cause environmental damage when water is contaminated by pollutants. Pollution can be significant when a sensitive industrial site is submerged. In the catchment areas of the Moselle, Saar and Nieds rivers, 39 Seveso (high threshold) sites have been identified, including 22 within the petrochemical platform of Carling - Saint-Avold and 43 wastewater treatment plants with a capacity of more than 10,000 population equivalent.

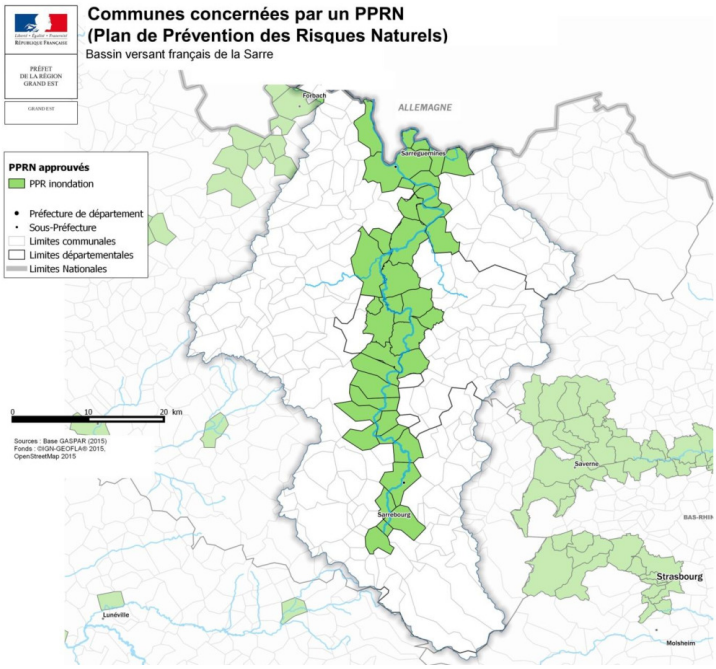


Figure 1.19: Zones highly exposed to flood risk identified in the PPRI (PPRI, 1998).

1.5 Conclusion

In this chapter, the different processes of the water cycle and the relationship between them were presented. We have thoroughly discussed the concept of a drainage basin, its different characteristics and its hydrological behavior. The geographical context of the study area was presented with a brief overview of the historical importance of the Saar river, playing a major role as part of the defense line implemented during World War II. The analysis of the geological environment of the Saar area showing the dominance of clay soils, proved the relevance of using the area to store surface water and flood the valley. The area is dominated by clay soils generating high impermeability. The meteorological data available in the Saar area from the Kappelkinger and Kaiserslautern stations, as well as the flow data from the Banque Hydro available at 11 gauging stations in the basin were analyzed providing a better understanding of the climatological environment and hydrological behavior of the basin. Finally, some of the most extreme historical flood events that have occurred in the Saar were investigated. This review of data sets and historical floods illustrated the vulnerability of the Saar basin to flooding, showing the high intensity frequency and return period recorded in this area where every two years a flood event is recorded in Sarreguemines.

Chapter 2

State of the art in watershed modeling

Contents

French Abstract

Un modèle hydrologique est un ensemble d'équations représentant les processus du cycle de l'eau et leurs relations. Le développement de ces modèles a permis une meilleure compréhension des phénomènes hydrologiques se produisant au niveau d'un bassin versant. Depuis le milieu des années 80 (Yang et al., 1998); (Rodriguez et al., 2008), des équations aux dérivées partielles non linéaires (EDP) et des paramétrisations physiques de plus en plus complexes prennent en compte les processus hydrologiques et la topographie de plus en plus finement.

On trouve dans la littérature des modèles hydrologiques ayant des propriétés et des champs d'application différents. Au cours du temps, les modèles unidimensionnels ont été remplacés par des modèles hydrodynamiques à deux dimensions simulant le lit majeur, puis la surface du bassin versant. Ces modèles peuvent être classés en fonction des processus physiques, de la distribution spatiale, de la dimensions du modèle ou la durée de simulation (Refsgaard, 1990).

La simulation de l'aléa inondation, dont les scénario d'événements extrêmes, nécessite la prise en compte de la topographie. Dans ce contexte, les systèmes d'information géographique (SIG) et les modèles hydrologiques sont complémentaires. Les Systèmes d'Information Géographique permettent de stocker, partager, récupérer et manipuler des données spatiales. Ils permettent également d'extraire d'un modèle numérique de terrain des caractéristiques utiles en modélisation hydrologique telles que le réseau hydrographique et les limites du bassin versant.

Dans ce chapitre, nous présentons les aspects des systèmes d'information géographique (Section 2.2) intéressant l'hydrologie (Section 2.3) au travers du traitement des modèles numériques de terrain (MNT). Ensuite, nous exposons plusieurs classifications des modèles hydrologiques pour justifier les choix du modèle d'onde diffusive pour réaliser des simulations avec des précipitation distribuée (Section 2.4).

2.1 Introduction

Hydrological models are mathematical representations of the water movement on or below the earth's surface. Since the mid-eighties (Yang et al., 1998) (Rodriguez et al., 2008), nonlinear partial differential equations (PDE) and additional physical parameterizations are considered to model more and more complex hydrological processes, by using an increasingly accurate representation of the topography. One-dimensional river models were replaced by two-dimensional hydrodynamic models operating in the riverbed. Rain-on-grid models allow to consider rainfall data over the entire watershed (David and Schmalz, 2021). Many hydrological models for watershed modeling exist in the literature, each with its own specificities and scope of application. These models can be classified according to different characteristics such as physical process, spatial distribution, model dimensions or time discretization (Refsgaard, 1990).

Best practices of hydrological models as the management of water resources, the forecasting of floods, the assessment of hazards (Fan et al., 2016); (Falter et al., 2016), the simulation of extreme event scenarios (Sorribas et al., 2016), and landscape modifications (Bayer and Collischonn, 2013) may require capturing the spatial distribution of various terrain-related data. In this context, the relationship between GIS and hydrological models is quite obvious.

Hydrological models take advantage of GIS tools for spatial data analysis and management. Geographic Information Systems (GIS) allow to store, share, retrieve and manipulate spatial data and the information directly or indirectly attached to them (name, occupation, flow, ...). They also allow to extract from a digital elevation model (DEM) valuable attributes for a hydrological simulation such as the hydrographic network and the characteristics of the watershed. According to La Berbera et al. (1993), the essential specificity of a GIS oriented towards hydrological applications is to be able to model the topography since it determines the flow.

In this chapter, we introduce the theoretical and practical aspects of Geographic Information Systems (GIS) and their application in hydrological modeling. We provide an overview of the general framework of hydrological modeling through a brief presentation of the role of numerical models in hydrology. Finally, the different possible classifications of hydrological models are discussed to justify our choice of a diffusive wave model in order to perform rain-on-grid simulations.

2.2 Geographic information system (GIS)

2.2.1 Definition of GIS

There are many definitions for GIS in the literature, all of which uses common terms related to cartography, spatial analysis, and database technology. One of the most common definitions is the one given by Laurini and Milleret-Raffort (1993) "Systems capable of storing, sharing, consulting and manipulating the objects represented on maps with their geometric description, as well as all the information attached to them".

The objective of a GIS is to provide a database that presents the spatial relationships between features in a given environment. This database is built using a set of graphical and non-graphical data sets (Antenucci et al., 1991). Graphical data contain cartographic elements in digital form. They are represented by 5 types of elements: point, line, area, pixel, and annotation (Burrough et al., 2015). Non-graphical data contains the characteristics and qualities of graphical data and their inter-relationships.

Building a GIS database is an expensive, time-consuming, and challenging step (Laurent, 2013). Data acquisition consists of transposing data from sources outside the GIS as from satellite imagery and photogram, into a compatible digital format. This process can be established with several methods as manual coding, digitization, and automatic vectorization of scanned images using suitable software (e. g. MapScan) (Petrescu Maftei, 2002).

GIS has shown a very satisfactory application in the natural and engineering sciences. It allows to manage large amounts of data and to cross-reference information from multiple datasets such as for example: water quality, geology and socio-economic conditions or soil type, slope and saturation level. The ability of GIS to superpose or overlay information improves the understanding of the environment being studied and can allow for more in-depth analysis. Prior to GIS, this process was very limited and applied by overlaying information using tracing paper (papier calque) (Laurent, 2013).

2.2.2 Brief history of GIS

One of the first uses of geographic data was the paper mapping of roads, property lines and water lines. With the enormous technological advancements in the 1890', the process of mapping perforated cards was automated. In 1960, the concept of digitization was introduced and GIS was presented as a software application for computerized mapping and satellite remote sensing (Petrescu Maftei, 2002).

The first GIS computer applications is the Canada Geographic Information System(CGIS), introduced in 1964 by an IBM team under the guidance of Roger Tomlinson. It was used to manage natural resources. Other projects were proposed at that time (Foresman, 1998). However, the first commercial software that is today the most popular and widely used was released by ESRI in 1978 under the name ARCINFO. In the following, we briefly introduce four of the best known and most used GIS software.

- **ARCINFO:** Created by ESRI (Environmental Systems Research Institute) in the United States, it is a GIS written in Fortran 77 and C. ARCINFO has shown a high accuracy in deriving spatially related information required to compute runoff hydrographs (Terstriep and Lee, 1989).
- **Topotoolbox:** Developped by Schwanghart and Kuhn (2010), it is a set of Matlab functions, dedicated to the analysis and geoprocessing of digital elevation models (DEM) in a non-GIS (Geographic Information System) environment. The tools are developed to support hydrological and geomorphological studies that involve the analysis of DEMs, material fluxes, and spatial variability of water, sediment, chemicals, and nutrients.
- **IDRISI:** Developed in Massachusetts, U.S.A, it was mainly used for processing images acquired by remote sensing. IDRISI was also applied to compute hydrologic parameters and creating input files that can be used directly by hydrological models (Rego, 1996) and in the management of groundwater quality (Nachtnebel et al., 1993). IDRISI is written in Borland Pascal and Visual Basic.
- **GRASS:** Created by US Army engineers, is a free and open source GIS software. GRASS is written in C. It was applied for geospatial data production, analysis, and mapping (Sonnentag, 2009). GRASS is the leading software in terms of functionalities and in open source GIS (Neteler and Raghavan, 2006).
- **QGIS:** Developed by the volunteers of the Open Source Geospatial Foundation (OS Geospatial), QGIS is coded in C++. It was mainly developed as an interface to easily access GRASS. QGIS offers a built-in Python command within the interface to perform computations. Plugins can be coded in Python and launched from QGIS.

2.2.3 GIS in hydrological modeling

The application of GIS in hydrology is not recent and has been going on for several years. It has been applied in several domains such as management of water resources, runoff modeling, risk assessment and decision-making. Warwick and Haness (1994) coupled ARCINFO with the HEC-1 hydrological model to test land cover and moisture conditions. Hammouri and El-Naqa (2007) have applied GIS for a rain-flow modeling of the Wadi Madoneh ungauged basin in Jordan. The main goal was to assess the potential of surface water for artificial groundwater recharge. Lee and Terstriep (1991) provided a GIS interface for the Agricultural Nonpoint Source Pollution Model (AGNPS) (Young, 1987).

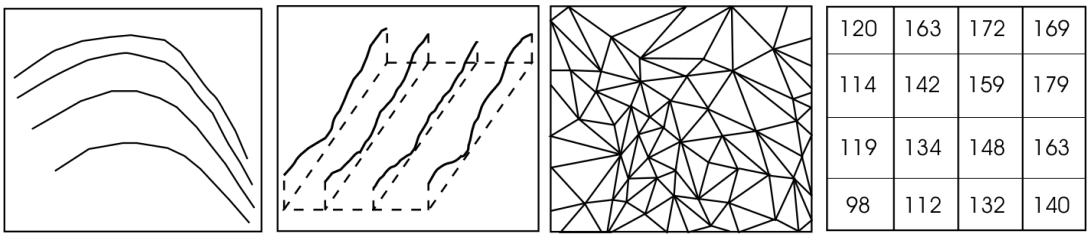
Distributed hydrological models (Section 2.4.2.2) aim to describe the water movement above (surface flow modeling) and below (subsurface flow modeling) the earth's surface. Therefore, distributed hydrological models are strongly dependent on the spatial analysis (Heywood et al., 2011). In this context, the application of GIS tools allowing to store, manage, analyze, capture, and describe spatial data in hydrological models is relevant. GIS can be used as a pre-processor providing input data required for the model simulation (slope calculation, watershed boundary...). GIS also allows to generate homogeneous areas in terms of hydrological responses, or so-called Hydrological Response Units (HRU), which can be defined by a cross-referencing of several datasets in the GIS (slope, lande use, vegetation, soil type...). Several studies have used GIS as a post-processor for hydrological models (Maidment et al., 1993), taking advantage of its high quality visual representations which improve the analysis of the results. For example, displaying water levels on a map is more understandable than processing a series of data in a table.

In hydrological modeling, the GIS must be able to provide, as much as possible, a fine representation of the terrain, since it is the feature that acts the most on the water flow behavior (La Berbera et al., 1993) (Laurent, 2013). The cartography elements of GIS are implemented in hydrology as follow:

- **Point:** It can correspond to a catchment outlet, meteorological and gauging stations;
- **Line:** It can be used to draw the flow path;
- **Polygon:** A bounded surface corresponding to a catchment area;
- **Raster:** Format of geographic data represented as a matrix of pixels.

2.3 Digital Elevation Model (DEM)

Modeling the relief of a terrain consists in defining a spatially referenced information in a geographic coordinate system (Sarrazin, 2012). For a long time, the relief of the terrain has been represented in several formats: dimensional points, contour lines, shading, or perspective views. In the last years, presenting elevation information in a digital format, which is now called digital elevation model (DEM), has gained in popularity, speed of generation, and accuracy due to the development of digital cartography and geographic information systems (Fig. 2.1). A DEM in raster or matrix format, which is the structure used almost systematically for modeling, has a defined spatial resolution in terms of the size of its unit cells.



(a) Contour lines (b) Profile lines (c) TIN format (d) Raster format

Figure 2.1: Topography presentation formats (Gomez, 2002).

In hydrological modeling, a sufficient knowledge of the topographic surface remains mandatory since the transfer of water masses is subjected to the force of gravity. Numerous DEM-based algorithmic works have been developed allowing to automatically extract morphometric and topographic features from the DEM useful for surface hydrology (Martz and Garbrecht, 1992) (D'Agostino et al., 1993) (Meisels et al., 1995). One of the best known and used in most GIS software is the D8 method (O'Callaghan and Mark, 1984), which computes the flow direction of each cell in a DEM raster in order to derive the hydrographic network. The accuracy of the hydrographic network extracted from the DEM depends on the scale of representation, the resolution, the quality, and the structure of the DEM (Sarrazin, 2012). The altimetry database in a DEM can be handled by a geographical information system software to detect hydrographic networks, delineate watersheds, and derive their characteristics as drainage length and direction, watershed area, shape, and average slope.

2.3.1 Data acquisition methods

Several data acquisition techniques exist for the construction of a DEM, by direct terrain measurement or remote sensing methods. The direct terrain measurement provides an accurate representation of the relief. However, it's a time-consuming and expensive method. Several practical methods can be applied as photogrammetry, radar interferometry, LiDAR, and digitization of curves on existing maps. The application of these techniques has evolved over time.

Early DEMs were based on photogrammetry. This technique uses methods from several disciplines as optics and projective geometry for matching images taken from different points of view to create DEM (Ouedraogo et al., 2014). At the scale of km^2, this technique is almost non-functional with poor resolution as it is not able to easily detect artificial structures that may affect the water flow behavior.

Modern DEMs are mainly based on the LiDAR (Light Detection And Ranging) method. This is a remote sensing technology based on the analysis of the properties of laser light. The distance from the sensor emitting the laser to an object or surface is given by measuring the delay between the pulse and the detection of the reflected signal. This information is combined with GPS data to create point clouds representing the terrain surface (Shan and Aparajithan, 2005). The measuring instrument is airborne in an airplane or helicopter. DEMs built from LiDAR data has a much higher accuracy than any other photogrammetric DEM (Lindsay and Evans, 2008) (Tarolli and Dalla Fontana, 2009) where it can provide DEM resolutions of less than 0.5m x 0.5m (Reutebuch et al., 2003). Several countries as Denmark (Danish Geodata Agency), Finland (National Land Survey of Finland), and Sweden (Swedish Mapping, Cadastral, and Land Registration Authority) have already a large dataset of DEMs created using LiDAR (Lidberg et al., 2017).

2.3.2 Resolution

Improvements in the knowledge of hydrology are being achieved partly due to the availability of superior hydrological maps obtained from high-resolution DEMs (Murphy et al., 2008). According to Grayson et al. (1993), the ability to understand the hydrological processes derived from the topography decreases if the DEM cells increases in size and exceeds the scale at which field observations are made. This is due to the risk of loss of information. The advantage of high resolution DEM is their ability to provide the geographical information of small scale features as small streams in forested areas (Dehvari et al., 2013) (Goulden et al., 2014) (Vaze et al., 2010) (Yang et al., 2014). Therefore, we may consider that the understanding of hydrological processes is mainly related to the DEM resolution (Sarrazin, 2012).

The application of high-resolution DEMs has introduced new challenges in hydrological modeling. First, the processing time of a high-resolution DEM is very long. This is due to the large amount of data points (Barnes et al., 2014). Second, in high-resolution DEM, the geographical information of structures as dams and road embankments will be recognized in the DEM data. In reality, water may be passing through the dams or underneath the roads. However, since these structures are slightly higher than their surrounding terrain, their geographical data added to the DEM will introduce them as artificial walls that blocks the water providing a misleading representation of the actual topography (Barber and Shortridge, 2005). DEM pre-processing methods can be applied to correct these defects, see Section 4.3.

2.3.3 Deriving DEM characteristics for watershed modeling

A DEM can therefore provide useful information for hydrological analysis. In the following, we present the method provided by GIS to extract hydrographic networks and watershed boundaries from a raster DEM. In a raster format, each cell of the DEM is defined by its x and y coordinates and its altitude information Z as shown in Figure 2.1(d). Extracting the hydrogaphic network starts by determining the flow direction of each cell by a slope computation.

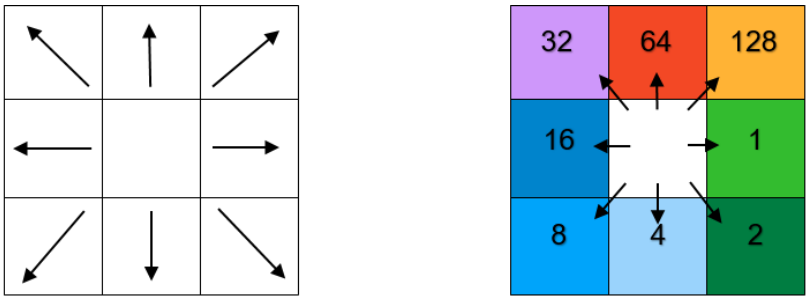


Figure 2.2: The 8 possible flow directions of a DEM cell with code color.

The majority of GIS tools apply the D8 method (O'Callaghan and Mark, 1984) to automatically route the flow from each raster grid cell of the DEM to one of its 8 adjacent cells. In GIS software as ArcGIS and QGIS, the 8 possible direction values are encoded as 1, 2, 4 ,8 ,16, 32 ,64 and 128 in a raster format, see Fig. 2.2.

78	72	69	71	58	49
74	67	56	49	46	50
69	53	44	37	38	48
64	58	55	22	31	24
68	61	47	21	16	19
30	53	34	12	11	12

(a) Elevation matrix

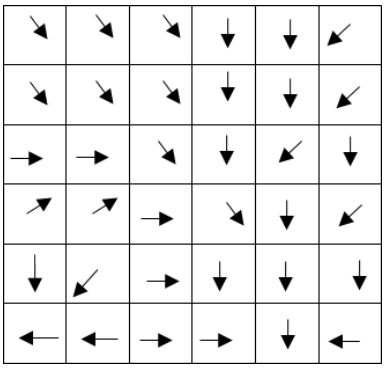
2	2	2	4	4	8
2	2	2	4	4	8
1	1	2	4	8	4
128	128	1	2	4	8
4	8	1	4	4	4
16	16	1	1	4	16

(b) Flow direction matrix

Figure 2.3: Elevation values encoded in a flow direction matrix.

The flow routing is based on selecting a flow direction for the DEM raster cells. This way, each cell receives one of the 8 possible flow direction values according to the direction with the greatest slope. The flow direction values are saved in a raster format, see Fig. 2.3. Despite its remarkable efficiency, the D8 algorithm has certain limits. This method does not allow to define a realistic flow direction in all contexts especially when the elevation of a cell is lower than its 8 neighboring cells. The existence of this type of cells is usually due to particularities in the terrain. Several methods as Filling and Breaching allow to correct DEM before starting the flow direction computations. These methods will be developed in detail in section 4.3.

The hydrologic flow is then accumulated from cell to cell along the directions calculated in Fig.2.3, providing a flow accumulation matrix where the values associated with each cell correspond to the number of cells it drains, see Fig. 2.4.



(a) Direction of flow for each cell

0	0	0	0	0	0
0	1	1	2	2	0
0	3	7	5	4	0
0	0	0	20	0	1
0	0	0	1	24	0
3	0	4	7	35	1

(b) Number of drained cells to each cell

Figure 2.4: Flow directions encoded in a flow accumulation matrix.

A hydrographic network can be deduced based on the matrix of accumulated flow, according to a minimum threshold of drained cells. Fig. 2.5 shows the hydrographic network derived for two different threesholds 20 and 4.

0	0	0	0	0	0
0	1	1	2	2	0
0	3	7	5	4	0
0	0	0	20	0	1
0	0	0	1	24	0
3	0	4	7	35	1

(a) Hydrographic network for a minimum threshold of 20 drained cells.

0	0	0	0	0	0
0	1	1	2	2	0
0	3	7	5	4	0
0	0	0	20	0	1
0	0	0	1	24	0
3	0	4	7	35	1

(b) Hydrographic network for a minimum threshold of 4 drained cells.

Figure 2.5: Deriving hydrographic network.

The watershed boundaries are determined based on the flow direction matrix. Once the watershed outlet is selected, all upstream cells flowing in the direction of the outlet cell are considered in the watershed area. Figure. 2.6 shows the perimeter bordering the cells draining the outlet cell with a flow accumulation value of 35.

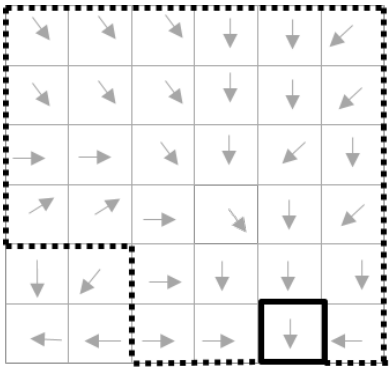


Figure 2.6: Watershed boundaries.

2.4 Hydrological models

Numerical models are considered as the standard tool to model a large selection of physical and environmental processes. A definition of a model given by Wheater et al. (2007) is "A simplified representation of the real-world system". Modeling, in general, is the process of describing a system based on input variables/parameters and initial/boundary conditions in order to provide a prediction of a system's behavior and a description of processes.

Numerical models can be introduced in an open-source framework offering the user access to the code, under a certain license, or in a closed source framework where the user does not have access to the code. In general, closed source software are commercial products that facilitates the user's ability to manipulate the model and avoid contact with the complex code behind it. Open-source software may enable large-scale replication of results, which is an essential requirement for science and beneficial to a larger audience than closed source codes. Several conditions can control the choice of a model: the objective, the experience of the user, and budget constraints. In general, open-source codes are applied in the educational and research field while closed source codes are more commonly used in the industry.

2.4.1 Motivations

The application of numerical models in the hydrological field allowed to enhance the understanding of hydrological processes. The history of hydrological models started with the rational method (Mulvany, 1850). A very simple empirical equation that expresses the time of concentration and computes peak flow. This model is well adapted for small basins to consider uniform rainfall. In the 1960s, the development of conceptual models provided a better representation of the hydrological processes at the watershed level. The different components of the hydrological cycle were presented as interconnected elements (reservoirs) (Dawdy and O'Donnell, 1965) (Crawford and Linsley, 1966). Modelers believed that for a real representation of the hydrological processes, models should rely only on parameters with physical interpretation. This type of model called physically-based started with the work of (Freeze and Harlan, 1969) and (Freeze, 1972).

In recent years, models have been developed to handle several characteristics such as watershed scale, groundwater flow, spatial variability and to address different objectives, for flood prediction, water quality evaluation, sediment and pesticide transport, study of possible future changes in climate and management of water resource. Very satisfactory results were given by Grillakis et al. (2010) in the application of the Hydrologiska Byrans Vattenbalansavdelning (HBV) model on flash flood case in Slovenia (Bergström, 1976) (Bergström et al., 1995). Borah and Bera (2004) has concluded that for soil erosion, sediment transport, and continuous flow simulation, the Soil & Water Assessment Tool (SWAT) is well applied (Arnold and Fohrer, 2005) (Neitsch et al., 2002). Devia et al. (2015) showed that the application of the MIKE SHE model (DHI, 1999) requires an extensive amount of data which may not be always available. Table 2.1 states five of the popular and most used models in research.

Model name	Model type	Spatial representation	Time	Spatial Dimension	Best use
HBV	Conceptual	Lumped Semi and fully distributed	Continuous event-based	1D-3D	Flood forecasting (Häggström and Lindström, 1990), water resources evaluation (Jutman, 1992), (Brandt and Bergström, 1994), simulating discharge for ungauged catchments (Seibert and Beven, 2009)
SWAT	Conceptual Physically-based	Semi-distributed, Distributed	Continuous	2D	Assess the impact of climate and land usechanges (Gosain et al., 2006), water and sediment circulation (Yang et al., 2009), runoff (Jeong et al., 2010), flash flood simulation (Boithias et al., 2017)
MIKE SHE	Physically-based	Semi-distributed, Distributed	Continuous	3D	Flash flood simulation (Sahoo et al., 2006), assess the impact of climate and land usechanges (Im et al., 2009), runoff (Ping et al., 2017)
HEC-HMS	Conceptual	Semi-distributed	Continuous, event-based	1D	Flood forecasting (Oleyiblo and Li, 2010), runoff (Halwatura and Najim, 2013) (Gebre, 2015) (Gumindoga et al., 2017)
TOPMODEL	Physically based	Distributed	Continuous	2D	Runoff and catchment hydrology (Hornberger et al., 1985), flood forecasting (Blazkov and Beven, 1997), forest ecology (Gao et al., 2017)

Table 2.1: Some of the most known and widely used hydrological models.

2.4.2 Classification

A brief survey of the literature shows that various classifications for hydrological models have been suggested (Dawdy and O'Donnell, 1965), (Clarke, 1973), (Todini, 1988), (Refsgaard, 1990) (Fig.2.7).

According to Refsgaard (1990), hydrological models can be classified regarding the extent of the physical processes into three categories (Section 2.4.2.1): empirical models based on equations, conceptual models based on simplified equations with physical and calibrated parameters, and physically based models considering only physical characteristics.

A second classification relies on the spatial variability of the parameters (Section 2.4.2.2) as follow: lumped models considering all the watershed study area as one unit, semi-distributed considering spatial variability at the sub-watershed scale, and fully distributed providing a network of small mesh elements to account for spatial variability.

Hydrological models can also be classified depending on their spatial dimension (Section 2.4.2.3). While some models are limited to river channel modeling, others allow for two-dimensions overland surface modeling and three-dimensions modeling taking into account the water flow in the sub-surface.

The hydrological input data considered in the modeling process can also be used to classify hydrological models (Section 2.4.2.4). While some models consider only the stream-bed of the study area as a computational domain which allows the use of flow data as model input, others consider the entire watershed as a computational domain which necessitates the use of precipitation data. The time scale can also be considered to classify the models as continuous or event-based (Section 2.4.2.5).

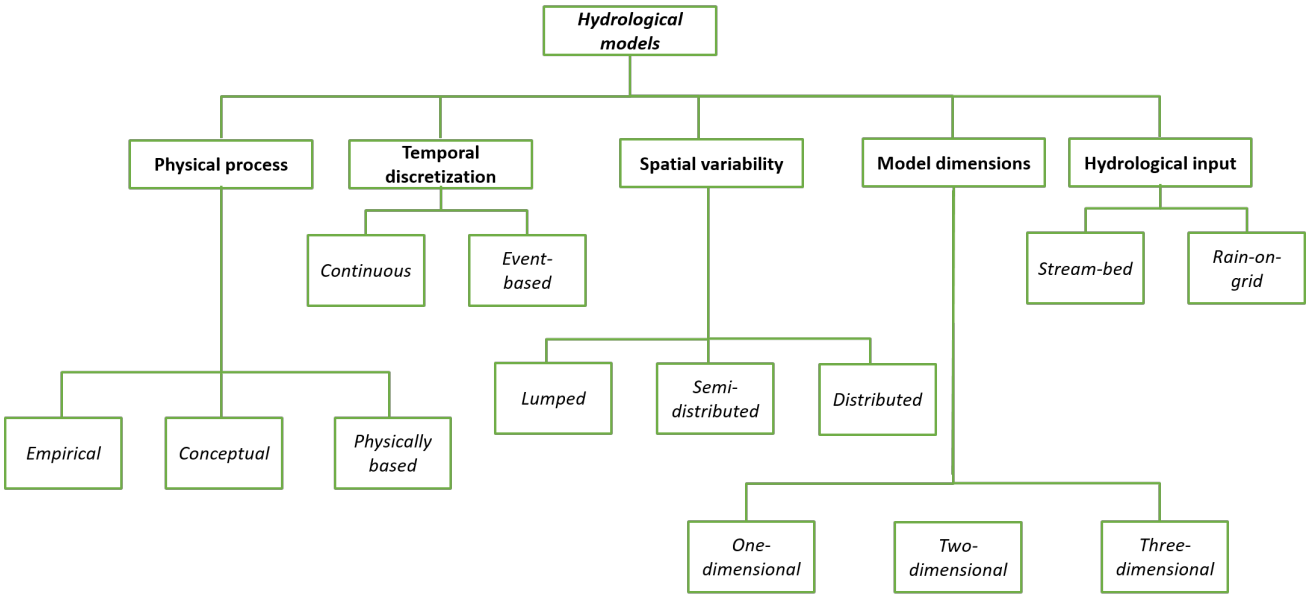


Figure 2.7: Classification of hydrological models.

2.4.2.1 Classification according to the physical process

- **Empirical models**

  Empirical models are described as a "Black Box". These models consist of a mathematical relationship between inputs and outputs $outputs = f(input)$ derived from a large data base of observation and experience regardless of the hydrological processes of the watershed or its physical characteristics. These models have their advantages. They are quite simple and require a unique parameterization, but this simplicity leads to the incapability to describe the physical processes of a watershed. The predictive ability of these models is satisfactory only in their area of validity and in conditions similar to those of which the model equations were derived.

- **Conceptual models**

  A conceptual model consists of three main concepts: water transfer from storage to atmosphere, runoff propagation over the surface, and water storage in the soil (Islam, 2011). These concepts are expressed in terms of interconnected elements (reservoirs) corresponding to the hydrological processes of a catchment. The model uses simplified (semi-empirical) equations based on parameters not only provided from the field (physical) but also through calibration.

- **Physically-based models**

  Physically-based hydrological models consist of an idealized representation of the real phenomena (Devia et al., 2015). The model uses mass, momentum, and energy conservation differential equations (Darcy's Law, Richard's equations, Saint-Venant) through discretization methods such as finite difference or finite element methods to model the water movement from and to the soil. This type of model, unlike the other two, is based only on state variables with physical characteristics (soil moisture, hydraulic conductivity, land use, soil cover ...) which enables a wide field of application. Due to this level of reliability provided by this type of model, the two hydrological models used in this research work are physically based models (Section 4.2).

2.4.2.2 Classification according to spatial variability

Refsgaard (1990) considers that hydrological models can be classified according to the spatial description at the watershed level as lumped, semi-distributed and fully distributed (Fig. 2.8). In general, conceptual models as HBV (Bergström, 1976) (Bergström et al., 1995) and HEC-HMS (Feldman, 2000) are lumped while physically based models as HEC-RAS (Brunner, 1995), MIKE-SHE (Refshaard et al., 1995) and SWAT (Manguerra and Engel, 1998) are distributed (Islam, 2011). However, due to the ambiguity of certain criteria, there is still some confusion about the belonging of this or that model to a given family.

(a) Lumped (b) Semi-distributed (c) Fully Distributed

Figure 2.8: Moderbach catchment subjected to different types of spatial variability.

- **Lumped models**

  In a lumped model, all the case study basin is considered as one unit. The spatial variability of physical parameters is not considered. Therefore, lumped models do not require long computation times (Sitterson et al., 2018).

  Assuming homogeneity in lumped models, all input/output data and parameters are average data. This may not properly represent large-scale watersheds and may lead to an over- or under-estimation of the computed runoff. For example, land use may affect the runoff process in a specific area. However, lumped models consider an average land use across the entire catchment which does not properly reflect the impact of land use on runoff in specific areas of the catchment. In these models, no specific flow can be calculated in the catchment area, the runoff is calculated at the catchment outlet only.

- **Semi-Distributed models**

  Semi-distributed models can be considered as lumped models applied in a quasi-distributed manner. The catchment is decomposed into a number of sub-units called Hydrologic Response Unit (HRU) considered "homogeneous" in terms of hydrologic properties (Gnouma, 2006). Spatial variability is considered at the sub-units scale. These sub-units are obtained by crossing topographical characteristics (slope, orientation, etc...) with land use, soil, and geological characteristics (Payraudeau, 2002). In these models, a runoff output can be computed at the outlet of each sub-unit.

- **Distributed models**

  The distributed or spatialized aspect allows to consider the spatial variability of the catchment characteristics as vegetation, land cover, land use, soil, and topography. Discretization is applied over the computational domain as a series of small triangular or square elements. Each cell has its own characteristics used to calculate the runoff value for it, taking into account its neighboring cells. This method provides detailed information about the runoff over the watershed at the small element scale, which allows to reflect the impact of the morphological characteristics of the catchment on the computed runoff. This type of model requires a long simulation time and efficient computational tools, especially

for large watersheds. However, model predictions are expected to be more accurate than lumped or semi-distributed models (Islam, 2011). Taking advantage of this level of accuracy, this research work involves two hydrological models that allow to consider the spatial variability of the physical parameters in a distributed manner over the watershed (Section 4.2).

2.4.2.3 Classification according to the model dimensions

- **One-dimensional models**

  Many efficient hydrological models as HEC-RAS (Brunner, 1995), MIKE 11 (Havno et al., 1995) and InfoWorks (Wallingford, 2003), were developed for river and floodplain modelling through a 1D channel routing model. These models rely on the one-dimension shallow water equation (Saint-Venant equation) and they are limited to model water flow in ditches, streams, channels, and rivers.

  Surface runoff due to groundwater exfiltration, precipitation on saturated surface, and exceeding the infiltration capacity, all these processes are not considered in this type of model. In general, 1D modeling is considered sufficient when the water flows in the channel and is constrained by its banks, but when the water exceeds the channel level and arrives in the floodplain, the water moves in two directions and therefore 2D modeling is required (Surwase et al., 2019).

- **Two-dimensional models**

  Two-dimensional hydrological models have been developed to overcome the limitations of one-dimensional models. Several hydrological models developed to perform 1D modeling have been upgraded to include 2D modeling capabilities such as HEC-RAS and InfoWorks-2D. In this type of model, the computational domain can cover the entire watershed area allowing to consider all sources of runoff. These models rely on the two-dimensional shallow water equation with either the Full Saint Venant equations (with added terms for turbulence modeling and Coriolis effects) or the Diffusive wave approximation (Section 4.2.1).

  However, some of these models are still limited to surface runoff modeling without considering groundwater flow. This type of model can be applied in situations where the subsurface is considered to be very impermeable and therefore the exchanges with the subsurface are negligible. This is consistent with the conditions in our study area where the dominance of clay in the soil layer provides an impermeable soil where infiltration can be neglected (Section 1.4.1). Hence, the application of a 2D models in this research work was favored (Section 4.2).

- **Three-dimensional models**

  Three-dimensional hydrological models overcome the limitations of two-dimensional models, as they take into account the interactions between surface runoff and groundwater flow (infiltration, percolation, exfiltration). Moreover, the integration of these interactions in hydrological models is necessary to properly describe the water cycle and the tight coupling between its various compartments.

The importance of modeling these interactions is reflected by the evolution of hydrological models such as TELEMAC (Janin et al., 1992) and Mike 3 (DHI, 2014) which were initially developed for 1D modeling and then improved to provide 2D and 3D modeling. Several other 3D models exist such as Delft3D (Manual, 1999), RSim-3D (Tritthart, 2005) and NIHM (Jeannot et al., 2018), most of them rely on the solution of the three-dimensional Richards equation to describe the dynamics of groundwater flow.

2.4.2.4 Classification according to hydrological inputs

- **Stream-bed models**

  Most flood studies limit the computational domain to the stream bed of the study area. The modeling is then performed using flow data as hydrological input to model the flood extent. The advantage of this modeling strategy is the short computation times required since the computational domain is restricted to the stream bed. However, this method ignores the impact of the watershed characteristics, such as soil parameters, land use and land cover on the hydrological behavior and limits the hydrological results (extent, discharge, velocity, depth) to the stream bed.

- **Rain-on-grid models**

  Rain-on-grid models aim to consider the entire watershed as the computational domain (David and Schmalz, 2021). Therefore, the modeling is performed using rainfall data associated to the entire watershed using a finite element or finite volume method which requires long computation times. Rain-on-grid models are generally distributed physically based models. This research work involves two rain-on-grid models to perform 2D simulations on large-scale basins.

2.4.2.5 Classification according to temporal discretization

Depending on the modeling objectives, the behavior of the basin can be reproduced only for particular events (typically floods) or continuously over a longer period of time including dry periods (non-events). Hydrological models can be classified as event-based models and continuous models (Wheater et al., 2007).

- **Event-based models**

  Event-based models as HEC-1 (Beard, 1981), aim to reproduce the basin's response in terms of flow during a heavy rain period. This kind of model, which is activated only for a specified period of time (flood event), needs to be able to specify the initial state of the basin according to the antecedent climatic conditions (Gnouma, 2006).

- **Continuous models**

  The main difference between event models and continuous models consists in the representation of the water cycle. Certain processes can no longer be neglected in continuous models, in particular evaporation and evapotranspiration (Gnouma, 2006). The most commonly used continuous models in the research are: MIKE-SHE, TOPMODEL and SWAT.

2.4.3 Validation

As a model is a simplification of a complex system, several approximations and assumptions are made to build the model. This implies uncertainties in the model, see Appendix A Model validation is therefore a necessary step to ensure a reliable performance. The comparison of model results with observed values allows to evaluate the limits of the model.

Several methods for validation have been the subject of much discussion and challenge over the past decade (Beven, 1989) (Bergström, 1991). Comparing the computed and observed outflow at the basin's outlet is often the only feasible option for validation in many cases (Refsgaard, 1990). The value of this point measurement is based on its temporal continuity and not on its spatial distribution. High water marks can also be used for validation. They correspond to the maximum water level observed during a flood event. For extreme events, the traces of the high water marks are surveyed and transformed into a coordinate system (X,Y,Z) to be archived or indicated by paint marks or plates. However, the reliability of this type of data is highly variable. They are often poorly referenced over time. Validation of distributed hydrological models only through this single location of the flow or high water mark is extremely limited in terms of the credibility of the spatialized aspect of the model. Distributed hydrological models allowing to compute water level, velocity and discharge at a spatial and temporal scale, need to be validated with observed flood maps. These data can be obtained from aerial photos, satellite images or constructed from observed data sets. One of the methods applied in flood modeling studies to compare computed and observed data at a spatial scale is the one based on the fit parameters (Di Baldassarre et al., 2009), (Horritt et al., 2007), (Quirogaa et al., 2016). This method is applied in chapter 5 to validate the historical simulations applied in the Saar region.

2.5 Conclusion

The ability of GIS tools to store and manage large spatial datasets has contributed to a better understanding of the hydrologic processes. In fact, coupling GIS with hydrologic models has facilitated model implementation through the ability to manage and provide model input data and improved model results by taking advantage of the efficient visualization capabilities. Among the different types of models presented in the chapter, physically-based distributed models take into account the spatial variability of physical parameters. Although this requires the discretization of the whole basin area, which leads to long simulation times, it allows to provide the model results (velocity, water depth and flow) at any point of the computational domain. This justifies our choice to use diffusive wave models in this research work to perform a 2D modeling allowing to consider the spatial variability of the physical parameters on the study area.

Chapter 3

A balanced drainage basin partitioning method

Contents

French Abstract

La modélisation hydrologique des grands bassins versants implique généralement un grand nombre de variables et un temps de calcul long. Les méthodes de décomposition de domaine ont été proposées pour faire face aux problèmes de mémoire et de temps de calcul (Kumar and Duffy, 2015) et pour implementer des algorithmes parallèles (Leandro et al., 2014).

Dans le cas d'un bassin versant, les méthodes de partition existantes proposent de diviser le cours d'eau aux confluences, puis en fonction de la longueur des tronçons de rivière, en ignorant toute information sur l'aire des sous-unités hydrologiques générées (Apostolopoulos and Georgakakos, 1997) (Vivoni et al., 2005) (Pontes et al., 2017). De plus, ces méthodes n'ont pas été conçues pour prendre en compte les structures de génie civil qui peuvent affecter le comportement hydrologique du bassin versant. Ceci a motivé le développement d'une méthode de partition des bassins versants qui prend en compte ces contraintes (Section 3.3) en considérant l'information d'accumulation des flux plutôt que des informations sur la confluence ou la longueur du tronçon.

Nous présentons dans ce chapitre un outil SIG appelé WatershedPart. Basé sur la Topotoolbox (Schwanghart and Kuhn, 2010), il permet de créer des partitions d'un bassin versant dont les sous-bassins sont d'aire très similaires (rapport de 1 à 2). Les données de cet outil sont un modèle numérique de terrain, un exutoire, et éventuellement des points correspondants aux ouvrages ou stations de mesure d'intérêt. Les performances évaluées pour le partitionnement du bassin du Moderbach montrent que notre méthode est supérieure aux méthodes de partition proposées dans la littérature (Section 3.4). Les performances en matière de simulation hydrologique sont évaluées dans les chapitres 4 et 5.

3.1 Introduction

The simulation of distributed physically-based hydrological models on large drainage basins usually involves a large number of variables and a long computation time. Domain decomposition methods were proposed to face both memory and time shortages (Kumar and Duffy, 2015) (Leandro et al., 2014). In recent years, the implementation of domain decomposition methods for hydrological models on parallel platforms has been considerably facilitated, mainly due to decreasing hardware costs and rapid development of computing power and performance.

Given a catchment, state-of-the-art partition methods (Apostolopoulos and Georgakakos, 1997) (Vivoni et al., 2005) (Pontes et al., 2017) propose to split the stream at confluences, then with respect to the reach length. Apostolopoulos and Georgakakos (1997) developed a parallel algorithm for flash flood prediction, showing the impact of the number of processors on the efficiency of parallel computing. The drainage basin was partitioned based on the segmentation of the natural water flow in the basin area. Vivoni et al. (2011) advanced the application of the distributed hydrological model "tRIBS" in a parallel framework using a basin partitioning based on the channel network segmentation.

The efficiency of parallel computing is strongly dependent on the distribution of the computational load over a multi-processor cluster (Kumar and Duffy, 2015). A spatial basin partitioning method can contribute in balancing the computational load by providing sub-basins of similar areas that are supposed to use the same size for mesh elements. This provides sub-basins with a similar number of unknowns. Therefore, they have comparable computational costs, which improves the computational performance of their simulation on a parallel platform. Partitioning methods proposed in the literature and applied in most research works, ignore any information about the area drained by a hydrological sub-unit. Furthermore, these methods were not intended to account for critical structures in the drainage basin area that can act on the hydrological behavior. This has motivated us to develop a drainage basin partitioning method that takes into account these conditions. We place a special emphasis on area balancing by considering the flow accumulation raster rather than the confluence and reach length information.

Starting from a DEM and an outlet, we introduce a GIS tool called WatershedPart. Based on the Topotoolbox functions (Schwanghart and Kuhn, 2010), this tool aims to partition a drainage basin with respect to its landscape particularities, such as engineered structures or gauging stations at stream level, and to ensure sub-basins of similar areas.

This chapter is organized as follows: first, a state of the art on the different GIS tools useful to carry out drainage basin partitioning is presented. The drainage basin partitioning methods applied in previous studies are discussed in order to highlight requirements and design opportunities. Next, the GIS tool WatershedPart is introduced. We present a drainage basin partitioning method that uses the flow accumulation information to produce an area-balanced basin partition. Finally, WatershedPart is used to partition the Moderbach basin and compare the results with the previous drainage basin partitioning methods. The performance of this domain decomposition method in hydrological simulation is evaluated in Chapters 4 et 5.

3.2 Drainage basin partitioning: tools and methods

GIS tools are widely used to support hydrological modeling, especially for distributed hydrological models requiring large amounts of data, by taking advantage of their ability to store and manage large datasets.

This section introduces and compares two efficient and well-known GIS tools, ArcGIS from ESRI and the TopoToolbox that provides a set of Matlab functions. Although both tools offer a wide set of functions for DEM analysis and watershed modeling, they are applied in different environments. Therefore, the commands, application, methods, and level of performances are different.

A brief state of the art on the previous drainage basin partitioning methods is also presented. Two of the most used methods for basin partitioning are compared and applied to the Moderbach basin showing their limitations.

3.2.1 GIS for watershed partitions

A detailed geographic and physical knowledge of the case study area is necessary to run a physically-based distributed hydrological model. The ability to model and handle the topography is mandatory since it is the major parameter that acts on the water flow.

GIS-based software are used in the research community and in engineering offices, for their interactive graphical interface and user-friendliness (Sahana et al., 2016). GIS provide tools and functionalities to apply a map processing approach to DEMs, in order to compute flow direction, flow accumulation and to derive drainage basin characteristics such as basin boundaries and stream network. As a result of advanced development in programming, software technologies, and desktop GIS capabilities some of these software may be driven by means of popular frameworks such as R, Python, or Matlab.

3.2.1.1 ArcGIS

ArcGIS is a GIS software developed by the American society ESRI (Environmental Science Research Institute), a pioneer in geographic information management in the early 2000 and the actual world market leader in GIS. ArcGIS provides the ability to manage and integrate data, perform advanced analysis, model processes, and display results on professional-quality maps. It is applied by a variety of scientific institutions and university departments to develop and illustrate groundbreaking research.

ArcGIS is well adapted to hydrological studies. It allows hydrologists to integrate a variety of data through several processes into a manageable system (Khatami and Khazaei, 2014). ArcGIS uses a number of tools and plugins such as the Hydrology toolset, ArcHydro, HEC-GeoHMS and HEC-GeoRAS to provide a geographical data analysis. The Hydrology toolset, for example, provides spatial analysis tools allowing to fill the sinks in a DEM, to compute the flow direction and flow accumulation, to delineate drainage basins and to derive stream networks (Khatami and Khazaei, 2014). The user-friendly interface have appealed to professionals in the hydrology sector (Sahana et al., 2016) since it allows to use maps, visualization, and analytical capabilities.

In the following we present the sequence of operations for drainage basin delineation in AcrGIS (Fig. 3.1).

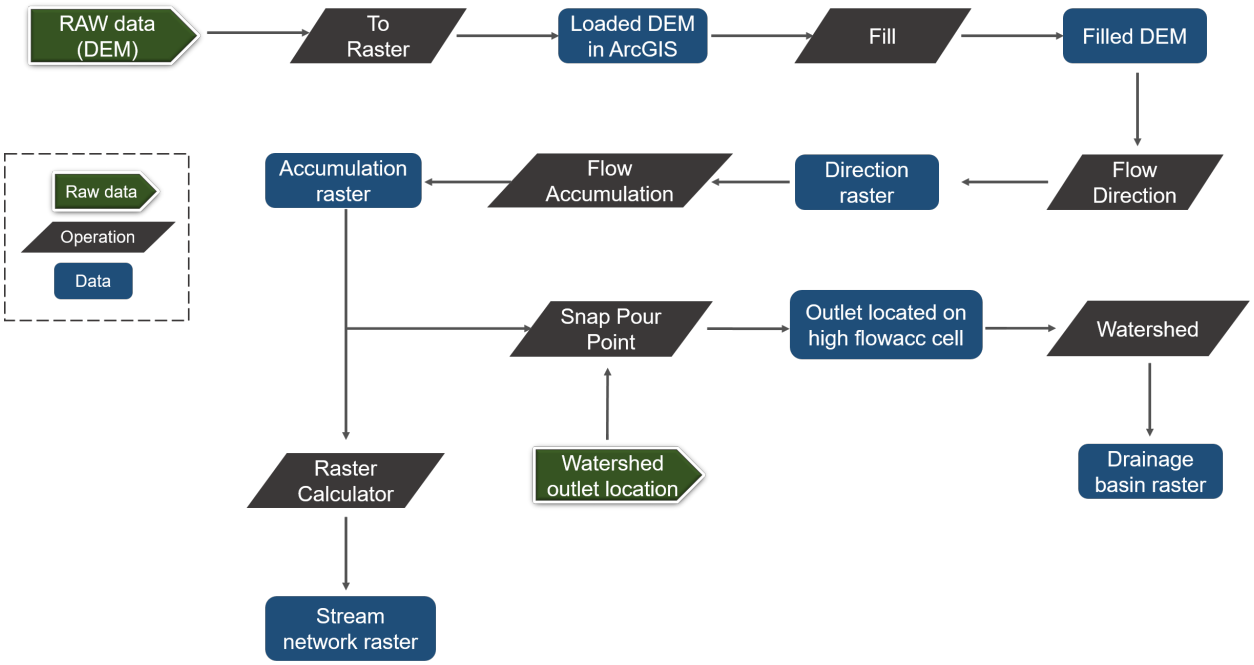


Figure 3.1: Basin delineation process using ArcGIS.

The process of watershed delineation starts by reading a DEM. The DEMs used for modeling are mainly saved in a raster format. When a DEM is stored in different formats as point shapefiles, floating-point value (.flt) files, or band interleaved by line (.bil), files it should be converted into a raster so it can be used in ArcGIS using the conversion tool ***To Raster***.

DEMs generally suffer from errors called sinks or pits, corresponding to depressions in the landscape or irregularities in the DEM. The correction of these sinks is mandatory to enable continuous flow towards the DEM edges. This correction process is provided in ArcGIS using the ***Fill*** function.

The flow direction of each cell is calculated based on the greatest slope of its eight neighboring cells using the ***Flow Direction*** tool, see Section 2.3.3. The flow accumulation is computed by accumulating the number of cells that flow into each of the DEM grid cells using the ***Flow Accumulation*** tool. The stream network is extracted based on a flow accumulation threshold using the ***Raster Calculator*** tool. The flow direction, flow accumulation and hydrographic network are saved in a raster grid format as presented in the previous chapter, see Figs. 2.3, 2.4, 2.5.

Drainage basins are delineated by determining all the cells draining into the outlet cell chosen by the user. Calculation of drainage basins may thus return a wrong basin delineation if the outlet cell is not located on a high flow accumulation cell. The ***Snap Pour Point*** function is applied to move the outlet point towards the highest flow accumulation cell, within a snap distance. This tool is applied to ensure the selection of points of high accumulated flow when delineating drainage basins. The delineation is executed using the ***Watershed*** tool.

ArcGIS is mainly designed for Windows. For Linux and Mac, it is possible to use ArcGIS in a virtualized environment. However, users will still be using a commercial closed source software with an expensive license. The ArcGIS features are accessible through the ArcGIS interface or through a Python script. Nowadays, the open source QGIS software is preferred by practitioners, notably within the French administration.

3.2.1.2 Topotoolbox in Matlab

GIS analysis in Matlab remains limited as compared to dedicated software like ArcGIS. The Topotoolbox, developed by Schwanghart and Kuhn (2010), has been proposed to overcome the limitations of the GIS Matlab toolbox useful for hydrology. The Topotoolbox is a set of Matlab functions useful for hydrological modeling in a non-GIS environment. It provides functions to manage and manipulate DEM, to compute flow direction and flow accumulation, to delineate drainage basins, and to derive stream network. The workflow of drainage basin delineation using the Topotoolbox functionalities is presented in Fig. 3.2 and listing 3.1 and illustrated on the Moderbach basin.

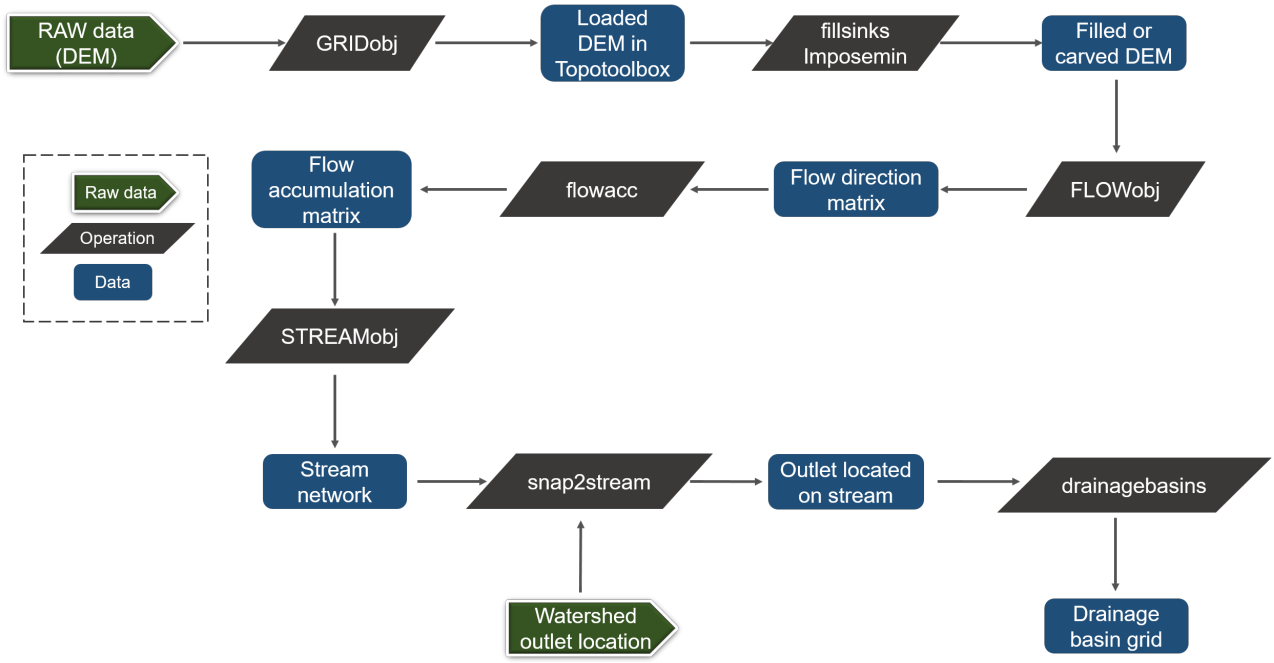


Figure 3.2: Basin delineation process using the Topotoolbox.

Listing 3.1: Drainage basin delineation.

```
% reading DEM
dem   =GRIDobj('DEM-file.asc');
% filling sinks in the DEM
dem   =fillsinks(DEM);
% computing flow direction
floDir=FLOWobj(dem,'preprocess','carve');
% computing flow accumulation
floAcc=flowacc(floDir);
% deriving stream network
stream=STREAMobj(floDir,floAcc>Delta);
% snapping basin outlet to the nearest stream
input =snap2stream(input);
% delineate drainage basin
db    =drainagebasins(floDir,input);
```

The process of basin delineation starts by reading a DEM. The Topotoolbox can read a DEM in a raster format, Ascii Grid or Geotiff, using the ***GRIDobj*** function (Fig. 3.3).

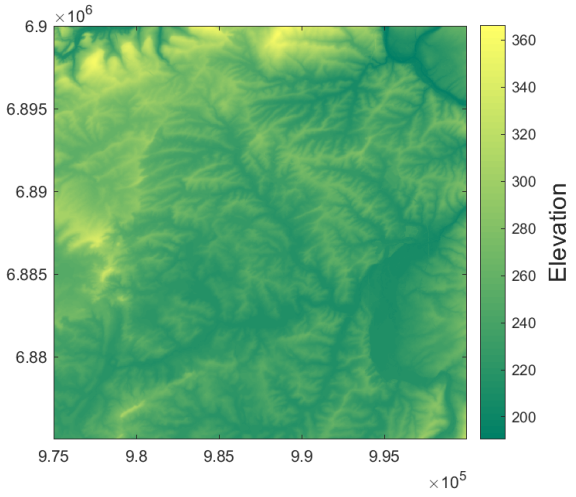
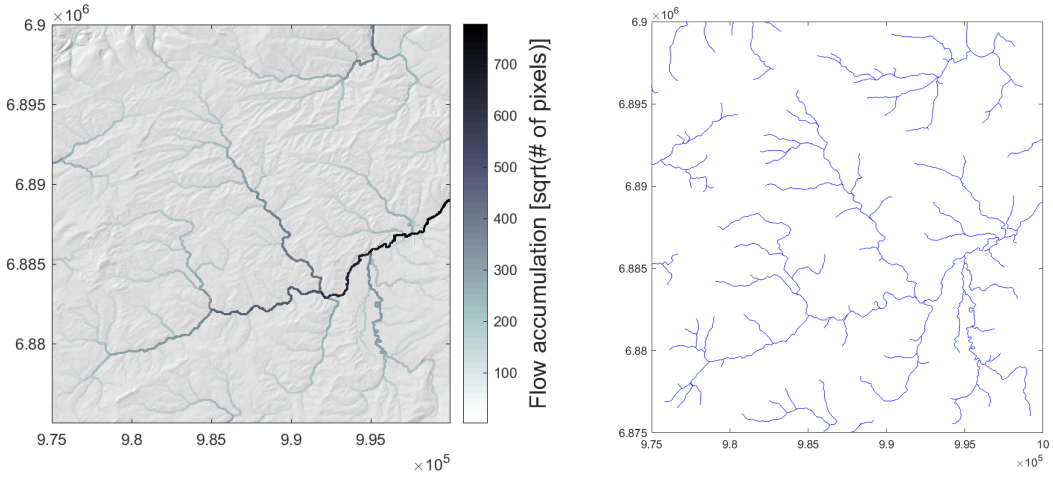


Figure 3.3: DEM of the Moderbach area (Coordinate system: Lambert-93).

For the correction of the sinks in a DEM, the ***fillsinks*** tool in the Topotoolbox is used to fill all the depressions. The Topotoolbox provides another tool to correct depressions in the DEM: the ***imposemin*** tool carves (decrease elevation) the necessary cells to ensure a continuous downward gradient. The difference between these two methods is developed in section 4.3.



(a) Flow accumulation (b) Hydrographic network

Figure 3.4: Deriving drainage basin characteristics (Coordinate system: Lambert-93).

The Topotoolbox computes the flow direction, flow accumulation and stream network using the ***FLOWobj***, ***flowacc*** and ***STREAMobj*** functions respectively, see Fig. 3.4, in the same way as in ArcGIS. These data are stored in a two column matrix format indicating, for the flow direction for example, the indices of "flow from" in one column vector and "flow to" in another column vector.

For drainage basins delineation, the ***snap2stream*** function is applied to snap the defined outlet to the nearest stream. The delineation is executed using the ***drainagebasin*** tool (Fig. 3.5).

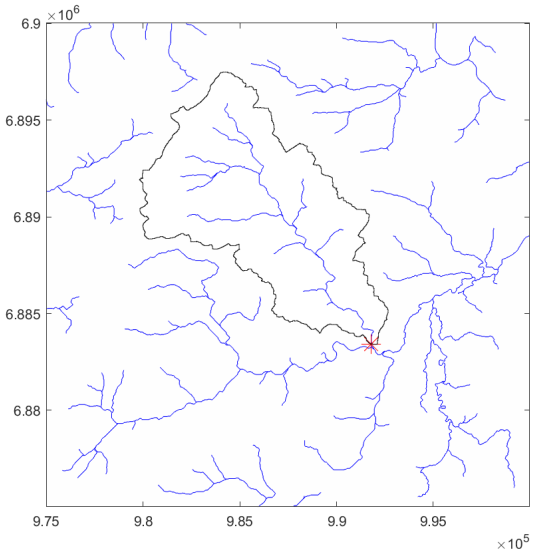


Figure 3.5: Delineation of the Moderbach basin (Coordinate system: Lambert-93).

As demonstrated in Schwanghart and Scherler (2014), the Topotoolbox was compared to other GIS tools available for watershed delineation as ArcGIS 10.1, TauDEM 5, SAGA GIS 2.1, and WhiteboxGAT 3.0.8 showing a better performance with respect to time and memory space management. The processing time to fill the sinks in a DEM, compute the flow direction or the flow accumulation, is less in MATLAB than it is in ArcGIS. Running the ***Fill*** tool, for example, in ArcGIS can be memory, CPU, and disk intensive. It can require up to four times the disk space of the input raster.

For all that reasons, the Topotoolbox was chosen for this research work. While the interface may not feel as intuitive to newcomers (Fig. 3.6) as ArcGIS, the functions of this toolbox can be easily modified. Although ArcGIS may have a better quality visualization process compared to the Topotoolbox, coupling the Topotoolbox with the image processing and mapping toolboxes improves the visualization and quality of the maps (Sahana et al., 2016). The basin characteristics extracted from each of the two GIS tools, such as basin boundaries and hydrographic networks, can still be read/used in the other tool by saving the data in a shapefiles format.

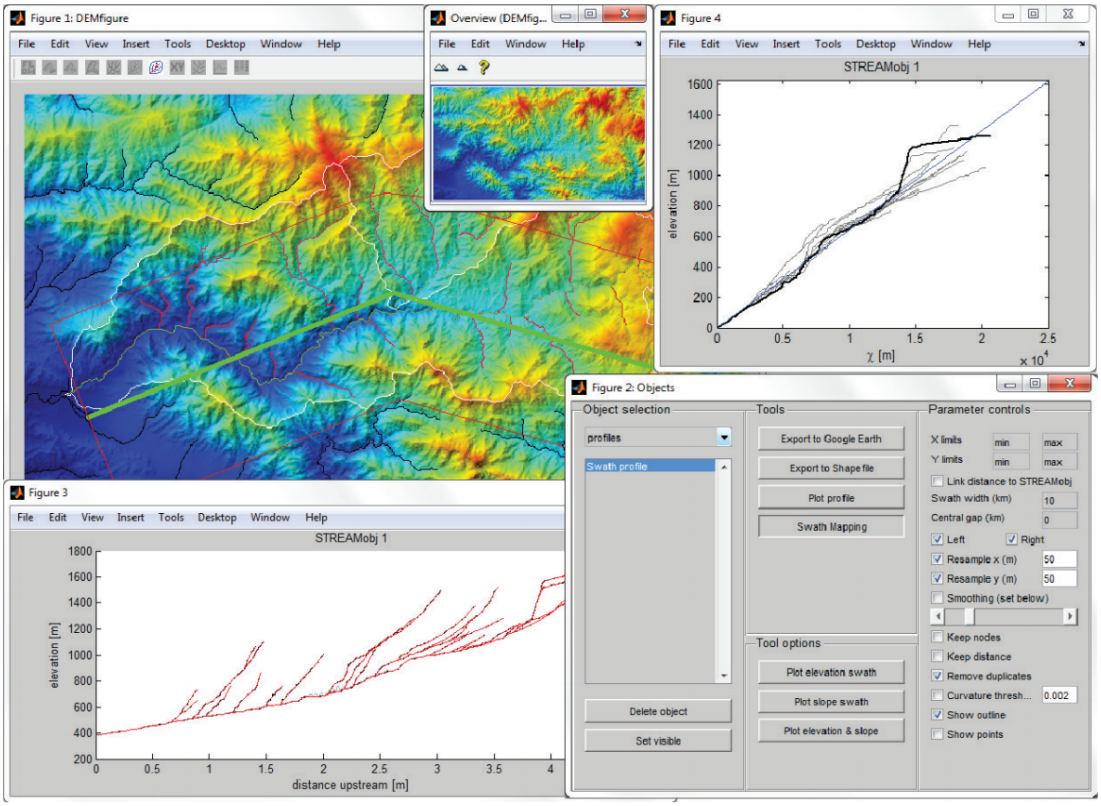


Figure 3.6: Example layout of topoapp, a graphical user interface for the Topotoolbox (Schwanghart and Scherler, 2014).

3.2.2 State-of-the-art partition methods

The most widely basin partitioning method applied in research splits the basin at the stream confluences. Due to its tree-like shape, the stream network can be segmented into river reaches. A river reach corresponds to a drainage segment located between two successive confluences or between a confluence and its outlet. A sub-basin is delineated as an incremental drainage area that contributed directly to a river reach that resides in it. Therefore, the number of sub-basins will be equal to the number of river reaches (Wang et al., 2009).

To define the required number of sub-basins, the stream network is derived with respect to a minimum drained cells threshold denoted by `Delta`. In other words, each DEM cell will be considered as a part of the stream network if it drains a number of cells greater than `Delta`. Therefore, the more the value of `Delta` decreases, the more reaches appear in the stream network providing more sub-basins. Figure 3.7 presents the nested partitions of the Moderbach basin produced for `Delta` of 20,000; 15,000 and 9,200. The characteristics of the nested partitions are reported in Tab. 3.1.

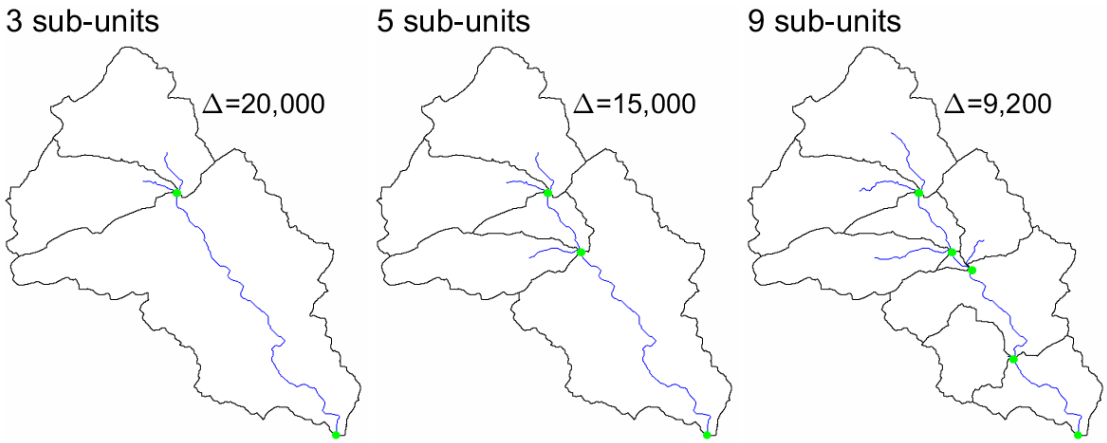


Figure 3.7: Confluence partitioning method applied to the Moderbach basins.

The wide use of this method is owed to its ease of application. Many software such as HEC-GeoHMS in ArcGIS (Doan, 2000) or the Topootolbox (Schwanghart and Kuhn, 2010), allows an automatic segmentation of the stream network, to perform a drainage basin partitioning at the confluences, see listing 3.2. The ***split*** function from the Topotoolbox was used to perform several confluence partitions of the Moderbach watershed using different flow accumulation thresholds (Fig.3.7). The ***split*** function allows to break the stream network at the confluences into river reaches and therefore a sub-basin will be associated to each reach.

Listing 3.2: Confluence partitioning method.

```
% reading DEM
dem         =GRIDobj('DEM-file.asc');
% computing flow direction
floDir      =FLOWobj(dem,'preprocess','carve');
% computing flow accumulation
floAcc      =flowacc(floDir);
% deriving stream network
stream      =STREAMobj(floDir,floAcc>Delta);
% splitting stream network at confluences
breakstream=split(stream);
% generating drainage basins with respect to confluences
db          =drainagebasins(floDir,breakstream);
```

As the number of confluences increases, there is an increasing probability for a very small sub-unit in the partition. Consequently, the ratio between the maximum and minimum areas increases. The application to the Moderbach basin shows that partitioning into 9 and 23 sub-basin provides very small sub-basins of 0.4 and 0.2 km^2, while the largest sub-basins have an area of 17.4 and 11.5 km^2 respectively, see Tab. 3.1.

The computation time for the confluence method corresponds to the time needed to proceed the DEM and to partition the basin using the split function.

Properties	Confluence partitions				Reach length partitions
Δ	20,000	15,000	9,200	3,100	3,100
#sub-units	3	5	9	23	34
Max. area	58.1	42.0	17.4	11.5	6.7
Min. area	13.6	4.9	0.4	0.2	0.2
Ratio	4.3	8.6	37.0	47.5	27.7
Mean area	29.7	17.8	9.9	3.8	2.6
Std dev.	24.9	14.2	5.6	2.7	1.5
Time (s)	3.7	3.8	3.9	4.2	10.9

Table 3.1: Characteristics of the confluence and reach length partitioning methods. Area expressed in km^2.

As an improvement to the confluence partitioning method, Pontes et al. (2017) uses the reach length as a criterion for decomposition, which reduces the sizes of the large sub-basins. This method is offered in IPH-HydroTools which is a plugin for MapWindowGIS, a free and open-source GIS software (Ames et al., 2008). We have implemented the reach length method in Matlab using the Topotoolbox functions considering a reach length condition of 2 km between two successive outlets. Figure 3.8 shows the improvement applied to the confluence method with the reach length method. Sub-basins concerned are filled in orange.

Listing 3.3: Confluence partitioning method.

```
% set the reach length threshold to 2 km
reachthreshold = 2000;
% count the number of sub-basins
% provided by the confluence method
nrDB = numel(db);
% for each sub-basin
for i = 1:nrDB
   % derive its flow direction
   cropedfloDir = crop(floDir,db(i));
   % isolate its stream (reach)
   reach =    STREAMobj(cropedfloDir,floAcc>Delta);
   % compute the length of the reach inside the basin
   dist = reach.distance;
   reachlength = max(dist);
   % compute the number of outlets needed to further
   % partition the reach
   NBoutlets=floor(reachlength/reachthreshold);
      % for each outlet to be added
      for i=0:NBoutlets
      % derive the outlet index based on the reach
      %  length threshold distance
      [outletdist,outletindice]= min(dist - reachthreshold*i);
      % derive the geographical coordinates of the outlet
      outletreachx=reach.x(outletindice);
      outletreachy=reach.y(outletindice);
      % add the new outlet to previous ones
      outletx = [outletx outletreachx];
      outlety = [outlety outletreachy];
      end
end
% generate drainage basins based on the confluences
% and reach length outlets
db=drainagebasins(floDir,outletx,outlety);
```

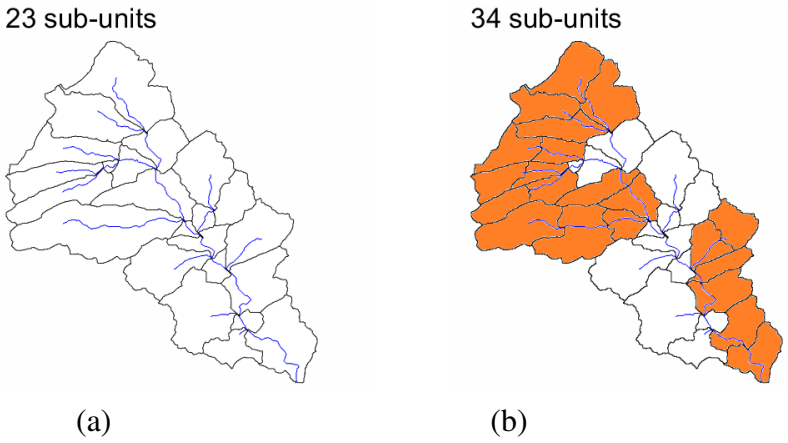


Figure 3.8: (a): Partition with the confluence method, (b): Partition with the reach length method (Bassins filled in orange correspond to the additional partitions generated by considering a reach length of 2 km).

The reach length method improves the confluence partitioning method. The large sub-basins provided by the confluence method were decomposed with the 2 km reach length condition allowing to reduce the maximum basin area from 11.5 to 6.7 km^2 and the standard deviation from 2.7 to 1.5 showing a closer distribution of sub-basin areas near their mean value. However, this method does not affect the size of the very small sub-basins. The smallest drainage basin of 0.2 km^2 is still there. The ratio between the maximum and minimum areas is still significant and the area balance is not satisfactory. Moreover, having very small basins for surface flow simulations may result in a spurious representation of the landscape due to the existence of a large number of engineered structures that can affect the hydrological behavior.

3.3 WatershedPart

The literature review and the decomposition performed on the Moderbach basin in the previous section shows that the former two methods, which may be viewed as zero-dimensional and one-dimensional since they are based on confluence points and reach lengths information, cannot guarantee an area-balanced partition of the watershed.

This has motivated the development of WatershedPart, a GIS tool that provides a drainage basin partitioning method delivering sub-basins of similar areas based on the flow accumulation infomation (Hariri et al., 2019). This partitioning method also allows to account for the topographical features of the terrain such as the engineered structures in the study area. WatershedPart is implemented in Matlab, relying on the Topotoolbox functions which have shown a better efficiency in terms of computation time and memory management compared to other GIS software (Section 3.2.1.2).

3.3.1 Positioning of WatershedPart

3.3.1.1 Sub-basins of similar area

Kumar and Duffy (2015) experimented the role of domain decomposition on the efficiency of parallel simulations in hydrological modeling. From a computational point of view, finite element models that are applied to sub-units with similar area are assumed to involve linear systems with a similar number of unknowns. Thereby, their solution has a comparable computational cost, and the domain decomposition method is more efficient. This implies the use of a drainage basin partitioning method allowing to partition a basin into sub-basins of similar areas having similar computational costs. However, it should be mentioned that the area is not the only parameter that can be used to equally distribute the computational load. Other parameters such as the slope, the basin shape and the elements size can affect the computational load.

As the drainage basin partitioning methods found in the literature do not guarantee the balance of the sub-basin areas, the resulting linear systems will generally be of very different sizes and the efficiency of the parallelization will be limited. As an improvement to that, we introduced a drainage basin partitioning method that ensures a balance of the sub-basin areas allowing to enhance the efficiency of parallelization.

3.3.1.2 Consider landscape particularities

As DEMs are only an approximation of the topography, they may underestimate the significant impact of engineered structures on the water flow. Structures that can act on the water flow and may yield water retention in the watershed area such as dams, bridges, ponds, or highways, should be considered during the partitioning procedure by setting outlets near to them. Note that landscape features that can act on the surface flow behavior can be identified by the intersection of stream and road networks. This process is presented in the Appendix B. Gauging devices may also be considered in the basin partition, allowing comparison of computed and observed flows at these locations.

3.3.1.3 Matlab framework

Matlab is a classical and well-documented scientific computing framework widely used in academic and engineering communities. Its well-organized documentation and visualization functionalities ease the fast development and prototyping of numerical methods. An important motivation for developing WatershedPart in the Matlab environment is the open-source aspect of the Topotoolbox. It allows to modify and enhance its functionalities in order to satisfy user objectives (Schwanghart and Kuhn, 2010). As an additional motivation, the Topotoolbox outperforms other GIS tools in terms of time and memory space management.

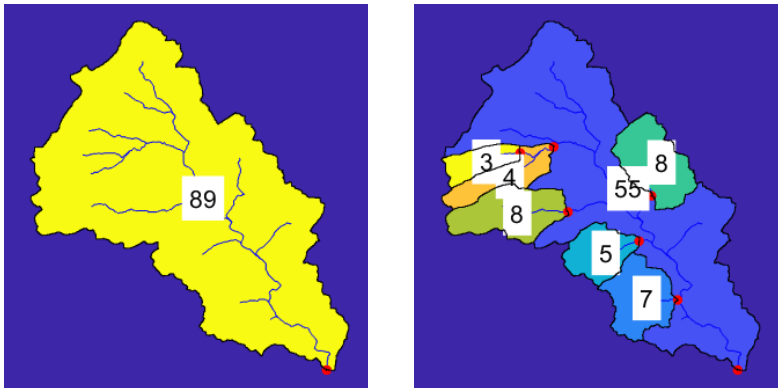
Although different GIS software with user-friendly graphical interface (GUI) provide tools useful for DEM analysis and watershed characterization, programming environments such as Matlab may be preferred since it already include large libraries for various computing tasks. Working in a non-GIS scientific computing framework may facilitate theoretical developments as well as the uptake by engineering students.

3.3.2 WatershedPart Workflow

WatershedPart articulates several functions of the Topotoolbox, providing a drainage basin partitioning method that satisfies landscape particularities and equally-sized sub-basins condition.

After proceeding with the drainage basin delineation as presented in Fig. 3.2 and in listing 3.1, the proposed method offers the possibility to choose user-defined outlets to proceed with a guided partition. Then, further decompositions are performed in order to ensure sub-basins of similar areas.

The Moderbach basin is selected to perform a guided area-balanced partition using WatershedPart. Once the Moderbach basin is delineated (Fig. 3.9(a)) following listing 3.1, the user-defined outlets that represent engineered structures of interest are introduced to proceed with the guided drainage basin partitioning method. The ***drainagebasins*** function will determine the area draining to each of these outlets. In our case study, the user-defined outlets are located near the six ponds in the Moderbach basin, see Fig. 1.7. This partition provides six sub-basins corresponding to the six ponds, and a large sub-basin comprising the main stream (Fig. 3.9(b)).



(a) Initial case study basin.

(b) Basin partition with respect to user-defined outlets.

Figure 3.9: Initial basin delineation and user-defined partition (The numbers displayed in the basins correspond to their surface).

The area-balanced partition method builds on the partition generated with the guided partition. It allows for further decompositions following listing 3.4, providing sub-basins of similar areas. Figure 3.10 presents the detailed workflow of the drainage basin partitioning method in WatershedPart.

Listing 3.4: Area-balanced partition.

```
% add user-defined outlets
outletPts=<user-def outlets>;
% generate guided partition
db=drainagebasins(floDir,outletPts);
% compute basin areas and select the largest basin
[maxSubA,imax]=polyarea(db);
% when a sub-basin area is greater than the max area
while (maxSubA>maxArea)
   % compute flow accumulation on the largest basin
   subFloAcc=flowacc(db(imax));
   % select the cell that drains half of its max flow
   [halFA,iSplitPt]=max(subFloAcc)/2;
   % derive the coordinates of this split cell
   [outletPts]=[outletPts coor(iSpliPt)];
   % generate a basin partition with the new outlet
   db=drainagebasins(floDir,outletPts);
   % check if a basin still exceeds the max area
   [maxSubA,imax]= polyarea(db);
end
```

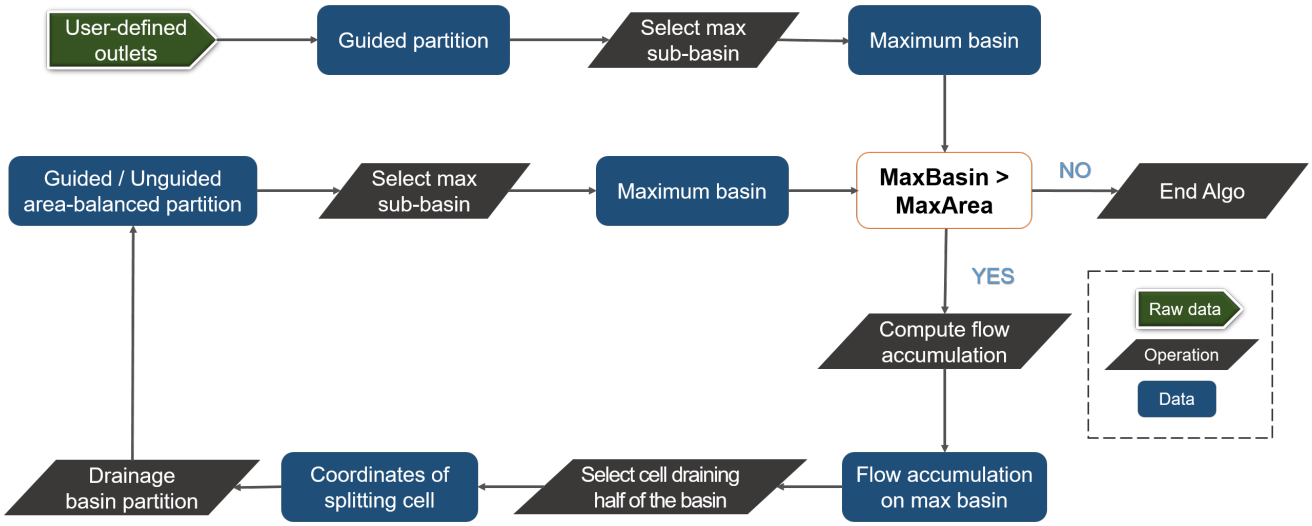
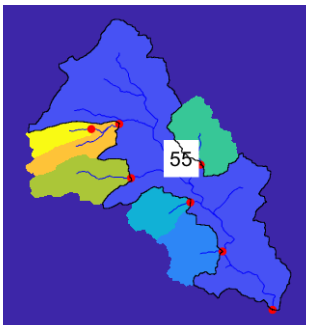
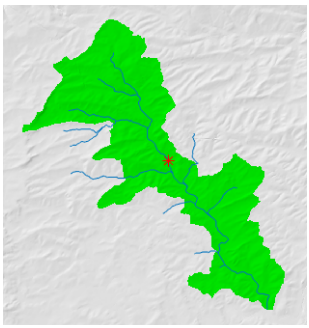


Figure 3.10: Area-balanced drainage basin partitioning method workflow.

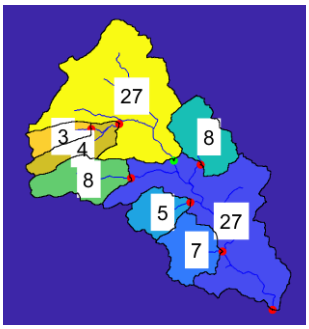
The areas of the sub-basins provided by the guided partition are computed to select the largest one with index (`imax`), see (Fig. 3.11(a)). The decomposition of the largest sub-basin into two equally-sized sub-basins is based on the flow accumulation information. Each cell of the flow accumulation raster provides the number of upstream cells it drains. This parameter is thus related to the area of its upstream basin. Partitioning the selected basin into 2 sub-basins of similar areas consists in determining the cell that drains half of the basin's surface. Therefore, after the selection of the largest sub-basin, its flow accumulation information is computed and a splitting point is determined by searching for the cell that drains half of that flow (`coor(iSplitPt)`), see (Fig. 3.11(b)). The ***drainagebasins*** function is applied to provide a new basin partition taking into account the determined outlet in order to decompose the largest sub-basin into 2 sub-basins of similar areas (Fig. 3.11(c)).



(a) Select the largest sub-basin.



(b) Determine the cell draining half of the basin.



(c) Partition the largest basin into 2 sub-basins of similar areas.

Figure 3.11: Area-balanced partition (The numbers displayed in the basins correspond to their surface).

This process is iterated until the desired stopping criterion is met. This can be based on the number of sub-units and/or on the sub-unit areas. In listing. 3.4, the process stops when all the sub-basins have an area less than a prescribed maxArea.

3.4 Results

In the following, WatershedPart is applied to generate nested partitions of the Moderbach basin.

3.4.1 Guided area-balanced drainage basin partitioning method

The guided method is carried out by starting from the 7 user-defined outlets (red points) that are located at the 6 pond outlets and the initial basin outlet. The generated sequence of nested partitions is illustrated in Fig. 3.12 and partition properties are reported in Tab. 3.2. Clearly, the ratio between the maximum and minimum areas is strongly influenced by the user-defined outlets. The ratio decreases from 18.3 to 9.1 because the partitions into 7 and 8 sub-basins involve very large sub-basins. However, further decompositions into 16, 32 and 64 with respect to the area balance condition (green points) reduces the ratio to range between 2 and 3. Note that the time of a few seconds required to generate these decompositions is very small compared to the computational time required to perform hydrological simulations using these decompositions.

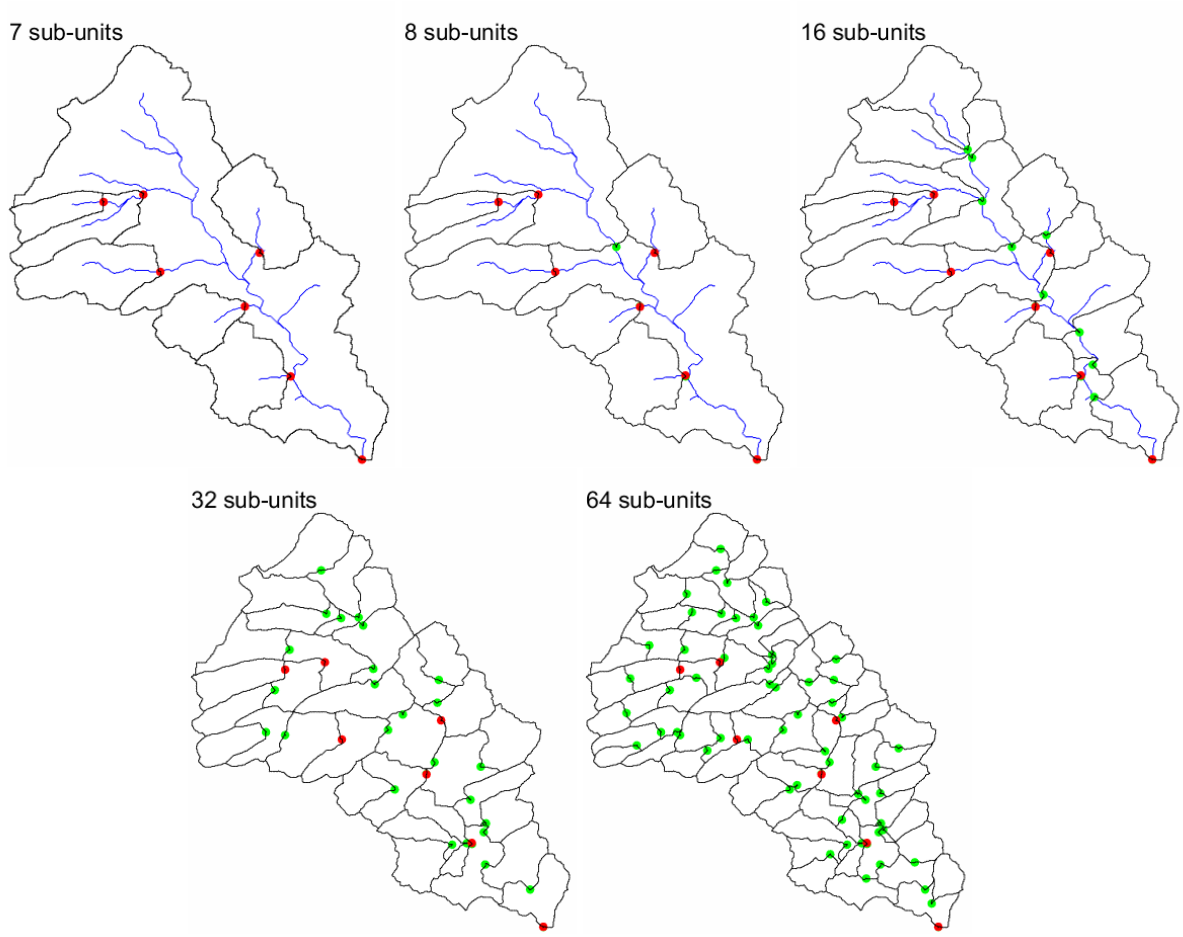


Figure 3.12: Guided iterative area-balanced drainage basin partitioning method (user-defined outlets in red and area balancing outlets in green).

Properties	Guided area-balanced partition				
#sub-units	7	8	16	32	64
Max. area	57.9	27.4	7.6	3.9	2.0
Min. area	3.0	3.0	2.9	1.7	0.7
Ratio	18.3	9.1	2.6	2.1	2.7
Mean area	12	11.1	5.6	2.7	1.4
Std dev.	18.7	10.2	1.6	0.6	0.3
Time (s)	4	6.2	9.5	14.9	26.7

Table 3.2: Characteristics of the guided area-balanced drainage basin partitioning method. Area expressed in km^2.

3.4.2 Unguided area-balanced drainage basin partitioning method

The unguided method only requires the initial basin outlet as an input. The nested partitions into a number of sub-basins that is a powers of 2 are plotted in Fig. 3.13. Corresponding properties are reported in Tab. 3.3. The different decomposition configurations show a very low ratio, ranging between 2 and 3. This method fulfills its objective by delivering sub-basins of similar area.

Properties	Unguided area-balanced partition					
#sub-units	2	4	8	16	32	64
Max. area	47.1	30.0	15.8	8.1	4.0	2.0
Min. area	42.1	17.1	8.1	3.8	2.0	0.7
Ratio	1.1	1.8	1.9	2.1	2.0	2.8
Mean area	44.6	22.3	11.1	5.6	2.8	1.4
Std dev.	3.5	5.5	2.7	1.3	0.6	0.3
Time (s)	6.2	7.2	8.9	11.4	16.9	28.9

Table 3.3: Characteristics of the unguided area-balanced drainage basin partitioning method. Area expressed in km^2.

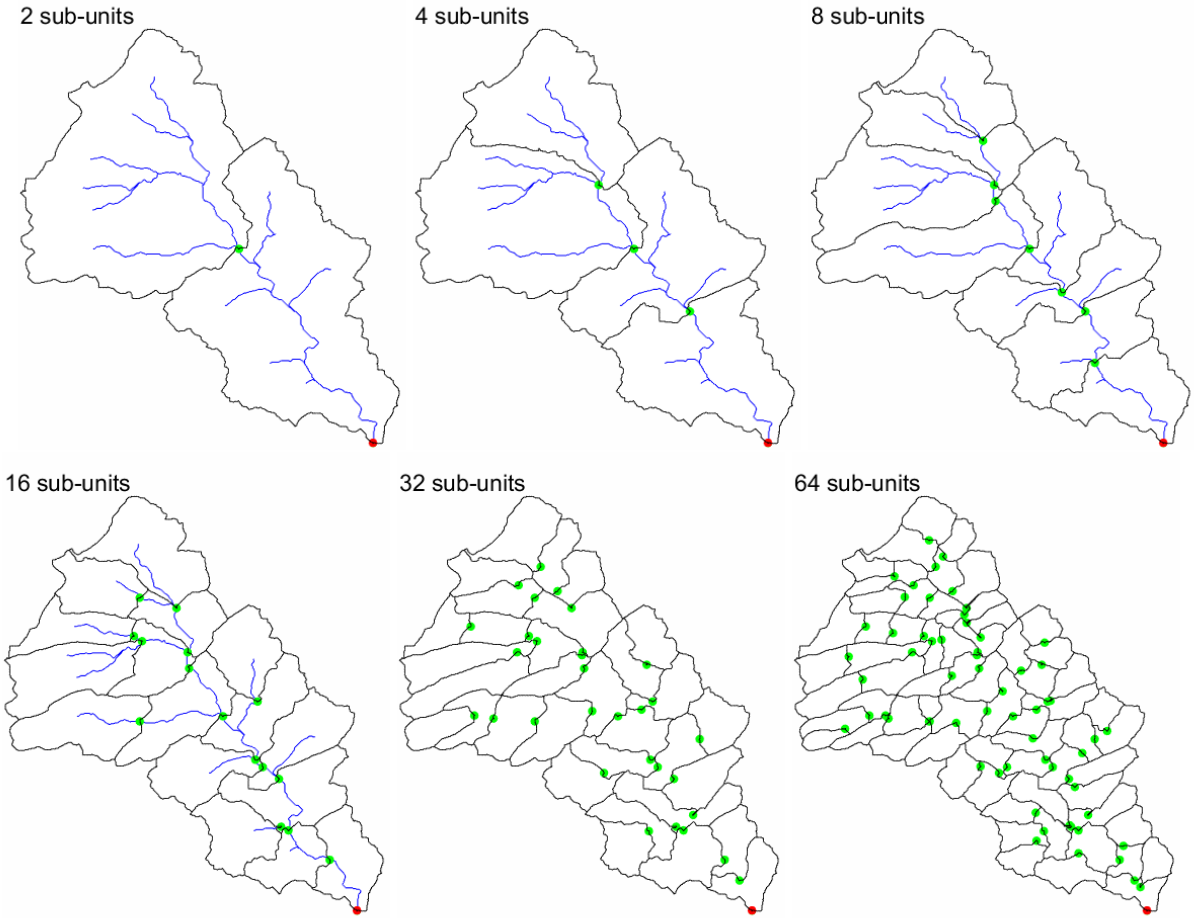


Figure 3.13: Unguided iterative area-balanced drainage basin partitioning method (user-defined outlets in red and area balancing outlets in green.

3.4.3 Comparison of partition methods

Domain decomposition methods (DDM) are applied to distribute the computational load. As discussed earlier in this chapter, the basic watershed partition offered by GIS tools allows to delineate sub-basins based on the confluence points. When the stream network presents nearby confluences, this partition method may create very small sub-basins. On the other hand, long reaches may result in large sub-basins. The confluence method may be improved by dividing the longest river reaches to delineate smaller sub-basins of the larger ones. However, this does not remedy the problem of very small sub-basins. These methods cannot guarantee an area-balanced partition of a watershed.

The area-balanced drainage basin partitioning method provided by WatershedPart enables to initiate user-defined outlets that correspond to landscape features, such as engineered structures or gauging stations to be taken into account in the decomposition (guided) or it can perform an area balanced drainage basin partitioning method without any need for user intervention (unguided).

The contribution of WatershedPart in the GIS framework for hydrological computations is the automation of the drainage basin delineation process in an iterative manner, resulting in a drainage basin partitioning method with equally-sized sub-basins. This condition is satisfied by considering the flow accumulation raster (2D information) rather than the reach length (1D information) or the confluence information (pointwise information) to partition the drainage basin.

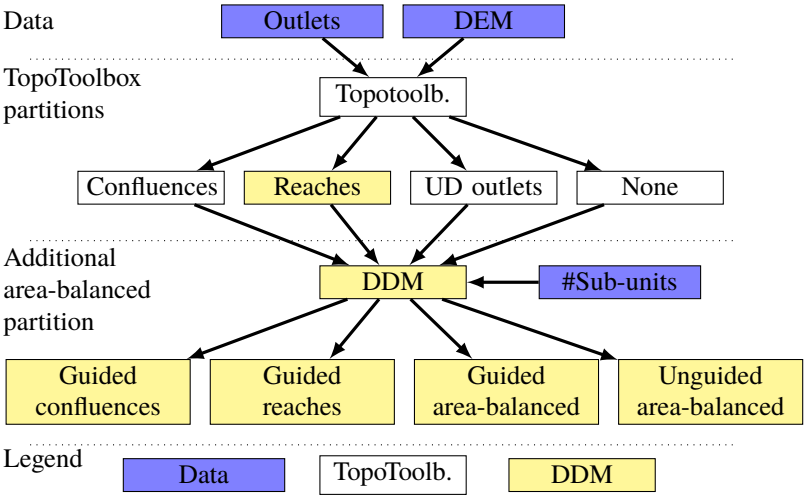


Figure 3.14: Drainage basin partitioning method workflow.

The basin partition methods are compared on the case study of the Moderbach watershed. Figure 3.15 plots the partitions performed with the two state-of-the-art methods (confluences, then a reach length of 2 km between two outlets) and the area-balanced methods proposed in WatershedPart (guided and unguided).

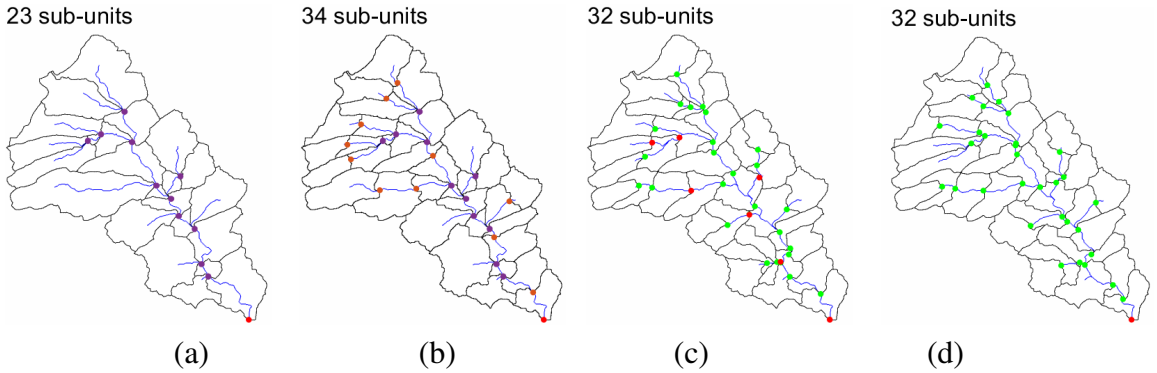


Figure 3.15: Comparison of area-balance of different basin partitioning methods a: Partition at confluences (magenta dots), b: Additional partition with respect to a reach length of 2 km (orange dots), c: Guided (red dots) area balanced (green dots) partition partition, d: Unguided area balanced partition partition (green dots).

The characteristics of the partitions are reported in Tab. 3.4. On the one hand, the confluence method and the reach length method produce very small sub-watersheds. In both cases, the ratio between the maximum and the minimum area is larger than 20. The balance of areas is not satisfactory. On the other hand, the ratio of 2 obtained for the partition methods provided by WatershedPart (guided and unguided) demonstrates the balance of the sub-watershed areas. Note that the computation times are very small as compared to the computation time needed for rain-on-grid simulations of these basins.

Properties	(a) Confluence	(b) Reach	(c) Guided	(d) Unguided
#sub-units	23	34	32	32
Max. area	11.5	6.7	3.9	4.0
Min. area	0.2	0.2	1.7	2.0
Ratio	47.5	27.7	2.3	2.0
Mean area	3.8	2.6	2.7	2.8
Std dev.	2.7	1.5	0.6	0.6
Time (s)	4.2	10.9	14.9	16.9

Table 3.4: Characteristics of the partitions displayed in Fig. 3.15 and computation times. Areas expressed in km^2.

To summarize, the two area-balanced methods provide similar performances when the number of sub-basins is large. This area-balanced partitioning method outperforms the partitioning method used in previous work in terms of providing a balanced partition. The ratio between the maximum and minimum area provided by WatershedPart is much smaller than the one provided by the confluence or the reach length drainage basin partitioning methods.

3.4.4 Improvements for simulation

Depending on the actual topography or its digitization, irregularities can appear in the shape of the basins delineated. When carrying out several basin delineations at different locations, it has been noted that for some basins, one or several cells near to the basin's outlet are considered isolated closed cells and independent from the drainage area. Obviously, defects with respect to the 4-neighbor connectivity have to be corrected before attempting any finite element modeling (Chen and Horng, 2010).

At the end of the partition process a connectivity check may be performed to correct the decomposition by moving some of the outlet points upward. This can be achieved automatically by inspecting the properties of each sub-unit by means of the `regionprops` Matlab function, for instance.

These defects appeared in both confluence and area-balanced basin partitioning method. It worth mentioning that these irregularities also appeared while performing domain partitioning in other GIS software as ArcGIS or QGis.

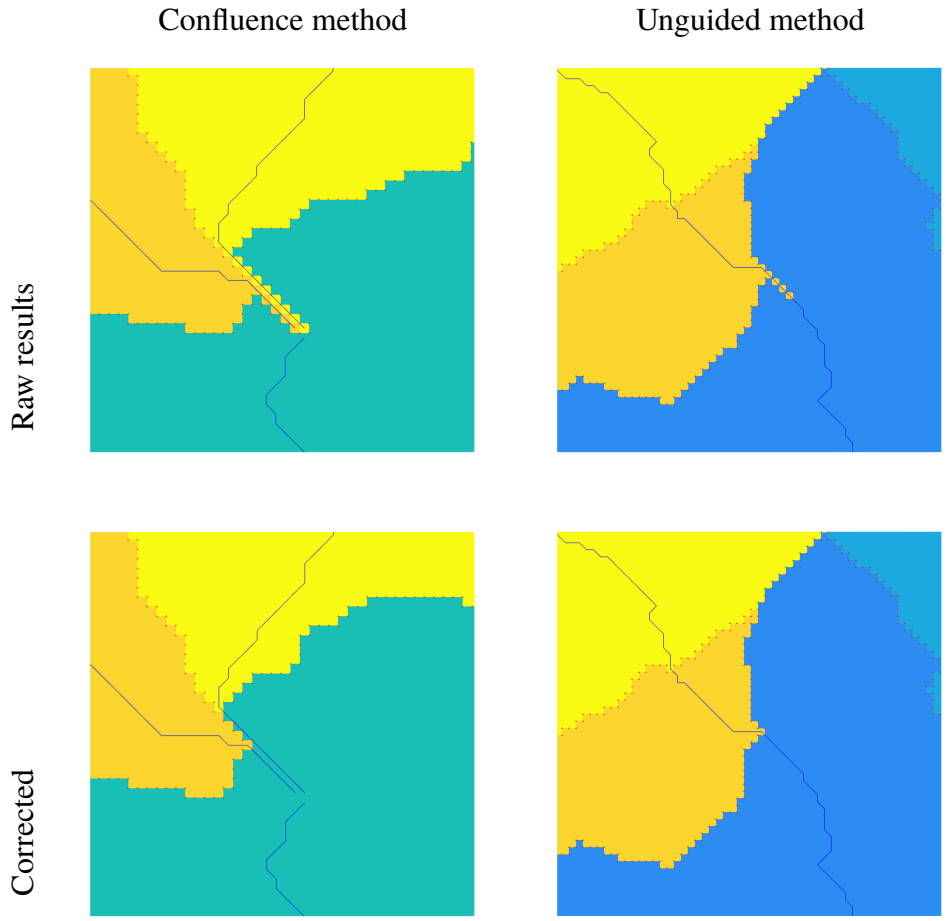


Figure 3.16: Isolated cells near to the outlet. Left: Decomposition at the confluences. Right: Unguided area-balanced partition.

3.5 Conclusion

Targeting domain decomposition for rain-on-grid simulations has motivated the design of the new partitioning method. To accommodate parallel computing, a special emphasis is put on area balancing by considering the flow accumulation raster (2D information) rather than the reach length (1D information) or the confluence information (point-wise information).

While software like ArcGIS or QGIS, providing graphical user interfaces GIS, are precious tools to support hydrological modeling, the Topotoolbox is a set of Matlab functions dedicated to DEM analysis, with a special focus on hydrology. The Topotoolbox was choosen to develop the new area-balanced drainage basin partitioning method due to its various advantages : (1) it provide tools useful for hydrological modeling (2) it works in a non-GIS environment, Matlab, which is familiar within research, education, and engineering communities (3) its open access aspect allows to modify and enhance its functions (4) it outperforms conventional GIS tools, in particular ArcGIS, in terms of time and memory management.

The partition method we proposed aims to account for engineered structures as user-defined outlets and consider small, but homogeneous hydrological sub-units. The proposed method ensures a balance of the sub-unit areas and obtaining a prescribed number of sub-units which improves the efficiency of the parallel simulations. This method allows to perform simulations with a reasonable number of unknowns per sub-unit which enables the opportunity for rain-on-grid simulations on very large watersheds, see Chapter 5.

Chapter 4

Rain-on-grid simulations on large drainage basins

Contents

French Abstract

Les modèles de précipitation distribuée sur un bassin versant impliquent la résolution d'équations aux dérivées partielles (EDP) non linéaires à l'aide d'une méthode par éléments finis ou par volumes finis, par exemple, discrétisées au moyen d'un maillage porté par la topographie de ce bassin. Le temps de calcul et les besoins en mémoire peuvent être importants car ils dépendent de la surface du bassin versant, de la taille des éléments de maillage, de la raideur des équations non linéaires discrètes et de la durée de la simulation.

Deux stratégies de calcul basées sur une répartition de la charge de calcul peuvent être mises en œuvre pour réduire les temps de simulation et parfois l'espace mémoire nécessaire à la simulation. La première stratégie utilise les capacités des processeurs multi-cœurs disposant d'une mémoire partagée. Dans ce cas, un solveur parallèle creux peut être utilisé pour diviser la matrice du système linéaire en sous-systèmes matriciels qui seront traités par les différents cœurs d'un même processeur. La deuxième stratégie est une méthode de décomposition de domaine mise en œuvre pour partitionner le domaine physique, ici le bassin versant, en un certain nombre de sous-bassins, puis organiser la résolution des sous-problèmes d'EDP dans chaque sous-bassin, et les nécessaires communications à leurs frontières. Ces sous-domaines peuvent être traités en parallèle sur un multi-processeur à mémoire distribuée sous réserve d'assurer la communication des données et la synchronisation entre les différentes simulations .

Les modèles de précipitation distribuée impliquent un maillage de la topographie du bassin versant représenté par un modèle numérique de terrain (MNT). N'étant qu'une approximation, il est souvent nécessaire de corriger ce MNT pour éliminer les dépressions et assurer un écoulement d'amont en aval sans obstacles. La présence de ponts sur les routes permettant l'écoulement n'est pas prise en compte dans un MNT. Ces ouvrages (routes ou barrages), dont l'altitude est supérieure à celle du terrain, agissent alors dans le MNT comme des structures artificielles qui empêchent l'écoulement de l'eau. Les outils disponibles dans les logiciels SIG ne sont pas toujours satisfaisants, créant des surfaces plates dans le MNT, ou exagérant la pente du cours d'eau localement (Lindsay and Creed, 2006). Nous présentons une nouvelle méthode hybride (Section 4.3.2.2) combinant les méthodes classiques pour corriger le MNT et construire des modélisations de l'écoulement de surface plus robustes.

Deux modèles hydrologiques, HEC-RAS (Brunner, 2016) et NIHM (Jeannot, 2018) sont présentés dans ce chapitre. Ces deux logiciels implémentent une modélisation de l'équation d'onde diffusive sur un bassin versant en présence d'une précipitation distribuée. D'un côté, HEC-RAS dispose d'un solveur parallèle creux et est principalement utilisé sur un processeur multicœur pourvu de Windows. Ce solveur a permis une accélération d'un facteur de 2,5 impliquant 6 cœurs pour la résolution du système linéaire du bassin de Moderbach (Section 4.2.2). De l'autre côté, NIHM a été modifié pour mettre en œuvre notre méthode de décomposition de domaine sur un multi-processeur. Les simulations réalisées avec NIHM et la méthode de décomposition de domaine ont montré des résultats satisfaisants par leur application sur le bassin de Moderbach. Cette stratégie a permis de réduire le temps de calcul d'un facteur 2 en passant d'une décomposition de 8 à 16 sous-bassins (Section 4.4.1). On observe cependant que la méthode des volumes finis telle qu'implémentée dans HEC-RAS est plus robuste numériquement et moins coûteuse en temps que la méthode des éléments finis implémentée dans NIHM pour des bassins versants de taille raisonnable.

4.1 Introduction

Performing rain-on-grid simulations on large drainage basins is not a common practice (Section 2.4.2.4). Most flood modeling works use flow or stage hydrographs as input data to model flood events (Quirogaa et al., 2016) (Surwase et al., 2019). In that case, the computational domain corresponding to the inundated area is limited to the river bed. This strategy is usually applied to reduce the computational time since it strongly depends on the computational area. However, the use of precipitation data to model floods can provide a more detailed description of the event since it considers the entire drainage basin as the computational domain. It allows for a more accurate consideration of properties that have a significant impact on surface flow such as soil parameters, land use and land cover over the entire basin.

Rain-on-grid software implement the solution of nonlinear partial differential equations (PDE) by means of a finite element or a finite volume method, for instance, and some structured or unstructured mesh that describes the watershed topography. Computing time and memory usage may be important as they depend on the catchment surface, the mesh size, and the stiffness of the nonlinear discrete equations to be solved in an iterative manner.

Two computing strategies may be implemented to balance the computational load in order to improve the performance of the model. On the one hand, a sparse parallel solver may be implemented to split the linear system into small blocks in order to distribute them on different cores. This approach is particularly well suited to a shared memory multi-core processor. On the other hand, a domain decomposition method, such as those presented in the previous chapter, may be set up to partition a catchment into a number of hydrological sub-units (Apostolopoulos and Georgakakos, 1997) (Vivoni et al., 2011). The PDE problem is decomposed into a number of boundary value sub-problems, accordingly. A domain decomposition method then organizes the solution by gathering the solutions computed for the sub-units and by setting relevant boundary conditions at the interfaces. This approach is efficient on a multi-processor cluster assuming a good balance of the computational effort.

The application of rain-on-grid models involves meshing the watershed topography represented by DEMs. As DEMs are only approximations of the topography, this may require a pre-processing method to correct DEMs prior to their application in any drainage analysis. Sinks and depressions in DEMs must be removed in order to provide a realistic representation of the landscape and the flow. Using the classic DEM pre-processing methods such as filling or carving sinks does not always provide satisfactory results (Lindsay and Creed, 2006). These methods may have some limitations, especially when the terrain contains a large number of engineered structures. In that case, the application of DEM pre-processing methods can create flat surfaces or distort the stream slope. Therefore, modeling surface flow over such terrain may require the use of more sophisticated methods in order to correct the DEM. In this chapter, we present a new hybrid method combining classic filling and breaching.

Two hydrological models are presented in this chapter. The well-known professional hydrological model HEC-RAS (Brunner, 2016) and an academic model called NIHM (Jeannot, 2018). While both models use the diffuse wave approximation to perform rain-on-grid simulations, the computational environments are different. HEC-RAS contains a sparse parallel solver that enhances the simulation performance by sharing the computational load over multiple cores. The availability of NIHM source codes favors the implementation of basin partitioning methods to perform parallel simulation of multiple sub-domains.

In this chapter, we first provide a detailed presentation of the two hydrological models HEC-RAS and NIHM to be used on a single processor. Next, the classic DEM pre-processing methods are discussed. We then present an existing hybrid method and propose another one to evaluate their hydrological performance compared to the classic methods. Finally, the implementation of domain decomposition in NIHM is presented assuming a distributed memory computer. The hydrological and computational performances of parallel simulations of NIHM are compared to those of HEC-RAS.

4.2 Hydrological models

Two hydrological models were applied in this research work, HEC-RAS and NIHM. Both models allow for 2D rain-on-grid simulations at the drainage basin scale by simplifying the 2D shallow water equation (Saint-Venant) with the diffuse wave approximation.

4.2.1 HEC-RAS

The Hydrologic Engineering Center's (HEC) River Analysis System (RAS), developed by the U.S. Army Corps of Engineers (USACE), is the best known and most used free hydraulic model. HEC-RAS is an integrated system of software. The system comprises a graphical user interface (GUI), separate hydraulic analysis components, data storage and management capabilities, graphics, mapping (RASMapper), and reporting facilities.

Released in July 1995 (Version 1.0), HEC-RAS was limited to perform 1D modeling mainly used for open channel flow analysis. In 2014, the HEC-RAS version 5.0.1 including 2D modeling capabilities was released (Brunner, 2014). Rain-on-grid modeling is implemented by means of nonlinear partial differential equations. HEC-RAS 2D is mainly used on Windows platforms and the set-up of a multi-domain simulation remains challenging.

The two-dimensional shallow water equations (continuity and momentum, respectively) are written in the watershed Ω as:

$$\frac{dh}{dt} + \nabla(uh) = R, \tag{4.1}$$

$$\frac{du}{dt} + (u\nabla)u + \frac{v_t}{h}(h\nabla u) + g\nabla(h+z) = gS_f\,, \tag{4.2}$$

where the unknowns h and u are the water depth and the flow velocity vector, respectively. The compound source/sink flux R allows to account for precipitation, evaporation and infiltration terms. Other parameters are the bed elevation z, the gravitational acceleration g, the horizontal eddy viscosity coefficient v_t, and the bottom friction coefficient S_f.

Assuming that the unsteady terms (the advection and viscosity) can be neglected, the diffusive wave approximation of the shallow water momentum equation (4.2) may be written as:

$$g\nabla(h+z) = gS_f. \tag{4.3}$$

A 2D model simulation in HEC-RAS needs the following data: topography data providing a physical description of the terrain (DEM), computational area, boundary conditions, and soil parameters (Manning coefficient).

The setup of a 2D HEC-RAS simulation starts by creating a terrain using a DEM that represents the topographical data of the case study area. The computational area is defined by a closed polygon within which the 2D computations will occur. This closed polygon represents the boundary of the supposed flooded area and can be imported as a shapefile or created using the editing tool of RAS-Mapper. Once the computational domain is defined, HEC-RAS provides an automated tool to build the computational mesh. The finite volume solution used in the 2D modeling allows for the use of both structured and unstructured mesh. The computational mesh generated can have elements with 3 to 8 sides each. It is recommended to consider a mesh grid size that matches the terrain grid size (Brunner et al., 2015).

Continuity and momentum equations are solved together with boundary conditions set on the watershed:

$$\partial\Omega = \partial\Omega_{in} \cup \partial\Omega_{out} \cup \partial\Omega_{side}. \tag{4.4}$$

These are used to model:

- a flow Q_{in} entering in the domain Ω at the inflow boundary $\partial\Omega_{in}$:

$$\iint_{\partial\Omega_{in}} u.nds = Q_{in}, \tag{4.5}$$

  where n is the unit normal vector to the boundary, s is the side boundaries and the integral is taken over the boundary surface on which the boundary condition applies,

- an outflow condition at the outlet $\partial\Omega_{out}$ implemented as a water surface gradient S_b:

$$\nabla(h + z).n = S_b, \quad \text{on } \partial\Omega_{out}, \tag{4.6}$$

- a normal flux condition on the side boundary Ω_{side}:

$$\nabla u.n = 0, \quad \text{on } \partial\Omega_{side} = \partial\Omega \setminus (\partial\Omega_{in} \cup \partial\Omega_{out}). \tag{4.7}$$

Finally, picking an appropriate and adequate computation timestep is very important for getting accurate results. The computation time interval or the time step selected should be small enough to allow water movement through the mesh grid cells. The HEC-RAS manual suggests trying several computational timesteps in order to choose the one providing stable results.

4.2.2 Simulation of HEC-RAS on a shared memory

Simulation of large basins via HEC-RAS, especially 2D rain-on-grid simulations, may be computationally intensive and requires long simulation time. However, the HEC-RAS 2D module has the advantage of using a sparse parallel solver with a shared memory. This parallel computing strategy aims to split the linear system to be solved, in order to solve smaller problems, using a multi-core processor.

The ideal number of cores used in a simulation depends on the size and shape of the computational area. Theoretically, as more cores are used the problem is broken down into smaller blocks. However, this will increase the communications and data transfer between the different blocks which is time-consuming (Brunner, 2016). This means that increasing the number of cores is not always in favor of reducing the computing time.

In the following, several tests are performed to experiment the impact of the number of cores involved on the computation time required in rain-on-grid modeling. Simulations are applied while varying the number of cores set to: 1, 2, 4, 6 and 8. Runtimes were recorded in minutes on a Intel Xeon 64 bits @3.80 GHz_6 processor with 16 GB RAM memory. Knowing that the machine used for simulations contains only 6 cores, the last simulation aims to test the computational performance of HEC-RAS if the user-defined number of cores to be involved exceeds the machine's capacity.

The workflow for running 2D rain-on-grid simulations in HEC-RAS is presented in Fig. 4.1. The simulations are applied to the Moderbach basin. The DEM of the Moderbach area is imported into HEC-RAS through RAS Mapper to model the terrain. The computational domain, which correspond to the entire drainage basin of the Moderbach, is imported in the RAS Mapper interface. The HEC-RAS mesh generation tool is used to build a $25m \times 25m$ mesh similar to the DEM resolution.

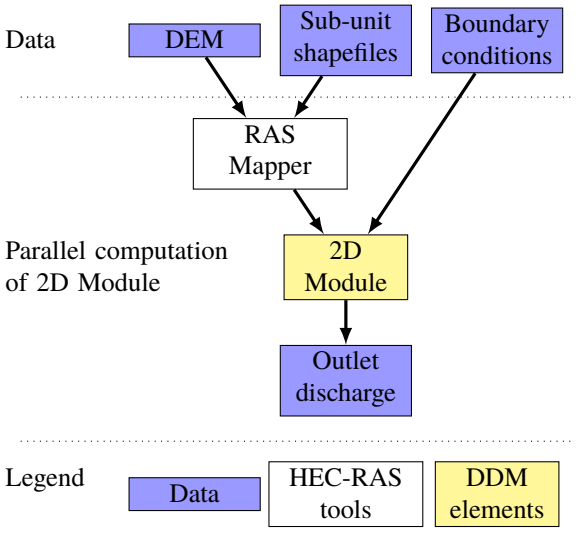


Figure 4.1: HEC-RAS workflow for a unique 2D rain-on-grid simulation.

Boundary conditions are added to the model considering a rainfall data of 80mm/day for 28 h then a dry period of 44 h (Fig. 4.2). These conditions were set as they allow to provide a rapid saturation of the Moderbach basin and then to drain the basin reaching a null flow at its outlet. This type of rainy episode has already occurred in the Moderbach basin. The Kappelkinger station recorded 76 mm on September 13, 2017. The normal slope at the basin outlet is set to 0.001 (Eq. 4.6) and the Manning coefficient is set to 0.03. Several tests are applied by varying the number of cores involved in the simulation. The computational time step is set to 10 seconds. The flood extent modeled in HEC-RAS is presented in Figure 4.4.

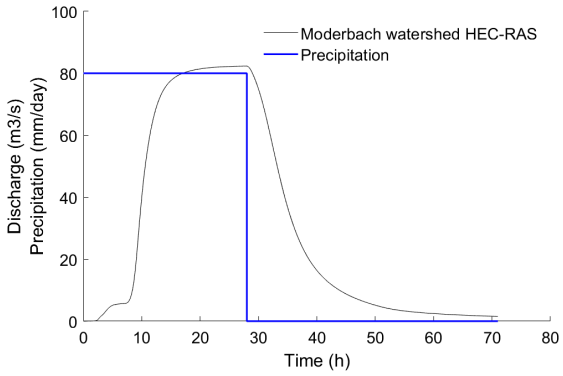


Figure 4.2: Outlet flow of the Moderbach watershed.

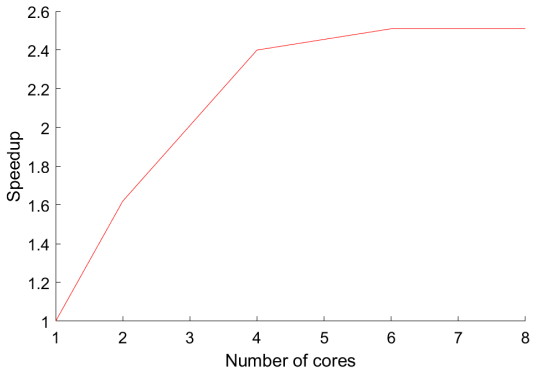


Figure 4.3: Speedup.

NB cores	Computation time (min)	Speedup
1	460.6	1
2	284.2	1.62
4	191.3	2.40
6	183.0	2.51
8	183.0	2.51

Table 4.1: Characteristics of HEC-RAS's parallel simulation on a shared memory.

Figure 4.2 shows the flow hydrograph computed at the outlet of the Moderbach basin. As expected, the basin reaches saturation after 28 h of rainfall. The flow curve decreases when the model is put to dry conditions until it reaches at null discharge at its outlet after 44 h. Performing this simulation several times while varying the number of cores involved in the simulation demonstrates the efficiency of the sparse parallel solver used in the HEC-RAS 2D simulations. The recorded execution times are presented in Tab. 4.1.

Figure 4.3 shows the speedup resulting from involving more cores in the simulation. The speedup is the ratio of the computational time required for the simulation using multiple cores to that of the simulation using a single cores. The speedup curve shows a rapid increase where the simulation involving 4 cores presents a speedup of 2.4. However, increasing the number of cores to 6 does not bring much improvement in terms of computation time. This is because,as mentioned earlier, increasing the number of cores involves more communication and data transfer which is a time consuming phase.

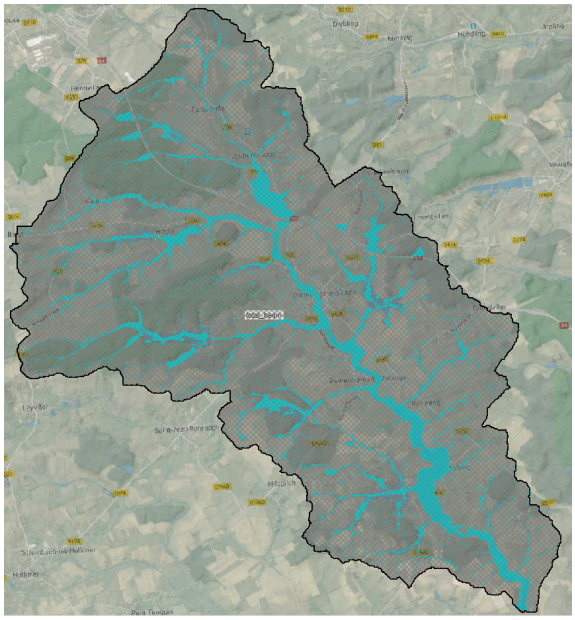


Figure 4.4: Moderbach simulation in HEC-RAS.

4.2.3 Normally Integrated Hydrological Model (NIHM)

The Normally Integrated Hydrological Model (NIHM) (Pan et al., 2015) is a physically-based spatially distributed hydrological model coupling a low-dimensional subsurface model, a 1-D river flow model, and a 2-D surface flow model. In this research work, we use the 2-D surface flow model of NIHM only (Jeannot, 2018).

For a more detailed description of the three compartments of NIHM as for additional information on the physical and mathematical formalism or numerical tests, readers are invited to refer to Pan et al. (2015), Weill et al. (2017), and Jeannot et al. (2018).

The 2-D surface flow model simplifies the Saint Venant equations (4.1) and (4.2) to a diffusive wave equation following the approach proposed by Di Giammarco et al. (1996). In an older version of NIHM (Pan et al., 2015) (Weill et al., 2017), the discretization has used a mixed hybrid finite element method (MHFEM) scheme. In the new version where the 2D surface flow model was integrated (Jeannot, 2018), the MHFEM scheme was replaced by the Crouzeix-Raviart finite element method (CRFEM) (Crouzeix and Raviart, 1973). It was shown that the CRFEM were equivalent to the MHFEM (Marini, 1985) (Vohralík and Wohlmuth, 2013). However, the formalism of CRFEM is simpler to implement. Both methods are finite element methods and therefore involve interpolation functions. The interpolation functions of the CRFEM are defined explicitly by their scalar value in the middle of the edges of the 3 sides in a triangle, while the interpolation functions of the MHFEM are local vectors defined by their integral properties on an element and its sides.

Contrary to HEC-RAS which includes a mesh generator and visualization tools, the NIHM hydrological model is supported by mSim and Gmsh in a preprocessing phase to generate the mesh and associated information. mSim is a finite element model (Kourakos and Harter, 2014), based on a set of Matlab functions, which provides the input file for the Gmsh mesh generator. Gmsh is freely available and can be run in Matlab.

4.2.4 Simulations with NIHM

The main objective of using NIHM in this research work is to implement the domain decomposition method developed in Chapter 3, in order to perform 2D rain-on-grid modeling on a large-scale basin through a parallel simulation of its sub-basins. This computational strategy is performed and presented later in this chapter in section 4.4.1 and compared to HEC-RAS simulations presented in section 4.2.2.

Before organizing a parallel simulation with the NIHM version of 2018 (Jeannot, 2018), we have considered to verify its hydrological performance by applying it to perform rain-on-grid simulations on small basins using one single core. A wide field inventory has been carried out on the Moderbach basin, comprising 6 ponds and 5 anthropogenic artifacts (dams and roads), which presents a good opportunity to evaluate the hydrological performance of a model in a particular landscape. The sub-basins considered for this simulation are those generated from the Moderbach decomposition into 8 sub-basins, see Fig. 3.12, using the guided area-balanced method presented in section 3.4.1. The precipitation event considers a sufficient amount, 80mm/day, to reach a rapid saturation of the sub-basins.

It is worth mentioning that these simulations were not performed to be compared with observed results. We intend to verify whether the flow hydrographs computed at the outlet of each sub-basin have a realistic behavior that matches the precipitation considered as an input of the simulation (80mm/day). At the outlet of each sub-basin, we expect an ascending flow curve that settles in the form of a plateau, once the sub-basin reaches saturation. This shape is expected since the input data of the simulated event presents a constant precipitation. Simulations are also performed with HEC-RAS in order to validate the results provided with NIHM.

The Digital Elevation Model (DEM) with a resolution of $25m \times 25m$ used in the simulation of the Moderbach sub-basins is subjected to two different methods, Filling and Breaching, in order to correct sinks and depressions in the DEM. These methods are discussed in Section 4.3.

4.2.4.1 Sub-basin of Diefenbach

The flow hydrograph computed at the outlet of all sub-basins of the Moderbach falls within the expectations, except for two sub-basin. The first sub-basin is the one at Diefenbach. Figures 4.5 and 4.6 show the outflow of the Diefenbach sub-basin (Fig.4.7) computed using the Filling and the Breaching method, respectively, to correct the DEM.

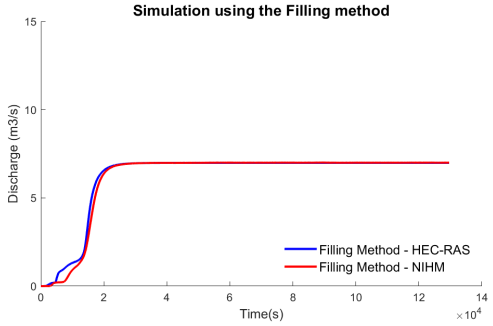


Figure 4.5: Outflow of the Diefenbach sub-basin computed using the Filling method.

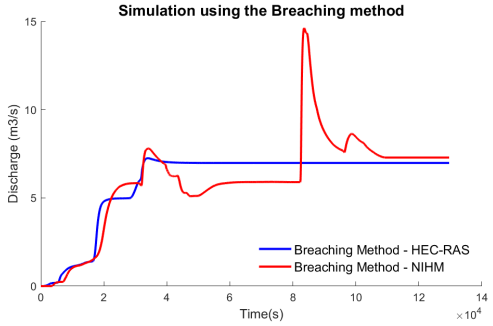


Figure 4.6: Outflow of the Diefenbach sub-basin computed using the Breaching method.

Using the filling method, the shape of the flow curves computed with NIHM and HEC-RAS are as expected, ascending curves that stabilize in the form of a plateau once the sub-basin is saturated (Fig. 4.5). Using the breaching method, the flow hydrograph computed at the outlet of Diefenbach sub-basin in NIHM, shows an unexpected peak. However, the simulation in HEC-RAS provided a normal flow hydrograph (Fig. 4.6).

The stable results provided by the HEC-RAS simulations using the breaching method (Fig. 4.6) are mainly owed to the FVM used to solve the 2D unsteady flow equation. However, NIHM uses the finite element method of Crouzeix-Raviart (CRFEM) (Crouzeix and Raviart, 1973). Indeed, it was well mentioned in the HEC-RAS manual that the FVM used to solve the 2D equation provides an increment of improved stability and robustness over other finite difference and finite element techniques. It also can handle subcritical, supercritical, and mixed flow regimes (flow passing through critical depth, such as a hydraulic jump).

The peak of the flow curve computed in NIHM using the breaching method (Fig. 4.6) can be explained by an accumulation of water in a storage area or in a depression in the terrain and a sudden evacuation of water once the storage area is full. To locate the water accumulation area, one solution was the visualization of the water depth all over the basin at the time step just before the appearance of the peak, see Fig .4.7. The visualization of the water depth all over the basin shows several cells with a high water depth, more than 2 m. These cells are located near two roads that cross the Diefenbach pond.

As a result of these simulations applied with two models using two different DEM pre-processing methods, we can state that HEC-RAS provided satisfactory results using both DEM pre-processing methods, filling and breaching. However, NIHM provided satisfactory results only when using the filling method. Applying the breaching method provided an unexpected behavior of the basin outflow. The limitations of the breaching method are further discussed in Section 4.3.1.2.

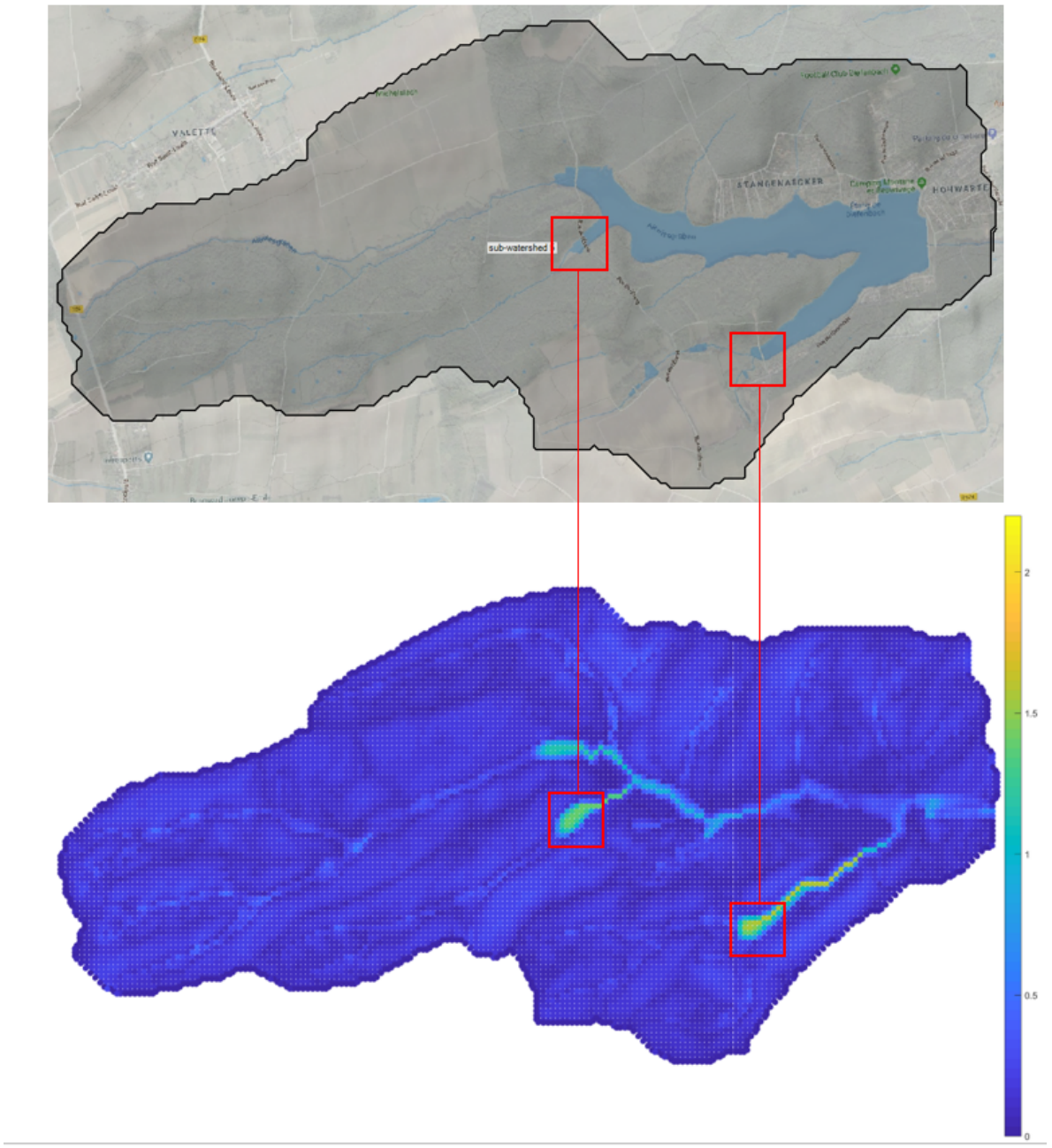


Figure 4.7: Visualization of water depth on the sub-basin of Diefenbach with a focus on high water level cells.

4.2.4.2 Sub-basin of Loupershouse

While we thought that the filling method is well adapted for the simulations performed in NIHM, this assumption was readily rejected. Figures 4.8 and 4.9 show the flow curves computed using the Filling and the Breaching method respectively, at the outlet of the sub-basin at Loupershouse (Fig. 4.10).

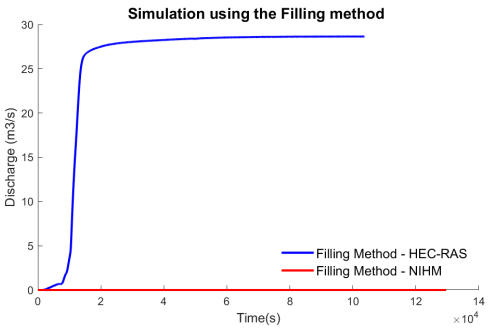


Figure 4.8: Outflow of the Loupershouse sub-basin using the Filling method.

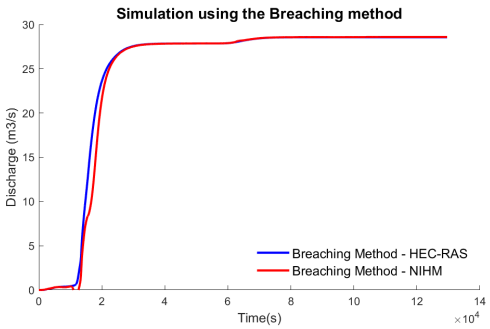


Figure 4.9: Outflow of the Loupershouse sub-basin using the Breaching method.

Using the filling method, the simulation in NIHM provides a null flow at the sub-basin outlet while the HEC-RAS simulation provided a realistic behavior of the flow curve (Fig. 4.8). This can be explained by the fact that the application of the filling method tends to generate flat areas in particular terrains where the calculation of the flow direction is problematic. The limitations of the filling method are further discussed in Section 4.3.1.1.

The application of the breaching method provided satisfactory results in both models NIHM and HEC-RAS.

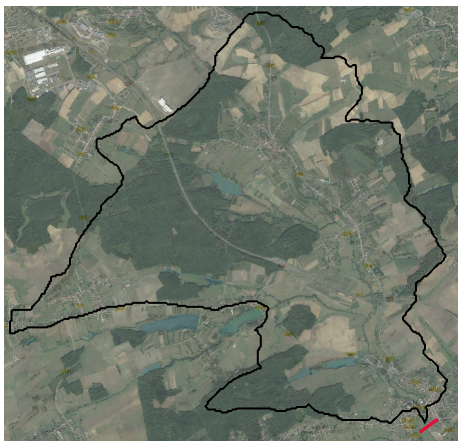


Figure 4.10: Sub-basin of Loupershouse.

To summarize, the simulations performed showed that the NIHM model has failed to correctly model the hydrological behavior on the Diefenbach sub-basin using the breaching method and on the Loupershouse sub-basin using the filling method. However, HEC-RAS provided satisfactory results in both sub-basins with either the filling method or the breaching method. Here, we can state that the filling and breaching methods may not be suitable for the hydrological model NIHM.

4.3 DEM pre-processing methods

Sinks or depression in a DEM are cells that have no downward slope to any adjacent cell because all adjacent cells have higher elevations (Fig. 4.11). The existence of sinks in a DEM is usually the result of limited resolution, poor quality data, interpolation errors, inaccuracies, actual depressions in the landscape, or artifacts such as bridges, ponds and road embankments (Martz and Garbrecht, 1999).

58	52	55	53	57	57
55	40	42	45	51	55
48	33	35	33	48	52
33	23	15	27	25	38
12	18	22	23	24	12
11	12	21	15	15	11

(a) Detection of a Sink

58	52	55	53	57	57
55	40	42	45	51	55
48	33	35	33	48	52
33	23	18	27	25	38
12	18	22	23	24	12
11	12	21	15	15	11

(b) Removing Sink

Figure 4.11: Detection and correction of a sink.

The existence of sinks presents a major limitation for the computation of the flow direction. This one is used to derive most of the DEM characteristics such as flow accumulation, hydrographic network, average slope and drainage basin boundary (Section 2.3.3). The computation of the flow direction is based on the D8 method which allows an automatic routing of the flow from cell to cell. The D8 method is not able to handle sinks in a DEM because its height value is less than all of its neighbors and therefore the flow direction cannot be attributed to any of its 8 adjacent cells. This is problematic for the application of drainage analysis methods based on overland flow simulation (Martz and Garbrecht, 1999).

The DEM must be corrected before it can be applied in a drainage analysis (Jenson and Domingue, 1988) (O'Callaghan and Mark, 1984). It is necessary to remove the sinks in order to ensure a continuous water movement across the digital terrain (Lidberg et al., 2017). Several sink removal methods have been proposed to improve the DEM quality. The failure of the filling and breaching methods to provide satisfactory results in the simulation performed with NIHM (Section 4.2.4.1 and 4.2.4.2) has motivated the need to use more advanced methods for sink removal. In the next section, we present and discuss the limitations of the classic methods. Then, we present two hybrid methods to overcome the limitations of the classic methods.

4.3.1 Classic methods

4.3.1.1 Filling

This widely used method allows filling sinks and topographic depressions by increasing the elevation of the cell in question to the elevation of the cell with the lowest value in the depression neighborhood (Fig. 4.12).

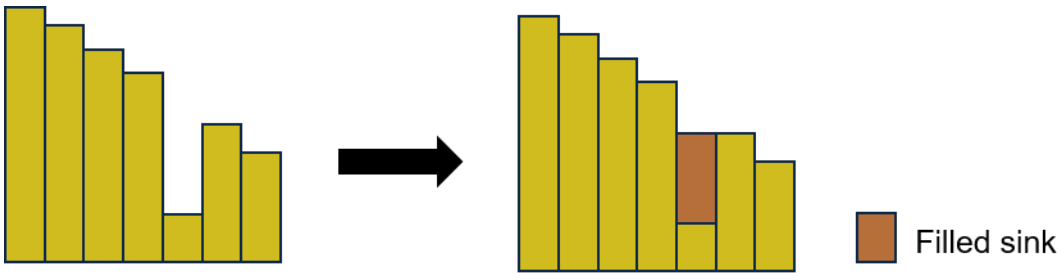


Figure 4.12: Profile view of a filled sink.

However, the filling method aiming to increase the elevation of sinks may replace the depression with a flat area (Fig .4.13). In that case, the calculation of the flow direction with the D8 method is problematic due to the appearance of parallel flow line, since all adjacent cells have the same height value. Therefore, the flow calculations in flat areas may be an issue. This was mostly the reason for which using the filling method with NIHM has failed to provide a realistic outflow (Fig. 4.8) for the Loupershouse sub-basin presented in Section 4.2.4.2.

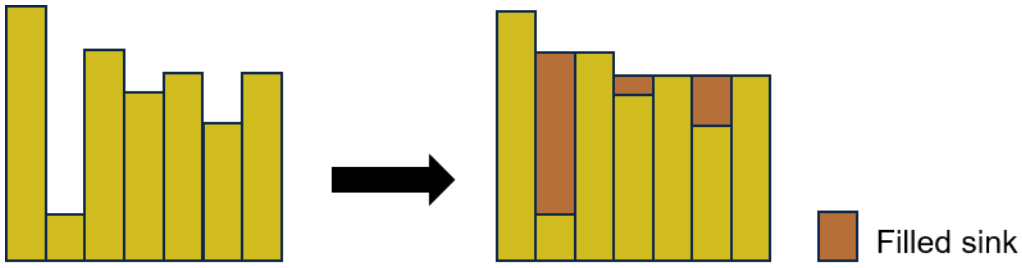


Figure 4.13: Limitation of filling sinks.

In our research work, filling the sinks is carried out with the ***fillsinks*** method proposed in the Topotoolbox presented in Section 3.2.1. It is worth mentioning that the time needed to fill one DEM varies between the different software used. The ***fillsinks*** method in the Topotoolbox outperforms other software by 8-10 times. However, processing speed is not the only important criterion, memory management may be better in other software (Schwanghart, 2014).

4.3.1.2 Breaching

The first breaching method described in the literature was given by Rieger (1993). This method aims to lower the elevation of the cell on the edge of the sink, allowing the depression to drain it in a continuous descending path (Fig. 4.14). Several studies have used the breaching method, also called depression breaching or depression carving (Martz and Garbrecht, 1998) (Soille et al., 2003) (Soille, 2004) (Metz et al., 2010).

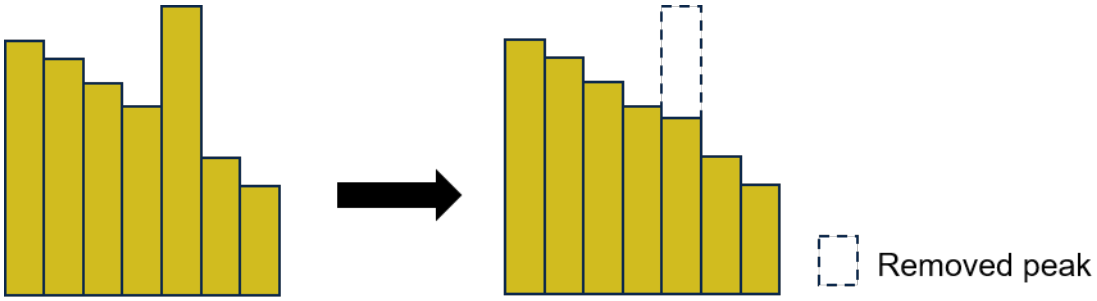


Figure 4.14: Profile view of a breached sink imposing a downward gradient.

The breaching method may be favored by practitioners for surface flow modeling as it brings less modification to the DEM than the filling methods (Soille, 2004) (Lindsay and Dhun, 2015) (Lindsay and Creed, 2005). The major differences between filling and breaching are presented in Appendix C. However, the breaching method has its limits. It is not well adapted for removing very deep sinks (Fig.4.15). In this case, the application of the breaching method may lead to very steep and unrealistic slopes and therefore the application of the filling method is more suitable (Lindsay, 2016). Deep sinks appear in general due to the existence of pothole wetlands, lakes, or ponds in the terrain landscape.

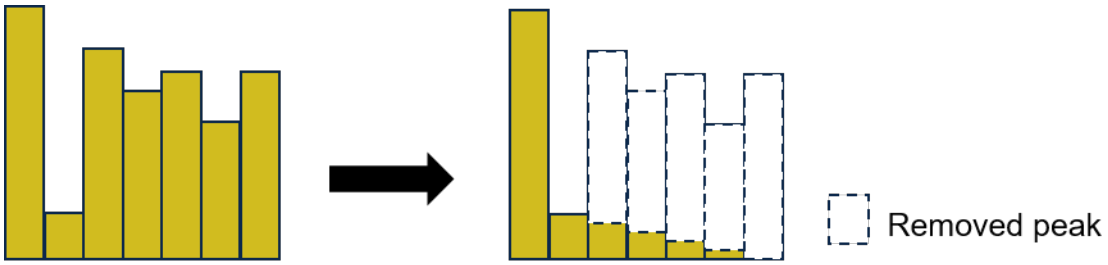


Figure 4.15: Limitation of breaching sinks.

For our research work, breaching sinks is carried out with the ***imposemin*** method proposed in the Topotoolbox. It allows carving in the DEM by imposing a minimum downward gradient along the stream to ensure a downstream slope (Schwanghart et al., 2013).

4.3.2 Hybrid methods

Although there are several methods to correct DEMs by a uniform treatment of all its sinks (Lindsay and Creed, 2006), such as the classic methods, the preference of filling or breaching each sink in the DEM allows for more improvement. This requires visual manipulation of the data and may necessitate field visits to examine the terrain features (dams, ponds,...). Several studies have proposed hybrid methods combining filling and breaching to correct sinks in a DEM. Lindsay and Creed (2005) evaluated the performance of several methods to remove sinks from a 5-m DEM and found that the hybrid method combining filling and breaching has caused fewer modifications to the spatial and statistical distribution of the terrain attributes. Poggio and Soille (2012) found that the hybrid method combining filling and breaching provided the most accurate hydrographic network compared to the other pre-processing methods.

In the following, we discuss an effective hybrid DEM pre-processing method that has been applied in previous works (Leach et al., 2017) and present the development of a new hybrid method designed to overcome the limitations of the classic methods outlined earlier in this chapter (Section 4.2.4).

4.3.2.1 Selective breaching

According to Lindsay (2016), deep sinks are problematic for breaching algorithms as illustrated previously in (Fig. 4.15). Here, breaching may create long and deeply incised channels. Deep sinks, which generally correspond to an underestimation of the elevation due to the existence of ponds, lakes or wetlands should be eliminated using the filling method. However, sinks that correspond to an overestimation of the elevation due to superior blockage of the flow should be treated using the breaching method.

Therefore, Lindsay (2016) introduced a hybrid method for DEM pre-processing called "selective breaching", allowing to remove sinks by combining filling and breaching. This method favors breaching sinks, except for the sinks that are deeper than a specified threshold depth. These deep sinks are removed using the filling method. This method is illustrated in Figure 4.16 and applied following listing 4.1. This method is applied for the simulation of the Diefenbach and Loupershouse basins, see Sections 4.3.2.3 and 4.3.2.4.

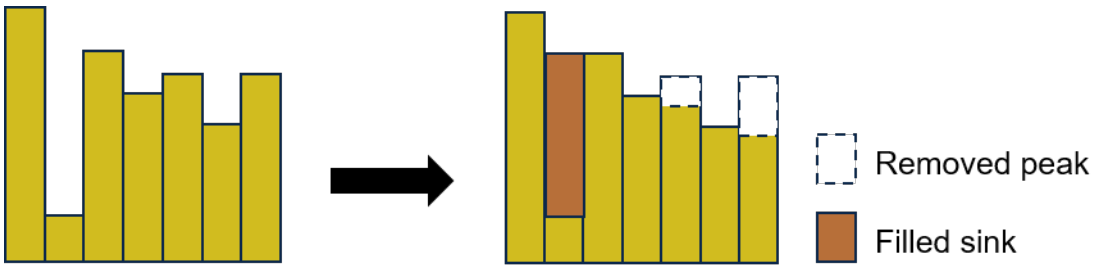


Figure 4.16: Selective breaching method.

Listing 4.1: Selective breaching method.

```
% Introduce the DEMs to be used
dem    =GRIDobj('DEM-file.asc');
% DEM subjected to Filling
Filleddem=fillsinks(DEM);
% DEM subjected to Breaching
Breacheddem=imposemin(DEM);
% For each cell of the DEM matrix
for i=1:size(dem,1)
   for j=1:size(dem,2)
   % if the sink depth is less than the threshold depth
   if Filleddem(i,j) - dem(i,j) < thresholddepth
      % the sink will be breached
      SelectiveBreachingdem(i,j)=Breacheddem(i,j);
   else
     % if the sink is very deep it will be filled
      SelectiveBreachingdem(i,j)=Filleddem(i,j);
  end
```

4.3.2.2 Outlet Breaching

After the application of the classic pre-processing methods on the Moderbach sub-basins and the identification of their limitations, we may deduce two conditions that a pre-processing method should respect in order to provide satisfactory results in drainage analysis. The first condition is to fill deep sinks caused by anthropogenic artifacts, pothole wetlands, lakes, or ponds since these sinks are not well treated with the breaching method (Fig. 4.15), and filling them may be more suitable. The second condition is to always breach the sinks at the location where a flow hydrograph needs to be computed. This is to ensure that no flat areas appear at this location (Fig. 4.13), preventing the calculation of flow direction which makes it impossible to derive the outflow of a basin.

In the following, we present a new hybrid method for DEM pre-processing called "Outlet breaching" (Hariri et al., 2021). This is a hybrid method combining filling and breaching to remove sinks. Due to the particularity of the Moderbach area containing an important number of anthropogenic artifacts and wetlands in the landscape, the method consists of satisfying the two conditions mentioned above as follow (Fig .4.17): (1) breaching the sinks near the basin outlet where a flow hydrograph needs to be computed (2) filling the rest of the sinks. This method is applied for the simulation of the Diefenbach and Loupershouse basins, see Sections 4.3.2.3 and 4.3.2.4..

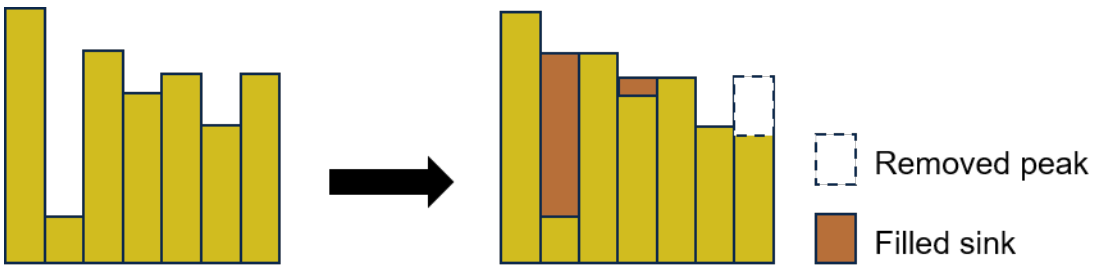


Figure 4.17: Outlet breaching method.

The method aims to apply a smooth transition between the sinks near the outlet subjected to breaching and the sinks far from the outlet subjected to filling. Therefore, we have chosen to use the cumulative distribution function of the normal distribution (Eq. 4.8). The value of $\phi(x)$ assigned to each sink will determine whether this sink is treated by the filling method or by the breaching method based on Equation 4.9.

The σ and μ parameters are set to 80 and 500 respectively which correspond to the blue curve (Fig .4.18). This cumulative distribution (blue curve) is chosen as it fulfills the conditions of the outlet breaching. If the sink is near the basin outlet, $\phi(x)$ will have a value very close to 0 and according to the Equation 4.9, the outlet breaching method will consider breaching that particular sink. If the sink is far from the basin outlet, $\phi(x)$ will have a value close to 1 and according to the Equation 4.9, the outlet breaching method will consider filling that sink. When the sink is neither too close nor too far from the basin outlet, $\phi(x)$ will range between 0 and almost 1 and therefore the sink will be treated by both filling and breaching relatively (Equation 4.9).

$$\phi(x, \sigma, \mu) = \frac{1}{2}(1 + erf(\frac{x - \mu}{\sigma\sqrt{2}})); \tag{4.8}$$

where σ the standard deviation, μ the mean and x the distance between the sink and the basin outlet are expressed in meters and erf is the error function.

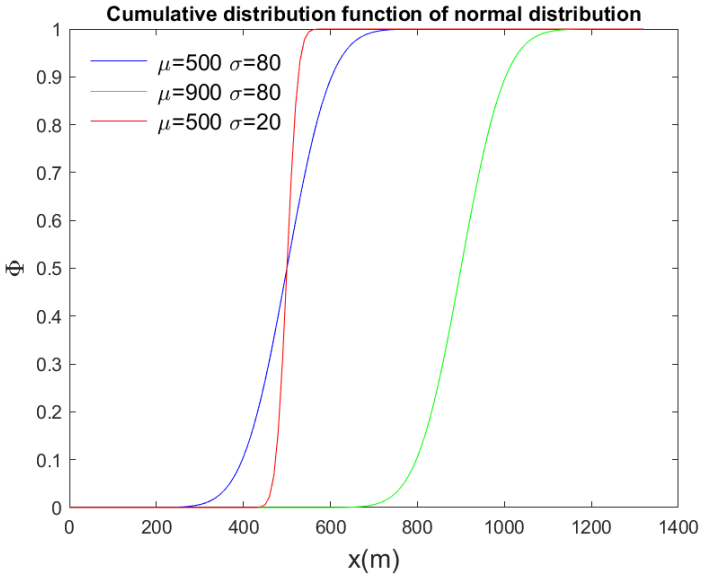


Figure 4.18: Cumulative distribution function of normal distribution.

$$Z_{OutletBreaching} = (1 - \phi)Z_{Breaching} + \phi Z_{Filling}; \tag{4.9}$$

where $Z_{OutletBreaching}$ is the matrix containing the elevation of the DEM grid cells after applying the outlet breaching method, $Z_{Breaching}$ is the matrix containing the elevation of the DEM grid cells after applying the breaching method and $Z_{Filling}$ is the matrix containing the elevation of the DEM grid cells after applying the filling method.

It is worth mentioning that other cumulative distribution functions may be chosen. For example, the red curve considering different values for σ and μ fulfills the conditions of this hybrid method while showing a fast transition between the filling and breaching. The green curve considers to breach the sinks within a radius of more than 600 m from the basin outlet. This could be applicable for larger basins.

4.3.2.3 Sub-basin of Diefenbach

In order to evaluate the hydrological performance of the hybrid methods, these are applied to perform rain-on-grid simulations on the two sub-basins of Diefenbach and Loupershouse where the classic methods failed to provide satisfactory results with NIHM (Section 4.2.4).

The first hybrid method "Selective breaching", is applied with a minimum filled depth of 2m. This means that sinks that require to increase their elevation by more than 2 m to reach the height of their neighbor cell will be subjected to filling. The rest of the sinks will be removed by applying the breaching method. The second hybrid method "outlet breaching" is applied with respect to the cumulative distribution function of normal distribution, blue curve (Fig. 4.18), allowing to breach the sinks near the basin outlet and to fill the rest of the sinks.

Figures 4.19 and 4.20 show the outflow at the sub-basin of Diefenbach (Fig. 4.7) computed using the Selective breaching and Outlet breaching methods respectively. Contrarily to the simulation of this sub-basin with the classic methods (Section 4.2.4.1), the simulation using the hybrid methods provides satisfactory results with NIHM showing the expected shape of the flow curves. Both hybrid methods achieved filling the deep sinks in the pond that were responsible for the unexpected behavior (peak) of the flow curve of in Figure 4.6.

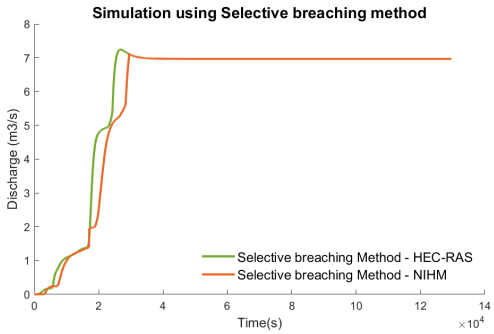


Figure 4.19: Outflow of the Diefenbach sub-basin using the Selective breaching method.

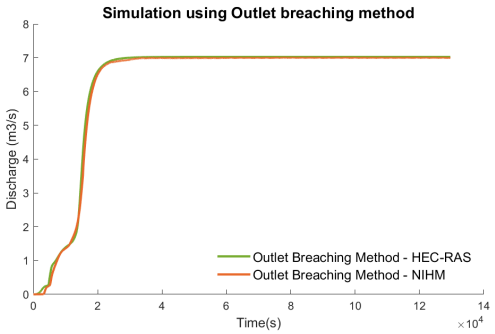


Figure 4.20: Outflow of the Diefenbach sub-basin using the outlet breaching method.

4.3.2.4 Sub-basin of Loupershouse

Figures 4.21 and 4.22 show the outflow of the Loupershouse sub-basin computed using the Selective breaching and outlet breaching methods respectively. While the simulation of this sub-basin in NIHM using the filling method (Section 4.3.2.4) failed to provide a realistic behavior showing a null discharge at the outlet (Fig. 4.8), the hybrid methods showed that there are well adapted for the simulation of this sub-basin. Both methods allowed to breach sinks near the outlet providing a flow hydrograph shape that corresponds to the precipitation data used.

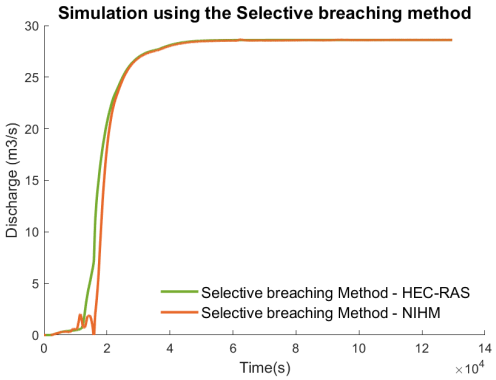


Figure 4.21: Outflow of the Loupershouse sub-basin using the Selective breaching method.

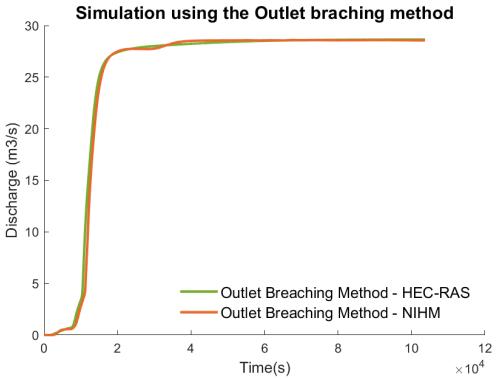


Figure 4.22: Outflow of the Loupershouse sub-basin using the Outlet breaching method.

To summarize, the results of the simulations show that for the hydrological model NIHM, the application of the hybrid methods to perform rain-on-grid modeling in the sub-basins of Diefenbach and Loupershouse provide satisfactory results. However, the application of the classic methods did not allow to correctly model the hydrological behavior in these two sub-basins. These simulations were also carried out with the HEC-RAS model, which showed a high level of robustness by providing satisfactory results on the two sub-basins, for both hybrid and classic methods.

4.4 Parallel simulation of NIHM

In this section, we couple the surface flow model of NIHM with a parallel tool called "ORWL" (Clauss and Gustedt, 2010) to perform parallel simulations of multiple sub-basins on a multi-processor cluster taking advantage of the domain partitioning method introduced in Chapter 3.

The hydrological and computational performances of this computing strategy is compared to the 2D simulations performed in HEC-RAS on a shared memory, presented in Section 4.2.2, by considering the Moderbach basin as a case study area to perform 2D rain-on-grid simulations. Runtimes were recorded in minutes on a Intel Xeon 64 bits @3.80 GHz_6 processor with 16 GB RAM memory.

The simulations performed in both models are subjected to the same conditions and input data (DEM, 2D equation, hydrological data and mesh size). However, the solvers for the 2D equations are different. HEC-RAS uses an implicit FVM allowing for the use of a structured or unstructured computational mesh while NIHM uses the finite element method of Crouzeix-Raviart (CRFEM) with a triangular irregular mesh (Crouzeix and Raviart, 1973).

4.4.1 Simulation of NIHM on a distributed memory architecture

The open-source aspect of NIHM has allowed us to apply a different parallel computing strategy than that applied with HEC-RAS. We propose to use a decomposed version of the drainage basin, taking advantage of the basin partitioning methods introduced in Chapter 3, to perform parallel simulations of multiple sub-basins on a multi-processor cluster. This parallel computing strategy required coupling NIHM with a parallel tool that provides communication and synchronization between the different sub-basin simulations.

ORWL "Ordered Read-Write Locks" (Clauss and Gustedt, 2010) is a programming approach that focuses on the management of shared resources. It is implemented to support computations performed in several tasks having temporal and data dependencies (Mansouri and Gustedt, 2016). The user needs to decompose the task to be performed into several interdependent tasks. ORWL provides an easy and predictable automated synchronization method between tasks for data-oriented parallel algorithms. The user does not need to manipulate processes or threads or to manage synchronization and communication between them (Clauss and Gustedt, 2010).

A balanced load distribution improves the efficiency of parallel computations (Kumar and Duffy, 2015). The decomposition of a drainage basin with the area balanced method, contributes in balancing the computational load by providing sub-basins of similar area that assume to have a similar computational cost.

We aim to execute several sub-basins of the Moderbach basin on a parallel platform using the 2D surface model of NIHM supported by ORWL. Therefore, the Moderbach decomposition configurations of 8 and 16 sub-basins are used (Fig. 4.23). These two decompositions, generated with the area-balanced method (Section 3.4.2), have a ratio of 1.9 and 2.1 between their maximum and minimum sub-basins respectively.

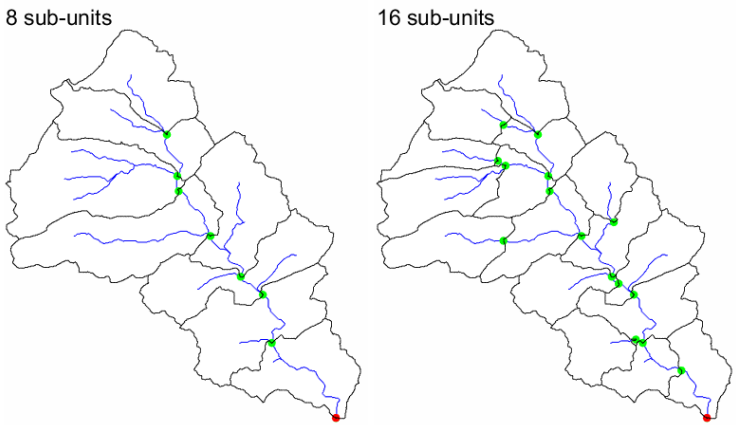


Figure 4.23: Unguided area-balanced basin partitioning method of the Moderbach in 8 and 16 sub-basins.

The parallel tool ORWL ensures communication and synchronization between the the different sub-basin where the flow computed at the outlet of a sub-basins will be used as input for the simulation of its downstream sub-basin (Fig. 4.24).

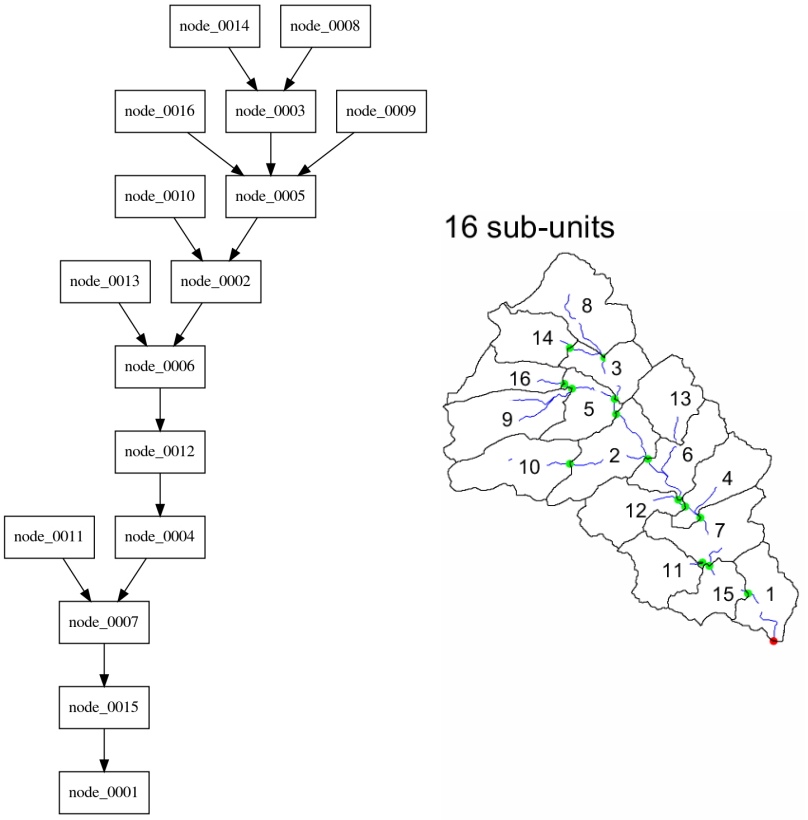


Figure 4.24: An illustration of the hierarchical communication relationship for the 16 sub-basins of the Moderbach supported by ORWL.

These simulations executed with NIHM use the same input data as those used for the simulation of the entire basin performed in HEC-RAS in section 4.2.2 such as DEM, precipitation data and Manning coefficient. Thus, the hydrological results and computational performance of the two models can be compared.

Figure 4.25 compares the Moderbach basin outflow derived from the 8 and 16 sub-unit decompositions simulated in NIHM with that derived from the entire basin simulated in HEC-RAS. The flow curves computed in NIHM show a high correspondence with the one calculated in HEC-RAS. These results confirm that the simulation of a decomposed version of a drainage basin is able to provide the same hydrological behavior as the simulation of the entire basin. The flow curves of the two models shows a slight offset just before reaching saturation and during the dry period of the event. The difference is illustrated in Fig. 4.26 which displays the relative error of the flow curves computed in NIHM with respect to the flow curve calculated in HEC-RAS. The error is so slight considering that the simulations are performed in two different models using different solvers.

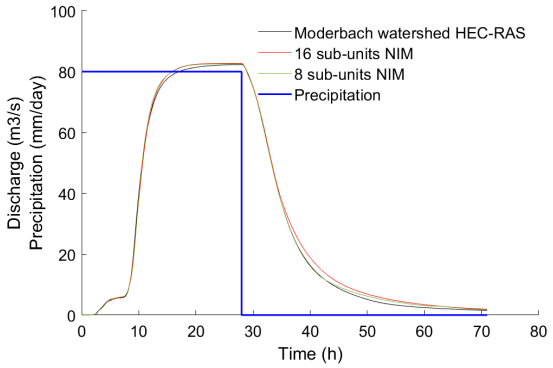


Figure 4.25: Outflow of the Moderbach partitioned in 8 and 16 sub-basins (computed with NIHM) and of the total basin (computed in HEC-RAS).

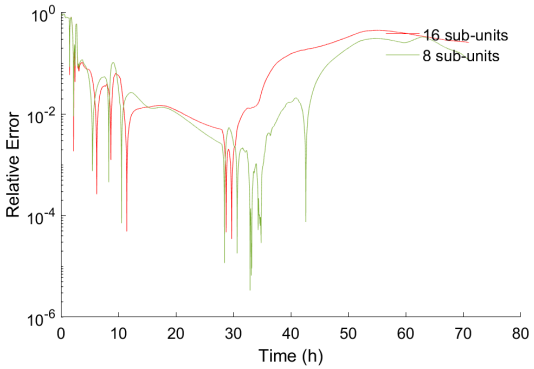


Figure 4.26: Relative error evaluated with respect to the outflow of the entire basin in HEC-RAS.

The computational results are reported in Tab. 4.2. The simulation of the Moderbach decomposed into 8 and 16 sub-basins require 60 h and 30 h respectively. This confirms that further decomposition of a basins allows providing the same hydrological behavior while requiring less simulation time. However, the time required for the HEC-RAS simulation recorded only 3h, which is much less compared to NIHM.

Model	HEC-RAS	NIM	
#sub-units	1	8	16
NB of mesh elements	145076	388612	390240
Computation time step (s)	10	1	1
Computation time (h)	3	60	30

Table 4.2: Characteristics of HEC-RAS and NIM parallel simulations.

Although both models use the same equation, the 2D equation solvers are different which justifies the large gap in the computing performance. The main parameter acting on the simulation time is the computational timestep considered in each model which strongly depends on the 2D equation solver. The Implicit FVM used in HEC-RAS allows for larger computational timesteps than other finite difference and finite element techniques as the one used in NIHM. A 10 s computation timestep was sufficient in HEC-RAS while NIHM required a 1s computational timestep in order to provide robust results. The 2D equation solvers also acts on the mesh generation. Even when using the same mesh size (25m), the number of mesh elements shows a significant difference between the two models. In NIHM a triangular irregular mesh is used while HEC-RAS allows for a structured or unstructured mesh with the possibility of having 3 to 8 sides for each mesh element. These are the main differences between the two hydrological models and the key parameters that are responsible for the superior computational efficiency in HEC-RAS compared to NIM.

4.4.2 Simulations on GRID5000

The simulation performed on the Moderbach basin (89 km^2) with the parallel framework NIHM-ORWL required long computation times using the local machine with 6 cores. The application to larger basins such as the Albe (409 km^2) or the Saar (1800 km^2) requires a higher level of computing power. Therefore, the simulations are carried out on the Grid5000 testbed (Grid5000, 2005). Grid5000 is a large-scale distributed testbed for research in grid computing, with a focus on parallel and distributed computing. The system is operated as a national network of clusters over 9 sites in France. Each site is provided with a 100 to 1000 CPUs cluster.

In the following, the three drainage basins of the Moderbach (89 km^2), the Albe (409 km^2), and the Saar (1800 km^2) are subject to parallel simulations in Grid5000. Several decomposition configurations are generated for each of these three basins. These decompositions are generated with the unguided area-balanced method allowing to provide sub-basins of similar area having similar computational cost. The decomposition configurations are tested in Grid5000 to perform a parallel simulation of the different sub-basins by sharing the computational load between several compute nodes. In our work, the nodes used are of the cluster 'grisou' having 2 x Intel Xeon E5-2630 v3 CPU of 8 cores/CPU, each, which means a total of 16 cores in each node. The simulations are performed while varying the number of nodes involved in the simulation to test their impact on the computation time.

The same rainfall event of 80mm/day for 28h and 44h under dry conditions (Fig. 4.2) modeled in the previous section on a local machine with 6 cores is now executed in the Grid5000 environment. However, for the simulations on Grid5000 the entire simulation period (72 h) would have required reservations of Grid5000 nodes for several days, which was not possible. Instead, we only simulated the first 1000 seconds of the event such that the computation results allows us to provide an approximation of the computation time for the entire period of 72 h. The measured execution times are presented in tables 4.3, 4.4, 4.5 and plotted in Fig. 4.27.

Starting with the Moderbach drainage basin, the decomposition configurations used are those of 1, 2, 4, 8, and 16 sub-basins (Table 4.3). The simulation of the total basin has failed. The computational domain and the number of elements are too big to be handled with only one core.

The simulation of the decomposition configurations 2, 4, 8, and 16 are performed with 1, 2, and 4 nodes. The results presented in table 4.3 show that the computation time for the simulation of the Moderbach decomposition configurations remained almost unchanged when varying the number of nodes applied in the simulation. Using one node containing 16 cores to handle 2, 4, 8, or 16 sub-basins is already more than enough. Not all computations of the 16 sub-basins can be active at the same time, since some of the sub-basins require information from other sub-basins. Using 16 cores, the parallel simulation is already saturated and increasing the number of cores in the simulation is irrelevant.

Computation time (h)	Moderbach decompositions				
Nb of nodes	1	2	4	8	16
1	-	0.92	0.53	0.24	0.08
2	-	0.92	0.53	0.22	0.08
4	-	0.92	0.52	0.22	0.07

Table 4.3: Computation time (h) for the simulation of the different Moderbach decomposition configurations.

In contrast to the increase in number of nodes, a decomposition of the basin leads to a significant gain in terms of computation time. While the simulation of the Moderbach decomposed into 2 sub-basin using 1 machine required 0.92 h, 8 and 16 sub-basins allow to reduce the computation time to 0.24 h and 0.08 h, respectively, that is by a factor of 3.83 and 10.81. This improvement was already visible in the simulations performed in Section 4.4.1 using a local machine. There, the parallel simulation of the Moderbach decomposed in 16 sub-basin provided an improvement in terms of the computation time compared to the simulation of the Moderbach decomposed in 8 sub-basins, see table 4.2. The computational efficiency of the simulations performed with the local machine is comparable to those on Grid5000.

In Section 4.4.1, the results presented correspond to the simulation of the entire event period of 72 h on a local machine, requiring a computation time of 60 h and 30 h for the Moderbach in 8 and 16 sub-basins respectively (Tab. 4.2). However, the simulation of the Moderbach in 8 and 16 sub-basins on Grid5000 are performed for only a simulation period of 1000 seconds, requiring a computation time of 0.24 h and 0.08 h, respectively. Therefore, the expected computation time required for the entire simulation of 72h on Grid5000 is 62 h and 22 h for 8 and 16 sub-basins, respectively.

For the basin decomposed into 16 sub-basins, the computation time in Grid5000 of 22 h is less than that required in the simulation using the local machine which was 30h. This is because the local machine only has 6 cores while the node used in the Grid5000 simulation has 16 cores. For the decomposition into 8 sub-basins, the computation times are very close 60 h and 62 h. This is because having more cores may not improve the computation time due to the saturation of the parallel simulation.

The simulation performed on the Albe basin failed to run the decomposition configuration in 4 sub-basins. Therefore, the basin is further decomposed into 32 and 64 sub-basins. These decompositions were performed using 2 and 4 nodes while the simulation using only one node has failed (Table 4.4).

Computation time (h)	Albe decompositions		
Nb of nodes	4	32	64
1	-	-	-
2	-	0.32	0.13
4	-	0.28	0.13

Table 4.4: Computation time (h) for the simulation of the different Albe decomposition configurations.

For the same number of nodes applied in the simulation, 2 or 4 nodes, the decomposition configuration of 64 sub-basins allows a decrease in the computation time by a factor of almost 2 compared to the decomposition configuration in 32 sub-basins. Increasing the number of nodes from 2 to 4 did not bring significant improvement due to the saturation of the parallel simulation as discussed earlier. Using 2 nodes having 32 cores for the simulation of 64 sub-basins is already enough since not all the 64 sub-basins can be executed at the same time. Some sub-basins will be waiting to receive the information from other sub-basins to be performed. Therefore, using 4 nodes having 64 cores to carry out the decomposition of 64 sub-basins does not bring significant improvement compared to the simulation using 2 nodes.

The Saar basin, the largest drainage basin in this study, is decomposed into 64 and 128 sub-basins. The simulation using only 1 node failed. Using two nodes allowed to simulate the Saar when decomposed into 128 sub-basins while with 4 nodes we succeeded to simulate with both decompositions 64 and 128. This simulation using 4 nodes shows that the decomposed configuration of 128 sub-basin reduces the computation time by a factor of 2.07 compared to the decomposition configuration of 64 sub-basins (Table 4.5).

Computation time (h)	Saar decompositions	
Nb of nodes	64	128
1	-	-
2	-	0.62
4	0.91	0.44

Table 4.5: Computation time (h) for the simulation of the different Saar decomposition configurations.

Contrary to the simulation performed in the Moderbach and the Albe, the simulation of the Saar decomposition configuration of 128 sub-basin using 4 nodes requiring 0.44 h provides a significant improvement in terms of computation times compared to the simulation using 2 nodes that required 0.62 h. This is because using 2 nodes having 32 cores to carry out 128 sub-basins would not saturate the simulation and increasing the number of cores to handle more sub-basin is possible.

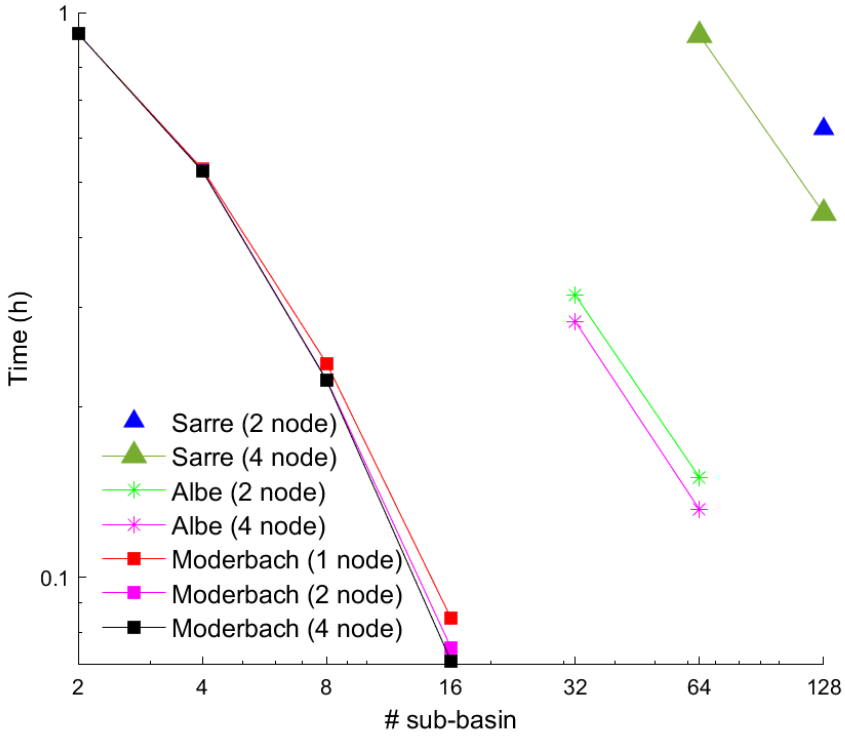


Figure 4.27: Computation time of the different decomposition configurations on Grid5000.

The results of the simulations performed on the Moderbach, the Albe, and the Saar demonstrate the efficiency of the parallel simulation of a decomposed drainage basin and the improvement brought in terms of the computing time. The computation times plotted in Figure 4.27 show again a superlinear speedup because of the size reduction of the linear system. The Moderbach basin have shown an important factor for decreasing the computation time where the decomposition of 16 sub basins is 10 times faster than the decomposition of 2 sub-basins while the decomposition of the Albe and the Saar provided a decrease of a factor of 2. The failure in the simulation of large drainage basins as the entire Moderbach and the Albe in 4 sub-basin, motivates the need to have a drainage basin decomposition methods allowing to provide sub-basin that can be handled by a given number of cores.

4.5 Conclusion

The differences between the two hydrological models HEC-RAS and NIHM, imposed different conditions on the correction of the DEM before its application in any drainage analysis. The application of conventional DEM pre-processing methods for sink removal, filling and breaching, was sufficient to provide robust results in the HEC-RAS simulations. This is mostly due to the FVM used to solve the 2D equation which provides an increased level of stability and robustness. However, the application of these methods in the simulation performed with NIHM did not provide satisfactory results for some of the sub-basins investigated in the Moderbach basin. NIHM is computationally weaker than HEC-RAS using the Crouzeix-Raviart finite element method (CRFEM) to solve the 2D equation. This lack of robustness was not considered as a major concern in the subsurface flow modeling work presented by Jeannot (2018) where the 2D surface flow from NIHM was used to support the subsurface model.

The failure of the classic methods to provide satisfactory results has motivated the use of more advanced methods that combines the classic filling and breaching. We applied the selective breaching method developed by Lindsay (2016) and introduced a new hybrid method called outlet breaching based on the limitations of the classic methods. The simulation of the Moderbach sub-basins using the hybrid methods provided satisfactory results and has shown that they are able to overcome the limitations of the classic methods in NIHM.

From a computational point of view, both parallel computing frameworks applied with HEC-RAS and NIHM proved to be efficient. Using the sparse parallel solver in HEC-RAS resulted in a significant improvement of the computing time, and running parallel sub-basins with NIHM-ORWL on a multiprocessor cluster showed a considerable speedup. However, the computational efficiency of the two models was different. It was expected that HEC-RAS would provide better computational efficiency than NIHM. In addition to the high level of hydrological robustness provided, the FVM in HEC-RAS allows for a larger computational time step and a mesh generation method that provides fewer mesh elements. All these parameters have contributed to reduce the computation time in HEC-RAS. The surface flow model of NIHM is currently under development at the Earth and Environment Institute of Strasbourg (ITES), in order to improve the linear and non-linear solvers.

The next chapter will present a parallel computing strategy that combines the two approaches presented in this chapter. The domain partitioning method will be implemented in HEC-RAS. This strategy will take advantage of splitting both, the drainage basin into sub-basins using the basin partitioning method and the linear system of each sub-basin using the sparse parallel solver of HEC-RAS.

Chapter 5

Modeling historical scenarios

Contents

French Abstract

L'utilisation des modèles de précipitation distribuée sur de grands bassins versants est contrainte par le maillage et le nombre d'inconnues. La revue bibliograhique effectuée par David and Schmalz (2021) rend compte d'articles scientifiques impliquant des modélisations comportant entre 5 000 à 3 000 000 d'inconnues et implémentées sur des ordinateurs appropriés (personnels ou parallèles). Dans HEC-RAS, les systèmes linéaires sont résolus à l'aide d'un solveur parallèle creux, ce qui permet une discrétisation impliquant quelques centaines de milliers d'éléments sur un processeur multi-cœur à mémoire partagée. Cependant, ces choix logiciels et architecturaux sont insuffisants pour réaliser une simulation de pluie distribuée sur un très grand bassin versant.

La modélisation des inondations dans la région de la Sarre est pertinente en raison de l'intensité et de la fréquence des inondations.En fait, En effet, la ville de Sarreguemines est confrontée à au moins un événement de crue tous les deux ans (Bonnet, 1997). Celle ci, exutoire de la partie française du bassin de la Sarre, est classée comme territoire à haut risque d'inondation. La modélisation de l'ensemble du bassin versant de la Sarre avec une discrétisation de 25m (qui est bien adaptée à la résolution de 25 m de nos MNT) génère environ 2 800 000 inconnues. Elle n'aboutit pas avec HEC-RAS en raison du dépassement de la capacité de mémoire. Ce problème a motivé le développement d'une stratégie de décomposition de domaine pour ces simulation 2D.

Les deux stratégies de calcul proposées dans le chapitre précédent sont combinées dans la section 5.2. Une partition du bassin versant est effectuée pour décomposer le bassin versant en sous-unités hydrologiques. Ensuite, chacun de ces sous-domaines est simulé en profitant d'un solveur parallèle creux. La méthode de décomposition de domaine, organisant la résolution des sous-problèmes et les communications entre sous-domaines, est automatisée à l'aide d'un code Virtual Basic for Applications (VBA) utilisant les fonctions de HECRASController pour récupérer les flux de sortie des bassins amont pour les réinjecter dans les bassins aval.

Cette nouvelle stratégie de calcul proposée dans HEC-RAS est validée dans la section 5.2 par une application sur le bassin du Moderbach apportant une accélération d'un facteur 2 lors de la mise en œuvre de la méthode de décomposition équilibrée proposée dans le Chapitre 3. Cette stratégie montre une supériorité par rapport aux autres méthodes de décomposition de domaine proposées dans la littérature en matière d'optimisation du temps de calcul.

Après validation, la section 5.3.1 présente l'application de cette stratégie pour la simulation des événements extrêmes et historiques sur la Sarre (partie française) qui draine une superficie de (1800 km^2), le sous-bassin de l'Albe (405 km^2) et le sous-bassin du Moderbach (89 km^2). Le triple événement de 2011 et les scénarios historiques de 1939-1940 ont été modélisés dans les bassins versants de l'Albe et du Moderbach respectivement. Les résultats sont en très bon accord avec les données de débit et les cartes de crues historiques. La simulation de la Sarre a été réalisée avec succès dans HEC-RAS en décomposant le bassin en 20 sous-bassins et en simulant chaque sous-bassin pour modéliser l'événement de crue de décembre 2010.

5.1 Introduction

The application of rain-on-grid models for the simulation of large-scale watersheds remains limited by the number of mesh elements in the computational domain. A review (David and Schmalz, 2021) of software, models, and parameterizations based on 20 very different rain-on-grid studies reports simulations comprising mesh elements ranging from 5,000 to 3,000,000, and run on appropriate (personal or parallel) computers. In HEC-RAS, linear systems are solved by means of a sparse parallel solver that can afford medium-range discretization involving a few hundred thousand of elements on a shared memory multi-core processor. However, these abilities are insufficient to perform rain-on-grid simulations on a very large-scale watershed.

Flood modeling in the Saar area is relevant considering the intensity and frequency of the floods. Moreover, the city of Sarreguemines, the outlet of the French part of the Saar basin, is classified as a territory of high flooding risk. In a first attempt, the simulation of the Saar basin of ($1800\,km^2$) using an element size of 25 m, (which is well adapted to the 25 m resolution of our DEM), involving about 2,800,000 elements in HEC-RAS failed since it exceeded the capacity of the available memory. Such an issue has motivated the development of a domain decomposition strategy in order to perform these 2D simulations of HEC-RAS.

The two computing strategies presented in the previous chapter can be combined. A domain decomposition method is required to partition the large-scale watershed into sub-basins. Then, each one of the sub-domains can be solved using the sparse parallel solver of HEC-RAS. The domain decomposition sequence organizing the resolution of sub-problems and communications between sub-domains is then automated by means of a Virtual Basic for Applications (VBA) driver written using the HEC-RASController functions. In this chapter, a proof of concept is proposed to evaluate the benefits of the domain decomposition method in HEC-RAS simulations.

The proposed strategy is applied to the modeling of historical and extreme flood events on three nested basins: the Moderbach ($89\,km^2$), the Albe ($405\,km^2$) and the Saar ($1800\,km^2$). The particularity of the Moderbach basin is that flooding the valley was not a sudden event but an expected and even more a planned event. We aim to model the flooding of the Moderbach valley caused by the ponds and dams built for this purpose in the period 1939-1940. In the Albe basin, the modeled event has a certain particularity showing three successive floods in a period of less than two months. In the Saar basin, we model one of the extreme events that occurred from 22 to 24 December 2010.

This chapter starts by presenting the methodology and tools used to implement the basin partitioning methods in HEC-RAS and to automate the sub-basins simulation process. Validation tests are performed to evaluate the gain brought by the implementation of the partitioning method in HEC-RAS. Then, this simulation strategy is applied to model the historical events in the Albe, Moderbach and Saar basins. Computed flow hydrographs and flood extents are compared to observed flow rates and flood maps.

5.2 Implementation of domain decomposition in HEC-RAS

The basin partitioning method is implemented in HEC-RAS through the RAS Mapper interface. The automation of this process using the HECRASController (Section 5.2.1) requires coupling a 1D model to the outlet of each 2D flow area (Section 5.2.2).

First, the sub-basins shapefiles are imported to implement an area-balanced domain decomposition method in HEC-RAS. Each sub-basin is meshed to discretize the 2D partial differential equations, either (4.1) and (4.2) or (4.1) and (4.3). Next, a 1D river model is added at the outlet of each sub-basin to automate the extraction of its outflow (Eq. 4.6) as proposed in (Goodell, 2018). Finally, the outflow of each sub-basin serves as an input to run its downstream sub-basin following Eq. 4.5.

In the following, we present the HECRASController (Section 5.2.1) used to automate the 2D HEC-RAS simulations requiring to couple a 1D model to the 2D flow areas (Section 5.2.2). The performance of this computing strategy is evaluated by an application on the Moderbach basin (Section 5.2.3).

5.2.1 HECRASController

The automation of HEC-RAS from an external program is possible using the HECRASController introduced in Goodell (2014). The HECRASController is an Application Programming Interface (API) designed to make transparent use of HEC-RAS to drive a sequence of simulations.

In a nutshell, HECRASController provides a collection of routines that allows to open a project, to select and to run the so-called plans, and to extract output data. This offers the opportunity to perform a complex calculation process involving several models. The routines are available in several programming languages, notably Visual Basic for Applications (VBA-in Excel).

The VBA script, see Fig. 5.1, organizes the automation of the computation sequence of multiple sub-basins to account for the natural water flow. It runs the 2D-1D models (Section 5.2.2) from the most upstream sub-basin to accumulate the streamflow at the most downstream sub-basin outlet. Each sub-basin calculation is carried out in one go, providing all the necessary input flow data to its downstream sub-basin.

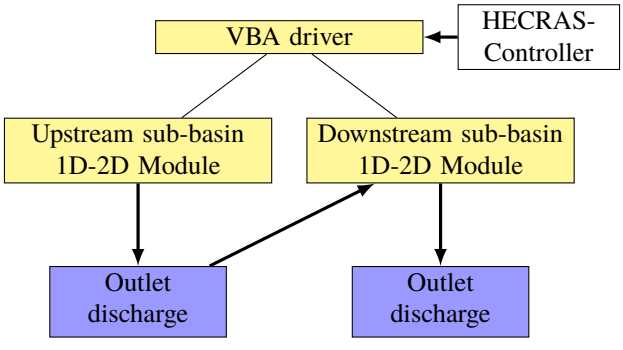


Figure 5.1: VBA algorithm for the automation of sequential domain decomposition with the HECRASController.

5.2.2 2D-1D modeling

The main challenge in the automation of 2D rain-on-grid simulations in HEC-RAS is that the HECRASController was mainly designed to deal with 1D problems. As proposed in (Goodell, 2018), connecting a 1D model to the outlet of the 2D flow area allows to automate the flow computation at the outlet of the sub-basin (Fig. 5.3). The outflow is computed at the river cross-sections using the HECRASController function **OutputDSS_GetStageFlow**. The computed outflow is used as inflow data for the simulation of its downstream sub-basin. As a downstream boundary condition, the 1D model uses a normal depth (slope) set to the most downstream cross-section.

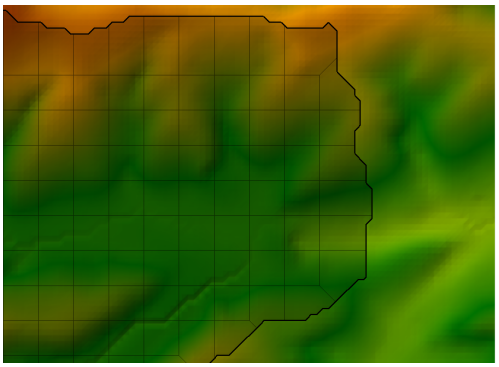


Figure 5.2: 2D model outlet in HEC-RAS.

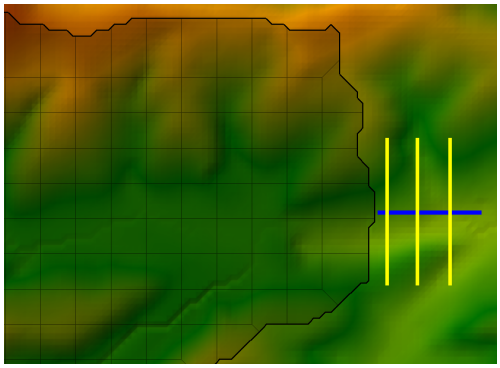


Figure 5.3: 2D-1D model in HEC-RAS (showing the stream in blue and the cross-section in yellow).

5.2.3 Application to the Moderbach watershed

5.2.3.1 Model setup

The performance of implementing basin partition methods in HEC-RAS is evaluated and analyzed through its application to perform rain-on-grid simulations on the Moderbach basin. In the following, two different basin partitioning methods are used and compared in order to test the impact of the partitioning method on the computational efficiency. The Moderbach decomposed into 3, 5, and 9 sub-basins using the confluence method, see Section 3.2.2, and the Moderbach decomposed into 2,4 and 8 sub-basins using the unguided area-balanced method, see Section 3.4.2, are implemented in HEC-RAS (Fig. 5.4).

We consider a flood event due to a precipitation of 80mm/day for 28 h allowing a rapid saturation and continues for 44 h with a dry condition in order to reach a null discharge at the basin outlet (Fig. 4.2). The downstream boundary condition at the 1D river is set to a slope of 0.001, estimated using RAS Mapper tools. The Manning coefficient is set to 0.03.

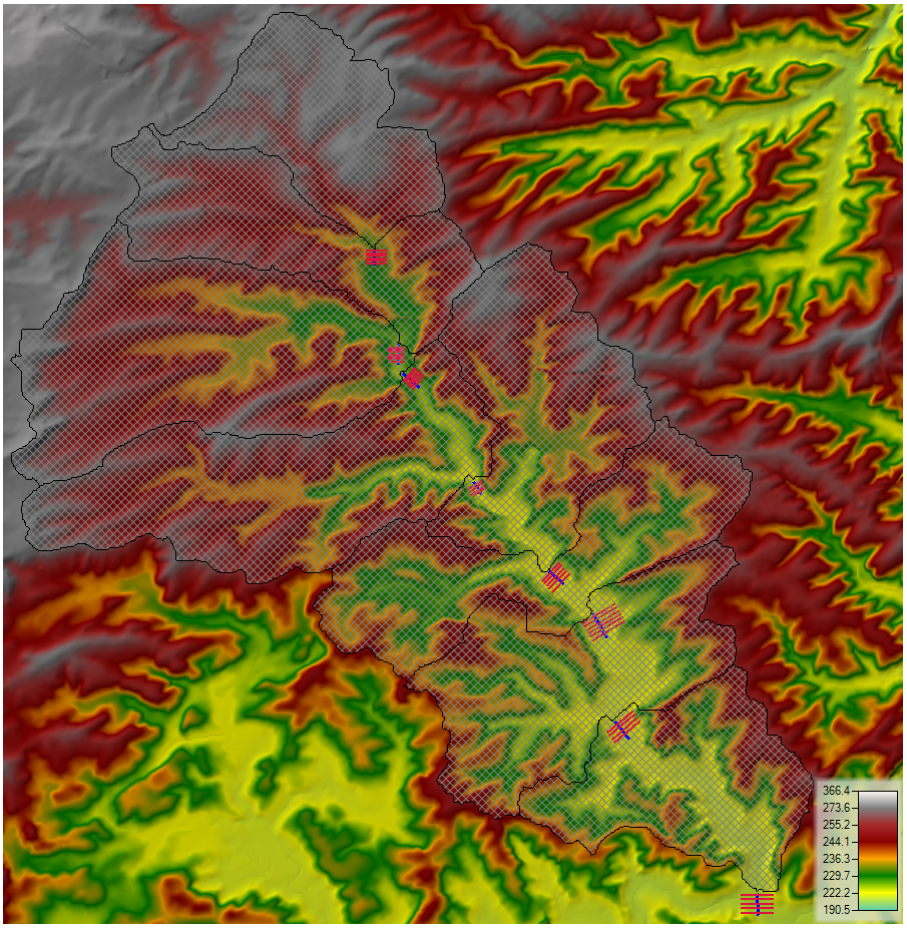


Figure 5.4: HEC-RAS model setup of the Moderbach in 8 sub-basins. Each sub-basin (2d flow area) is coupled to a 1D model (river in blue and cross sections in red).

Figures 5.5 and 5.6 compare the VBA algorithms that organize the sequential simulation of the Moderbach decomposed into 8 sub-basins with the unguided area-balanced method and 9 sub-basins with the confluence method respectively. Each black arrow connecting two rectangles corresponds to an execution performed by the VBA code using the HECRASController functions. On the one hand, the partitions generated by the confluence method allows each sub-basin to strictly have none or two inflows. On the other hand, the partitions provided by the unguided area-balanced method allows for sub-basins having none or multiple inflows (1, 2, 3 or more). For this partition into 8 sub-basins, we have 0 or 1 inflow whereas the partition into 16 sub-basins had possibilities of 1, 2 and even 3 inflows for a sub-basin (Fig. 4.24).

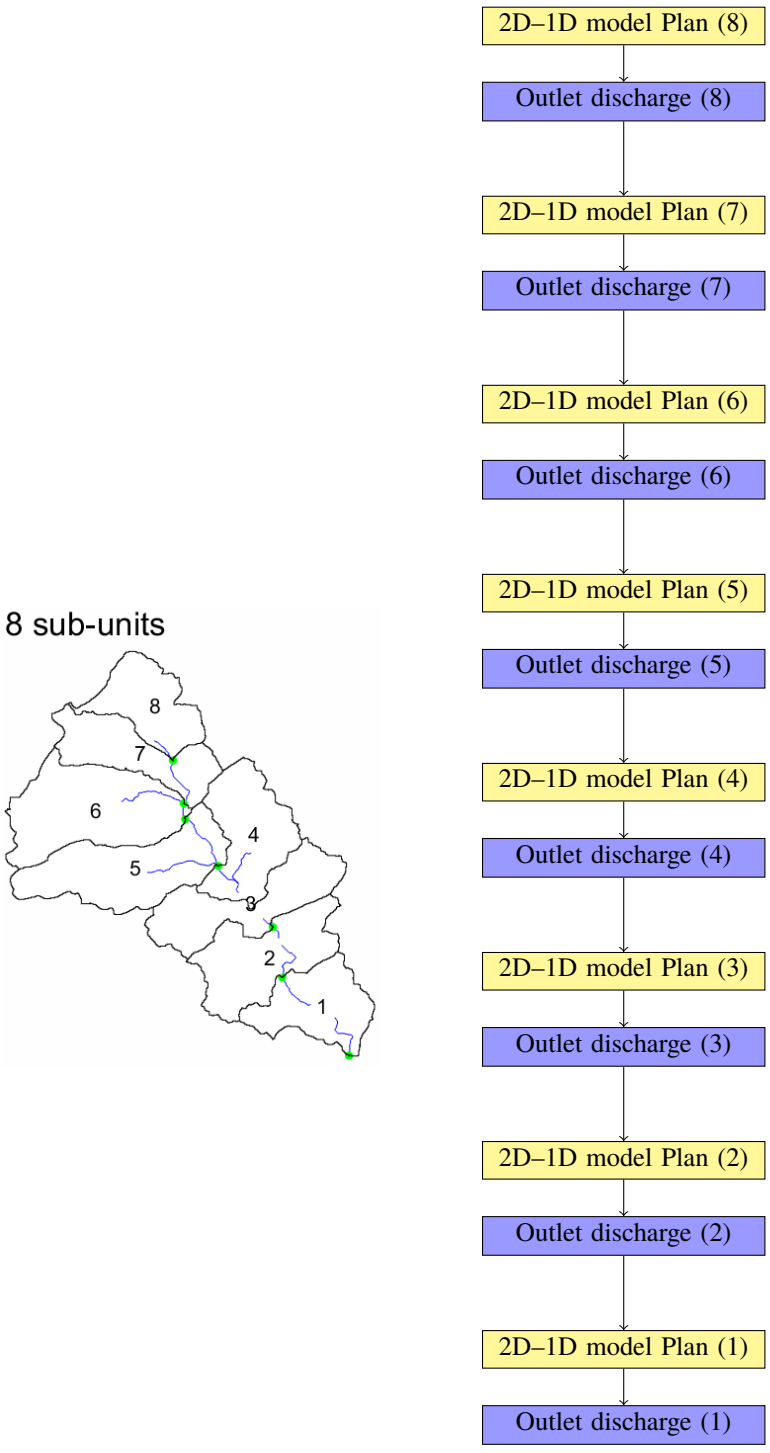


Figure 5.5: VBA algorithm organizing the sequential simulation of the 8 sub-basins of the Moderbach generated with the unguided area-balanced method.

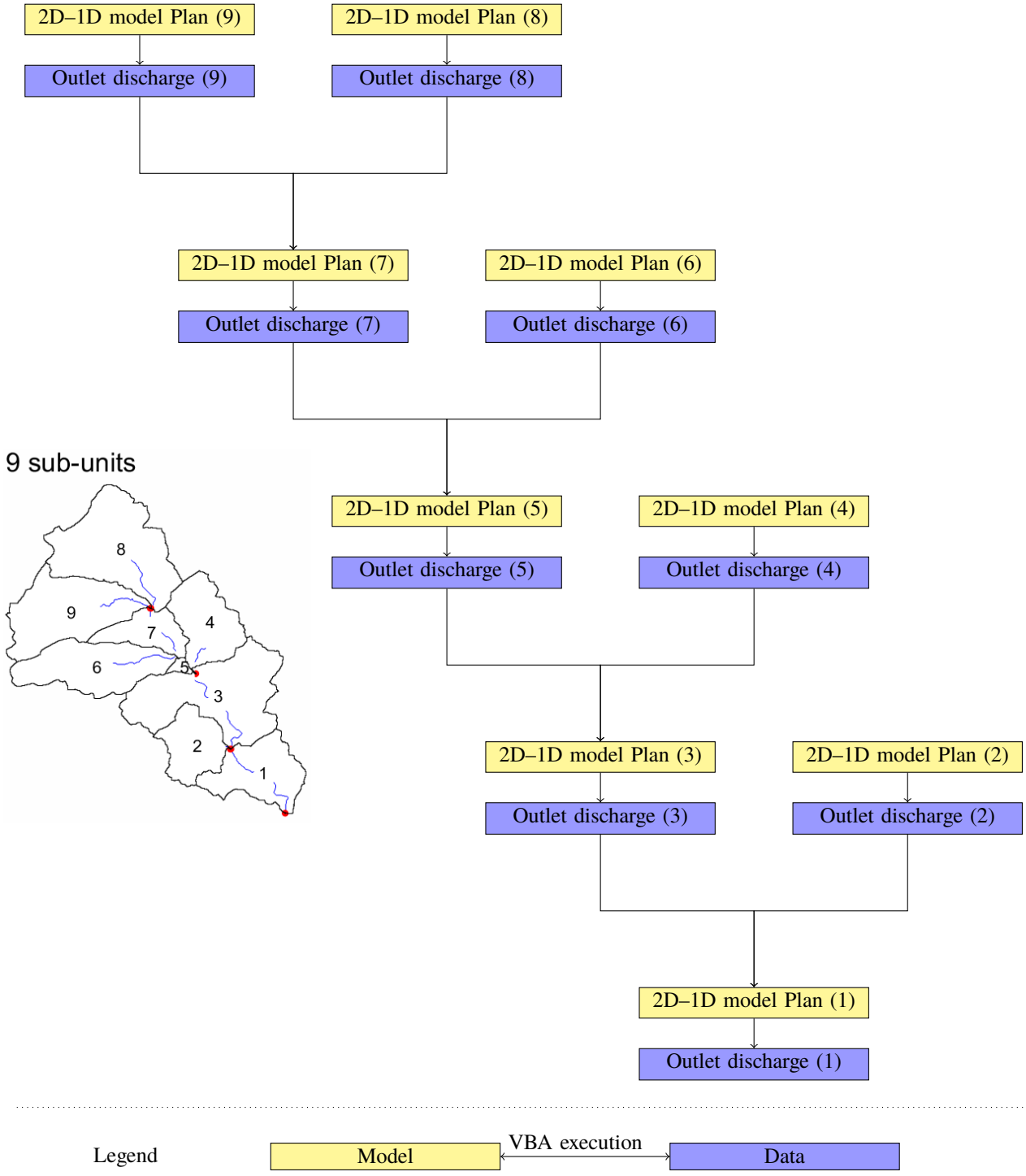


Figure 5.6: VBA algorithm organizing the sequential simulation of the 9 sub-basins of the Moderbach generated with the confluence method.

5.2.3.2 Hydrological results

The process of the sequential simulation of multiple sub-basins in HEC-RAS is validated by the ability of the decomposed basin configurations to provide the same hydrological behavior as the simulation of the entire basin. This hydrological behavior is represented by means of the basin outflow. Fig. 5.7 and Fig. 5.9 compare the flow hydrographs computed at the outlet of the Moderbach decomposition configurations 2, 4, 8 and 3, 5, 9 respectively with the flow hydrograph computed in the simulation of the entire basin. The outflow computed in the decomposed basin configurations, with both partitioning methods, shows a strong agreement with that computed in the entire basin simulation. The relative errors displayed in Fig. 5.8 and Fig. 5.10 exhibit very small discrepancies. These may be due to the value of the slope parameters, to topographical effects, and to the partition method since the null flux boundary condition (Eq. 4.7) is applied onto the boundaries $\partial\Omega_{side}$ for the catchment and the subunits, respectively.

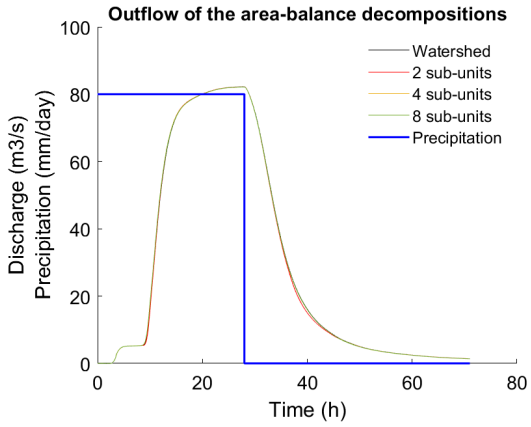


Figure 5.7: Comparing the flow hydrograph of the basin decomposed in 2,4 and 8 with the one of the total basin.

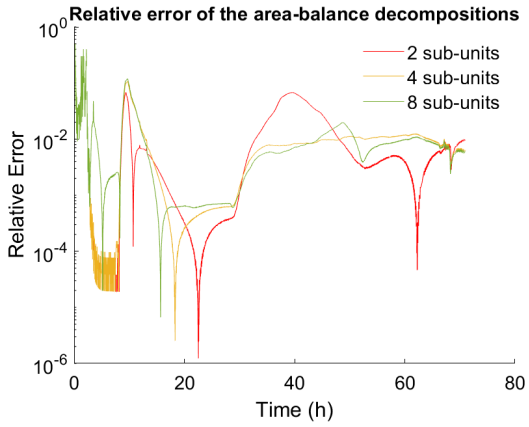


Figure 5.8: Relative error of the outflow of the basin decomposed in 2, 4 and 8 evaluated with respect to the total basin.

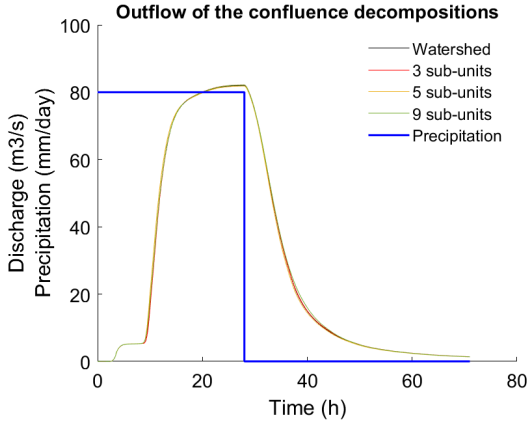


Figure 5.9: Comparing the flow hydrograph of the basin decomposed in 3,5 and 9 with the one of the total basin.

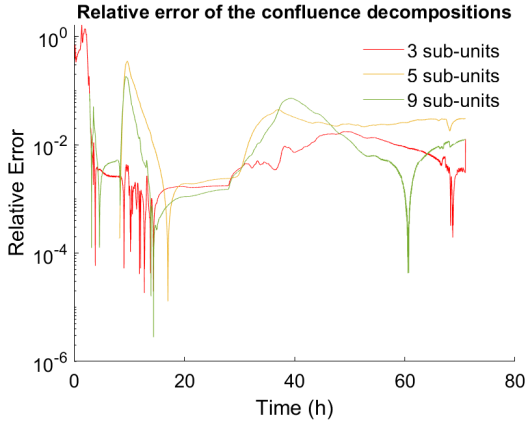


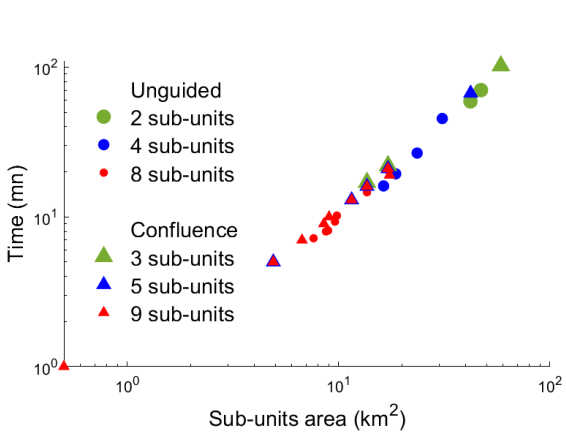
Figure 5.10: Relative error of the outflow of the basin decomposed in 3, 5 and 9 evaluated with respect to the total basin.

5.2.3.3 Computational efficiency

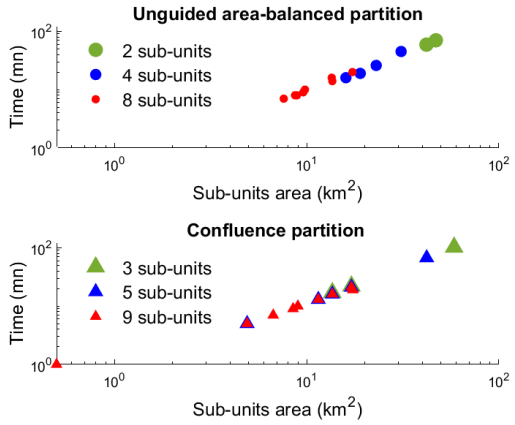
To evaluate the performances of the sparse parallel solver of HEC-RAS and the general behavior of the basin partition method, an overview of the computation times with respect to the sub-unit areas and the partition methods is provided in Fig. 5.11.

Triangles and filled circles are used for the unguided area-balanced partition method and the confluence partition method, respectively. As expected, points are almost aligned in this log-log representation, see Fig. 5.11.a. The computation time and the sub-basin areas are linked by a power law, the slope of which reflects the performance of the HEC-RAS sparse parallel solver.

The differences in the solution times for the sub-basins of very similar areas indicate that the latter is a key parameter, but not the only one in this nonlinear solver. Indeed, Fig. 5.11.b exhibits three colored clusters (filled circles of the area-balanced method) with very similar area and computing time. In contrast to that, the dispersion of the triangular points is much higher for the confluence method as very small sub-units are created.



(a) Computation time



(b) Clusters presentation

Figure 5.11: Computation time per basin for the unguided area-balanced method (triangles) and the confluence method (filled circles).

The main characteristics of these solution times are presented in Tables 5.1 and 5.2. The total time and the speedup are plotted in Fig. 5.12 against the number of sub-basins. The speedup is the ratio of the computation time for the decomposed basin to that of the entire basin. We see that for both partitioning methods, the decrease in computing time is multiplicative. For example, the times for the unguided balanced partition (Tab. 5.1) show a speedup of 1.4 for the simulation of the configuration with 2 sub-basins and 2 for 8 sub-basins. For the confluence partition (Tab. 5.2) we have a speedup of 1.30 for 3 and 1.8 for 9 sub-basins.

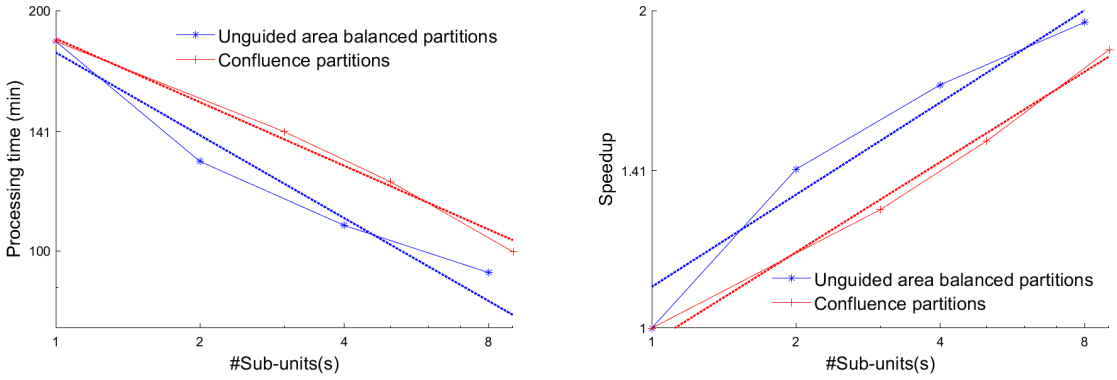
Properties	Unguided partitions			
#Sub-units	1	2	4	8
Max. time	183.0	70.3	45.4	20.1
Min. time	183.0	59.1	16.1	7.2
Ratio	1.0	1.2	2.8	2.8
Mean time	183.0	64.7	26.9	11.7
Std dev.	-	7.9	13.1	4.7
Total time	183.0	129.4	107.6	93.9
Speed-up	1.0	1.4	1.7	2

Table 5.1: Computation time and speedup data for the HEC-RAS simulations of unguided partitions

Properties	Confluence partitions			
#Sub-units	1	3	5	9
Max. time	183.0	102.0	66.8	21.3
Min. time	183.0	17.1	4.7	0.8
Ratio	1.0	6.0	14.3	25.5
Mean time	183.0	47.1	24.3	11.1
Std dev.	-	47.6	24.5	6.7
Total time	183.0	141.3	121.6	99.7
Speed-up	1.0	1.3	1.5	1.8

Table 5.2: Computation time and speedup data for the HEC-RAS simulations of confluence partitions

As we have seen, the computation of the linear system generally grows super-linear in the size of the parts, and thus with Jensen's inequality (Jensen, 1906) the sum of the times of the linear systems of all parts is much smaller than for the linear system of the entire basin. Additionally, Fig. 5.12 shows regressions of the form ax^b, where b is -0.27 and -0.32, respectively. This clearly shows the advantage of using basin partitioning over the simulation of the entire basin in hydrological modeling, but also that the unguided area-balanced partition method improves the confluence method by an important factor.



(a) Time (b) Speedup

Figure 5.12: Basin partitioning method-HEC-RAS computing times.

5.3 Historical and extreme scenarios

The second part of this chapter presents a selection of the historical extreme flood events that occured in the Saar area, each with its own particularity and interest. The events are modeled by taking advantage of the improvement brought to the parallel computating in HEC-RAS, presented earlier in this chapter. Three case study basins are considered in this chapter the Albe, the Moderbach and the Saar.

5.3.1 Albe: December 2011 - January 2012

The first scenario to be modeled in the Albe basin (409 km^2) corresponds to the period of December 2011 - January 2012. This period has a certain particularity since it has experienced three flood events in less than two months.

Simulation of the entire basin in HEC-RAS may require a long simulation time. Therefore, this scenario takes advantage of the strategy implementing domain partitioning in HEC-RAS to perform sequential simulation of multiple sub-basins instead of running the simulation on the entire basin.

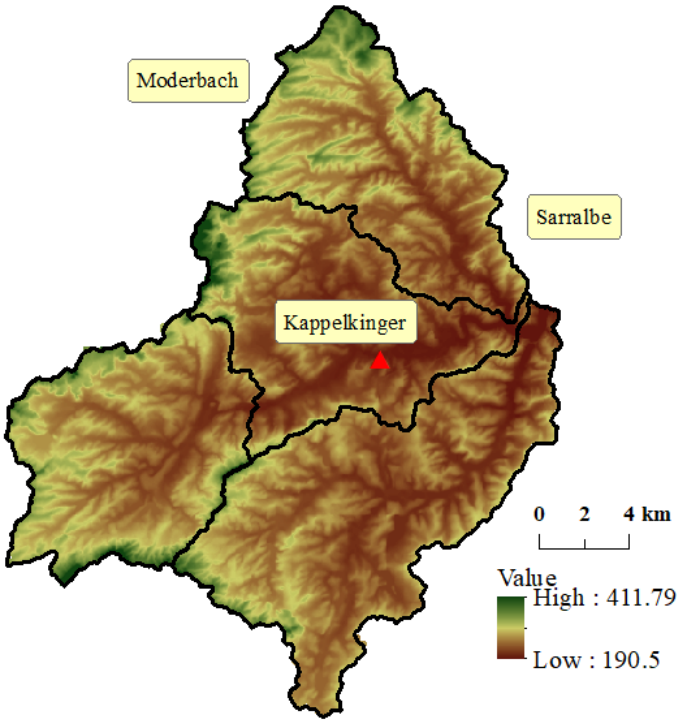


Figure 5.13: The Albe basin decomposed to 4 sub-basins using the guided area-balanced method.

The Albe basin is decomposed into 4 sub-basins using the guided area balanced method, which allows to consider the Moderbach watershed as a sub-basin of this decomposition (Fig. 5.13). The area and the computation time of the 4 sub-basins are presented in the table 5.3.

The computation time for the sub-basins should not be compared with that of the Moderbach or Saar simulations presented later in this chapter. Indeed, the modeled event duration are different. This allows to associate different timestep conditions to provide mapping and outlet hydrograph data to each simulation depending on the duration of the modeled event. For this simulation seeking to model a 60-day period, a timestep of 1 day for the HEC-RAS conditions mapping output and hydrograph output is sufficient to provide a detailed description of the event and to provide results on a daily basis to be compared with the observed daily flow data.

The input data required for this simulation are the precipitation data provided by the Kappelkinger station, located in the basin (Fig. 5.14). The total amount of precipitation recorded in in this period of December 2011 - January 2012 is 193 mm.

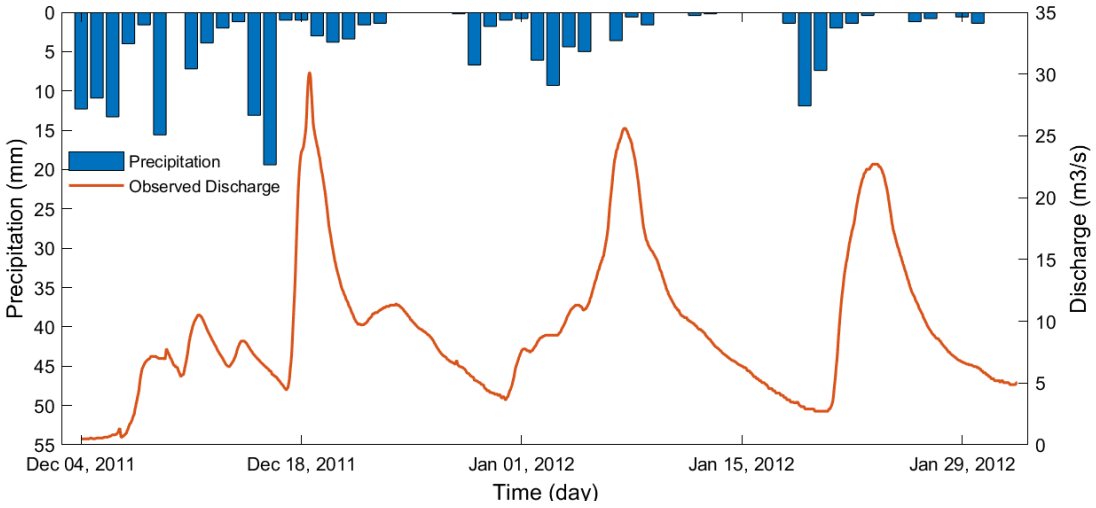


Figure 5.14: Precipitation and observed discharge in the Albe basin for the period of December 2011 - January 2012.

Since no flood map is provided for this event showing the inundated area, the outflow of the basin is considered to validate the model results. The discharge computed at the outlet of the Albe basin is compared to the observed discharge provided by the gauging station l'Albe a Sarralbe (Station A9192060) and available on the Banque Hydro website `http://www.hydro.eaufrance.fr/` (Fig. 5.14).

Albe's sub-basins	1	2	3	4
Area (km^2)	89.19	130.02	94.42	92.09
Computation time (h)	12.05	16.24	12.26	11.98

Table 5.3: Area and computation time for the Albe sub-basins

The computed and observed outflows are presented in Fig. 5.15. The differences in water volume are presented in Tab. 5.4. For the first peak, the simulation overestimates the observed data, the computed discharge is almost two times the observed discharge. The simulated water volume exceeds the observed data by 51%. We assume that this excess of water is because HEC-RAS do not take into account the infiltration process. Therefore, all the precipitation is modeled as runoff which overestimates the surface flow.

Flooding event period	4-31 December	1-18 January	18-30 January
Observed Water volume (m^3)	5400.75	4662.40	3487.19
Computed water volume (m^3)	10922.97	4732.16	2950.78
Relative error %	51	1	18

Table 5.4: Water volume for computed and observed data.

The excess of precipitation during the first flood event has a major impact on the soil condition, leading to a saturation of the upper soil layers. The clayey nature of the soil in this area, see Section 1.4.1, allows for a rapid saturation. Considering these soil conditions for the second and third flood events, the infiltration process can be considered negligible since the soil is saturated, which corresponds to the HEC-RAS model conditions. This explains the strong match of the computed and observed flow hydrographs for the second and third flood events with a water volume variation of 1% and 18% respectively.

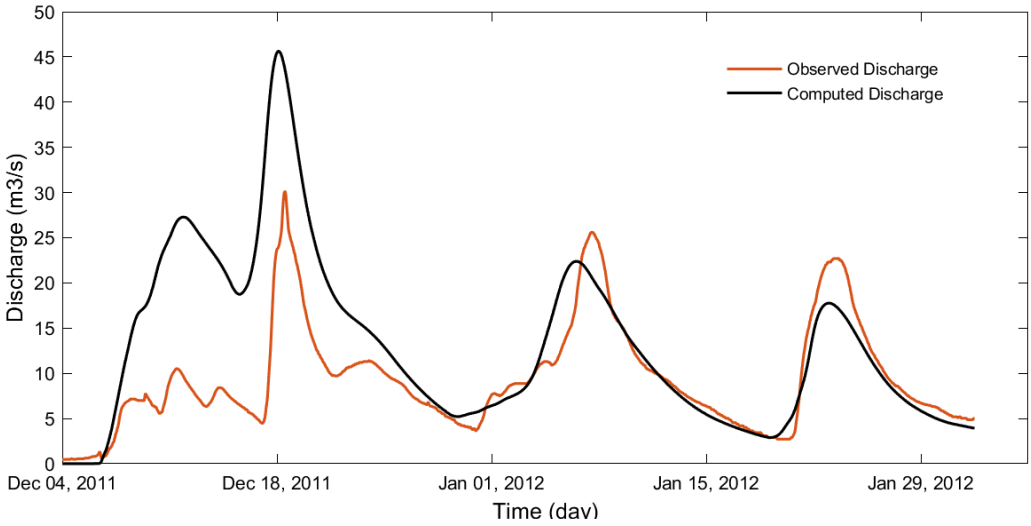


Figure 5.15: Computed and observed discharge at the outlet of the Albe watershed.

5.3.2 Moderbach: 1939 - 1940

The long occupation of the French territory during the First World War brought to light the need to make the national territory secure and inviolable. In 1919, the French leaders gathered in the Superior Council of War to come up with a territorial defense strategy.

The return of the Alsace-Lorraine had given the country a new border to protect (Marque, 1989). Pétain wanted to build a continuous front, an organized battlefield on the border, as he had done in 1918 in Champagne (Oldham, 1997). However, the idea of a Chinese wall did not please everyone due to the difficulty of building a continuous fortification along the border while the country was not able to pay for it. Marshal Joffre proposed to create a defense system of discontinuous fortified areas. At that time, the Minister of War Painlevé is considered the actual initiator of the defensive fortified region project. André Maginot, Painlevé's successor, voted the law of January 1930 to fund the project (Marque, 1989). The plan was to build a defense system for the entire northern, eastern and southern border, including Corsica, focusing on the German border in three regions: in front of Metz (Fig. 5.16), on the northern border of Alsace, and in front of Belfort.

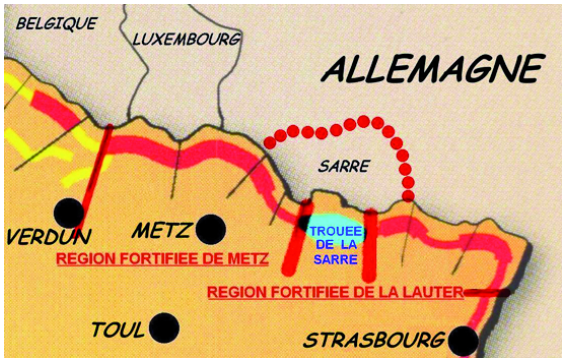


Figure 5.16: Line of discontinuous fortifications in front of Metz (Becker, 2007).

A coherent defensive waterline was created in the Moderbach basin in about two years (1932-1934), from Hoste to Sarralbe over a distance of 16 km. This defense line was based on six reservoirs and five dams (Fig. 5.17).

This strategy was implemented by building a combination of reservoirs and dams to store and increase the water level in the valley. Figure 5.18 shows the capacity of the reservoirs, the dams and the longitudinal profile of the Moderbach basin valley. An impressive documentation work on this project was done by Paul Marque (Marque, 1989). These documents were handed over to Philippe Keuer who agreed to share them with us as part of this research work.

The Hoste-Haut and Hoste-Bas ponds are intended to fill the Loupershouse forebay. The Diefenbach pond was created to fill the forebay of Puttelange-les-Farschvillier. The Welschof pond, which is the only one located on the left bank of the river facing the enemy, was mainly intended to fill the Rémering-les-Puttelange forebay and contributed to fill the forebay of Richeling with the pond of Marais de Rémering. Finally, the pond at Hirbach is used to fill the Holving forebay. The characteristics of the ponds and forebays are presented in the table 5.6.

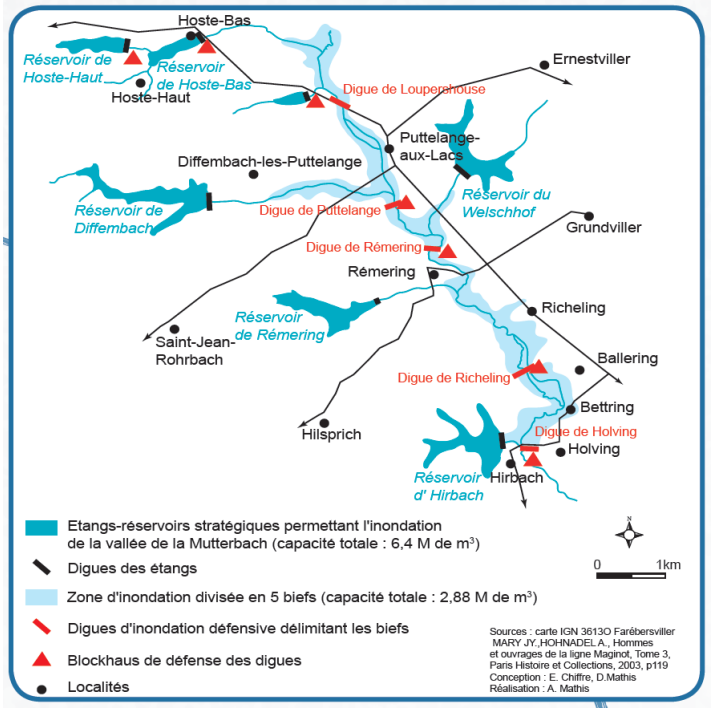


Figure 5.17: Defense water system in the Moderbach basin (Mathis et al., 2017).

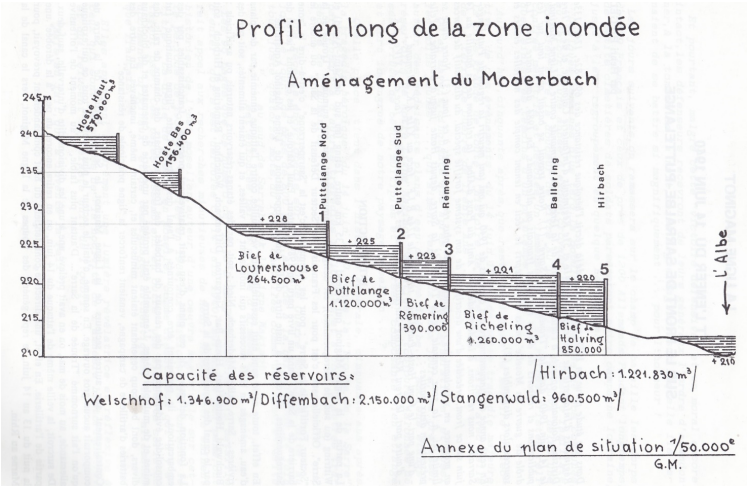


Figure 5.18: Dams and ponds of the Maginot line in the Moderbach basin (Marque, 1989).

Nowadays, these ponds have become ponds for fishermen and tourists. They are regularly surveyed to detect any possible problems. The Diefenbach pond, which is the largest of the ponds of the aquatic Maginot line, has been dry for 3 years now. According to the president of the SILMA (intercommunal union of the aquatic Maginot line), during a regular maintenance phase of the Diefenbach pond a communication problem between the pond and the dike was detected. This necessitated to drain the pond from the end of December 2018. The maintenance work will begin on April 1, 2022 with a budget of 1.5 to 2 million euros willing to fill the pond again in the end of 2022 (Chambru, 2021).

5.3.2.1 Scenario modeling and validation

These hydro-systems have become the main defensive systems in this area. They mobilize the water resources to spread voluntary flooding in the valley. Originally, the plan was to open and release the water from the five reservoirs to flood the valley in 36 hours against the attack of the German troops (Marque, 1989). In 1939, the heavy rainfall of October allowed to flood the valley taking advantage of the dams built on the main river.

In this section, the two flooding scenarios are investigated. Scenario 1 corresponds to the military plan established to flood the valley by the evacuation of the ponds. Scenario 2 corresponds to the inundation of the valley due to the heavy precipitation of October 1939. Both scenarios are modeled in HEC-RAS.

In order to validate the results of the modeled scenarios, two flood maps showing the flood extent in the Moderbach basin are used. The first flood map we have found in the library of the ASCOMEMO association, was designed by the Germans. The information about the inundated area was added to the map in June 1940 (Fig. 5.19). The second map was designed by the French army's geographic service. The information on the inundated area was added in September 1944 on Etat Major's map created in 1921 (Fig. 5.20).

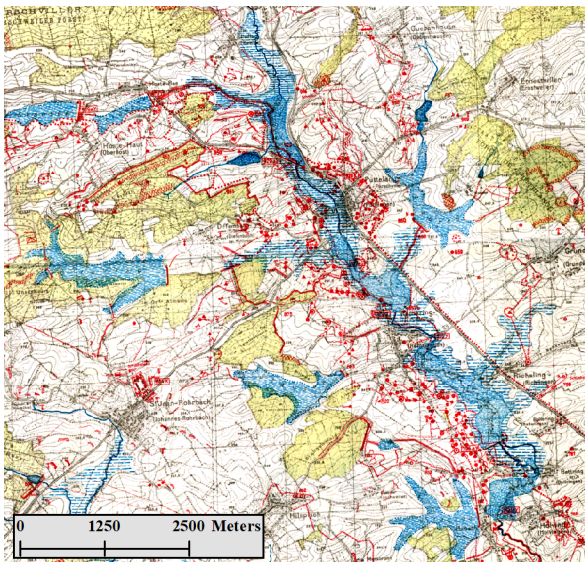


Figure 5.19: Flood map of 1940.

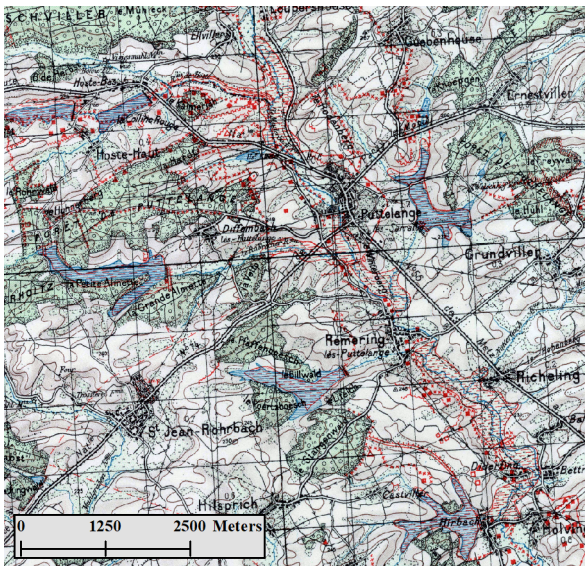


Figure 5.20: Flood map of 1944.

Prior to comparing the HEC-RAS simulation results with the flood maps, we thought it might be worthwhile to compare the flood maps between them since they were taken from two different sources. Thus, we started by georeferencing the two maps in the same coordinate system (Lambert 93) using QGIS.

Validation of the hydrological modeling results can be carried out with the fit parameter (Horritt et al., 2007) (Di Baldassarre et al., 2009) (Quirogaa et al., 2016). These parameters F1 and F2 has been shown to provide reliable results for flood modeling studies, allowing the comparison of model results with observed or constructed flood data.

The equations of F1 and F2 are based on the comparison between the observed flood (FO) and the simulated flood (FS):

$$F1 = \frac{A}{A + B + C}, \tag{5.1}$$

$$F2 = \frac{A - B}{A + B + C}, \tag{5.2}$$

where A is the flood extent correctly predicted by the model ($FS \cap FO$) which is the common flood extent of FS and FO, B is the overpredicted flood extent by the model which corresponds to the difference between the computed flood extent and the common flood extent (FS\A) and C is the underpredicted flood extent by the model which corresponds to the difference between the observed flood extent and the common flood extent (FO\A). The surface of a flood extent is expressed as S(flood extent)

These parameters give the degree of accuracy of the model: F1 varies between 0 and 1 and F2 between -1 and 1. Values close to 1 indicate higher accuracy.

Table 5.5 shows the computation of the fit parameters calculated to compare the two flood maps. In this case, both flood extents are observed data. Therefore, we will consider the 1940 flood map as our reference map and the flood extent on this map will correspond to FO while the one on the 1944 map will correspond to FS. A represents the common flood extent in both maps. B represents the area where the flood extent of the 1944 map overestimates the 1940 map, and C represents the area where the 1944 flood extent underestimates the 1940 map.

The ratio of the flood extents derived from the two maps is 0.46. The values of F1 and F2 are respectively 0.41 and 0.38 showing a correspondence between the two flood maps.

S(1940 map) (km^2)	S(1944 map) (km^2)	Ratio 1940/1944 -	S(A) (km^2)	S(B) (km^2)	S(C) (km^2)	F1 -	F2 -
5.41	2.50	0.46	2.31	0.18	3.10	0.41	0.38

Table 5.5: Fit parameters for comparing the two flood images.

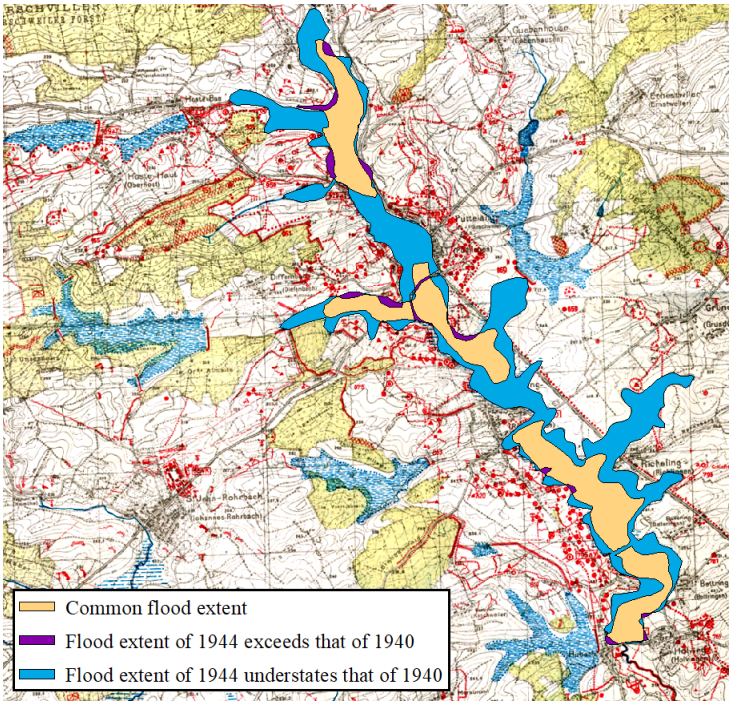


Figure 5.21: (A) Common flood extent (orange), (B) flood extent of 1944 exceeding that of 1940 (magenta), (C) flood extent of 1940 exceeding that of 1944 (blue) on the 1940 map.

The Figure 5.21 illustrates the differences between the flood extents of the two maps. The common flood extent A (orange) with an area of 2.32 km^2, is very close to the area of the flood extent in the 1944 map of 2.5 km^2. The flood extent of the 1944 map that exceeds the flood extent of the 1940 map, which corresponds to B (magenta), have a very small area of 0.18 km^2. Thus, the flood extent of the 1944 map is almost entirely contained within the flood extent of the 1940 map.

A large part of the flood extent presented on the 1940 map does not appear on the 1944 map, which corresponds to parameter C showing an area of 3.1 km^2. This indicates that the observations presented in the two maps have not considered the same conditions.

5.3.2.2 Scenario 1: Flood due to ponds evacuation

This scenario aims to model the flood extent generated by the evacuation of the 5 ponds in the Modebach basin to flood the valley. The modeled flood extent will be compared to the flood extents derived from the two flood maps designed by the Germans (1940) (Fig. 5.19) and by the French (1944) (Fig. 5.20).

The computational domain corresponding to the supposed flooded area is considered as the major bed of the valley with the connection to the downstream of each of the 5 reservoirs releasing water. Since the computational area is not too large (Fig. 5.22), the application of a basin partitioning method is not required. The computational domain was created using the 2D flow area editing tool of RAS-Mapper. A mesh grid is generated all over the computational domain considering a 25m mesh size (similar to the resolution of the DEM of the case study area).

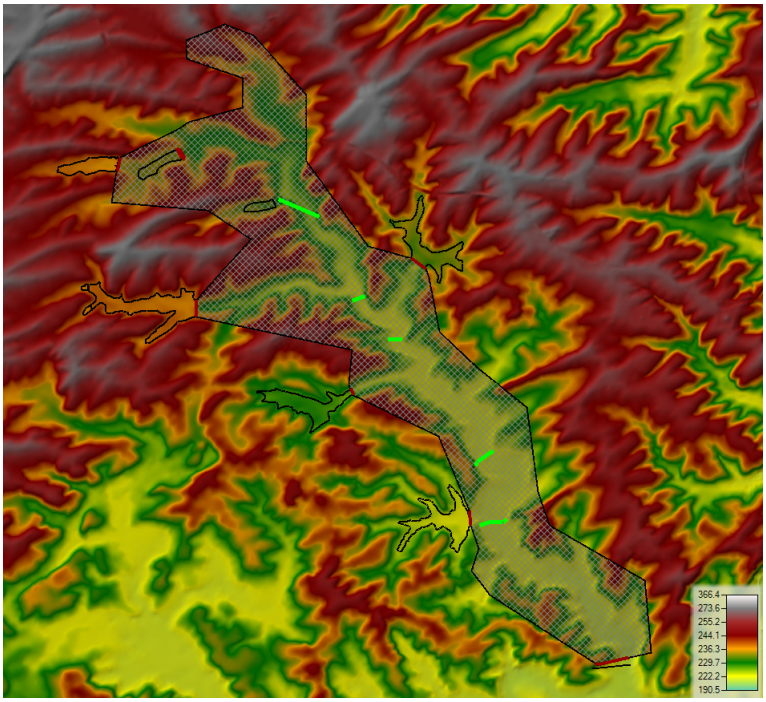


Figure 5.22: Computational domain for the ponds evacuation scenario (The dams of the defensive system built on the main river are shown in green).

The hydrological conditions of this scenario consider the evacuation of the 5 ponds in 36 hours. Therefore, 5 upstream boundary conditions of the type “flow hydrograph” were located respectively at the: Etang Hoste bas, Etang Diefenbach, Etang Welschof, Etang Marais, and Etang Hirbach. The simulation is performed for 36 hours which was the time set in the military plan to flood the valley.

According to the book of Paul Marque "La ligne maginot aquatique", the Hoste-bas reservoir would release a flow of 1.9 m^3/s to fill the Loupershouse forebay. Indeed, releasing this flow rate for 36 hours would generate a water volume equal to the maximum capacity of the Loupershouse forebay (264 500 m^3). As no information is available on the flow rates assigned to evacuate the 4 remaining ponds, we will consider this relationship between the flow rate of the ponds and the maximum capacity of the forebays to determine the flow rates of the 4 remaining ponds. Each pond must generate a flow allowing to fill, at the maximum, its corresponding forebay in 36 h. Table 5.6 shows the flow rates set to evacuate the 5 ponds.

Forebay	Loupershouse	Puttelange	Rémering	Richeling	Holving
Forebay capacity (m3)	264 500	1 120 000	390 000	1 260 000	850 000
Pond	Hoste Haut-Bas	Diefenbach	Welschof	Rémering	Hirbach
Ponds capacity (m3)	735 400	2 150 000	1 346 900	960 500	1 220 000
Ponds discharge (m3/s)	2.0	8.6	3.0	9.7	6.5

Table 5.6: Characteristics of the ponds and forebays of the Moderbach basin.

It should be mentioned that the flow rates set to evacuate the Diefenbach and Réremering ponds are considered relatively high. It is possible that the plan was to fill the large capacity of their corresponding forebays with the help of other reservoirs. For example, we know that the Welschof pond was not only intended to fill the Rémering-les-Puttelange forebay but also contributed to filling the Richeling forebay. Since the information is not clear on how this process was established, we stick to the flow rates we proposed in table 5.6 to model this scenario.

The simulation of this scenario in HEC-RAS required a computing time of 8.4 minutes. It was expected to have a short computing time since the computational domain is limited to the river bed. Figures 5.23 and 5.24 show the flood extent modeled in HEC-RAS (green hatched area) on top of the flood extent in the map from 1940 (blue hatched area) and 1944 (red hatched area) respectively. Table 5.7 presents the calculations of the fit parameters for the 1940 and 1944 flood maps, showing a good correspondence between the computed and observed flood extents. The fit parameters F1 and F2 show good values of 0.39 and 0.34 with respect to the 1940 flood map and 0.5 and 0.25 with respect to the 1944 flood map. The ratio of the computed flood extent to the 1944 flood extent is very close to 1, with a value of 0.98. However, the ratio is less than 0.5 with the 1940 flood map.

Flood image	S(FO)	S(FS)	S(FS)/S(FO)	S(A)	S(B)	S(C)	F1	F2
-	(km^2)	(km^2)	-	(km^2)	(km^2)	(km^2)	-	-
1940	5.41	2.46	0.45	2.21	0.24	3.10	0.39	0.34
1944	2.50	2.46	0.98	1.65	0.81	0.84	0.50	0.25

Table 5.7: Fit parameters for scenario 1 with respect.

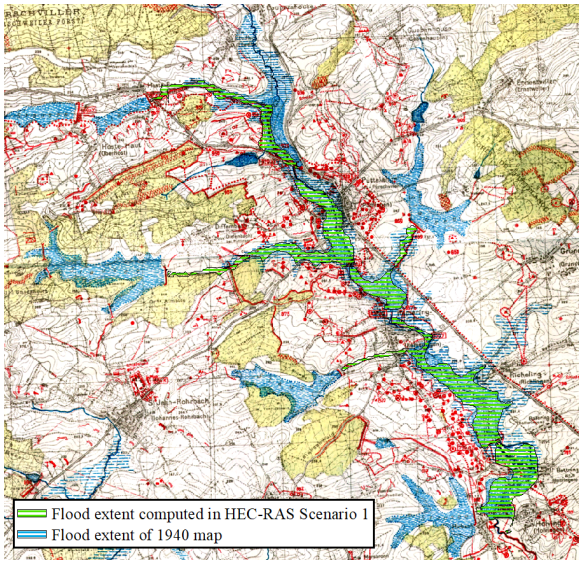


Figure 5.23: Scenario 1: HEC-RAS flood extent on the flood map of 1940.

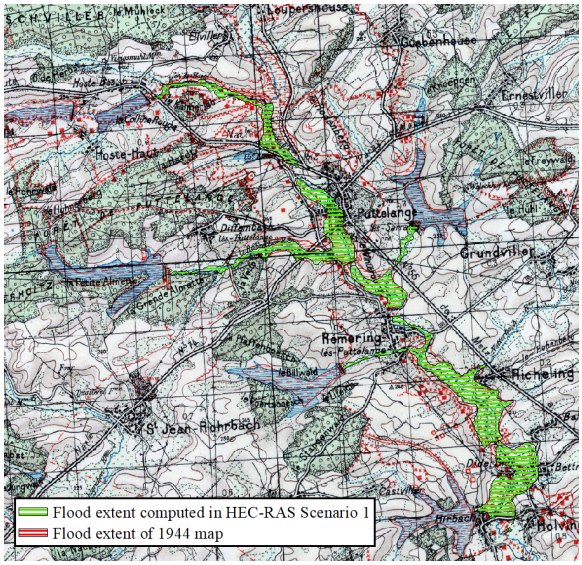


Figure 5.24: Scenario 1: HEC-RAS flood extent on the flood map of 1944.

Figures 5.25 and 5.26 show the common (orange), overpredicted (magenta), and underpredicted (blue) flood extents by the HEC-RAS model compared to the 1940 and 1944 flood maps respectively. The simulation was not able to correctly reproduce the flood extent in the most upstream part of the Moderbach basin in the Loupershouse reach. In both flood maps, this is an underpredicted flood extent (blue). For the rest of the basin valley, the model provided a good prediction of the flood extent in the 1944 map, whereas in the 1940 map the flood extent was poorly predicted. The blue polygon corresponding to the flood extent underpredicted by the model is dominant, as shown in Fig. 5.25.

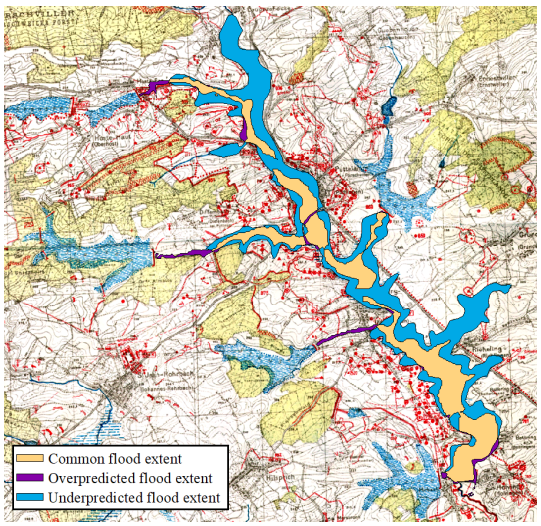


Figure 5.25: Scenario 1: Common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1940 flood map.

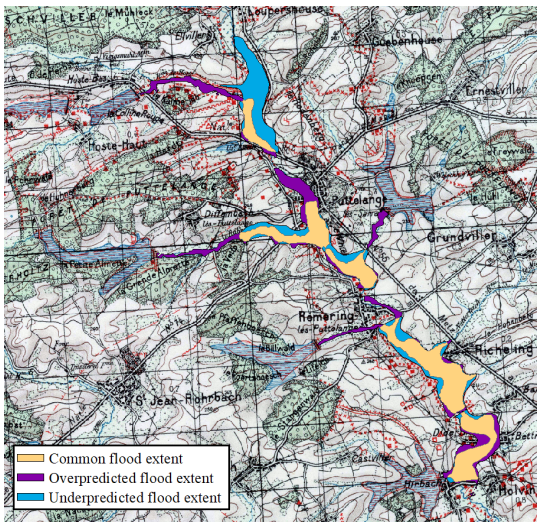


Figure 5.26: Scenario 1: Common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1944 flood map.

The figure below shows the hazard map produced in RAS Mapper in HEC-RAS. The flooded areas with a water depth over $2m$ are mainly located in the near upstream of the 5 dams on the main stream river. This illustrates the role of dams built on the main river to retain and increase the water level in order to flood the valley.

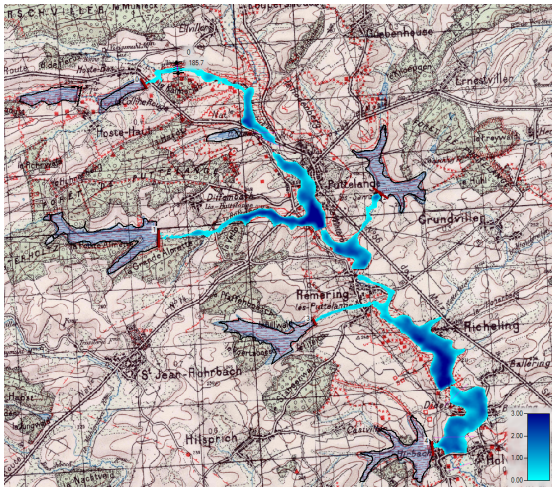


Figure 5.27: Flood hazard map.

5.3.2.3 Scenario 2: Flooding due to heavy rainfall

This scenario aims to model the flooded area in the Moderbach basin resulting from the heavy rainfall of October 1939. The modeled flood extent will be compared to the flood extents derived from the two flood maps of 1940 (Fig. 5.19) and 1944 (Fig. 5.20).

The simulation of this scenario requires considering precipitation data as hydrological input for the model, unlike scenario 1 which considered flow rates. The computational domain subjected to rainfall will correspond to the Moderbach basin with the outlet at the Holving dam, the last dam in the valley to retain water. Therefore, the Moderbach decomposition configuration into 8 sub-basins using the unguided area balanced method, see Fig. 5.4, is used to perform a sequential rain-on-grid simulations on its sub-basins except for the most downstream sub-basin (Fig. 5.28). The last sub-basin is not considered in this scenario because it is outside of the required computational domain (downstream of Holving dam).

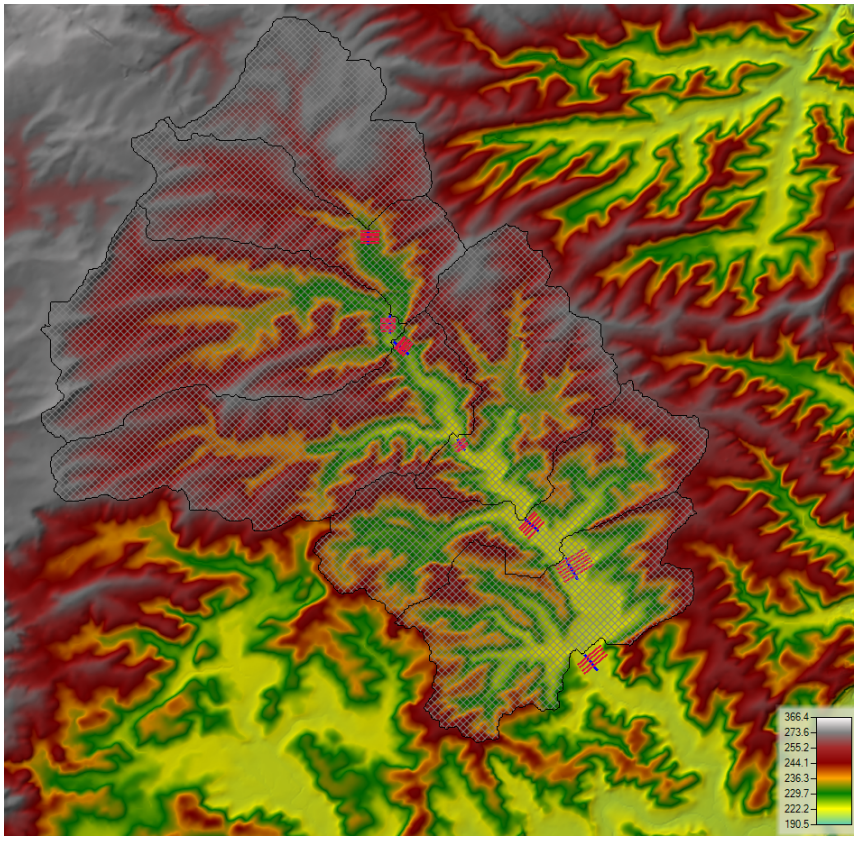


Figure 5.28: Computational domain for the precipitation scenario decomposing the Moderbach in 7 sub-basins where each sub-basin (2d flow area) is coupled to a 1D model (river in blue and cross sections in red).

The 7 sub-basins are imported into the RAS-Mapper as 2D computational domains. Each of them has been coupled, at its outlet, to a 1D river model in order to allow the automation of the sequential simulation using the HECRASController functions as presented in Section 5.2.2.

A mesh grid is generated on all the computational domains considering a mesh size of 25m (similar to the DEM resolution of the study area).

The precipitation amount recorded in October 1939 on the north east boundary of France was exceptional (Fig. 5.29). The station in Kaiserslautern has recorded 212 mm in October 1939 (Klein Tank et al., 2002) while the average precipitation for the month of October calculated over the period 1901-2018 is 48 mm (Fig. 5.30).

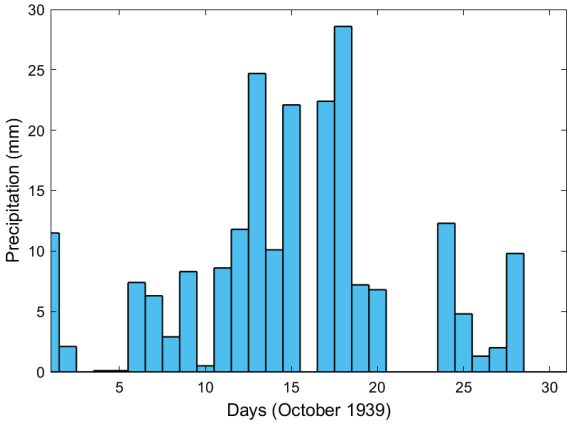


Figure 5.29: Daily precipitation for October 1939.

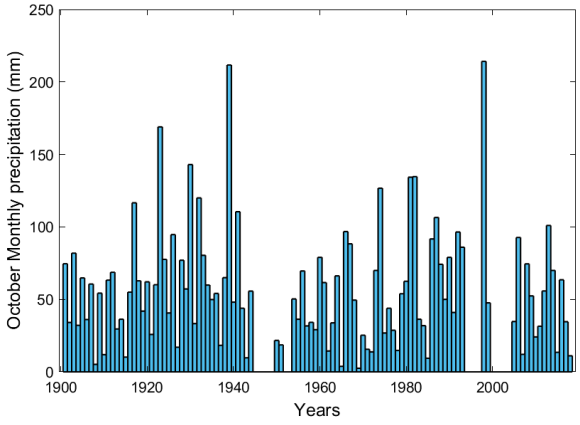


Figure 5.30: October monthly precipitation (1901-2018).

The precipitation data from 1 to 30 of October 1939 are added to the HEC-RAS model to perform a 2D rain-on-grid simulations. The computation time registered to perform the sequential simulation of these 7 sub-basins was 82.6 min with a timestep of 1 min. The flood extent calculated in HEC-RAS is compared to the 1940 (Fig. 5.31) and 1944 (Fig. 5.32) maps. Table 5.8 shows the fit parameter calculations with respect to the two flood maps.

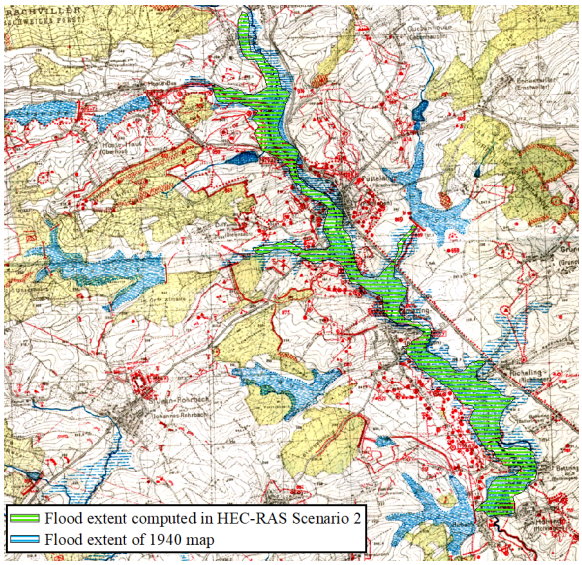


Figure 5.31: Scenario 2: HEC-RAS flood extent on the flood map of 1940.

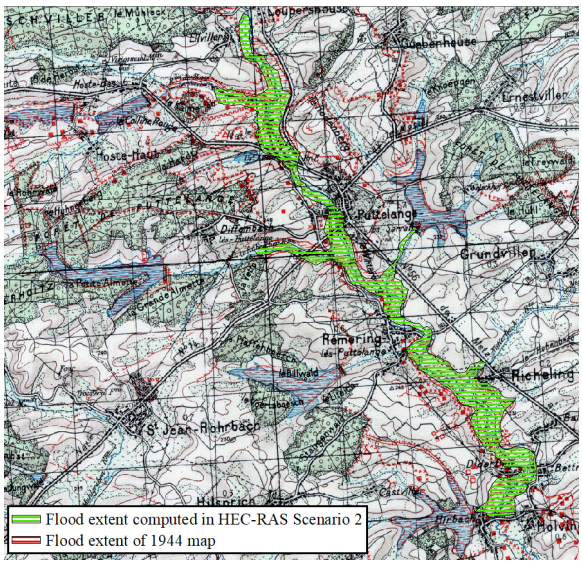


Figure 5.32: Scenario 2: HEC-RAS flood extent on the flood map of 1944.

The flood extent computed in this scenario covers an area of 3.1 km^2 which is larger than that computed in the first scenario modeling basin evacuation showing a flood extent of 2.46 km^2. This means that the heavy rainfall in October 1939 was more than sufficient to achieve the flooding required in the military plan. The fit parameters computation of F1 and F2 provided high values of 0.51 and 0.47 for the 1940 flood map and 0.59 and 0.30 for the 1944 flood map. This indicates a significant correspondence between the computed and observed floods extents

Flood image	S(FO)	S(FS)	S(FS)/S(FO)	S(A)	S(B)	S(C)	F1	F2
-	(km^2)	(km^2)	-	(km^2)	(km^2)	(km^2)	-	-
1940	5.41	3.10	0.57	2.87	0.22	2.54	0.51	0.47
1944	2.50	3.10	1.24	2.07	1.02	0.42	0.59	0.30

Table 5.8: Fit parameters for scenario 2.

Figures 5.33 and 5.34 show the common (orange), overpredicted (magenta), and underpredicted (blue) flood extents by the model HEC-RAS compared to the 1940 and 1944 flood maps respectively. The simulation of this scenario has significantly underpredicted the extent of flooding in the 1940 flood map where the blue polygon is dominant, see Fig. 5.33, with an area of 2.54 km^2. However, the simulation slightly overpredicted the extent of the 1944 flood. The parameter B corresponding to the flood extent overpredicted by the model (magenta polygon) has an area of 1.02 km^2 equal to almost half the common area.

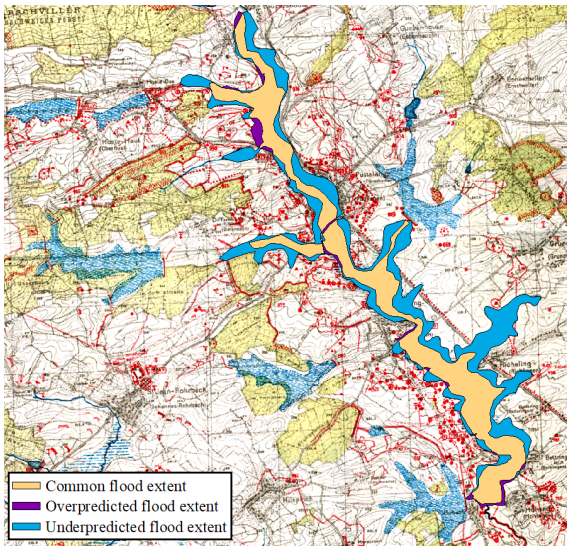


Figure 5.33: Scenario 2: common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1940 flood map.

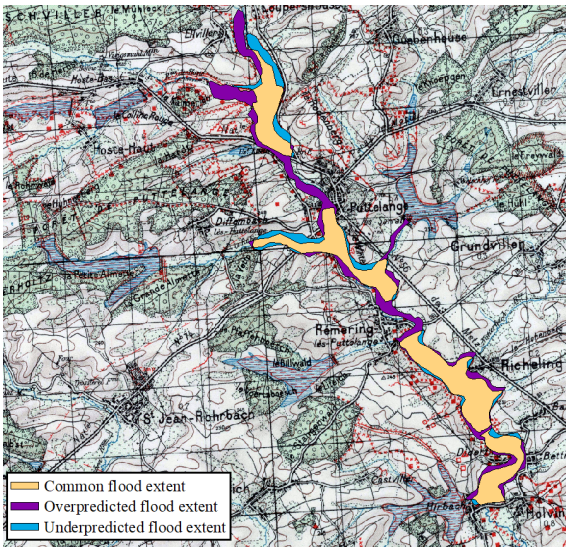


Figure 5.34: Scenario 2: common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1944 flood map.

5.3.2.4 Comparing scenarios

One of the most significant differences when comparing the two scenarios is seen in the forebay of Loupershouse upstream of the valley.

On the flood map for 1940, the observed flood extent covers the Loupershouse forebay. On the one hand, the results from the simulation of the first scenario, modeling the evacuation of the 5 ponds, show that the blue polygon corresponding to the underpredicted flood extent is dominant in this area (Fig. 5.35). In other words, this simulation did not succeed to model the flood extent in the Loupershouse forebay. On the other hand, the results from the simulation of scenario 2, modeling the heavy rainfall of October 1939, show that the orange polygon corresponding to the common flood extent reaches the Loupershouse forebay (Fig. 5.36). This means that this simulation allows to model the flood extent in this area.

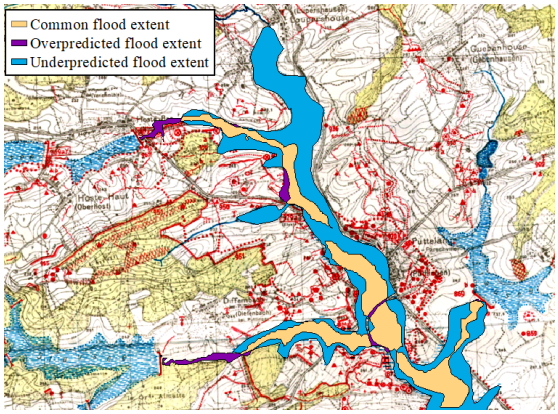


Figure 5.35: Scenario 1 on the Loupershouse forebay: common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1940 flood map.

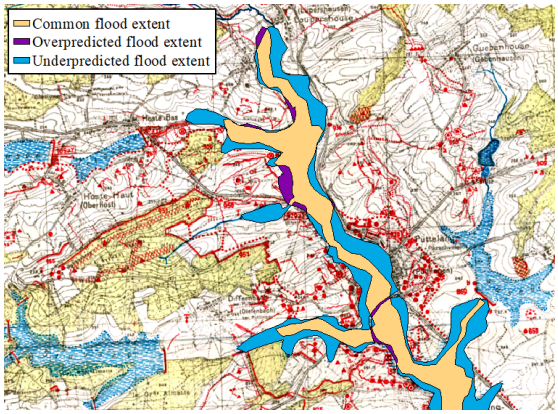


Figure 5.36: Scenario 2 on the Loupershouse forebay: common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1940 flood map.

This was also noted when comparing the results of scenarios 1 and 2 to the 1944 flood map. The simulation of scenario 1 underestimates the flood extent in the Loupershouse forebay (Fig. 5.37), whereas scenario 2 succeeds in modeling the flood extent in this area (Fig. 5.38).

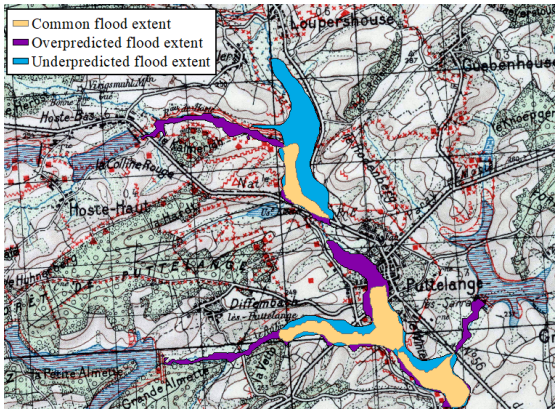


Figure 5.37: Scenario 1 on the Loupershouse forebay: common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1944 flood map.

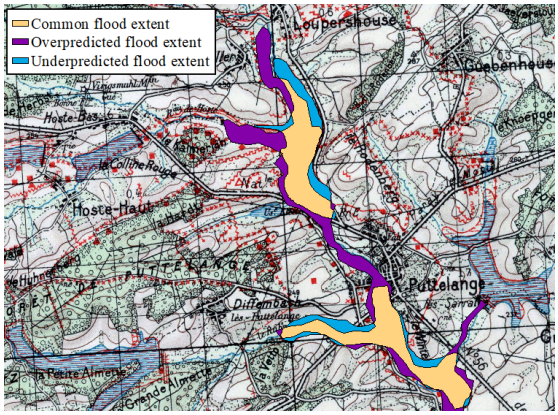


Figure 5.38: Scenario 2 on the Loupershouse forebay: common (orange), overpredicted (magenta) and underpredicted (blue) flood extent compared to the 1944 flood map.

These results are reasonable since the military plan aiming at flooding the valley by evacuating the ponds (scenario 1) would not be interested in flooding the upstream part of the valley. Therefore, we can state that these flood maps (1940 and 1944) showing an observed flood extent may not correspond to the flooding generated by the evacuation of the Moderbach ponds but mostly to the flooding caused by the heavy rainfall of 1939.

5.3.3 Saar: December 2010

The last scenario to be modeled is the flood event of December 2010. The case study area for this simulation is the French part of the Saar basin (1800 km^2). Twelve rasters issued from `BD Alti®25m - V2` are concatenated into a unique raster comprising 12.10^6 cells.

In a first attempt, we designed a HEC-RAS plan modeling the Saar catchment with an element size of 25 m, *i.e.* involving about 2,800,000 elements, to verify the occurrence of a memory overflow in HEC-RAS. Such an issue has motivated the development of the basin partition method in HEC-RAS simulations.

Eleven gauging stations are distributed along the main stream, the data of which may be found on `http://www.hydro.eaufrance.fr/`. The catchment is split into 20 sub-basins using the guided area balanced method to account for the 11 gauging stations located on the main river (red dots). Other sub-basins outlets are represented using green dots (Fig. 5.39). The sub-basins range from 11 km^2 to 164 km^2 in area, 17,253 to 262,183 in mesh elements, and 15 to 405 minutes in computational time, respectively. The simulation was completed in 86 hours considering a 10 s timestep.

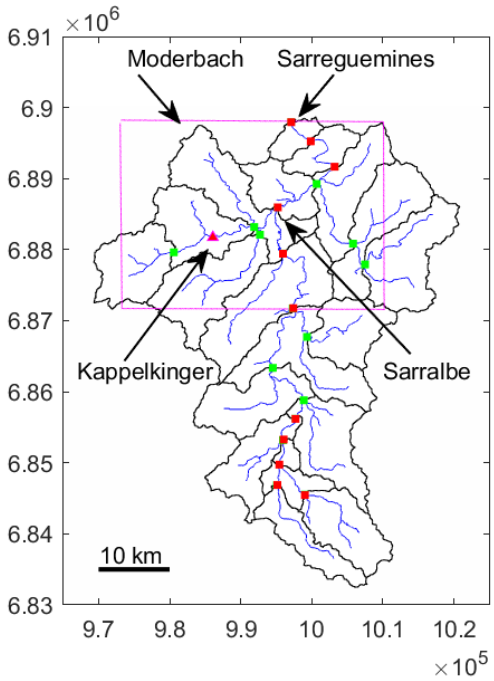


Figure 5.39: Area balanced partition of the French Saar catchment into 20 sub-basins at gauging stations (red dots) and additional points (green dots). The magenta box delineates the observation area.

We select the stations located in the two main cities, namely Sarralbe (Station A9200100) and Sarreguemines Centre (Station A9311050) to compare computed and measured discharges for the extreme flood event that occurred from 22 to 24 of December 2010. A flood map and the related shapefiles produced by the SERTIT (Service régional de traitement d'images et de télédétection, Strasbourg) from SAR Cosmo-Skymed images (5 m resolution) are used for comparison with our simulation results (Fig. 1.18). It should be noted that these images tend to underestimate urban and forest areas. In addition, permanent water bodies (rivers and ponds) and very small water objects (less than 100 m^2) were not reproduced.

When looking at the precipitation data for early December, we notice that the flood event of 22-24 December was preceded by heavy precipitation a few days before (Fig. 5.40). The heavy precipitation during the period of 6-9 December recorded 50 mm at the Kappelkinger gauging station. This event had certainly saturated the soil. Therefore, the simulation of the rainfall event of 22-24 December did not dramatically overestimate the event, as in the case of the first extreme event modeled in the Albe basin (Section 5.3.1). This makes the soil situation consistent with the HEC-RAS condition, which does not consider any infiltration effect.

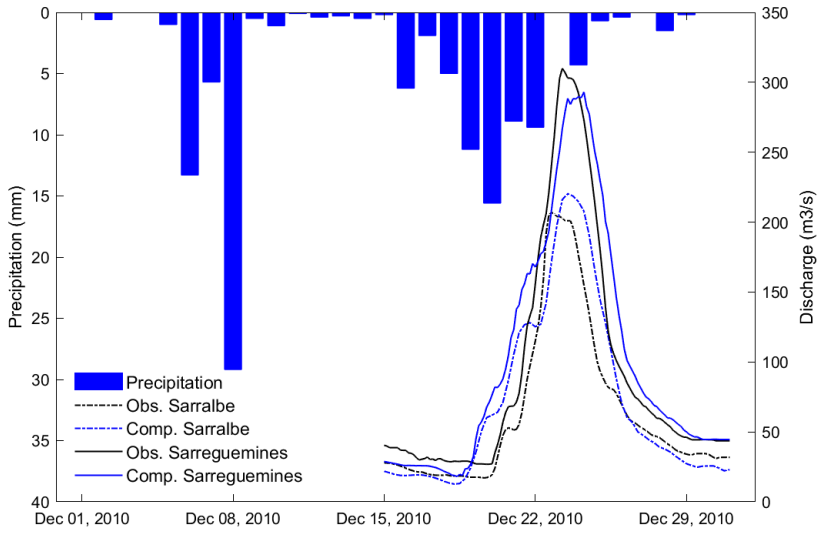


Figure 5.40: Precipitation, observed and computed discharges.

Precipitation data, discharge data, and results (at Sarralbe and Sarreguemines) plotted in Fig. 5.40 are in very good agreement. In particular, the difference in the water volume (Tab 5.9) between calculated and observed data (about 11.9% and 9.1%, respectively) comes in support of the quasi-imperviousness assumption.

Sub-basin	Observed water volume	Computed Water volume	Relative error
-	(m^3)	(m^3)	%
Sarralbe	24519.9	27856.2	11.9
Sarreguemines	35554.2	39095.3	9.1

Table 5.9: Water volume of computed and observed data.

The observed flood extent (FO) and simulated water surface (FS) for the 22, 23 and 24 of December are compared in Fig. 5.41, 5.42, and 5.43, respectively. The partition points, including the gauging stations are plotted as dots.

The common flood extent (A=$FS \cap FO$) plotted in orange, the over-predicted flood extent by the model (B=FS\FO) plotted in magenta, and the under-predicted flood extent by the model (C=FO\FS) plotted in blue, highlight the differences between the simulated and observed flood extents.

In contrast to the Moderbach case study, where the fit parameters were useful to compare the simulated extent with the one observed in the 1939 and 1944 flood maps, the computed flood extent in Figures 5.41, 5.42, and 5.43 shows large differences with the observed data on the map of 2010. As discussed earlier, the forested areas and water bodies in the Saar watershed are not considered in the observed flood map. Therefore, the calculated flood extent takes into account large areas that are not captured in the observed flood which makes it difficult to compare.

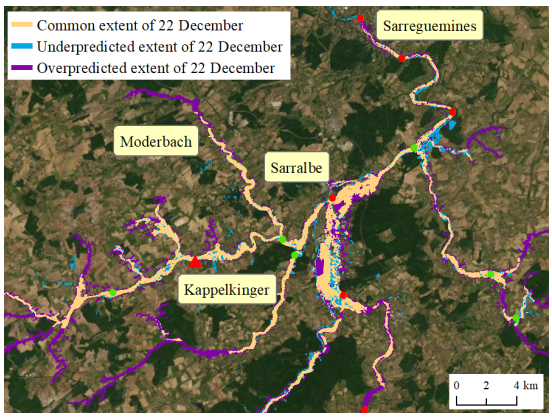


Figure 5.41: HEC-RAS flood extent on the flood map of December 22, 2010.

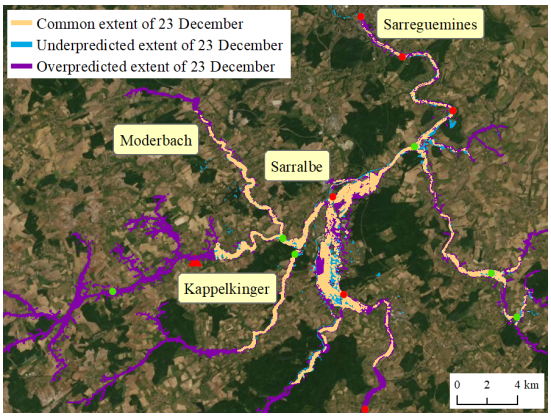


Figure 5.42: HEC-RAS flood extent on the flood map of December 23, 2010.

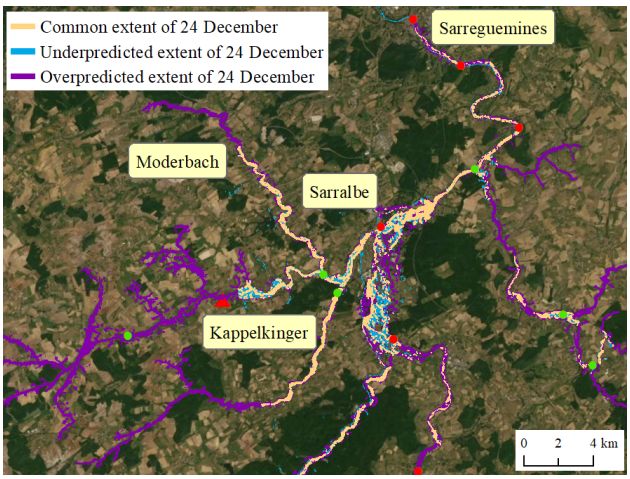


Figure 5.43: HEC-RAS flood extent on the flood map of December 24, 2010.

5.4 Conclusion

This chapter presents the modeling for historical and extreme flood events in the Saar area by performing 2D rain-on-grid simulations. Unlike most flood studies, the simulations presented in this chapter consider the entire drainage basin as a computational domain in order to provide a detailed description of the flood characteristics all over the basin. A selection of historical extreme flood events in the Albe, Moderbach and Saar river basins were performed and compared with observed data.

The extreme event modeled over the Albe basin was particular, showing three successive heavy rainfall events in less than two months. The simulation of this event allowed a better understanding of the effect of the soil condition on surface flow modeling. The dominant existence of clay in the study basin allows for rapid saturation of the soil, which enhances the velocity and propagation of surface flow. Indeed, the soil type in this area was one of the main reasons for the implementation of the water defense system at this location.

The inundation of the Moderbach basin was modeled by considering two scenarios. The first one corresponded to flooding the valley by the evacuation of the ponds and the second scenario corresponded to the flooding generated by the heavy rainfall of October 1939. The model results were compared to two historical flood maps showing that the flood extent information added to these maps corresponds to the flood generated by the heavy rainfall of October 1939 and not to the flood generated by the evacuation of the Moderbach ponds in 1940.

The last historical event on the Saar basin aimed at modeling the heavy precipitation of 22-24 December 2010. Due to a heavy precipitation event that occurred earlier in December, the soil was saturated during the period of December 22-24, which is suitable for HEC-RAS model conditions that do not consider the effect of infiltration. Therefore, the model provided satisfactory results showing good agreement with the observed data.

From a computational point of view, 2D modeling of large-scale basins such as those in the Saar area is both intensive and expensive, requiring long simulation times. The proposed strategy using basin partitioning methods in HEC-RAS allows for sequential simulation of multiple sub-basins on a local machine while other studies tend to run the entire basin in one go on a cluster system. To facilitate its implementation and reduce user intervention and workload, this process was automated by a VBA script. The application of this strategy to model the Moderbach basin showed a significant improvement in the computational time of the 2D simulations in HEC-RAS.

General Conclusion

The ambition of this research project was to carry out rain-on-grid simulations on a large basin so as to model the historical and extreme floods of the Saar basin that played a major role in the protection of the north-eastern borders of France during the Second World War. The challenge was to consider the drainage basin as a computational domain making use of the precipitation data assigned to the entire basin. As the simulation of a distributed hydrological model over a large-scale basin is known to be time and computationally intensive, the thesis proposes a methodology to reduce the computation time of 2D rain-on-grid simulations on large-scale basins. This is based on a domain decomposition method allowing to balance the computational load in a computing environment.

This work required a review of historical documents and flood data for the Saar basin to obtain a solid understanding of its historical role and hydrological behavior. The data analysis clearly showed that flooding is frequent in this region. For example, in Sarreguemines, a flood event is recorded every two years. Moreover, this city is classified as a high-risk flood area. Two historical flood maps drawn by the Germans in 1940 and by the French in 1944 showing the extent of flooding in the Moderbach basin and the flood map resulting for the extreme flood event of December 2010 provided by SERTIT-ICube were used to validate the simulation results generated in this project.

This research work developed a domain decomposition method to be implemented in hydrological modeling. The method proposes to partition a drainage basin into sub-basins of similar areas in order to balance the computational load. This has been proven to enhance the efficiency of parallel computing in hydrological modeling. The initial choice of the hydrological model NIHM was mainly done on the availability of its source code. For the sake of robustness, the HEC-RAS software was firstly used for validation purposes, as it uses the diffusive wave equation, the same 2D equation than NIHM. Comparing hydrological results in the two models highlighted the limitations of NIHM.

The domain decomposition method was implemented in NIHM supported by the parallel tool ORWL allowing for parallel rain-on-grid modeling over multiple sub-basins. The application of this computing strategy showed that it can provide the same hydrological behavior for the simulation of a entire or a decomposed watershed, while the decomposed watershed requires less computation time. The improvement brought to the performance of the hydrological model NIHM by means of this parallel computing strategy has been verified.

As HEC-RAS showed a better hydrological performance than NIHM, we then develop a new strategy for implementing basin partitioning in HEC-RAS. We have automated the sequential simulation of multiple sub-basins in HEC-RAS using the HECRASController functions. The application of this strategy on the Moderbach basin showed that sequential rain-on-grid simulation of multiple sub-basins reduced the computational time compared to rain-on-grid simulation of the entire basin at one go. This strategy allowed to carry out rain-on-grid modeling of large scale basins such as the Albe ($405\,km^2$) and the Saar ($1800\,km^2$) considering a 25m grid.

Firstly, the simulation carried out on the Albe basin validated our assumption that the soil may

be considered as quasi impervious in this area. The simulation results showed that after a heavy precipitation, the soil is saturated and the observed data showed a strong correspondence with the data calculated in HEC-RAS that does not take into account the infiltration process. Secondly, the modeling of the Moderbach inundation considering flooding of the valley by the evacuation of the ponds and the heavy precipitation of October 1939 showed good results compared to the two flood maps of 1940 and 1944. Thirdly, the simulation carried out in the Saar basin aimed to model the flood event of December 2010. The results were compared to the flood maps provided by SERTIT-Icube and to the observed discharges available on Banque Hydro, showing a strong agreement.

The main objective of this research work aiming in modeling the historical extreme inundations in the Saar region was accomplished through the parallel computing methods proposed. These has allowed to perform rain-on-grid simulations on large scale basin as the Saar (1800 km^2) involving about 2,800,000 elements with a 25m mesh size.

This thesis raises a number of exciting research prospects. Several aspects could be further explored from a spatialized modeling perspective. This study used a 25m resolution DEM that was subjected to classic or hybrid methods to correct sinks by filling, breaching, or a combination of both methods. Using a higher resolution DEM with a finer mesh and performing a more detailed correction process through field visits to investigate and explore the various landscape features can improve the model results. Refining parts of the subdomains may lead to more partitioning complex strategies. Comparing HEC-RAS with NIHM highlighted the limitations of NIHM and the improvements needed. Currently, a new version of NIHM is under development at the Institute of Earth and Environment of Strasbourg (ITES), in order to improve the linear and non-linear solvers. The implementation of basin partitioning in HEC-RAS that we proposed, considers a sequential simulation of the different sub-basins. Although this process is automated, further improvements can be introduced to this method. Running HEC-RAS on Linux will allow independent sub-basins that do not require information transfer between them to be executed in parallel on a multi-processor cluster. This would greatly improve the computational efficiency of the simulation.

References

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., et al. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *Fao, Rome*, 300(9):D05109.

Ames, D. P., Michaelis, C. D., Anselmo, A., Chen, L., and Dunsford, H. (2008). Mapwindow gis.

Andre, M. (2001). Cours d 'hydrologie générale. *Ecole Polytechnique Fédérale de Lausanne,(EPFL), Suisse*.

Antenucci, J. C., Brown, K., Croswell, P. L., Kevany, M. J., and Archer, H. (1991). *Geographic Information Systems: a guide to the technology*. Elsevier.

Apostolopoulos, T. K. and Georgakakos, K. P. (1997). Parallel computation for streamflow prediction with distributed hydrologic models. *Journal of Hydrology*, 197(1-4):1–24.

Arnold, J. G. and Fohrer, N. (2005). Swat2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes: An International Journal*, 19(3):563–572.

Baldassarre, G. D. and Montanari, A. (2009). Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences*, 13(6):913–921.

Barber, C. P. and Shortridge, A. (2005). Lidar elevation data for surface hydrologic modeling: Resolution and representation issues. *Cartography and Geographic Information Science*, 32(4):401–410.

Barnes, R., Lehman, C., and Mulla, D. (2014). Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models. *Computers & Geosciences*, 62:117–127.

Bayer, D. M. and Collischonn, W. (2013). Deforestation impacts on discharge of the ji-paraná river–brazilian amazon. *Climate and land surface changes in hydrology. 359th ed. Gothenburg: IAHS Publication H*, 1:327–332.

Beard, L. R. (1981). *HEC-1 flood hydrograph package: Users manual*. US Army Corps of Engineers, Water Resources Support Center.

Beck, M. B. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water resources research*, 23(8):1393–1442.

Becker, B. (2007). La trouée de la sarre point faible de la ligne maginot ?

Beguería, S., Vicente-Serrano, S. M., Reig, F., and Latorre, B. (2014). Standardized precipitation evapotranspiration index (spei) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International journal of climatology*, 34(10):3001–3023.

Bergholt, D. and Lujala, P. (2012). Climate-related natural disasters, economic growth, and armed civil conflict. *Journal of Peace Research*, 49(1):147–162.

Bergström, S. (1976). *Development and application of a conceptual runoff model for Scandinavian catchments*. Springer.

Bergström, S. (1991). Principles and confidence in hydrological modelling. *Hydrology Research*, 22(2):123–136.

Bergström, S. et al. (1995). The hbv model. *Computer models of watershed hydrology.*, pages 443–476.

Berz, G. (2000). Flood disasters: lessons from the past - worries for the future. In *Proceedings of the Institution of Civil Engineers-Water and Maritime Engineering*. Thomas Telford Ltd.

Beven, K. (1989). Changing ideas in hydrology-the case of physically-based models. *Journal of hydrology*, 105(1-2):157–172.

Blazkov, S. and Beven, K. (1997). Flood frequency prediction for data limited catchments in the czech republic using a stochastic rainfall model and topmodel. *Journal of Hydrology*, 195(1-4):256–278.

Boithias, L., Sauvage, S., Lenica, A., Roux, H., Abbaspour, K. C., Larnier, K., Dartus, D., and Sánchez-Pérez, J. M. (2017). Simulating flash floods at hourly time-step using the swat model. *Water*, 9(12):929.

Bonnet, F. (1997). Communes. *Annuaire des villes de France https://www.communes.com/ville-sarreguemines*.

Borah, D. K. and Bera, M. (2004). Watershed-scale hydrologic and nonpoint-source pollution models: Review of applications. *Transactions of the ASAE*, 47(3):789.

Bourgin, F. (2014). *Comment quantifier l 'incertitude prédictive en modélisation hydrologique? Travail exploratoire sur un grand échantillon de bassins versants*. PhD thesis, Doctorat Géosciences Ressources Naturelles et Environnement, Spécialité Hydrologie, AgroParisTech.

Brandt, M. and Bergström, S. (1994). Integration of field data into operational snowmelt-runoff models: Paper presented at egs xvii general assembly (edinburgh, gb–april 1992). *Hydrology Research*, 25(1-2):101–112.

Brouquisse, F., Cornet, J., Descatoire, I., Faure, J., Goutx, D., and Tekatlian, A. (2007). Guide méthodologique pour le pilotage des études hydrauliques.

Brunner, G. (2016). Hec-ras river analysis system, 2d modeling user's manual version 5.0. *Davis: US Army Corps of Engineers, hydrologic engineering center*.

Brunner, G. W. (1995). Hec-ras river analysis system. hydraulic reference manual. version 1.0. Technical report, Hydrologic Engineering Center Davis CA.

Brunner, G. W. (2014). Combined 1d and 2d modeling with hec-ras. *US Army Corps of Engineers*.

Brunner, G. W., Piper, S. S., Jensen, M. R., and Chacon, B. (2015). Combined 1d and 2d hydraulic modeling within hec-ras. In *World Environmental and Water Resources Congress 2015*, pages 1432–1443.

Burrough, P. A., McDonnell, R., McDonnell, R. A., and Lloyd, C. D. (2015). *Principles of geographical information systems*. Oxford university press.

CETMEF, C. (2001). Etude de sensibilité des modélisations 1d - etape 1 :. *recensement et quantification des sources d'incertitude externes*.

Chambru, C. (2021). Etang de diefenbach : la digue fragilisée sera bien réparée courant 2022. *LE REPUBLICAIN LORRAIN https://www.republicain-lorrain.fr/environnement/2021/07/06/etang-de-diefenbach-la-digue-fragilisee-sera-bien-reparee-courant-2022*.

Chen, R.-J. and Horng, S.-J. (2010). Novel scan-ca-based image security system using scan and 2-d von neumann cellular automata. *Signal Processing: Image Communication*, 25(6):413–426.

Clarke, R. (1973). A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of hydrology*, 19(1):1–20.

Clauss, P.-N. and Gustedt, J. (2010). Iterative computations with ordered read–write locks. *Journal of Parallel and Distributed Computing*, 70(5):496–504.

Crawford, N. H. and Linsley, R. K. (1966). Digital simulation in hydrology'stanford watershed model 4. *Stanford University Department of Civil EngineeringStanford, CA United States 94305*.

Crouzeix, M. and Raviart, P.-A. (1973). Conforming and nonconforming finite element methods for solving the stationary stokes equations i. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 7(R3):33–75.

D'Agostino, V., Stanghellini, M., and Trisorio-Liuzzi, G. (1993). A fortran-77 program for preliminary extraction of drainage networks based on a dem. *Computers & Geosciences*, 19(7):981–1006.

David, A. and Schmalz, B. (2021). A systematic analysis of the interaction between rain-on-grid-simulations and spatial resolution in 2d hydrodynamic modeling. *Water*, 13(17):2346.

Dawdy, D. R. and O'Donnell, T. (1965). Mathematical models of catchment behavior. *Journal of the Hydraulics Division*, 91(4):123–137.

Dehvari, A., Heck, R. J., et al. (2013). Effect of lidar derived dem resolution on terrain attributes, stream characterization and watershed delineation. *International Journal of Agriculture and Crop Sciences (IJACS)*, 6(13):949–967.

Devia, G., Ganasri, B., and Dwarakish, G. (2015). A review on hydrological models, aquatic procedia, 4, 1001–1007.

DHI (1999). Mike she danish hydraulic institute. *water movement: user manual*.

DHI (2014). Mike 3 flow model fm. hydrodynamic module-user guide. *Danish Hydraulic Institute*.

Di Baldassarre, G., Schumann, G., and Bates, P. D. (2009). A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. *Journal of Hydrology*, 367(3-4):276–282.

Di Giammarco, P., Todini, E., and Lamberti, P. (1996). A conservative finite elements approach to overland flow: the control volume finite element formulation. *Journal of Hydrology*, 175(1-4):267–291.

Doan, J. (2000). Geospatial hydrologic modeling extension hec-geohms. *US Army Corps of Engineers, Institute for Water Resources, Hydrologic Engineering Centre, Davis, CA*.

EPRI (2011). Évaluation préliminaire des risques d 'inondation2011bassin rhin. *Ministére de l ' Écologie, du Développement durable, des Transports et du Logement*.

Falter, D., Dung, N., Vorogushyn, S., Schröter, K., Hundecha, Y., Kreibich, H., Apel, H., Theisselmann, F., and Merz, B. (2016). Continuous, large-scale simulation model for flood risk assessments: proof-of-concept. *Journal of Flood Risk Management*, 9(1):3–21.

Fan, F. M., Collischonn, W., Quiroz, K., Sorribas, M., Buarque, D., and Siqueira, V. (2016). Flood forecasting on the tocantins river using ensemble rainfall forecasts and real-time satellite rainfall estimates. *Journal of Flood Risk Management*, 9(3):278–288.

Feldman, A. D. (2000). *Hydrologic modeling system HEC-HMS: technical reference manual*. US Army Corps of Engineers, Hydrologic Engineering Center.

Ficchi, A. (2017). *An adaptive hydrological model for multiple time-steps: Diagnostics and improvements based on fluxes consistency*. PhD thesis.

Flavien, R. and Laurent, P. (2014). Réglement de surveillance, de prévision et de transmission de l 'information sur les crues du service de prévision des crues rhin-sarre (r.i.c.). *DREAL Alsace Direction Régionale de l ' Environnement, de l 'Aménagement et du Logement d 'Alsace*.

Foresman, T. W. (1998). *The history of geographic information systems: perspectives from the pioneers*, volume 397. Prentice Hall PTR Upper Saddle River, NJ.

Freeze, R. A. (1972). Role of subsurface flow in generating surface runoff: 1. base flow contributions to channel flow. *Water Resources Research*, 8(3):609–623.

Freeze, R. A. and Harlan, R. (1969). Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of hydrology*, 9(3):237–258.

Gao, J., Holden, J., and Kirkby, M. (2017). Modelling impacts of agricultural practice on flood peaks in upland catchments: An application of the distributed topmodel. *Hydrological Processes*, 31(23):4206–4216.

Gebre, S. L. (2015). Application of the hec-hms model for runoff simulation of upper blue nile river basin. *Hydrology: Current Research*, 6(2):1.

Gilli, E., Mangan, C., and Mudry, J. (2008). *Hydrogéologie: objet, méthodes, applications*. Dunod.

Gnouma, R. (2006). Aide à la calibration d'un modèle hydrologique distribué au moyen d'une analyse des processus hydrologiques: application au bassin versant de l'yzeron. *LGCIE INSA de Lyon, Lyon. 448p*.

Gomez, E. (2002). Modélisation intégrée du transfert de nitrate à l'échelle régionale dans un système hydrologique. *Application au bassin de la Seine., Ecole Nationale Supérieure des Mines de Paris*.

Goodell, C. (2014). *Breaking the HEC-RAS Code: A User's Guide to Automating HEC-RAS*. h2ls.

Goodell, C. (2018). Powerful feature of hec-ras exposed. *https://icewarm.com.au/australian-water-school/short-courses/course/powerful-hec-ras-features-exposed/*.

Gosain, A., Rao, S., and Basuray, D. (2006). Climate change impact assessment on hydrology of indian river basins. *Current science*, pages 346–353.

Götzinger, J. and Bárdossy, A. (2008). Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resources Research*, 44(12).

Goulden, T., Hopkinson, C., Jamieson, R., and Sterling, S. (2014). Sensitivity of watershed attributes to spatial resolution and interpolation method of lidar dems in three distinct landscapes. *Water Resources Research*, 50(3):1908–1927.

Grayson, R., Blöschl, G., Barling, R., and Moore, I. (1993). Process, scale and constraints to hydrological modelling in gis. *IAHS PUBLICATION*, pages 83–83.

Gresillon, J. (2004). Inondations. In *Aléa et enjeux au regard du changement climatique, Actes de colloque, Colloque ONERC, Collectivités locales et changement climatique: quelles stratégies d'adaptation*.

Grid5000 (2005). A large scale, reconfigurable, controlable and monitorable grid platform. *In Proceedings of the 6th IEEE/ACM International Workshop on Grid Computing https://www.grid5000.fr/mediawiki/index.php/Grid5000:Home*.

Grillakis, M., Tsanis, I., and Koutroulis, A. (2010). Application of the hbv hydrological model in a flash flood case in slovenia. *Natural Hazards and Earth System Sciences*, 10(12):2713.

Gumindoga, W., Rwasoka, D. T., Nhapi, I., and Dube, T. (2017). Ungauged runoff simulation in upper manyame catchment, zimbabwe: Application of the hec-hms model. *Physics and Chemistry of the Earth, Parts A/B/C*, 100:371–382.

Häggström, M. and Lindström, G. (1990). Application of the hbv model for flood forecasting in six central american rivers.

Halwatura, D. and Najim, M. (2013). Application of the hec-hms model for runoff simulation in a tropical catchment. *Environmental modelling & software*, 46:155–162.

Hammouri, N. and El-Naqa, A. (2007). Hydrological modeling of ungauged wadis in arid environments using gis: a case study of wadi madoneh in jordan. *Revista mexicana de ciencias geológicas*, 24(2):185–196.

Hargreaves, G. H. and Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied engineering in agriculture*, 1(2):96–99.

Hariri, S., Gustedt, J., Weill, S., and Charpentier, I. (2021). A hybrid breaching-filling method for sink removal adapted to parallel hydrological simulations. In *EGU 2021*.

Hariri, S., Weill, S., Gustedt, J., and Charpentier, I. (2019). Pairing gis and distributed hydrological models using matlab 2. In *CAJG-2nd Conference of the Arabian Journal of Geosiences*.

Havno, K., Madsen, M., Dorge, J., et al. (1995). Mike 11-a generalized river modelling package. *Computer models of watershed hydrology.*, pages 733–782.

Henine, H. (2005). *Interfaçage entre un modèle hydrologique et un modèle hydrodynamique au sein d'un système d'information intégré sous web incluant les SIG*. PhD thesis, Alger, Ecole Nationale Polytechnique.

Heywood, I., Cornelius, S., and Carver, S. (2011). An introduction to geographical information systems. 446 s.

Hornberger, G., Beven, K., Cosby, B., and Sappington, D. (1985). Shenandoah watershed study: Calibration of a topography-based, variable contributing area hydrological model to a small forested catchment. *Water Resources Research*, 21(12):1841–1850.

Horritt, M., Di Baldassarre, G., Bates, P., and Brath, A. (2007). Comparing the performance of a 2-d finite element and a 2-d finite volume model of floodplain inundation using airborne sar imagery. *Hydrological Processes: An International Journal*, 21(20):2745–2759.

Huang, L., He, B., Han, L., Liu, J., Wang, H., and Chen, Z. (2017). A global examination of the response of ecosystem water-use efficiency to drought based on modis data. *Science of the Total Environment*, 601:1097–1107.

Im, S., Kim, H., Kim, C., and Jang, C. (2009). Assessing the impacts of land use changes on watershed hydrology using mike she. *Environmental geology*, 57(1):231.

Islam, Z. (2011). A review on physically based hydrologic modeling. *University of Alberta: Edmonton, AB, Canada*.

Janin, J., Lepeintre, F., and Pechon, P. (1992). Telemac-3d: A finite element code to solve 3d free surface flow problems. In *Computer modelling of seas and coastal regions*, pages 489–506. Springer.

Jeannot, B. (2018). *Modélisation hydrologique intégrée de bassins versants fortement transitoires: développement d'outils numériques et applications*. PhD thesis.

Jeannot, B., Weill, S., Eschbach, D., Schmitt, L., and Delay, F. (2018). A low-dimensional integrated subsurface hydrological model coupled with 2-d overland flow: Application to a restored fluvial hydrosystem (upper rhine river–france). *Journal of hydrology*, 563:495–509.

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta mathematica*, 30(1):175–193.

Jenson, S. K. and Domingue, J. O. (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric engineering and remote sensing*, 54(11):1593–1600.

Jeong, J., Kannan, N., Arnold, J., Glick, R., Gosselink, L., and Srinivasan, R. (2010). Development and integration of sub-hourly rainfall–runoff modeling capability within a watershed model. *Water Resources Management*, 24(15):4505–4527.

Jutman, T. (1992). Production of a new runoff map of sweden. In *Proceedings of Nordic Hydrological Conference*, pages 4–6.

Khatami, S. and Khazaei, B. (2014). Benefits of gis application in hydrological modeling: A brief summary. *VATTEN–Journal of Water Management and Research*, 70(1):41–49.

Klein Tank, A., Wijngaard, J., Können, G., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., et al. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the european climate assessment. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 22(12):1441–1453.

Klemes, V. (2002). *3 Risk analysis: The unbearable cleverness of bluffing*. Cambridge University Press Cambridge, UK.

Kourakos, G. and Harter, T. (2014). Vectorized simulation of groundwater flow and streamline transport. *Environmental modelling & software*, 52:207–221.

Kumar, M. and Duffy, C. (2015). Exploring the role of domain partitioning on efficiency of parallel distributed hydrologic model simulations. j hydrogeol hydrol eng 4: 1. *of*, 12:2.

La Berbera, P., Lanza, L., and F, S. (1993). Hydrologically oriented gis and application to rainfall-runoff distributed modelling : case study of the arno basin. *HydroGIS 93: Application ofGeographic Information Systems in Hydrology and Water Resources*, 9:171–17.

La Jeunesse, I. (2016). Changement climatique et cycle de l 'eau-impacts, adaptation, législation et avancées scientifiques. *Vecteur Environnement*, 49(3):70.

Laurent, F. (2013). *Outils de modélisation spatiale pour la gestion intégrée des ressources en eau: Application aux Schémas d'Aménagement et de Gestion des Eaux*. PhD thesis.

Laurini, R. and Milleret-Raffort, F. (1993). Les bases de données en géomatique. Technical report, Hermès.

Leach, J., Lidberg, W., Kuglerová, L., Peralta-Tapia, A., Ågren, A., and Laudon, H. (2017). Evaluating topography-based predictions of shallow lateral groundwater discharge zones for a boreal lake-stream system. *Water Resources Research*, 53(7):5420–5437.

Leandro, J., Chen, A., and Schumann, A. (2014). A 2d parallel diffusive wave model for floodplain inundation with variable time step (p-dwave). *Journal of Hydrology*, 517:250–259.

Lee, M. T. and Terstriep, M. L. (1991). Applications of gis for water quality modeling in agricultural and urban watersheds. In *Hydraulic Engineering*, pages 961–965. ASCE.

Lidberg, W., Nilsson, M., Lundmark, T., and Ågren, A. M. (2017). Evaluating preprocessing methods of digital elevation models for hydrological modelling. *Hydrological Processes*, 31(26):4660–4668.

Lindsay, J. B. (2016). Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models. *Hydrological Processes*, 30(6):846–857.

Lindsay, J. B. and Creed, I. F. (2005). Removal of artifact depressions from digital elevation models: towards a minimum impact approach. *Hydrological Processes: An International Journal*, 19(16):3113–3126.

Lindsay, J. B. and Creed, I. F. (2006). Distinguishing actual and artefact depressions in digital elevation data. *Computers & Geosciences*, 32(8):1192–1204.

Lindsay, J. B. and Dhun, K. (2015). Modelling surface drainage patterns in altered landscapes using lidar. *International Journal of Geographical Information Science*, 29(3):397–411.

Lindsay, J. B. and Evans, M. G. (2008). The influence of elevation error on the morphometrics of channel networks extracted from dems and the implications for hydrological modelling. *Hydrological Processes: An International Journal*, 22(11):1588–1603.

Linkov, I. and Burmistrov, D. (2003). Model uncertainty and choices made by modelers: Lessons learned from the international atomic energy agency model intercomparisons. *Risk Analysis: An International Journal*, 23(6):1297–1308.

Maidment, D. R. et al. (1993). Gis and hydrologic modeling. *Environmental modeling with GIS.*, pages 147–167.

Manguerra, H. and Engel, B. (1998). Hydrologic parameterization of watersheds for runoff prediction using swat 1. *JAWRA Journal of the American Water Resources Association*, 34(5):1149–1162.

Mansouri, F. and Gustedt, J. (2016). Le modèle de programmation orwl pour la parallélisation d'une application de suivi vidéo hd sur architecture multi-coeurs. In *Conférence d'informatique en Parallélisme, Architecture et Système (COMPAS)*.

Manual, D. D. D. (1999). User manual delft3d-flow. *WLjDelft Hydraulics, The Netherlands*.

Marini, L. D. (1985). An inexpensive method for the evaluation of the solution of the lowest order raviart–thomas mixed method. *SIAM journal on numerical analysis*, 22(3):493–496.

Marque, P. (1989). *La Ligne Maginot aquatique*. FeniXX.

Martz, L. W. and Garbrecht, J. (1992). Numerical definition of drainage network and subcatchment areas from digital elevation models. *Computers & Geosciences*, 18(6):747–761.

Martz, L. W. and Garbrecht, J. (1998). The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models. *Hydrological processes*, 12(6):843–855.

Martz, L. W. and Garbrecht, J. (1999). An outlet breaching algorithm for the treatment of closed depressions in a raster dem. *Computers & Geosciences*, 25(7):835–844.

Mathis, D., Chiffre, E., and Weimerskirch, G. (2017). Les hydrosystèmes défensifs: des paysages militaires atypiques de la ligne maginot. *Revue de Géographie historique*.

Matott, L. S., Babendreier, J. E., and Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resources Research*, 45(6).

McMillan, H., Krueger, T., and Freer, J. (2012). Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, 26(26):4078–4111.

McVicar, T. R., Rui, L., Walker, J., Fitzpatrick, R. W., Changming, L., et al. (2002). Regional water and soil assessment for managing sustainable agriculture in china and australia. *Regional water and soil assessment for managing sustainable agriculture in China and Australia.*

Meisels, A., Raizman, S., and Karnieli, A. (1995). Skeletonizing a dem into a drainage network. *Computers & Geosciences*, 21(1):187–196.

Metz, M., Mitasova, H., and Harmon, R. (2010). Accurate stream extraction from large, radar-based elevation models. *Hydrology & Earth System Sciences Discussions*, 7(3).

Mulvany, T. (1850). On the use of self-registering rain and flood gauges in making observations of the rainfall and flood discharges in a given catchment. *Transactions and Minutes of the Proceeding of the Institute of Civil Engineers of Ireland, Session*, 1.

Murphy, P. N., Ogilvie, J., Castonguay, M., Zhang, C.-f., Meng, F.-R., and Arp, P. A. (2008). Improving forest operations planning through high-resolution flow-channel and wet-areas mapping. *The Forestry Chronicle*, 84(4):568–574.

Musy, A. and Laglaine, V. (1992). *Hydrologie générale: cours polycopié: section génie rural, environnement, mensuration*. Ecole polytechnique fédérale de Lausanne Département de génie rural Institut d 'aménagement des terres et des eaux.

Nachtnebel, H., Fuerst, J., and Holzmann, H. (1993). Application of geographical information systems to support groundwater modelling. *IAHS PUBLICATION*, pages 653–653.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Srinivasan, R., and Williams, J. R. (2002). Soil and water assessment tool. *Users Manual. Version 2005. GSWRL Report 02-02, BRC Report 2-06, Temple, Texas, USA.*

Neteler, M. and Raghavan, V. (2006). Advances in free software geographic information systems. *Journals of Informatics*, 3(2).

O'Callaghan, J. F. and Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. *Computer vision, graphics, and image processing*, 28(3):323–344.

Oldham, P. (1997). *The Hindenburg Line*. Pen and Sword.

Oleyiblo, J. O. and Li, Z.-j. (2010). Application of hec-hms for flood forecasting in misai and wan'an catchments in china. *Water Science and Engineering*, 3(1):14–22.

Ouedraogo, M., Degré, A., and Debouche, C. (2014). Synthèse bibliographique: le modèle numérique de terrainde haute résolution, ses erreurs et leur propagation. *Biotechnologie, Agronomie, Société et Environnement*, 18(3):407.

Pan, Y., Weill, S., Ackerer, P., and Delay, F. (2015). A coupled stream flow and depth-integrated subsurface flow model for catchment hydrology. *Journal of Hydrology*, 530:66–78.

Payraudeau, S. (2002). *Modélisation distribuée des flux d'azote sur des petits bassins versants méditerranéens*. PhD thesis.

Petrescu Maftei, C. (2002). *Etudes concernant les écoulements superficiels: modélisation spatialisée de l'écoulement sur le bassin de Voinesti-Roumanie*. PhD thesis, Montpellier 2.

Ping, J., Yan, S., Gu, P., Wu, Z., and Hu, C. (2017). Application of mike she to study the impact of coal mining on river runoff in gujiao mining area, shanxi, china. *PloS one*, 12(12):e0188949.

Poggio, L. and Soille, P. (2012). Influence of pit removal methods on river network position. *Hydrological Processes*, 26(13):1984–1990.

Pontes, P. R. M., Fan, F. M., Fleischmann, A. S., de Paiva, R. C. D., Buarque, D. C., Siqueira, V. A., Jardim, P. F., Sorribas, M. V., and Collischonn, W. (2017). Mgb-iph model for hydrological and hydraulic simulation of large floodplain river systems coupled with open source gis. *Environmental Modelling & Software*, 94:1–20.

PPRI (1998). plan de prevention du risque d 'innondation de la vallée de la sarre. *Service de la navigation de strasbourg*.

Quirogaa, V. M., Kurea, S., Udoa, K., and Manoa, A. (2016). Application of 2d numerical simulation for the analysis of the february 2014 bolivian amazonia flood: Application of the new hec-ras version 5. *Ribagua*, 3(1):25–33.

Refsgaard, J. C. (1990). Terminology, modelling protocol and classification of hydrological model codes. In *Distributed hydrological modelling*, pages 17–39. Springer.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., and Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process–a framework and guidance. *Environmental modelling & software*, 22(11):1543–1556.

Refshaard, J., Storm, B., et al. (1995). Mike she. *Computer models of watershed hydrology.*, pages 809–846.

Rego, F. (1996). Use of geographic information systems andhydrologic and hydraulic modelling in the definition of flooded areas. *Master of Science Thesis Instituto Superior Técnico, Lisbon*.

Reutebuch, S. E., McGaughey, R. J., Andersen, H.-E., and Carson, W. W. (2003). Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian journal of remote sensing*, 29(5):527–535.

Rieger, W. (1993). Automated river line and catchment area extraction from dem data. *International Archives of Photogrammetry and Remote Sensing*, 29:642–642.

Roche, M. (1963). Hydrologie de surface, gauthier-villars éditeur. *Paris, 431p*.

Rodriguez, F., Andrieu, H., and Morena, F. (2008). A distributed hydrological model for urbanized areas–model development and application to case studies. *Journal of Hydrology*, 351(3-4):268–287.

RRP (2011). Donnees ouvertes- extrait base de données régionale des sols d'alsace -. *Référentiel Régional Pédologique (RRP) Alsace*.

Sahana, V., Madhusoodhanan, C., and Eldho, T. (2016). Automation of watershed delineation & characterization: A comparison between arcgis and matlab tools. *National Conference on Water Resources & Flood Management with special reference to Flood Modelling*.

Sahoo, G., Ray, C., and De Carlo, E. (2006). Calibration and validation of a physically distributed hydrological model, mike she, to predict streamflow at high frequency in a flashy mountainous hawaii stream. *Journal of Hydrology*, 327(1-2):94–109.

Sarrazin, B. (2012). *MNT et observations multi-locales du réseau de drainage d'un petit bassin versant rural dans une perspective d'aide à la modélisation spatialisée*. PhD thesis, Université de Grenoble.

Schwanghart, W. (2014). Topotoolbox matlab based software for topographic analysis.

Schwanghart, W., Groom, G., Kuhn, N. J., and Heckrath, G. (2013). Flow network derivation from a high resolution dem in a low relief, agrarian landscape. *Earth Surface Processes and Landforms*, 38(13):1576–1586.

Schwanghart, W. and Kuhn, N. J. (2010). Topotoolbox: A set of matlab functions for topographic analysis. *Environmental Modelling & Software*, 25(6):770–781.

Schwanghart, W. and Scherler, D. (2014). Topotoolbox 2–matlab-based software for topographic analysis and modeling in earth surface sciences. *Earth Surface Dynamics*, 2(1):1–7.

Seibert, J. and Beven, K. J. (2009). Gauging the ungauged basin: how many discharge measurements are needed? *Hydrology and Earth System Sciences*, 13(6):883–892.

Seibert, J. and Vis, M. J. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, 16(9):3315–3325.

Shan, J. and Aparajithan, S. (2005). Urban dem generation from raw lidar data. *Photogrammetric Engineering & Remote Sensing*, 71(2):217–226.

Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Avant, B., and Muche, M. (2018). An overview of rainfall-runoff model types. *U.S. Environmental Protection Agency (EPA)*.

SLGRI (2017). Strategie locale de gestiondes risques d 'inondationdu bassin versant de la sarre.

Soille, P. (2004). Optimal removal of spurious pits in grid digital elevation models. *Water Resources Research*, 40(12).

Soille, P., Vogt, J., and Colombo, R. (2003). Carving and adaptive drainage enforcement of grid digital elevation models. *Water resources research*, 39(12).

Sonnentag, O. (2009). Neteler, m., mitasova, h., 2008. open source gis a grass gis approach, springer, ny, usa, isbn 978-0-387-35767-6, 406pp., usd 99.00, cdn 128.95, eur 81.95, hard-bound. *Computers and Geosciences*, 35(11):2282–2282.

Sorribas, M. V., Paiva, R. C., Melack, J. M., Bravo, J. M., Jones, C., Carvalho, L., Beighley, E., Forsberg, B., and Costa, M. H. (2016). Projections of climate change effects on discharge and inundation in the amazon basin. *Climatic Change*, 136(3-4):555–570.

Surwase, T., SrinivasaRao, G., Manjusree, P., Begum, A., Nagamani, P., and JaiSankar, G. (2019). Flood inundation simulation of mahanadi river, odisha during september 2008 by using hec-ras 2d model. In *Proceedings of International Conference on Remote Sensing for Disaster Management*, pages 851–863. Springer.

Tarolli, P. and Dalla Fontana, G. (2009). Hillslope-to-valley transition morphology: New opportunities from high resolution dtms. *Geomorphology*, 113(1-2):47–56.

Terstriep, M. L. and Lee, M. T. (1989). Regional stormwater modeling, q-illudas and arc/info. In *Computing in civil engineering: Computers in engineering practice*, pages 338–345. ASCE.

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical review*, 38(1):55–94.

Todini, E. (1988). Rainfall-runoff modeling past, present and future. *Journal of hydrology*, 100(1-3):341–352.

Tritthart, M. (2005). *Three-dimensional numerical modelling of turbulent river flow using polyhedral finite volumes: Dissertationsschrift*, volume 193. Inst. für Wasserbau u. Ingenieurhydrologie, Abt. Ingenieurhydrologie.

Vangenuchten, M. T., Rolston, D., and Germann, P. (1990). Transport of water and solutes in macropores-preface.

Vaze, J., Teng, J., and Spencer, G. (2010). Impact of dem accuracy and resolution on topographic indices. *Environmental Modelling & Software*, 25(10):1086–1098.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7):1696–1718.

Viviani, C., Dousset, S., Hariri, S., and Charpentier, I. (2020). Les étangs-réservoirs de la ligne maginot aquatique : un socio-écosystème durable ? *Revue Zones Humides Infos*.

Vivoni, E. R., Mascaro, G., Mniszewski, S., Fasel, P., Springer, E. P., Ivanov, V. Y., and Bras, R. L. (2011). Real-world hydrologic assessment of a fully-distributed hydrological model in a parallel computing environment. *Journal of Hydrology*, 409(1-2):483–496.

Vivoni, E. R., Mniszewski, S., Fasel, P., Springer, E., Ivanov, V., and Bras, R. (2005). Parallelization of a fully-distributed hydrologic model using sub-basin partitioning. *Eos Trans. AGU*, 86:52.

Vohralík, M. and Wohlmuth, B. I. (2013). Mixed finite element methods: implementation with one unknown per element, local flux expressions, positivity, polygonal meshes, and relations to other methods. *Mathematical Models and Methods in Applied Sciences*, 23(05):803–838.

Wallingford, S. (2003). Infoworks rs: An integrated softwaresolution for simulating flows in rivers in channels and onfloodplains. *http://www.wallingfordsoftware.com/products/infoworks*.

Wang, H., Li, T., Gao, J., Fu, X., and Wang, G. (2009). Binary-tree coding for drainage network of large-scale basins. *Journal of Hohai University: Natural Sciences*, 37(5):499–504.

Warwick, J. and Haness, S. (1994). Efficacy of arc/info gis application to hydrologic modeling. *Journal of Water Resources Planning and Management*, 120(3):366–381.

Weill, S., Delay, F., Pan, Y., and Ackerer, P. (2017). A low-dimensional subsurface model for saturated and unsaturated flow processes: ability to address heterogeneity. *Computational Geosciences*, 21(2):301–314.

Wheater, H., Sorooshian, S., and Sharma, K. D. (2007). *Hydrological modelling in arid and semi-arid areas*. Cambridge University Press.

Yanai, R. D., See, C. R., and Campbell, J. L. (2018). Current practices in reporting uncertainty in ecosystem ecology. *Ecosystems*, 21(5):971–981.

Yang, D., Goodison, B. E., Metcalfe, J. R., Golubev, V. S., Bates, R., Pangburn, T., and Hanson, C. L. (1998). Accuracy of nws 8" standard nonrecording precipitation gauge: Results and application of wmo intercomparison. *Journal of atmospheric and oceanic technology*, 15(1):54–68.

Yang, P., Ames, D. P., Fonseca, A., Anderson, D., Shrestha, R., Glenn, N. F., and Cao, Y. (2014). What is the effect of lidar-derived dem resolution on large-scale watershed model results? *Environmental modelling & software*, 58:48–57.

Yang, Q., Meng, F.-R., Zhao, Z., Chow, T. L., Benoy, G., Rees, H. W., and Bourque, C. P.-A. (2009). Assessing the impacts of flow diversion terraces on stream water and sediment yields at a watershed level using swat model. *Agriculture, ecosystems & environment*, 132(1-2):23–31.

Young, R. A. (1987). Agnps, agricultural non-point-source pollution model: a watershed analysis tool. *Conservation research report (USA). no. 35.*

Annexes

Appendix A

Uncertainties in hydrological modeling

Contents

A.1 Introduction

Modeling physical processes may be difficult due to the complexity of the system or to the limited availability of data. Using certain assumptions based on experience, reconstructed data, estimated and calibrated parameters, generally implies uncertainties in the model(Götzinger and Bárdossy, 2008) (Brouquisse et al., 2007). Uncertainty analysis comes in the last phase of the modeling process. It is applied after the calibration and validation of the model. It aims to analyze, understand and quantify the difference between the observed and the computed data. The importance of the uncertainty analysis remains in improving the confidence of the decision-maker to undertake a certain problem, according to the risk evaluated in an uncertain environment (Bourgin, 2014).

In the water cycle, the observation of hydrological processes is limited and the identification of its physical characteristics is imperfect (Bourgin, 2014). Hydrologists are aware that the approximations considered in the modeling process affect the precision of the model outputs (Baldassarre and Montanari, 2009). Several contributions have been introduced to discuss uncertainties in hydrological modeling (Beck, 1987) (Linkov and Burmistrov, 2003) (Refsgaard et al., 2007) (Matott et al., 2009). Uncertainties do not only depend on the hydrological model processes (runoff, evapotranspiration, infiltration) or physical characteristics (soil parameters, land use, land cover. Klemes (2002) discuss the influence of political and social factors on uncertainties.

A.2 Source of uncertainties

Uncertainty can be considered as natural uncertainty, also called stochastic uncertainty, or epistemic uncertainty. Natural uncertainty results from natural variability and can be examined by probabilistic statistical methods. Epistemic uncertainty is difficult to quantify. It results from a lack of knowledge in the spatial heterogeneity or groundwater flow for example (McMillan et al., 2012).

Baldassarre and Montanari (2009) classifies uncertainties in hydrological modeling according to their source as follow: (1) uncertainties in the observations used as input data (2) uncertainties in parameter estimation (3) uncertainties in the structure of the model itself. In the following, we briefly present these different sources of uncertainty.

- **Uncertainties in the observation data**

  The input data of a hydrological model, which are derived from field observations such as precipitation and flow, may have different levels of uncertainty. These uncertainties can be caused by the poor condition of the measuring equipment used, by the method of measurement applied, or by human errors.

  In any case, the precipitation data will always be considered uncertain since these data, based on point observation from one or more measuring stations located in the study area, are used to estimate an approximation of the average water depth over the entire study area. Differently, the flow data are not measured directly. The flow rates are calculated using a measured water level by means of rating curves which may introduce uncertainty into the flow data. The accuracy of the measurement is 1 to 10 cm for the height and 10 to 20% for the flow. For the high water marks which can be used as a validation parameter for modeling extreme events, the average order of magnitude of their accuracy is 20 to 30 cm in altimetry. These numbers on the accuracy of the data come from the Technical Center for Maritime and River Studies (CETMEF, 2001).

- **Uncertainties on the estimated parameters**

  Some parameters can be estimated in hydrological models due to the difficulty or lack of data needed to calculate them, such as the Manning coefficient or other parameters of the soil and vegetation. The problem is significant in conceptual models that uses a calibration process for the estimation of a large number of parameters. The estimation of these parameters is based on adjusting the model response to fit the actual observed data. Uncertainty about these parameters results from the estimation process which can provide a combination of different parameter sets.

- **Uncertainties on the theory and the structure of the model**

  Theoretically, some processes remain poorly understood and are not taken into account in physically based models. This is for example the case of preferential flow in macropores (Vangenuchten et al., 1990). Moreover, even if the theories are well established, the implementation of their equation, sometimes very complicated, may be difficult. The uncertainty associated with the model structure reflects the limited ability of a hydrological

model to perfectly represent the physical processes. This may be due to the simplified structure of the model compared to the complexity of the modeled system, its resolution, or its numerical implementation.

A.3 Researchers practices for uncertainty

Even though the sources of uncertainty are well defined in the literature, the understanding of the interest and relevance of considering uncertainties in research studies and in the industrial field remains a critical challenge. Yanai et al. (2018) has conducted an online survey with the participation of more than 140 researchers to find out the current practices in reporting uncertainties by researchers in ecosystem ecology addressing soil, vegetation, precipitation, and stream flows. The aim was to provide an idea about how researchers consider uncertainties in their work in terms of reporting uncertainties.

Of all the sources of error proposed, the uncertainty provided from the discharge calculation based on the stage height measurement is the most reported by the researchers. However, most respondents answered that they do not often report measurement uncertainties of precipitation data. They affirm that they have sufficient knowledge about how to report and quantify this uncertainty and they admit that it is often important to be considered. The measurement of this source of uncertainty could therefore be possible but rarely performed (Yang et al., 1998). For more details on the most commonly reported uncertainties in hydrological and environmental analyses and how researchers deal with missing or below detection values, and whether these methods are standardized at their sites, readers may refer to (Yanai et al., 2018).

Appendix B

Detect landscape features for basin partition

Contents

B.1 Introduction

The guided drainage basin partitioning method proposed by WatershedPart allows to consider engineered structures that can act on the hydrological behavior of a basin. These engineered structures should be defined by their geographical coordinates in order to add outlets at these locations to proceed with a guided partition. In section 3.4.1, the partition of the Moderbach basin has considered the 6 ponds in the basin due to the importance of these structures in retaining water volume and the elevation difference upstream and downstream of these ponds. However, the user may be interested in considering all the landscape features in the basin area.

The intersection of the river stream network with the road network in the basin will allow to detect the structures acting on the stream. The road network of the basin can be downloaded from open street maps in a shapefile format while the river network is extracted from the DEM using a GIS tool.

In the following, we present a simple method allowing to determine and derive the geographical coordinates of all the landscape features, that can act on the hydrological behavior of the watershed as to block, store or manage water passage through dams or culverts. Once the geographic coordinates of these structures are detected, they will be introduced as user-defined outlets to perform a guided partition.

B.2 Refine the road network

Prior to crossing the road and river networks, the road network must be carefully checked to eliminate roads with no significant importance. The figure below shows the stream network derived in a storage area of the Moderbach watershed. The stream intersects sailing and fishing outposts . These intersection points are of no interest and should not be considered in the basin partitioning process.

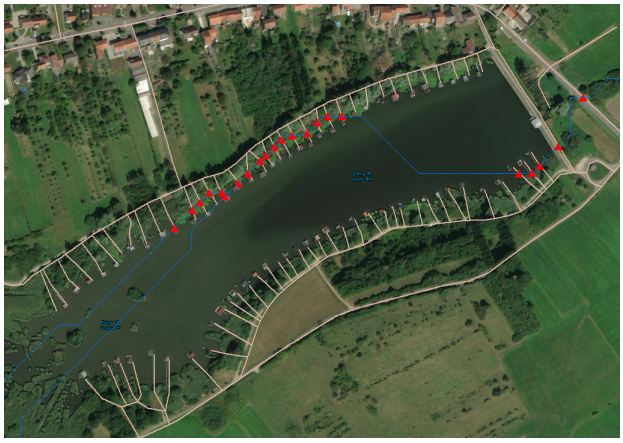


Figure B.1: Intersection of stream network with irrelevant pathways.

B.3 Removing nearby points

Once the road network is checked, we proposed to intersect it with the stream network. In some locations, several intersections can occur in few meters, see Fig. B.2. considering these nearby intersections in the basin partitioning procedure would not be valuable, since the scale of the basins will be way larger than the small distance between these consecutive intersections.

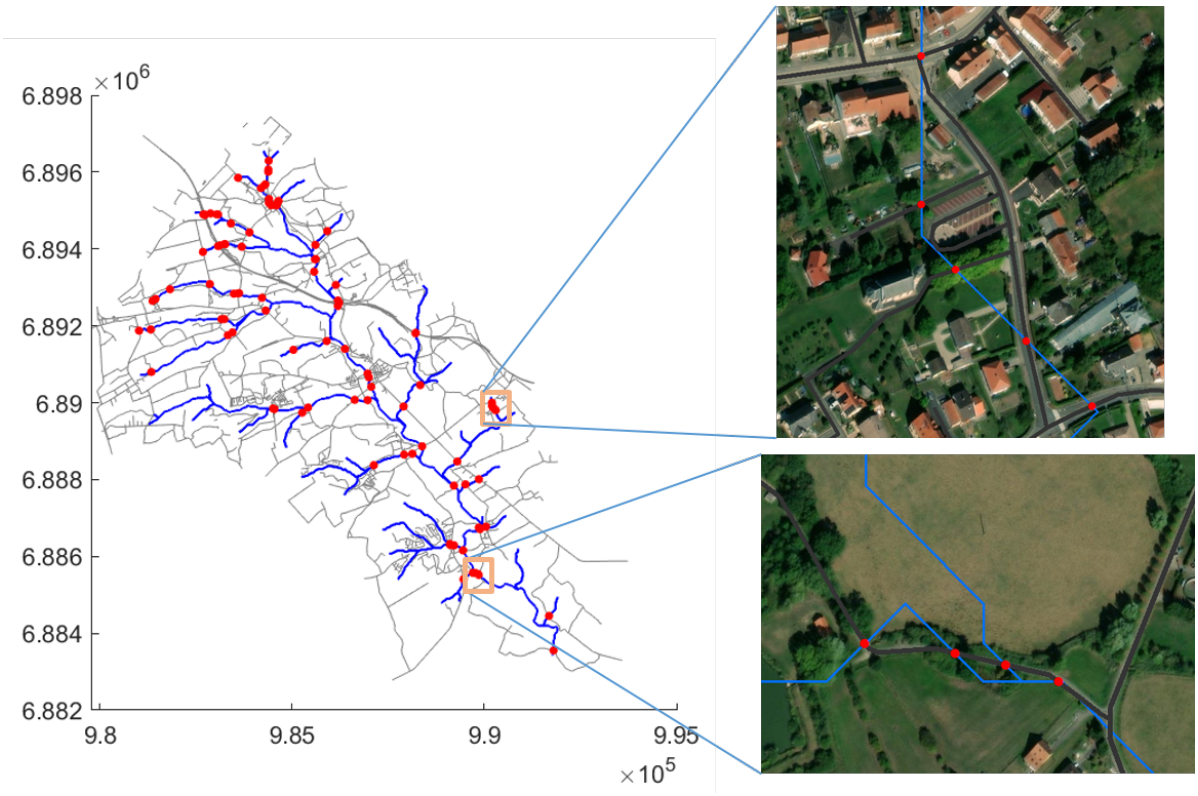
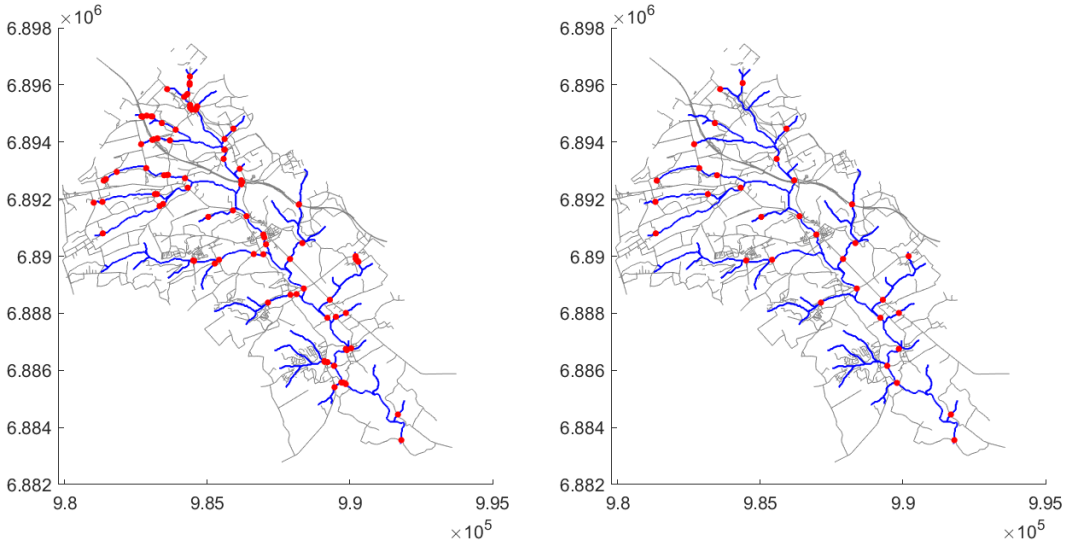


Figure B.2: All intersection points between road and stream network.

When all the intersection points are detected, we eliminate the points that are close to each other within a given threshold distance. The minimum distance between two outlets is set to 20 times the resolution of the DEM corresponding to a distance of 500 m. Figure B.3a shows all the intersection points of the road and stream network while figure B.3b shows the final intersection points after elimination of close points.



(a) All points (b) Points with 500 minimum distance

Figure B.3: Intersection points between road and stream network.

The final stage of this work belongs to the user to select the intersection points to be considered in the basin partitioning method. This method allows to generate a text file containing the number and geographical coordinates of each intersection point. It is up to the user to consider all or some of the intersection points to partition the basin. The choice mainly depends on the geographical importance of the points. Intersection points in the main stream near the outlet may have a major impact on the hydrological behavior of the basin while intersection points near the basin boundary are less important.

Figure B.4: Basin partitioning using user-defined intersection points.

Appendix C

Comparing classic pre-processing methods

Contents

C.1 Introduction

The use of DEM pre-processing methods such as classic filling and breaching, will certainly introduce significant changes to the original DEM. The objective of the following is to evaluate and compare the evolution of the DEM when applying the classic DEM pre-processing methods. The original DEM is compared to the filled and breached DEMs with regards to the stream network longitidunal profile, the absolute volume of the DEM and the hydrological behavior against obstacles.

C.2 Stream network longitudinal profile

Filling the depressions in the DEM or imposing a downward gradient path in the stream bed using the breaching method will certainly change the river profile level. We present a simple application on the longest stream in the Moderbach watershed, to show the modifications applied to the stream profile after preprocessing the DEM using these two methods.

Figure. C.1 shows the longitudinal profile of the main stream of the Moderbach in the initial DEM, in the DEM subjected to the filling method and in the DEM subjected to the breaching method.

Both classic methods aim at achieving the same objective, to establish a downward continuous flow path in the DEM raster, while the approach is different. Filling the sinks consist of elevating the height of depressions while the breaching method forces to reduce the height of some cells with high elevations in order to provide a continuous downward flow path.

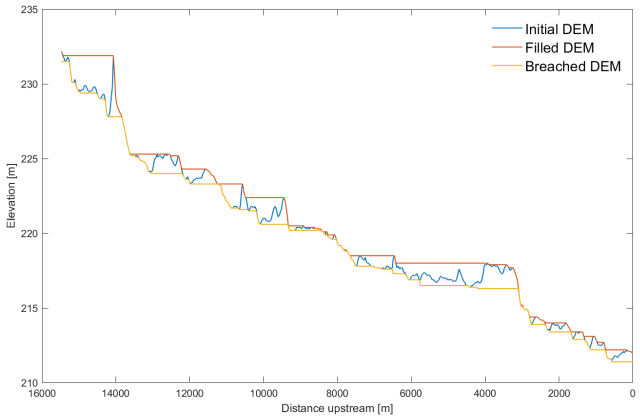


Figure C.1: Stream network longitudinal profile of different DEMs.

C.3 Absolute volume of the DEM

The volume added to or extracted from a DEM due to the application of a sink removal methods, has been widely used to evaluate the impact of these methods on the DEM (Lindsay and Creed, 2005) (Poggio and Soille, 2012).

In the following, we compute and compare the energy required to modify the DEM using each of the two classic methods, filling and breaching.

Using the fill sinks method: the energy required to fill the DEM is expressed in terms of the volume added to the initial DEM to fill all the sinks and depressions (Fig. C.2):

$$V_{filled} = \sum_{i=1}^{N} (DEM_{filled}.Z - DEM_{initial}.Z) * DEM.Cellesize^2 = 27.10^6 m^3 \quad \text{(C.1)}$$

Using the breaching method: the energy required to carve the DEM is expressed in terms of the volume carved from the initial DEM to ensure a continuous downward flow path (Fig. C.3):

$$V_{carved} = \sum_{i=1}^{N} (DEM_{initial}.Z - DEM_{carved}.Z) * DEM.Cellesize^2 = 5.10^6 m^3 \quad \text{(C.2)}$$

N is the number of cells in the DEM, $DEM_{initial}.Z$ is the elevation of the cell in the original DEM, $DEM_{carved}.Z$ is the elevation of the cell in the breached DEM, $DEM_{filled}.Z$ is the elevation of the cell in the filled DEM and $DEM.Cellesize$ is the dimension of the cell in DEM (DEM resolution).

The volume added to the DEM to fill all the depressions is 5 times larger than the one carved from the initial DEM to ensure a continuous downward gradient. Which means that fewer changes and modification are applied to the DEM using the breaching method. Obviously, the fill sinks method will lead to major changes in the DEM which may risk to no longer correspond to the real landscape.

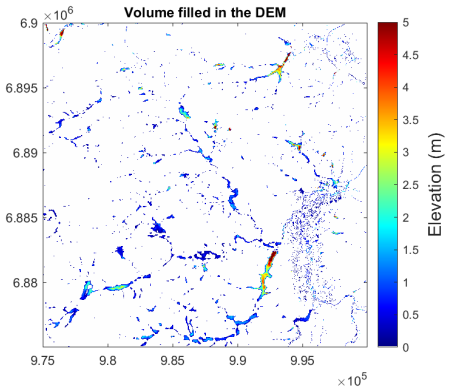


Figure C.2: Volume added to fill the DEM.

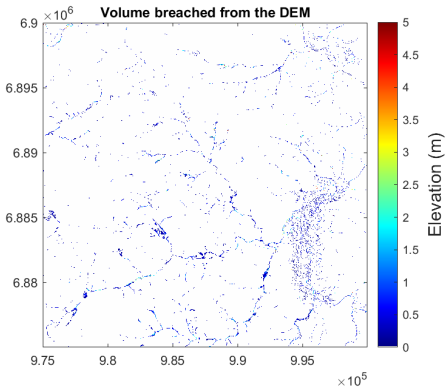


Figure C.3: Volume breached from the DEM.

C.4 Hydrological behavior across obstacles

DEM pre-processing methods may act on the water flow behavior across anthropogenic artifacts that may exist in the landscape as dams or road embankments. The modeler should be aware that the DEM is not able to consider the information of the culverts/spillways allowing the evacuation of the flow through the dam or underneath the road. In a DEM, these structures behave more like a wall that blocks the water and prevents it from passing to the other side.

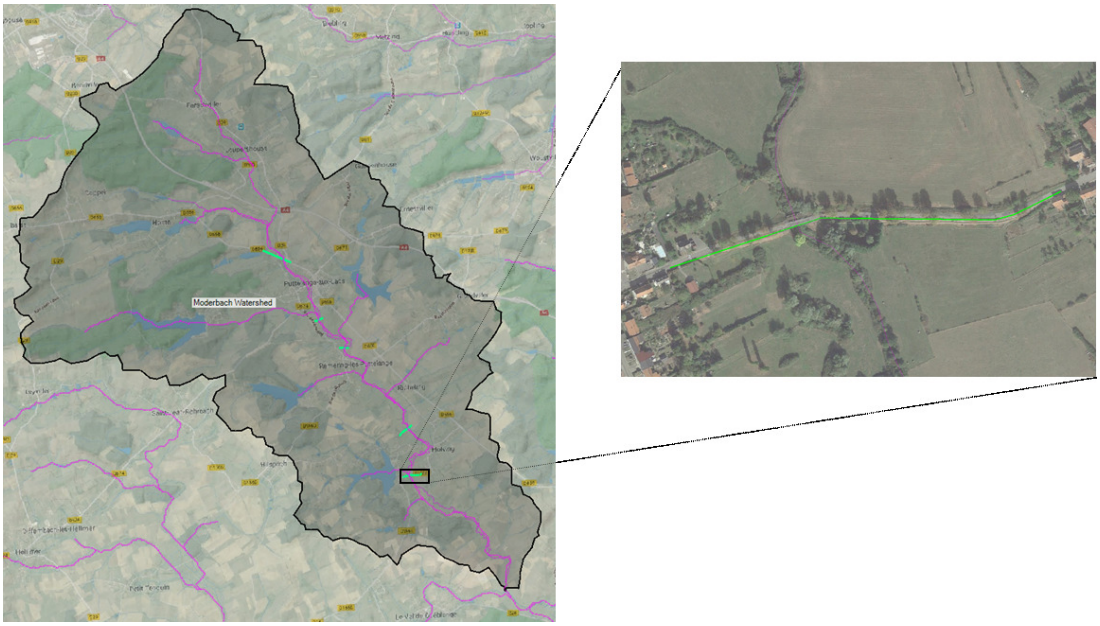


Figure C.4: Dam downstream of the Holving's dike.

In the DEM of the Moderbach area, we focus on the dam structure near Holving (Fig. C.4). The aim is to run a hydrological simulation in this area to compare the behavior of the water flow through the dam in the three following simulations: simulation using the initial DEM, simulation using the DEM subjected to the fill sinks method, and simulations using the DEM subjected to the breaching method. The conditions of the simulation applied are inspired from the event of

1940, where the ponds on the Moderbach river were evacuated to flood the valley for military reasons (to defend the German troops). In the following, we suppose the evacuation of the Holving pond. For this application, in order to provide high-resolution visualization, the simulations are performed using HEC-RAS. The question to be answered here is: "Which of the three simulations applied is the most accurate for routing the water through the dam ?"

The visualization of the surface flow generated in the three simulations applied using the three different DEMs in HEC-RAS, shows the impact of the DEM pre-processing methods on the behavior of the simulated water flow across the dam. In the initial DEM, these structures act as high obstacles not allowing the flow to pass through and the water will be accumulated upstream of the dam (Fig. C.5). This can be acceptable if the modeled event considers that the dam is always closed during the simulation time. The use of the Fill sinks method as a preprocessor for the DEM, will increase the height of the cells upstream and downstream of the dam. This will create a flat area and the water will flow all over as if the dam doesn't exist anymore (Fig. C.6). Breaching the DEM seems to be the closest to reality. It enabled the water to flow through the structure (Fig. C.7).

Figure C.5: HEC-RAS simulation using the initial DEM.

Figure C.6: HEC-RAS simulation with the filled DEM.

Figure C.7: HEC-RAS simulation with the breached DEM.

Université de Strasbourg

Sleimane Hariri

msii ÉCOLE DOCTORALE

Modélisation des inondations historiques de la Sarre

Résumé en Français

Les modèles de précipitations distribuées considèrent les données de précipitations sur un bassin versant comme un terme source dans leurs équations aux dérivées partielles non linéaires. Comme leur discrétisation comprend un maillage de la topographie du bassin versant, leur solution peut échouer en raison d'un manque de capacité de mémoire pour les grands bassins versants. La volonté de réaliser des simulations de précipitations distribuées sur des bassins de grande échelle a motivé la conception d'une nouvelle méthode de partitionnement. Nous mettons particulièrement l'accent sur un équilibre de la charge de calcul pour améliorer son efficacité. Pour une utilisation générale, les performances de la méthode de partitionnement des bassins sont notamment illustrées avec le logiciel HEC-RAS.

Comme cas d'étude, nous considérons le bassin versant de la Sare (partie française) qui draine une superficie de (1800 km^2). L'application de la nouvelle méthode de partitionnement pour effectuer des simulations de précipitations distribuées sur ce bassin à grande échelle montre un meilleur équilibre pour les aires des sous-bassins. En conséquence, la charge de calcul requise pour une résolution utilisant (2 800 000) mailles distribuées de manière équilibrée entre les sous-domaines est considérablement réduite. Les simulations d'événements extrêmes et historiques montrent un très bon accord avec les données observées telles que les données de débit, des cartes de crue (2010), des cartes d'inondations historiques (1940 et 1944).

Mots clés : équilibre des surfaces, HEC-RAS, inondations historiques, modèles hydrologiques, modèle de pluie distribuée, partitionnement des bassins versants.

Résumé en anglais

Rain-on-grid software accounts for precipitation over a watershed as a source term in its nonlinear partial differential equations. As their discretization comprises a mesh of the watershed topography, their solution may fail due to memory shortage for large watersheds. Targeting rain-on-grid simulations on large scale watersheds has motivated the design of a new basin partitioning method. We place a special emphasis on balancing the computational load in order to improve its efficiency. For the sake of general usability, the performance of the basin partitioning method is notably exemplified by using the HEC-RAS software.

As a case study, we consider the Saar catchment (French part) that drains an area of (1800 km^2). The application of the new partition method to perform rain-on-grid simulations on this large scale basin demonstrates a better equilibrium for the sub-basin areas. As a consequence, the computational load that is required for the solution using (2 800 000) mesh elements balanced between the subdomains is substantially reduced. Simulations of extreme and historical events are in very good agreement with observed data such as discharge data, a flood extent map (2010) or with historical flood maps (1940 and 1944).

keyword: Area balancing, HEC-RAS, Historical floods, Hydrological models, Rain-on-grid, Watershed partitioning.