Development of an integrated approach in hydrogeophysics for a multi-scale study of karst reservoirs : the case of the Fourbanne watershed, Jura mountains

Anthony Abi Nader

To cite this version:

Anthony Abi Nader. Development of an integrated approach in hydrogeophysics for a multi-scale study of karst reservoirs : the case of the Fourbanne watershed, Jura mountains. Hydrology. Université Bourgogne Franche-Comté, 2023. English. NNT : 2023UBFCD070 . tel-04617102

HAL Id: tel-04617102
https://theses.hal.science/tel-04617102

Submitted on 19 Jun 2024

THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE

PREPAREE A l'UNIVERSITE DE FRANCHE-COMTE

Ecole doctorale n° 554

Environnements-Santé (ES)

Doctorat de Géodynamique des Enveloppes Supérieures

Spécialités : Hydrogéophysique, Sismologie

Par

Anthony Abi Nader

Développement d'une approche intégrée en hydro-géophysique pour l'étude multi-échelle des réservoirs karstiques : le cas du bassin versant de Fourbanne (Massif du Jura, France)

Thèse présentée et soutenue à Besançon, le 11 décembre 2023

Composition du Jury :

Mme Doubre, Cécile	Physicienne CNAP, Université de Strasbourg, France	Rapportrice
M. Champollion, Cédric	Maître de Conférences, Université de Montpellier, France	Rapporteur
M. Malet, Jean-Philippe	Directeur de Recherche CNRS, Université de Strasbourg, France	Examinateur et Président
M. Stehly, Laurent	Physicien Adjoint CNAP, Université Grenoble Alpes, France	Examinateur
M. Valley, Benoît	Professeur, Université de Neuchâtel, Suisse	Examinateur
Mme Albaric, Julie	Maîtresse de Conférences, Université de Franche-Comté, France	Directrice de thèse
M. Sue, Christian	Professeur, HDR, Université de Franche-Comté, France	Directeur de thèse
M. Pohl, Benjamin	Directeur de Recherche CNRS, Université de Bourgogne, France	Invité
M. Steinmann, Marc	Maître de Conférences, Université de Franche-Comté, France	Invité

Acknowledgements

In this part, you're going to see a more playful aspect of me. For those who know me best, they know that I am always joking around and being silly. I dedicate this part to all the people who made my PhD possible. Get ready, because things will get a bit dramatic in the following paragraphs. So, welcome to my personal diary!

This thesis is the result of three years of work at the University of Bourgogne Franche-Comté in the Chrono-Environment laboratory, in collaboration with the Biogeosciences laboratory (University of Dijon) and EOST (University of Strasbourg). It was carried out thanks to the funding from the Bourgogne Franche-Comté Region and the OSU THETA.

I would like to thank, above all, my thesis advisors, Julie Albaric and Christian Sue, as well as my co-advisors Marc Steinmann and Benjamin Pohl (aka old Ben), for giving me the opportunity to work on such a thesis project—a multidisciplinary project that combines scientific domains and involves several days (weeks, months) of fieldwork, all aimed at advancing science in favor of water resources. It was my utmost honor and pleasure to learn, fail, and then get back on my feet (literally, in the cave), growing and maturing during these three years. The PhD was a comprehensive experience in which I was able to grow in several key areas: scientifically, relation-ally, physically, and emotionally. This journey not only enhanced my research skills and deepened my knowledge, but it also strengthened my ability to forge meaningful relationships and collaborate effectively. Physically, the demands of fieldwork challenged me (yes I am talking about going about 100 meter underground!), while emotionally, navigating the highs and lows of intense research with limited time. I wouldn't have done all of the above without my PhD advisors.

Julie, I want to thank you for always being there in all of the ups and downs of my journey. You were a great mentor on the scientific, professional and personal scales. You were a great support for me and I wouldn't have done it without you. I am utterly proud and honored for having done this path with you. You always had your way to encourage and guide me. Thank you for entrusting me with this project and all of the instrumentations envolved, and thank you for introducing me to the world of hydrogeophysics and for giving me the opportunity to attend numerous national and international scientific meetings throughout my thesis. I want to thank Marc for his insight and time (through countless meetings) in the hydrogeological aspect of my thesis, without you I wouldn't be able to master it and apply it in the seismological interpretation.

Thank you Ben for your insight in climatology and for you continuous positivity and motivation, you had always your way to lift someone up! Finally, thank you to Christian for your trust and many words of encouragement, as well as your wise guidance especially towards the end of my thesis (the ruff part!).

I am deeply grateful to Cécile Doubre and Cédric Champollion for their keen interest in reviewing my work. I also want to thank Laurent Stelhy, Benoît Valley et Jean-Philippe Malet for serving as examiners on my thesis committee and for the valuable insights they contributed to my research. I am glad for all the fruitfuil discussions that occured on the day of my theis defense.

I also extend my gratitude to Clément Hibert and Jean-Philippe Malet for hosting me in their lab and introducing me to state-of-the-art geophysical techniques, including artificial intelligence and fiber optic monitoring.

I am deeply thankful to the speleologists team whose assistance was instrumental in completing a significant portion of this thesis. Their support in gaining access to the underground gallery, conducting surveys, and ensuring our safety during challenging fieldwork was invaluable. They consistently demonstrated a high level of responsiveness, motivation, and scientific curiosity, making them not only essential collaborators but also delightful companions. A special acknowledgment goes to Denis Motte for patiently addressing all my inquiries and providing indispensable maps and data.

I also thank Jordan De-Molo and Mariantonietta Longobardi for their contributions in the project and providing valuable scientific and fieldwork support.

I acknowledge all the colleagues in the Geology Department at Chrono-environnement, Helene Celle, Nicolas Carry, Sophie Denimal, Catherine Bertrand for the fruitful discussions we had concerning my results and interpretations, as well as for their time and availability in discussing ’rock’ and ’water’ issues.

Now, moving on to a more challenging part of these acknowledgments: I want to thank the people who supported me throughout these three years. This includes both new friendships and those I have long known, all of whom provided constant emotional support. I am glad to say that I have made friendships that will forever mark me and experienced moments that I will never forget.

To my Chrono-environnement family, including all the PhD students, postdocs, interns... my journey would not have been the same without you. I want to thank PRINCESS LEIA for being my soul sister and making these years much better, filled with love and laughter. I also want to express gratitude to Jordy and Alex (my office colleagues and forever friends, and by

the way, it was MY office) for all the moments we shared together in Room -116L. There, we supported each other, laughed, danced, sang, or simply enjoyed each other's company, making the burdens much more bearable. I am equally grateful to Lise, Laurie, Lorraine, and Noémie for their beautiful impact on me and for infusing the lab's atmosphere with joy. You are beautiful human beings inside and out! Special thanks to Josho for always being the most supportive and thoughtful, and for patiently listening to my complaints and offering the right words of comfort. I also want to extend my gratitude to Olivier, Kilian, Jean-Baptiste, Milva, Jonathan, Thomas, Victor, Emma, Céline, Amandine, Mellina, Romain, Clara, Cyril, Pierre... the list goes on! Sharing the thesis experience with people facing similar challenges made everything much easier. We formed an incredible support system for each other during the three-year roller coaster. And of course thank you for all the nights spent in bars and all the house parties we had!

I am also deeply grateful to all my laboratory colleagues (he list migh get really long) for the wonderful moments we shared together: the countless coffee breaks, the lunches we enjoyed together, and the various activities and ceremonies we participated in (like the BBQ!). I thank you for all the discussions, laughs, coffees, beers, and everything in between!

I want to thank Théo for his continuous support during the ups and downs, I couldn't have done it without you. I want to thank Bassam and Nawal with whom my journey started at Besançon, they were the first silver lining. I want to thank Magalie, Nancy and Cynthia for their presence in my life and the comfort they were giving since so many years.

I want to thank my best friends, Alain, Anna, Rawad, Salwa, and Vanessa. They stood by me even though I was sometimes (often) unbearable with all my complaints. They were always there to lift me up and provide a safe space where I could scream, cry, and laugh. They gave me unconditional care and love.

And last but certainly not least, I want to thank my mom, Katia, dad, Fadi, and two sisters, Joyce and Joëlle, for everything they have provided me with. Their encouragement, belief in me, and unwavering support have been invaluable throughout. They stood by me during my lowest moments and celebrated with me during my triumphs. They never doubted my abilities and always provided the motivation I needed. Being away from you was not easy, and the circumstances were not always in our favor. However, I am proud to say that your son is now a PhD.

Abstract

Understanding the functioning of karst aquifers is essential for effective groundwater resource management and environmental protection. The intrinsic variability in karst terrain arises from the dissolution of soluble rocks like limestone, resulting in an intricate and interconnected network of fractures, sinkholes, caves, underground conduits, and surface depressions. This study focuses on the Fourbanne karstic watershed (Jura Mountains, France), hosting one main underground conduit accessible through a 20 m drilled shaft. This thesis offers various complementary approaches to better characterize the studied site while drawing upon different principles and analysing jointly geophysical and hydrogeological data. Two main axes are developed: (1) study of the hydrodynamic functioning of the underground river and (2) study the 2D watershed's response to rainfall in terms of saturation and flow patterns. The underground river is continuously monitored by two long term broadband seismometers (located at the surface and underground), hydrogeological probes and a fiber optic cable for temperature measurements. We first propose a new approach based on a machine learning algorithm, the Random Forests (RF), and seismic noise recordings, to predict the underground river water height. We use data recorded from both of the long term seismometers and yield a prediction with an efficiency above 95% and 53% for data from the underground and surface seismometers, respectively. The RF model proves to be a reliable method for remote monitoring of the water heights, even with weak river-induced signals. In a second step, we use Fiber Optic Distributed Temperature Sensing to survey temperature changes in the karst reservoir to improve our understanding of its recharge processes. We investigate in detail the temperature response of the river and one of its affluent to atmospheric changes based on a 6 months hydrogeophysical dataset. We show that precipitation events induce change in the river temperature. When the water height increases, it flows into adjacent conduit(s), pushing out long residence water stored in a reservoir/sink into the affluent. A local recharge is feeding the affluent, potentially with water infiltrating from the epikarst through ∼E-W fracture system. In the second axis, we use the seismometer located underground and a dense seismic array of 60 nodes, deployed in the framework of this project for a period of 4 months, and compute single station cross-component correlations to recover seismic velocity variations. The objective is to monitor the long term groundwater variations using 4 years data recorded at the underground seismometer, and investigate the groundwater circulation on the watershed scale as direct response to precipitation using the seismic array. The long term analysis reveals seasonal variations generated by change in saturation between wet and dry seasons. Velocity variation and precipitation are concomitant, indicating fast infiltration mechanisms in highly fractured media. The seismic array yields anisotropic results, giving insight into the complex and heterogeneous nature of the studied site. The results show preferential flow patterns direction's are highly variable within the karst. Finally, the results of this thesis provide new insights into the monitoring of karst reservoirs through hydrogeophysical techniques, highlighting the crucial role these techniques may play in studying such heterogeneous systems.

Résumé

Comprendre le fonctionnement des aquifères karstiques est essentiel pour une gestion efficace des ressources en eau. La variabilité intrinsèque des terrains karstiques résulte de la dissolution de roches solubles telles que le calcaire, ce qui donne lieu à un réseau complexe et interconnecté de pertes, grottes, conduits souterrains et de dolines. Ce travail se concentre sur le bassin versant karstique de Fourbanne (montagnes du Jura, France), qui abrite un conduit souterrain principal, accessible par un puit artificiel de 20 m. Différentes approches complémentaires sont proposées pour mieux caractériser le site étudié en s'appuyant sur l'analyse conjointe de données géophysiques et hydrologéologiques. Deux axes principaux sont développés : (1) l'étude de l'hydrodynamique de la rivière souterraine et (2) l'étude 2D des variations de saturation et des transferts d'eau en réponse aux précipitations, à l'échelle du bassin versant. Le monitoring continu de la rivière souterraine est assuré par plusieurs instruments: deux sismomètres (situés à la surface et dans le conduit souterrain), des sondes hydrologiques et une fibre optique pour la température. Nous proposons tout d'abord une nouvelle approche basée sur un algorithme d'intelligence artificielle, le Random Forests (RF) et le bruit sismique, pour prédire la hauteur d'eau de la rivière souterraine. Nous utilisons les données enregistrées par les sismomètres permanents et obtenons une prédiction avec une efficacité de 95% et 53% avec les sismomètres souterrain et de surface, respectivement. Le modèle RF s'avère être une méthode fiable pour le suivi à distance des hauteurs d'eau, même avec de faibles signaux induits par la rivière. Dans un deuxième temps, nous utilisons la méthode FO-DTS (Fiber Optic Distributed Temperature Sensing) pour étudier les changements de température dans le réservoir karstique et améliorer notre compréhension des processus de recharge. Nous étudions en détail la réponse de la température de la rivière et de l'un de ses affluents aux changements atmosphériques en utilisant des données hydrogéophysiques enregistrées pendant 6 mois. Nous montrons que les précipitations induisent des changements dans la température de la rivière. Lorsque la hauteur d'eau augmente, elle s'écoule dans le(s) conduit(s) adjacent(s), repoussant dans l'affluent l'eau stockée depuis longtemps dans un réservoir. Une recharge locale alimente l'affluent, potentiellement avec de l'eau infiltrée depuis l'épikarst à travers le système de fractures ~E-W. Concernant le deuxième axe, nous utilisons le sismomètre situé dans le conduit et un réseau sismique dense de 60 nodes, déployé pour une période de 4 mois. A partir de l'autocorrélation du bruit sismique ambiant, nous mesurons les variations relatives de vitesse des ondes sismiques (dv/v). L'analyse des données long terme (4 ans) issues de la station permanente révèle des variations saisonnières générées par le changement de saturation entre les saisons humides et sèches. Les variations de vitesse et les précipitations sont concomitantes, ce qui indique des mécanismes d'infiltration rapide dans des milieux très fracturés. L'autocorrélation, en utilisant le réseau dense, génère des variations de vitesse anisotropes, donnant un aperçu de la nature complexe et hétérogène du site étudié. Les résultats montrent que les directions d'écoulement préférentiel sont très variables dans le karst. Finalement, les résultats de cette thèse offrent de nouvelles perspectives pour le suivi des réservoirs karstiques à l'aide des techniques hydrogéophysiques, soulignant le rôle prépondérant que ces techniques peuvent jouer dans l'étude de systèmes fortement hétérogènes.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Karst Aquifers

According to Ford and Williams (2007), approximately 25% of the world's population depends on groundwater extracted from karst aquifers. In France, they occupy 35% of the territory's surface area (see Fig. 1.1), and contribute in 40% of the total drinking water supply (Chen et al., 2017). More precisely, in Franche-Comté region, 80% of the water intakes for drinking water supply are drawn from karstic aquifers (Bakalowicz, 1999).

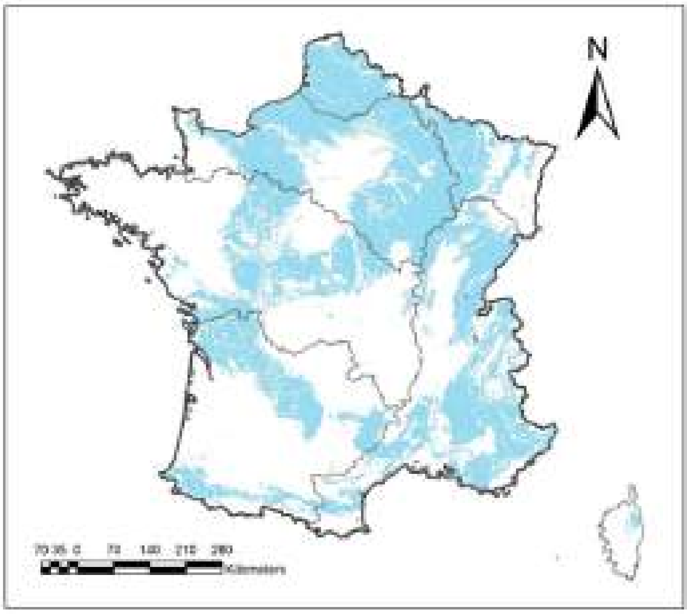


Fig. 1.1 Distribution of karstic carbonate rocks in France (Maréchal and Rouillard, 2020).

Karst is a geological formation found in carbonate rocks, such as limestone and dolomite. Water infiltrates, enriched with carbon dioxide (CO_2), and flows through fractures and weak zones in the rock, gradually dissolving the rocks forming preferential flow paths and underground drainage networks (Fig. 1.2), generally called secondary porosity (Bakalowicz, 2005; Dörfliger et al., 2010; Ford and Williams, 2007). This transformation occurs over a short duration on the geological scale: from a few thousand to several tens of thousands of years. By creating different sized voids within the massif, karst systems become highly heterogeneous environments with varying degrees of permeability.

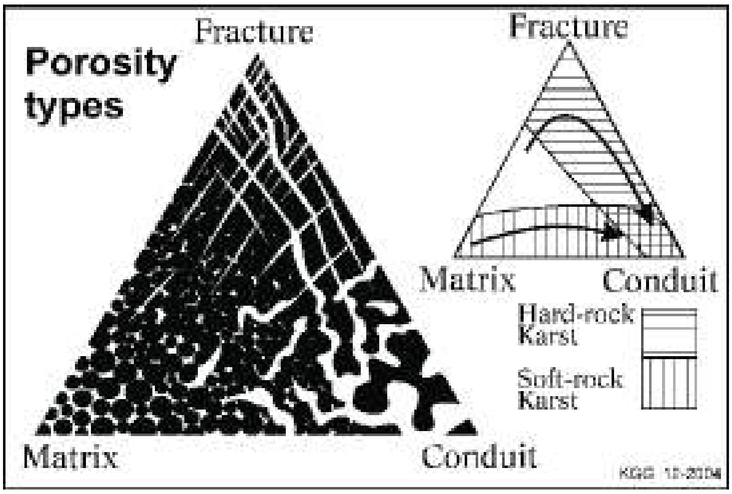


Fig. 1.2 Karst porosity. Left: the three types of porosity encountered in karstic medium. Right: evolution of the dominant porosity in hardrock and soft-rock karsts. (Grimes, 2006).

Figure 1.3 is a conceptual sketch of karst aquifers in the Jura mountains, modified from Bichet and Campy (2008). Karst aquifers operate through a complex interplay of sinkholes, fractures, and matrix flow. Sinkholes formed from the surface, as well as fault, allow rapid infiltration of water into the karst system. Within the matrix, the rock itself can hold and transmit water through pore spaces. The impermeable marl layer forms an aquitard, where the flow of water is restrained. Limestone at the level of the outlet is saturated and the network is completely submerged.

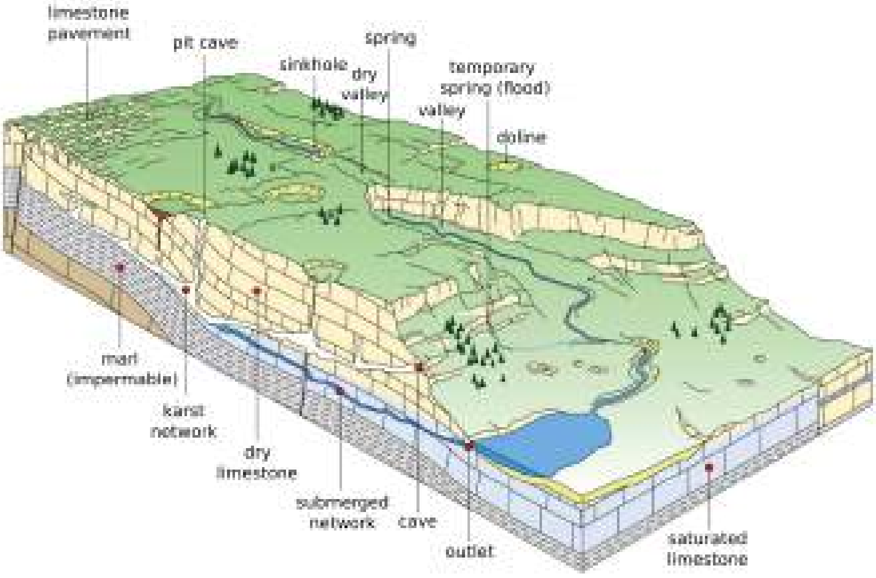


Fig. 1.3 Conceptual sketch of karst environment in the Jura Mountains (modified from Bichet and Campy, 2008).

A karstic aquifer is vertically divided into three zones: the epikarst, the unsaturated or vadose zone, and the saturated or phreatic zone (Fig. 1.4). The epikarst corresponds to the superficial part of the karst. This relatively thin layer is highly permeable due to surface physico-chemical processes (Blavoux and Mudry, 1986). The vadose zone is an important area for recharge, as precipitation and other surface water sources can infiltrate into the karst aquifer through this unsaturated zone. Water in the vadose zone can eventually flow down to the saturated zone, recharging the groundwater supply (Hartmann et al., 2014). The boundary between the saturated and unsaturated zones fluctuates depending on the occurrence and intensity of precipitations.

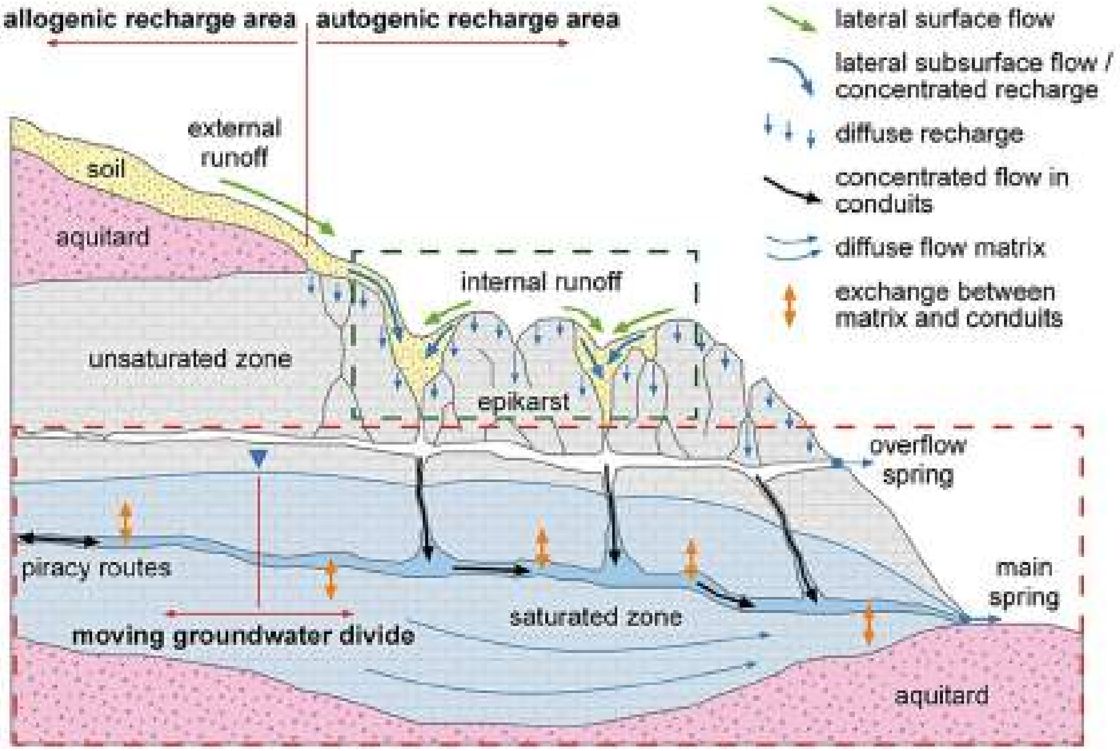


Fig. 1.4 Conceptual model of a karst system including all characteristic karst processes; dark green and red dashed lines represent the soil/epikarst and the groundwater subsystems (Hartmann et al., 2014).

High spatial variability can be encountered in the subsurface of karst aquifers, as a result of the multiple scaled porosities and different permeabilities of the medium. The latter leads to a spatial-temporal variability in the occurrence of hydrological processes. Hydrogeological monitoring and other conventional methods deliver punctual observations and global indications on the aquifer, with limited information on its structure and spatial distribution of its heterogeneities. Geophysical methods come in complementary, providing non-invasive ways to gain insights into the subsurface structure of karst environments. These methods can also help to identify potential pathways as well as to better understand saturated and unsaturated zones' dynamics and the interaction between both.

1.2 Hydrogeophysics for Aquifer Monitoring

Hydrogeophysics is a multidisciplinary field presenting a cross-over between geophysical methods applied in hydrological context to investigate and characterize subsurface water resources and related phenomena. Electrical Resistivity Tomography (ERT) is for example commonly used for studying aquifers, as it can give insights on hydraulic conductivity (Vogelgesang et al., 2019), porosity (Binley and Kemna, 2005), and aquifer structure (Torrese, 2019). Ground-penetrating radar can provide imagery of the shallow subsurface through transmission of electromagnetic waves that reflect back to antenna when collision with contrasting dielectric material (Doolittle et al., 2006). In karst media, this method allowed the detection of karstic features such as air-filled and sediment-filled cave, as well as delineating their geometry (Bermejo et al., 2020). Gravimetry methods allow to monitor water transfer and storage in karst reservoirs (Champollion et al., 2018; Fores et al., 2017). This work relies on passive seismic techniques and distributed temperature sensing measurements, only recently applied to karst medium (e.g., Caldwell et al., 2020; Fores et al., 2018).

1.2.1 Fiber Optic Distributed Temperature Sensing

Fiber Optic Distributed Temperature Sensing (FO-DTS) is a technology used to measure temperature variations along the length of an optical fiber cable. It relies on the principles of optical physics to obtain continuous profile of the temperature distribution along the fiber cable. This method has shown its strength in multiple applications such structural health monitoring in concrete structures, leakage monitoring in dams, oil and gas application, fire detection, atmospheric parameters monitoring (Ukil et al., 2012). FO-DTS was used to survey and inspect hydrological processes such as stream-aquifer interactions, groundwater discharges, hyporheic exchange (Day-Lewis et al., 2008). It was also employed to estimate thermal conductivity and groundwater flux in porous media at a very high sampling resolution, for a better understanding of the variability of groundwater flows (Simon et al., 2021). In karstic media, FO-DTS was applied to study the interaction between stream temperature, streamflow and groundwater exchange in an arid karst environment in Texas, USA (Caldwell et al., 2020). They recorded data for a period of 3 years (see Fig. 1.5 for few monthly examples), that were then used in land surface models to construct extended (30-year) water temperature records. Their model revealed a long-term warming trend for water temperature, indicating the susceptibility of these vulnerable systems to climate change.

1.2.2 Environmental Seismology

Environmental seismology consists in analysing seismic vibrations originating and/or affected by processes occurring outside the solid earth. Alternatively, they can be influenced by changes in external environmental factors, such as temperature, hydrology, or human activities, which disrupt their propagation within the solid earth.(Larose et al., 2015). Using ambient seismic

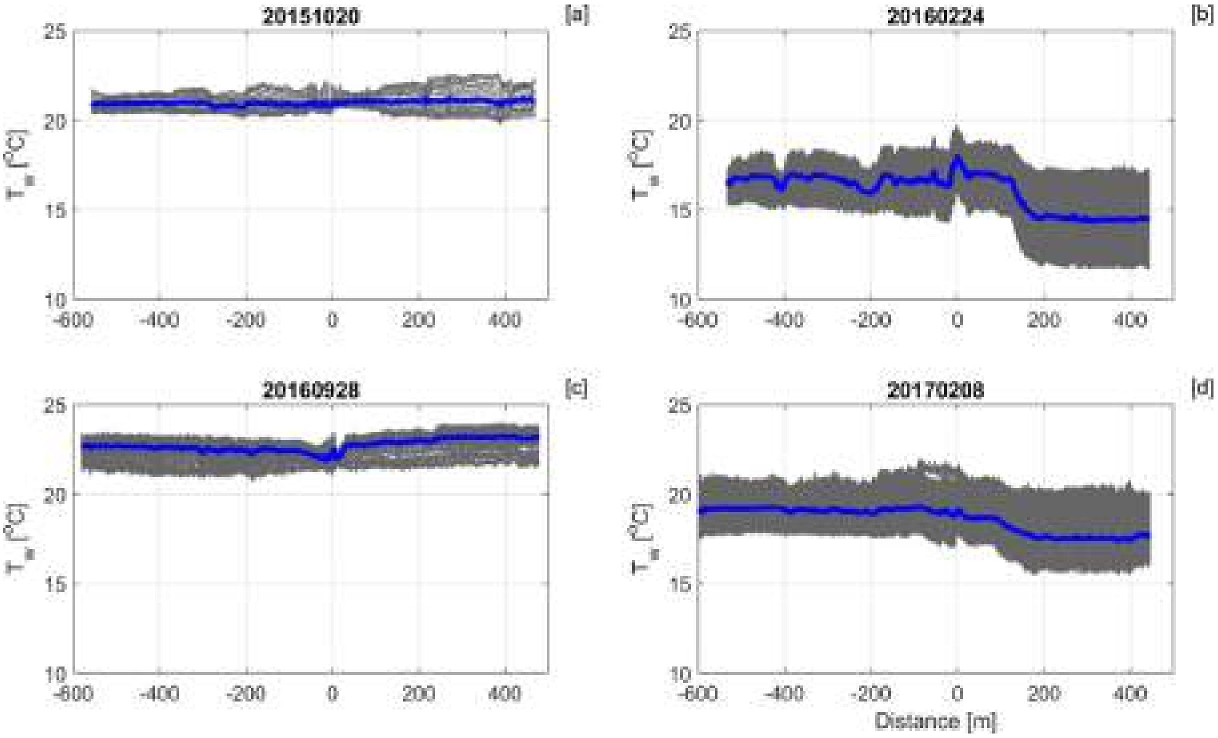


Fig. 1.5 Fiber optic DTS measurements for [a] October 2015, [b] February 2016, [c] September 2016, and [d] February 2017. Gray lines indicate 6-minute measurements aggregated into mean (blue) for the measurement period. The origin distance (0 m) indicates the junction of the 400 m upstream and -600 m downstream cable. (from Caldwell et al., 2020).

noise we can track a variety of natural processes such as landslides (Mainsant et al., 2012a), river hydrodynamics (Burtin et al., 2008), glacier erosion (Gimbert et al., 2016), permafrost warming (Albaric et al., 2021), and other natural hazards (Larose et al., 2015). Precipitations induce changes in the hydro-mechanical properties of the ground that can be precisely measured through seismic recordings (Karato and Wong, 1995), in particular with seismic cross-correlation methods (Sens-Schönfelder and Wegler, 2006). Fores et al. (2018) used seismic cross-correlation to study infiltration and saturation changes in a karstic site in the Larzac plateau (France). Using a simple velocity model, the authors demonstrated that the recorded velocity variations occurred in a depth of tens of meters. They observed a delayed response between velocity variation and precipitation, that is related to slow infiltration rate (Fig. 1.6).

Vidal et al. (2021) also used ambient seismic noise cross-correlation on a karstic site in Northern Italy. They measured changes in seismic velocity occurring within the first 4 km of depth, caused by water recharge/discharge cycles in karst aquifers, transmitting large pore-pressure changes. In addition, in their case study, the spatial distribution of the main velocity variations do not correspond with the spatial distribution of rainfall (Fig. 1.7). This is related to the role that play geological formations at the surface in the hydrological loading: sites with less permeable formation will induce surface run-off, limiting groundwater drainage in depth.

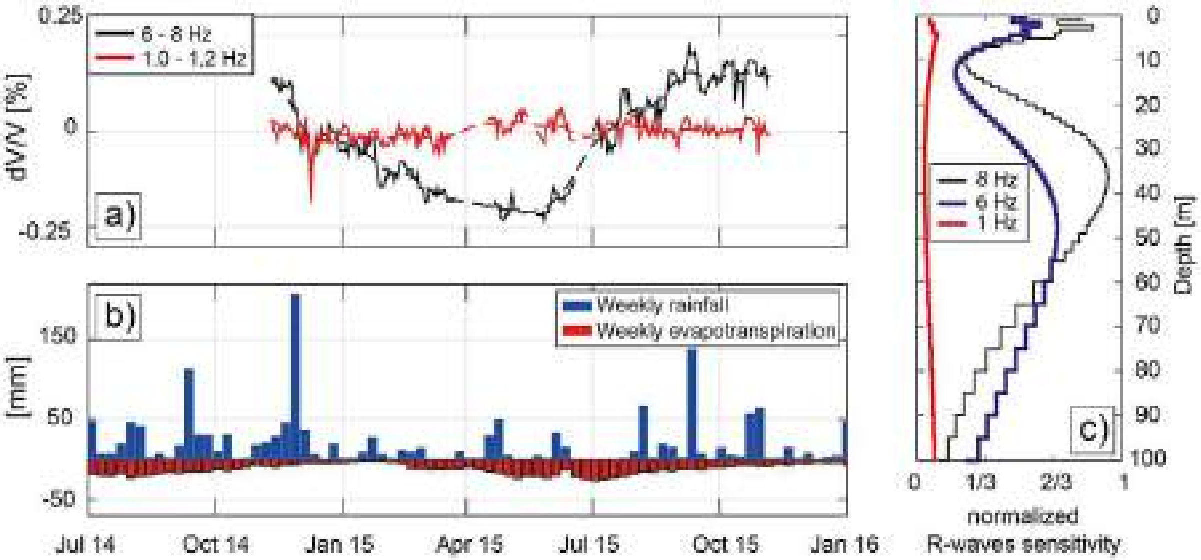


Fig. 1.6 Relation of rainfall precipitation vs. seismic-velocity perturbation. (a) Seismic velocity changes (dV/V) using ambient seismic noise, after a band-pass filtering at 1 to 1.2 Hz (black line) and 6 to 8 Hz (red line); (b) weekly rainfall and evapotranspiration between July 2014 and January 2016; and (c) sensitivity kernel at 1, 6, and 8 Hz computed by the software developed by Herrmann (2013) (from Fores et al., 2018).

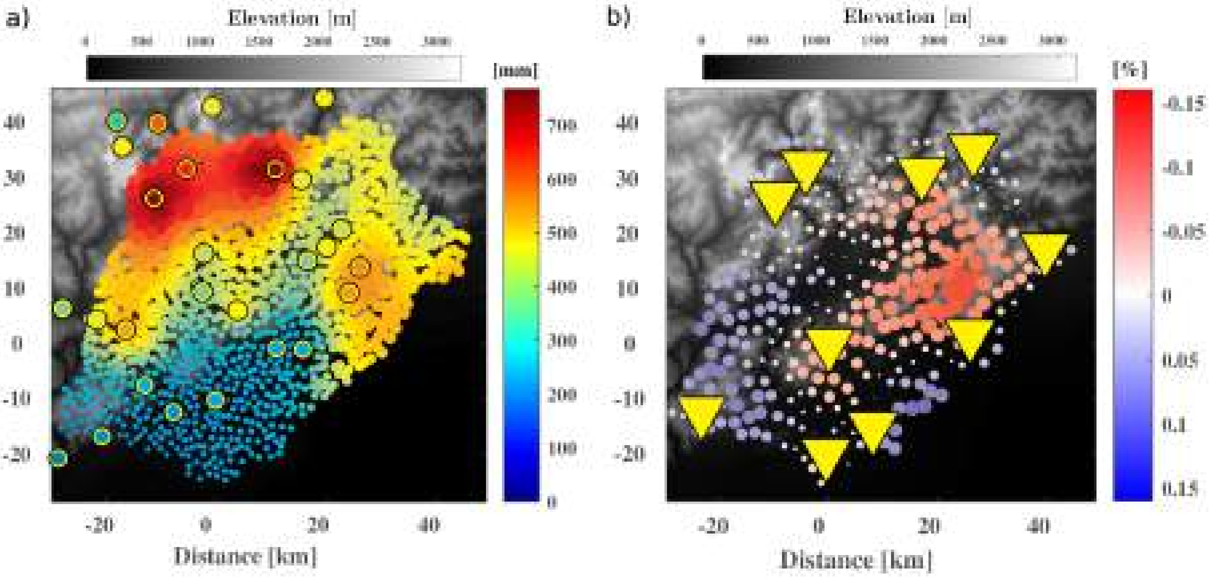


Fig. 1.7 Relation of rainfall precipitation vs. seismic-velocity perturbation. a) Kriging result from cumulated precipitation measurements from pluviometric stations (coloured circles) on 25/11/2018 over the previous 30 days. b)Averaged mapping result of the relative seismic-velocity variations (dv/v) with 10 stations (yellow triangles) on 25/11/2018 over the previous 7 days. (modified from Vidal et al., 2021).

1.3 Objectives

This thesis aims to better understand and characterize water transfer within a karst aquifer in Jura. It starts from the scale of the main conduit driving an underground river in the unsaturated zone and heads to the scale of the watershed. An approach combining hydrogeological, geophysical and machine learning have been developed to optimise the remote monitoring of these inaccessible, heterogeneous and anisotropic reservoirs. The manuscript is structured as follow:

Chapter 1: Studied Site and Related Observatories

In this chapter we present the targeted karst aquifer (Fourbanne Watershed, Jura Mountains, France), its geographical and geological characteristics and what we know about its hydrogeological functioning based on previous studies. We also present the related observatories and the instrumentation used in this project. In fact, the studied site is a part of a national observatory, with continuous hydrological monitoring since 2014, and seismological monitoring since late 2018. We also present the surveys done in the framework of this thesis project and their corresponding fieldwork.

Chapter 2: Remote Monitoring of the Underground River's Hydrodynamics

In this chapter we focus on processes within the underground river. Divided into two main sections, this chapter presents innovative methods to monitor remotely the underground river' water height and recharge. In the first section we present a machine learning based method, the Random Forests, applied on continuous seismic data, to forecast the underground river water height, using both seismometers located underground near the river and at the surface. The second section consists in using fiber optic cable deployed in the underground river to survey and analyse changes in the river's temperature at high spatial and temporal resolutions in response to precipitation events, and better understand lateral inflows and recharge mechanisms.

Chapter 3: Seismic Noise Autocorrelation to Track Groundwater flows within the watershed

In this chapter we study the response of the whole watershed to precipitation through seismic noise analysis. We use a permanent seismometer to study long term variations and a dense array of 60 seismometers for a shorter term (4 months) 2D response. Using seismic noise autocorrelation, we compute the seismic velocity variation (dv/v) beneath every station. Comparing the dv/v with rainfall data, we discuss about changes in saturation in the unsaturated part of the aquifer, and try to understand its recharge mechanisms, as well as the interactions between saturated and unsaturated parts.

Chapter 2

The Fourbanne karst aquifer: a hydrogeophysical observatory

2.1 Geological and Hydrogeological Settings of the Fourbanne Karst Aquifer

2.1.1 Regional Geology: The Jura Mountains

The studied site is the Fourbanne watershed located in Jura mountains, eastern France (Fig. 2.1). The Jura mountains are a sub-alpine mountain range, located north of the Western Alps and mark the majority of the border between France and Switzerland, extending over 360 km from the Rhône River to the Rhine, and ranging in width from a few kilometers to 65 km between Besançon (France) and Neuchâtel (Switzerland) with a total area of 22,686 km² (Cholet, 2017).

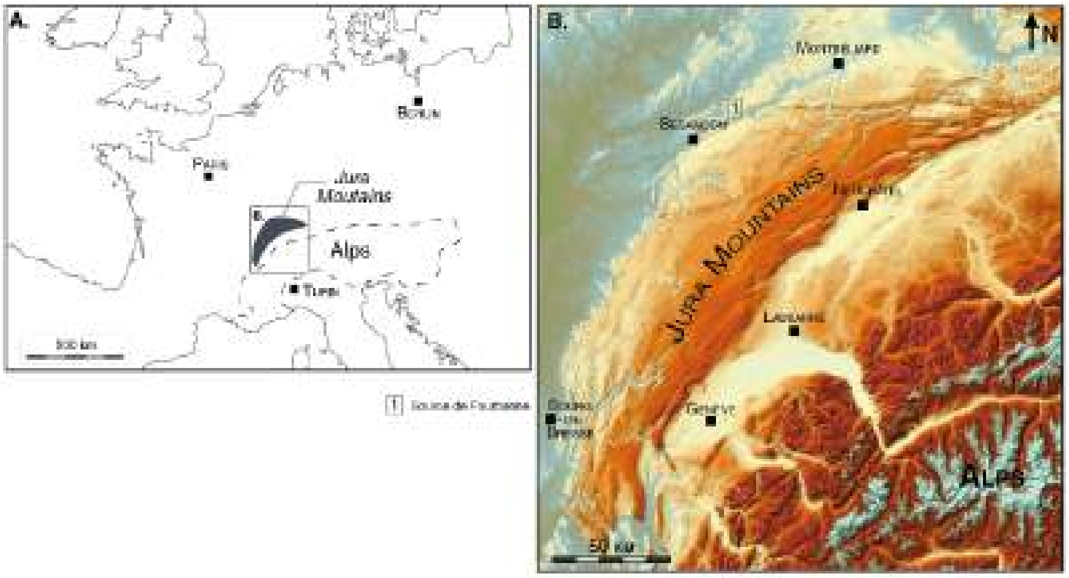


Fig. 2.1 The Jura mountains. a) Geographical location and b) IGN topographical map showing the location of the studied site, from Cholet (2017).

The Jura massif is surrounded by tertiary and quaternary sedimentary depressions, with the Rhine graben to the north, the Bresse graben to the west, and the Molasse basin of the Swiss plateau to the southeast (Fig. 2.2). It is composed of three main units (Sommaruga, 1996): the internal folded Jura, the external plateau Jura and the most external tabular Jura, where the Fourbanne watershed is located.

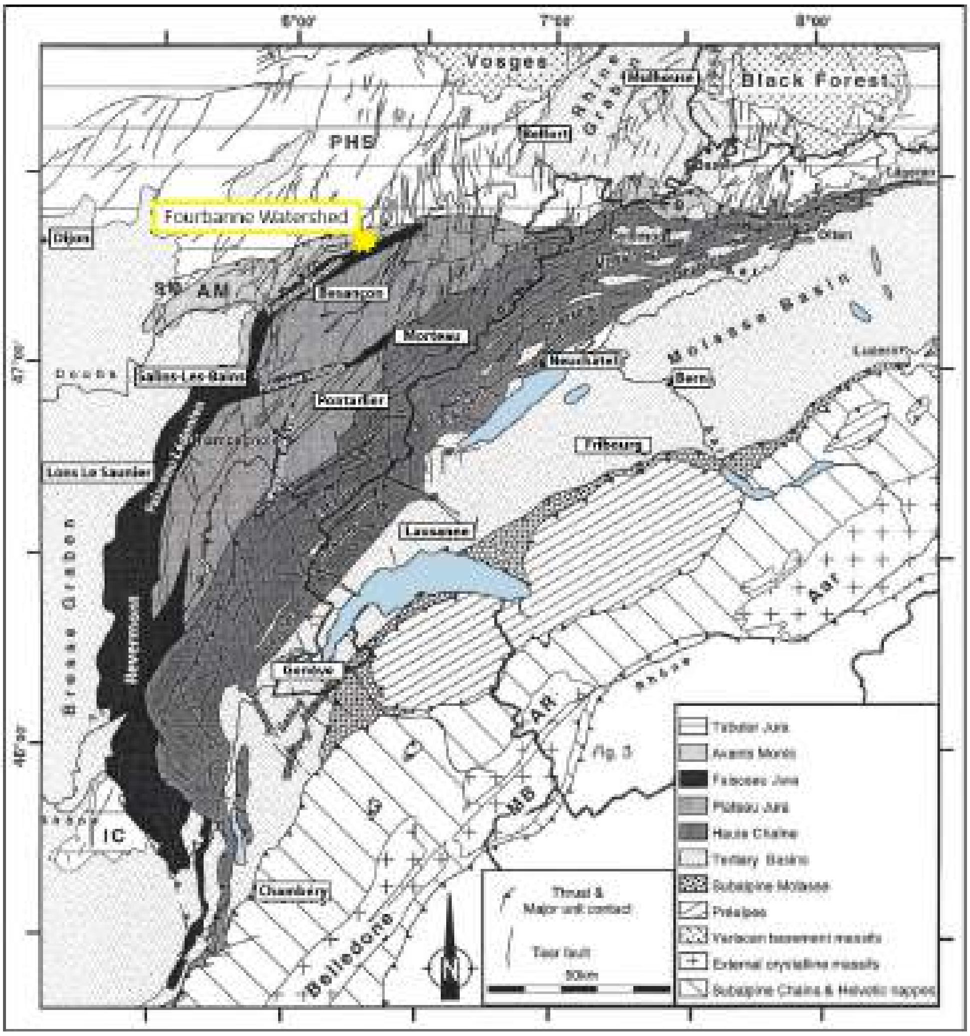
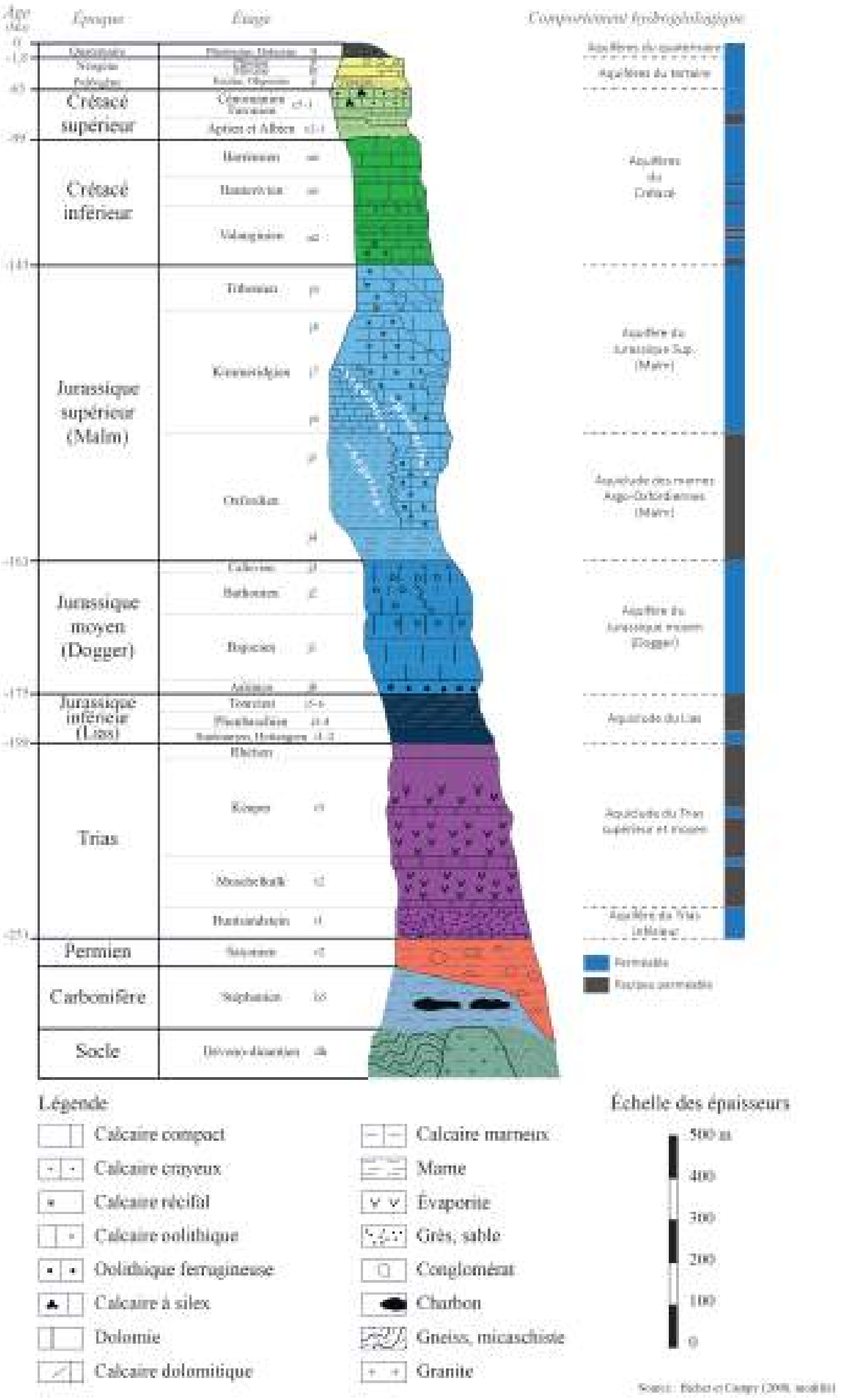


Fig. 2.2 Main tectonic units of the Jura massif and the position of the Fourbanne Watershed (from Sommaruga, 1996).

The Jura mountains are mainly composed of Mesozoic deposits (dating from -251 to -65 million years ago), particularly in the Jurassic and Cretaceous series, and to more recent deposits from the Cenozoic era. The entire stratigraphic series of the Jura Mountains and the associated aquifers and aquicludes are illustrated in Figure 2.3.
The Jura mountains are highly karstified: 80% of drinking water and almost all river water originates from karst springs. The large limestone series of the Middle (or Dogger) and Upper Jurassic (or Malm) are the main karst aquifers, they are underlained respectively by the marly-argillaceous liassic and oxfordian series acting as aquicludes.



Source : Bichet et Campy (2008, modifié)

Fig. 2.3 Stratigraphic sequence characteristic of the Jura Massif from Bichet and Campy (2008), linked to the main hydrogeological features.

2.1.2 Description and limits of the watershed

The Fourbanne Watershed is a site of the Jurassic Karst Observatory, which is a local node of the national network "SNO Karst" (labellized by the National Institute for Earth Sciences and

Astronomy - INSU) (Jourde et al., 2018) and the french network of Critical Zone Observatories OZCAR (Gaillardet et al., 2018). It has been the target of two PhD studies in hydrogeology by Charmoille (2005) and Cholet (2017).

It is located at the intersection of major regional geological structures, which are the "faisceau bisontin", the Jura plateau, and the Avant Monts (Fig.2.2). The Doubs valley is situated in the northeastern extension of the "faisceau bisontin," which separates the Jura plateau from the Avant Monts zone. The "faisceau bisontin" is a folded and faulted zone with a width of approximately 5 km. Between Besançon and Baume-les-Dames, at the location of the outlet at Fourbanne, the "faisceau bisontin" is reduced to a very narrow and highly deformed area. The watershed is spanning over 30 km². The limits were set through geological field investigations, aerial photographs and digital elevation model analysis and finally artificial tracing. Charmoille (2005) has first delimited the basin as shown in Figure 2.4.

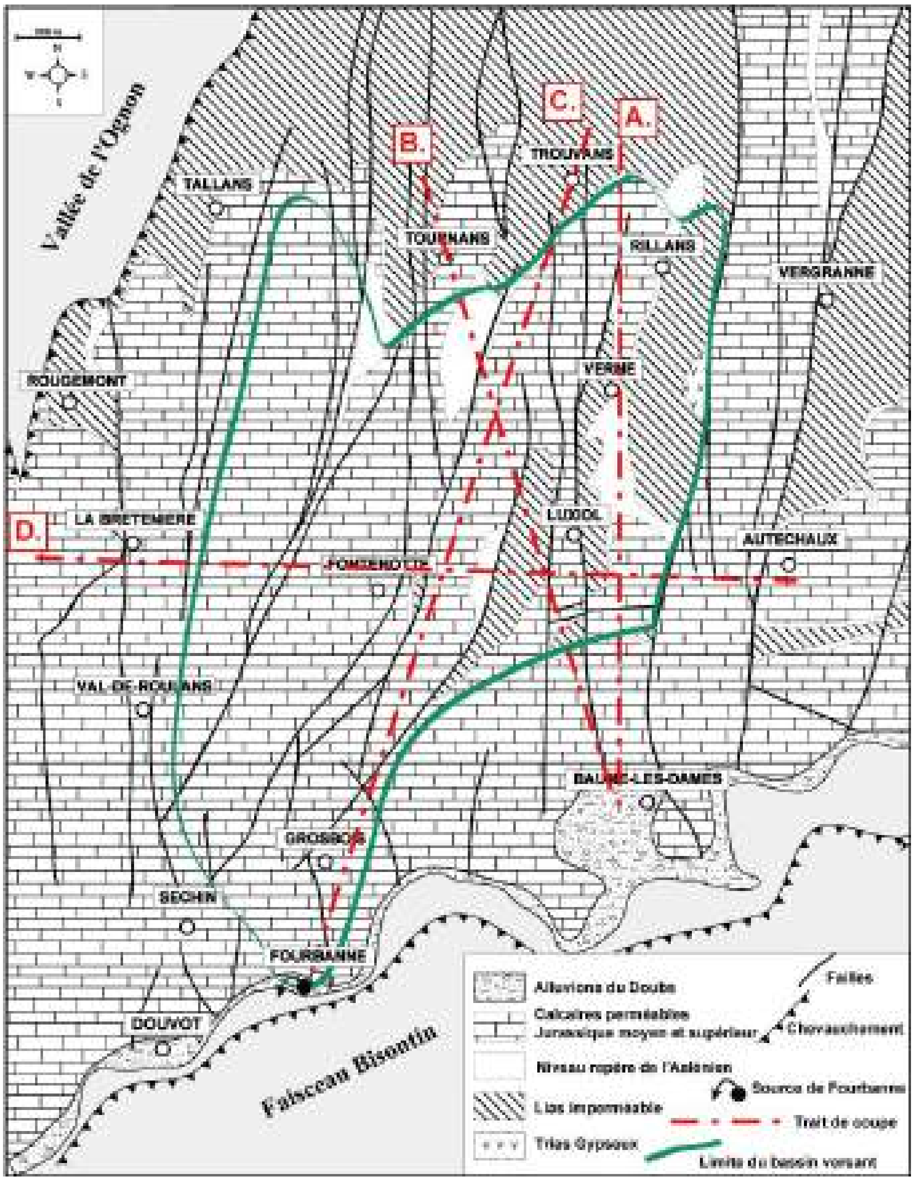


Fig. 2.4 Fourbanne watershed limits (green line) according to Charmoille (2005), along with the geological formations and fault lines (from, Charmoille, 2005).

These limits have been modified by Cholet (2017) based on new tracer tests completed with those of cavers, master's students and engineering firms (Fig. 2.5). Complementary geological field work also allowed to adjust the western limit of the watershed.

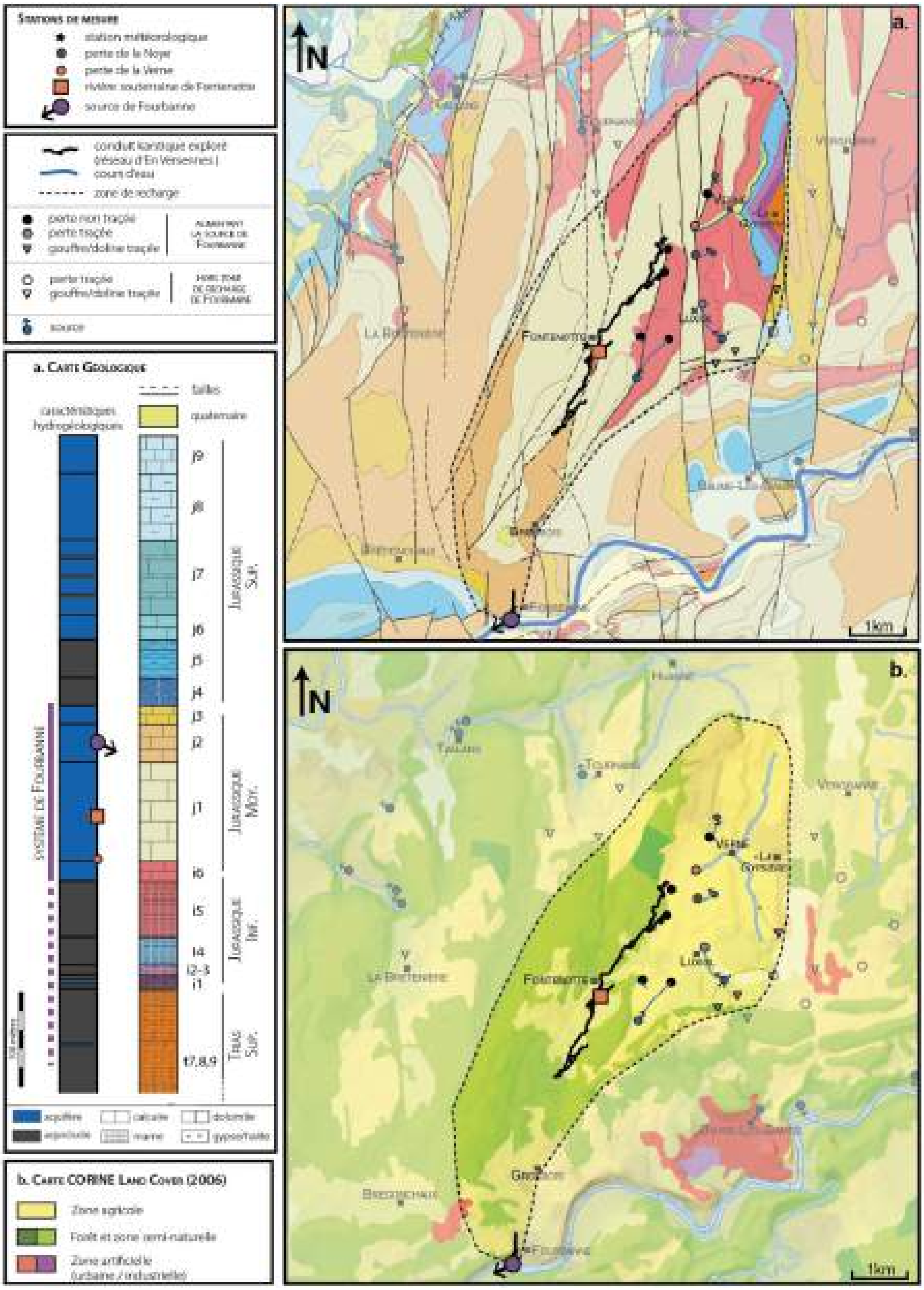


Fig. 2.5 Description of the Fourbanne karst watershed through (a) the geological map and (b) the land use map, (from, Cholet, 2017).

The watershed geology predominantly consists of limestones from the Middle Jurassic era, intersected with a network of north-south meridian faults. These faults expose the marl layers mostly from the Lower Jurassic and the Triassic (Lias) at the surface (Fig. 2.5a and 2.3). Cholet et al. (2017) showed that the limestones between the villages of Fontenotte and La Bretenière are slightly dipping to the west, facilitating part of the flow in that direction. On the other hand, the NNE-SSW trending fault network between the villages of Tournans and Val-de-Roulans appeared to act as a hydraulic barrier and was therefore considered by Cholet et al. (2017) as the western boundary of the Fourbanne watershed .

The maximum elevation of the basin is about 450 m, with 262 m at it's outlet (Fourbanne) where it connects by 120 m stream to the Doubs River. The watershed lands are mainly agricultural (57%), with the rest of the watershed covered by forests and small urban areas (2.5b). The aquifer is a vauclusian type, due to the presence of of a 25 m deep siphon below the outlet, and is characterised by allochthonous recharge through several sinkholes that drain surface streams. Figure 2.6 is a hydrogeological map of the Fourbanne watershed, taken from Cholet (2017). This figure first showcases the different geological units and faults across the watershed, as well as its hydrological main components : the main explored ~N-S conduit, the swallow holes, the location of the cave access and the outlet spring.

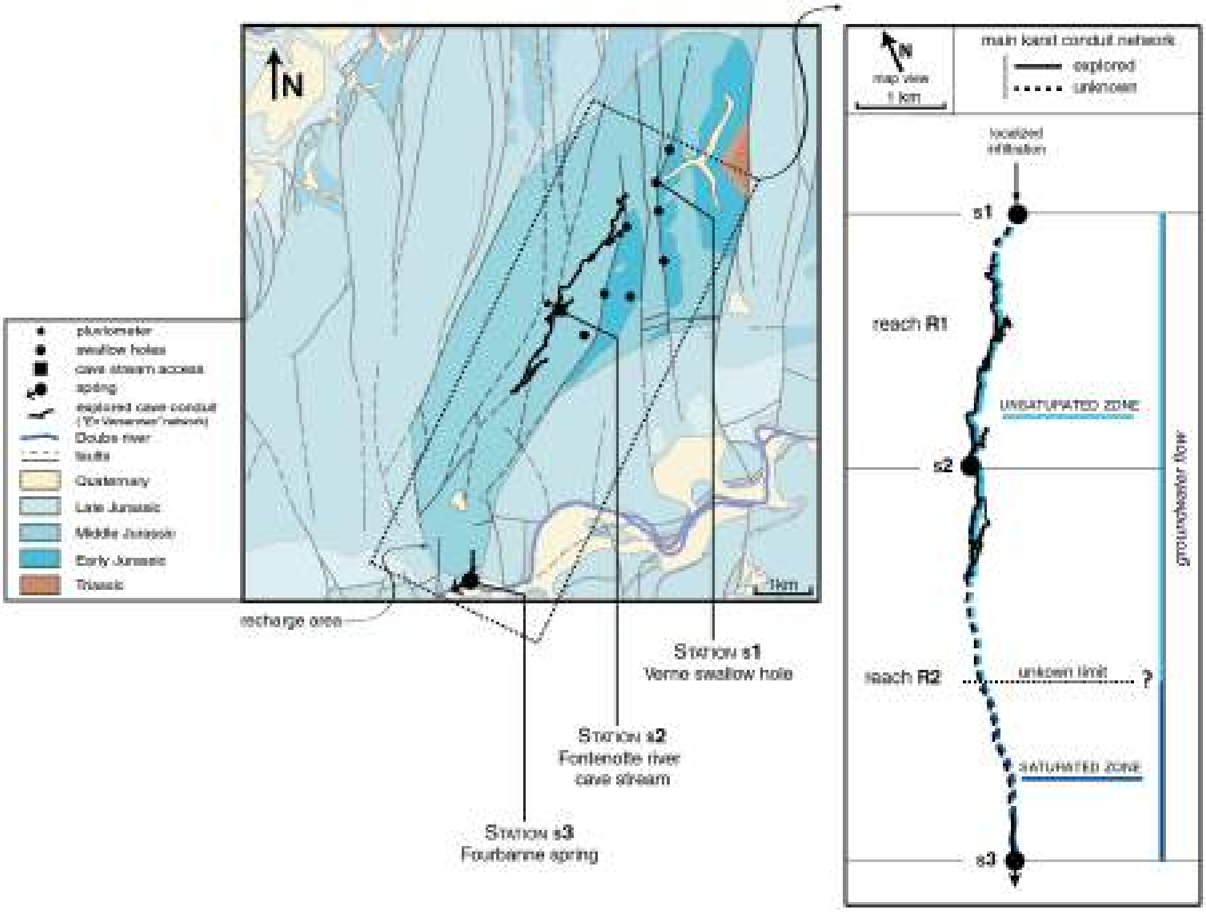


Fig. 2.6 Hydrogeological map of the Fourbanne karst system with the main geological layers, existing faults and swallow holes, and a tropographical trace of the main explored conduit network (from Cholet, 2017).

The presence of marls and limestones on the surface manifests into two distinct aspects of flow: the marly impermeable layer, contributing to surface runoffs, and the karstified limestone layers with their well developed underground conduit network. Eight major swallow holes with sinking streams are located within the basin, governing the punctual recharge of the aquifer. Most of these sinkholes are located at normal fault-controlled shale-limestone contacts. According to tracer tests, the system primarily receives recharge from the En-Versenne (Fontenotte) and Verne swallow holes, which are situated at fault-controlled marl-limestone contacts approximately 9 km North-East of the Fourbanne spring. The main underground conduit of the system is accessible to speleologists via 3 access points: a natural entrance in Luxiol (The En-Versenne cave), an artificial entrance through a 20 m vertical well drilled at Fontenotte (Fig. 2.7), and an underwater entrance at the outlet at Fourbanne (Fig. 2.8).

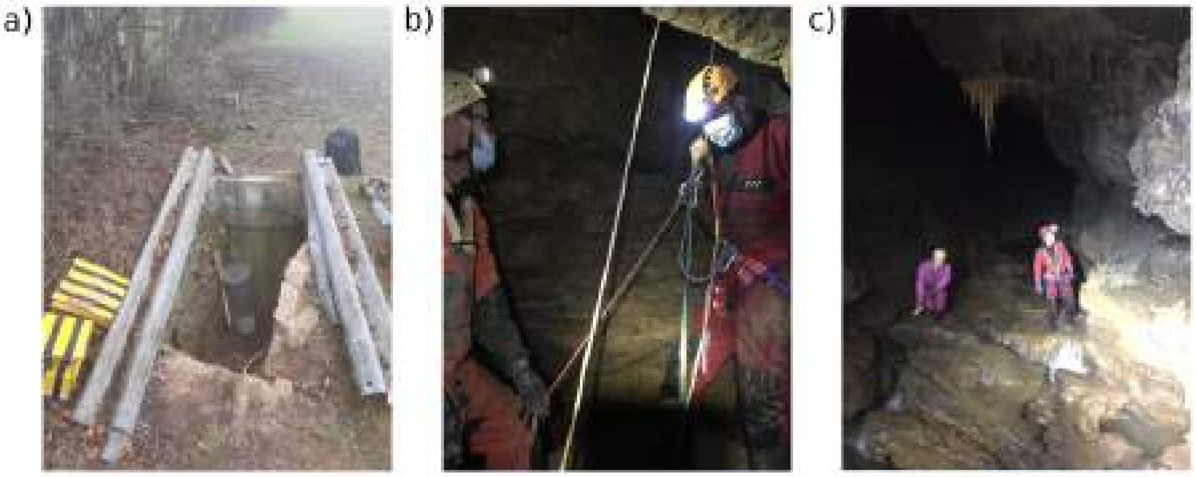


Fig. 2.7 Photos of the Fontenotte area. a) Surface entrance for the artificial well giving access to the underground conduit. b) Underground access through abseiling. c) Inside the karstic conduit.

Fig. 2.8 Underwater entrance at the outlet (Fourbanne), (photo from Marlon Mendonça Dias).

2.1.3 Underground Karst Conduit

Speleologists of the ASDC (Association Spéléologique du Doubs Central) have extensively surveyed the unsaturated part of the system, mapping an underground conduit over 9 km. Cave divers have also explored 2 km of the saturated zone, starting from the outlet's entrance. A straight line distance of about 2 km stays unexplored between the the northern and southern sections of the conduit, which marks the transition between unsaturated and saturated parts of the aquifer (Fig. 2.9). This limit, highlighted in Figure 2.6 (right), is not exactly known and varies depending on the piezometric level of the saturated zone, whose position is mainly controlled by normal faults that lower the Dogger limestones to the south. Actually, Cholet (2017) showed that the limit between the saturated and unsaturated zones fluctuates depending on the baseflow condition. Baseflow is the portion of the streamflow that is sustained between precipitation events, during a flood event the total flow will be the summation of the base and flood components. Cholet (2017) showed that the saturated part extends to about 1.7-2.2 km upstream of the outlet.

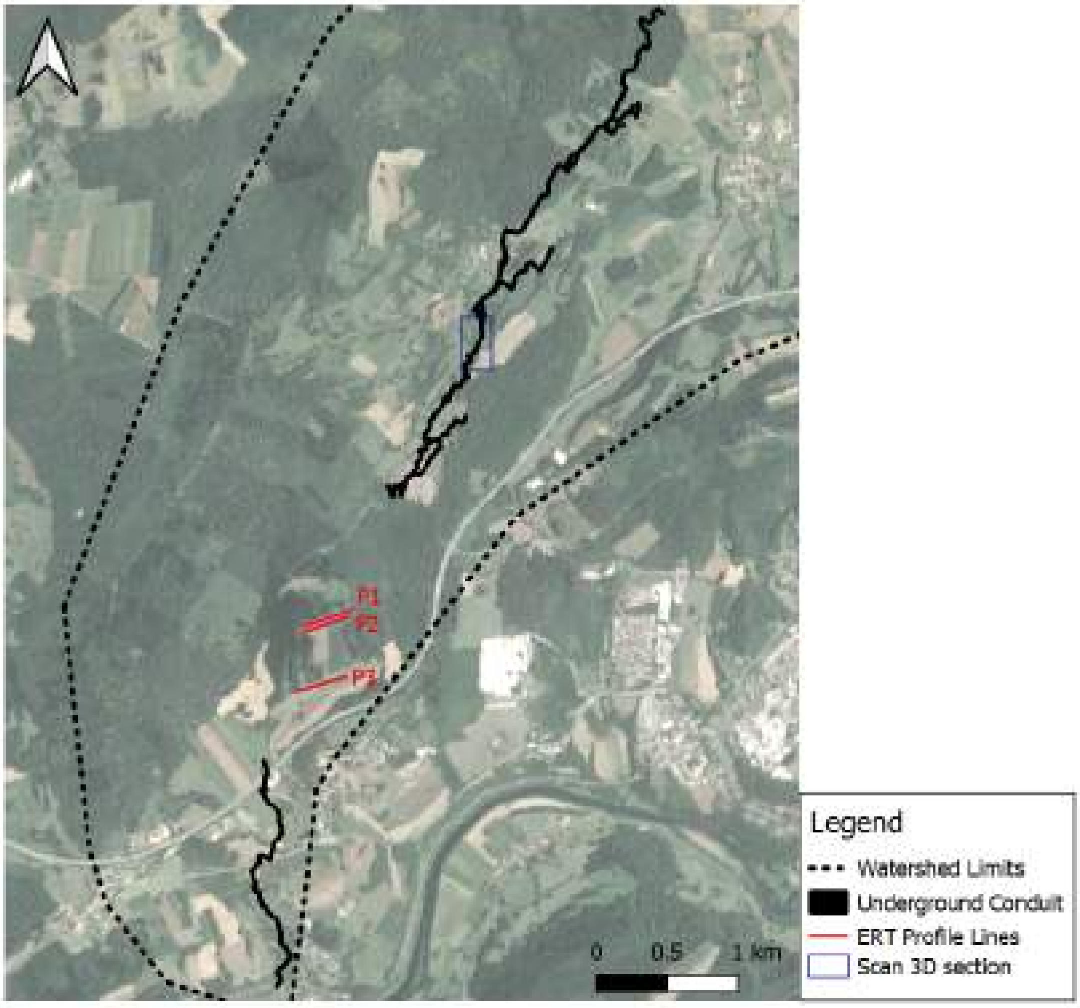


Fig. 2.9 Sketch of the underground conduit, provided by Denis Motte (ASDC). The blue rectangle indicated the portion of the conduit which is instrumented by FO-DTS and 3D scanned by Geotopo. The red lines correspond to the locations of ERT profiles.

Charmoille (2005) showed that the structure of the underground conduit network is governed by two main families of tectonic fractures: the first family with a north-south direction and another with a southwest-northeast direction. The main orientation of the explored network varies between N20°E and N30°E (Fig. 2.10).

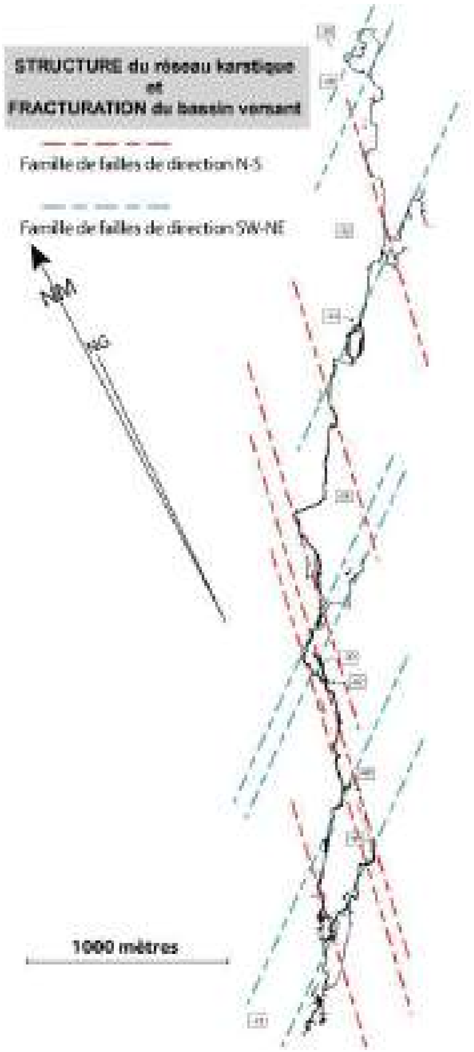


Fig. 2.10 The observed fractures in the Avants-Monts area and the drawing of the unsaturated section of the undeground conduit, from Charmoille (2005)

Three electrical resistivity tomography (ERT) profiles have beeen carried out in Grosbois, south of the watershed. This method consist in injecting electrical currents into the ground using electrodes, measuring resultant potential differences, and enables the mapping of subsurface resistivity distribution. The mathematical equations on which is based the method as well as the principle of the inversion process are developed and explained in Loke (2001). The main objective was to locate fossil gallery that may be used to supply divers with oxygen tanks. The other objective was to investigate the continuity of the underground conduit and eventually better understand its configuration and connections. The 3 profiles were located based on the presence of depressions (or dolines) at the surface, which might be indicators to presence of highly karstified zones or cavities. The factors behind the genesis of these depressions are the

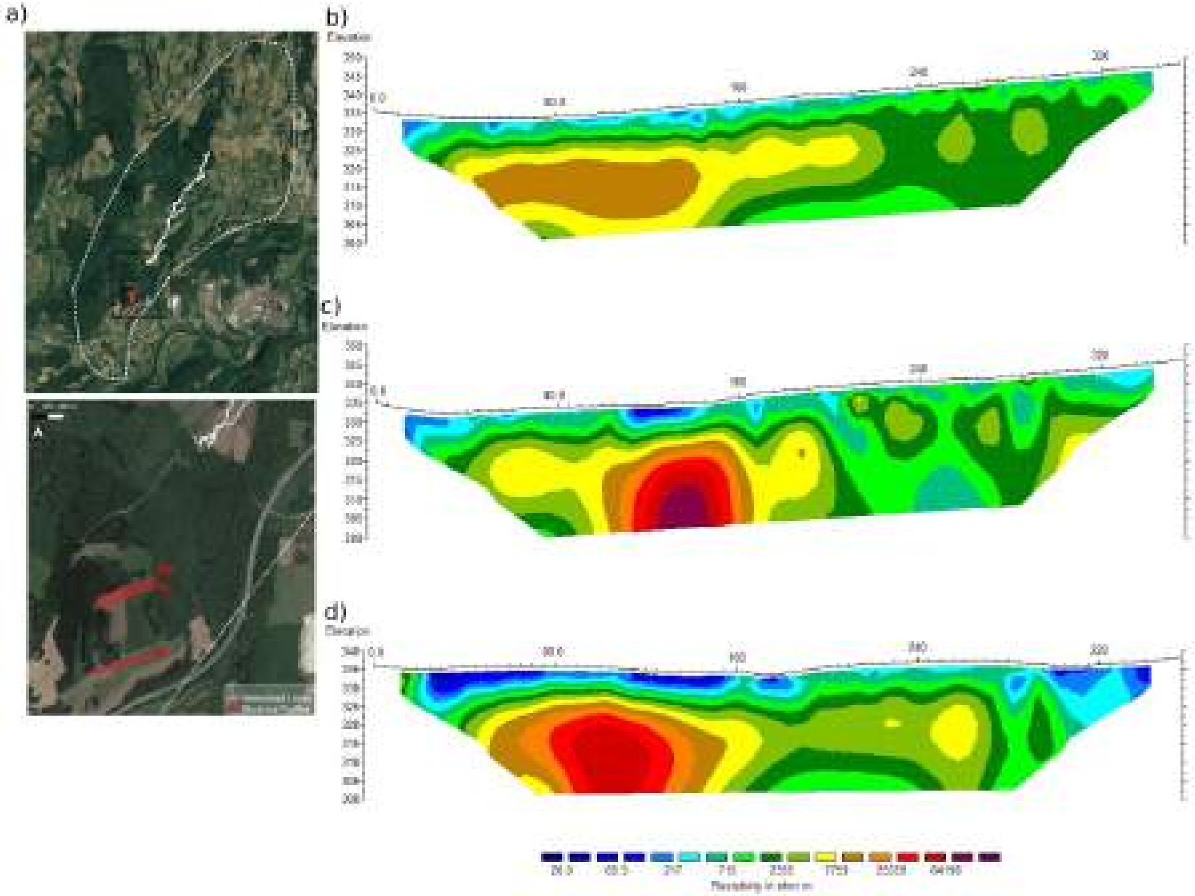


Fig. 2.11 ERT profiles around Grobois. a) Location of the 3 profiles. b) Resistivity sections resulting from the 2D inversion for P1 c) P2 d) and P3 profiles. The elevation is in m.

same as those responsible for the karstification of carbonate formations: Water seeps in through cracks and fissures in the rock and then, through dissolution the cracks will become wider and the soil at the surface settles forming depressions. Each of the lines has a length of 355 m, constructed by deploying each time 72 electrodes with 5 m spacing. The adapted configuration is the Wenner-Schlumberger configuration: it involves injecting a small electrical current into the ground through two outer electrodes and measuring the resulting potential difference between two inner electrodes. Data processing was done with Res2Dinv software (Loke, 2012). The lines positions as well as the obtained resistivity sections are illustrated in Figure. 2.11.

The electrical resistivity values for karst formations can vary widely depending on factors such as the composition of the rock, clay content, the presence of water, and the degree of karstification, values might range between 20 and 6000 Ω.m for example (Cheng et al., 2019). Generally, karst formations are characterized by high electrical conductivity (thus low electrical resistivity) due to the presence of water-filled voids and fractures. However, karst fillings and sediments can have a resistivity that ranges from around 100 to several thousand Ω.m, depending on the composition and water content (Muchaidze, 2008). Resistivity profiles reveal a relatively high-resistivity zone in each profiles (in particular P2 and P3), with values raging between 20 000 and 100 000 Ω.m which may be interpreted as air-filled cavities. Low resistivity zones above them may indicate the presence of fractures and joints connected to the air-filled zone. The latter indicates that the underground conduit in Fontenotte potentially continues its path in the downstream section of the watershed and passes through Grosbois, which might generate fast water transfer mechanisms with the upstream part of the watershed and the superficial part of the karst.

2.1.4 Hydrogeological Response to Precipitations

Three hydrological monitoring stations were deployed since 2013 to measure continuously the level and the temperature of water: one station was installed at the sinkhole in Verne (Fig. 2.12) at the entrance of the aquifer, a second station in the underground conduit accessible through the drilled shaft in Fontenotte and a third station at the outlet at Fourbanne (s1, s2 and s3 in Fig. 2.6). The probes at Fontenotte and Fourbanne also measure electrical conductivity and turbidity. In addition, a meteorological station (Campbell BW200) was installed at Fontenotte to measure rainfall, temperature and wind speed (Fig. 2.13). The data from these probes were used by Cholet (2017) to study and model the hydrogeological behavior of the watershed.

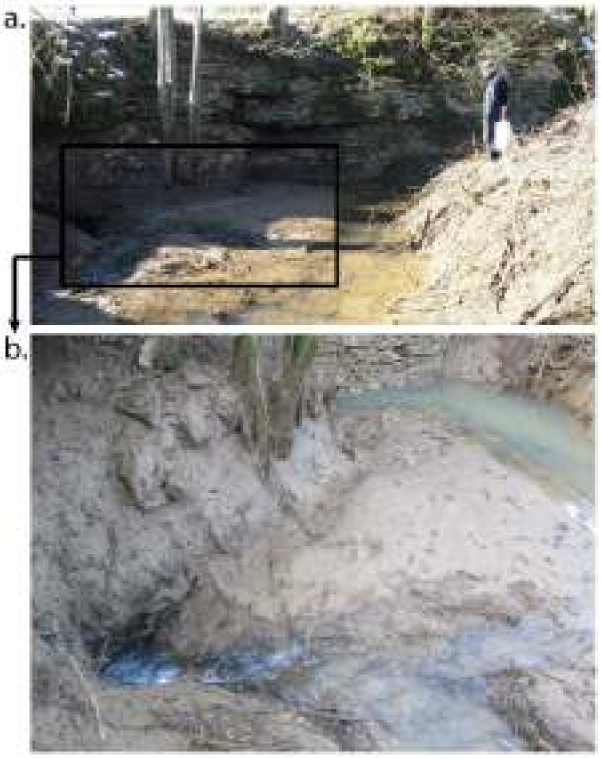


Fig. 2.12 Verne sinkhole during dry season (from Cholet, 2017).

The aquifer's response to precipitation can be understood through the analysis of hydrogeological data (Fig. 2.14). For instance, flood and low-water periods can be detected through the monitoring of the water level in the underground network. The temperature variation of the river depends on the precipitation temperature: during cold periods, the precipitations are colder than the underground river, leading to a decrease in the temperature of the latter. Hence, the opposite will happen during hot weathers (Fig. 2.14). In the same way, EC increases by concentration effect during low flow period while it generally decreases by dilution effect during high flow periods.

Fig. 2.13 Meteorological station deployed at Fontenotte (SNO Karst).

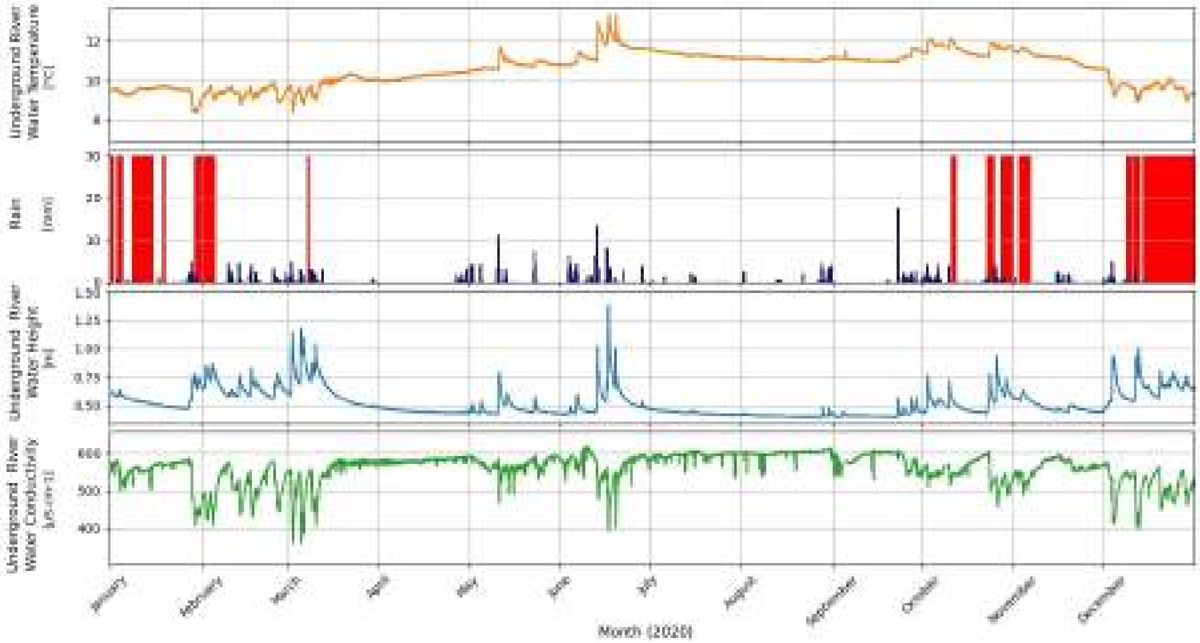


Fig. 2.14 Hydrological data collected from the CTD deployed in the underground river and measuring water temperature, height and electrical conductivity, and from the Campbell measuring precipitations, for the year 2020. The red bars indicate periods where the Campbell battery was off.

Cholet (2017) built a hydrogeological functionning scheme of the Fourbanne karst aquifer showing lateral exchanges between the main conduits and the neighbouring compartments for 3 hydrogeological conditions (Fig. 2.15): a major precipitation happening during a high flow period (Fig. 2.15a), a minor precipitation happening during a low flow period (Fig. 2.15b), and finally a major precipitation happening during an extremely dry period (Fig. 2.15c).

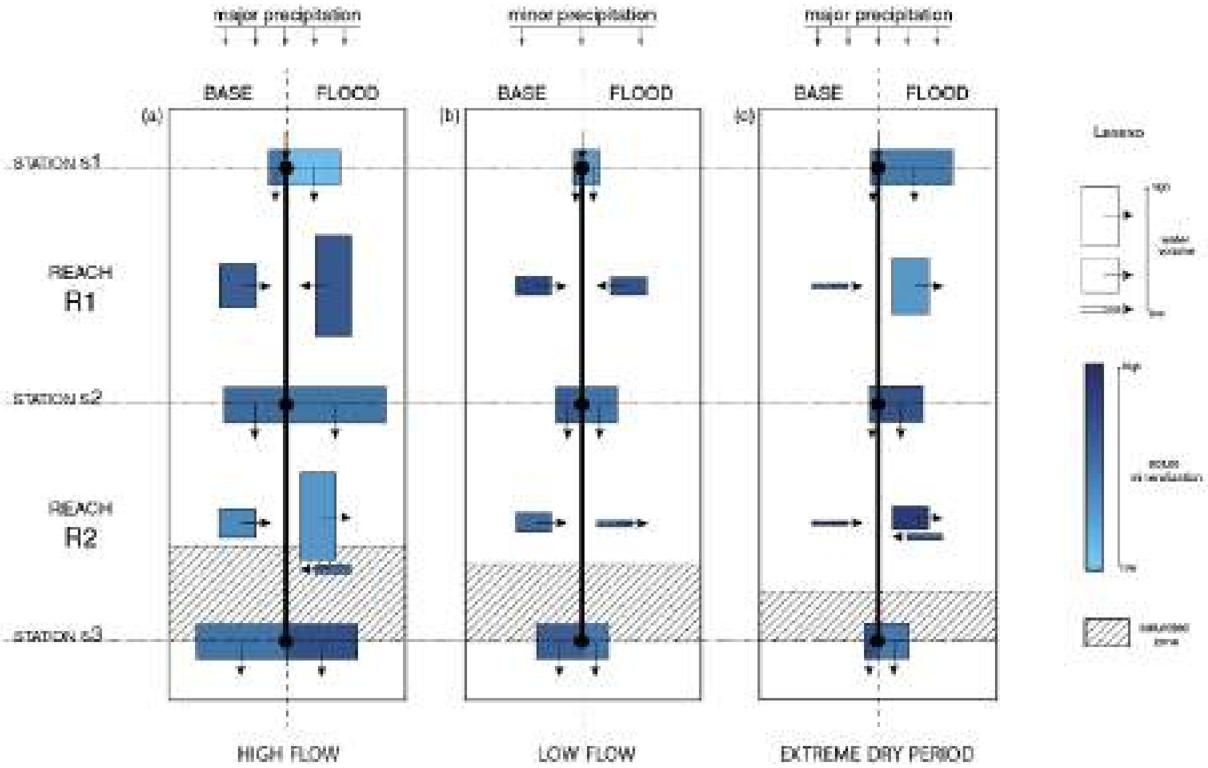


Fig. 2.15 Hydrogeological functioning scheme of the Fourbanne karst aquifer, showing the contributions of lateral exchanges in terms of volume and mineralization for three hydrological conditions, at the level of the three stations and between stations, for both base and flood components, from Cholet (2017).

The conduit is divided into two reaches: the first one between station s1 (Verne) and s2 (Fontenotte) and the second between s2 and s3 (Fourbanne) (see Fig. 2.6). Reach 1 is entirely located in the unsaturated zone, whereas reach 2 covers the lower part of the unsaturated zone and the entire saturated zone. The functional scheme describes lateral exchanges (in- and out-flows) during base flow and flood flow conditions, in terms of water volume and mineralization. A highly mineralized water indicates that it has been stored longtime in the aquifer. The figure also shows the moving limit between the unsaturated and the the saturated zone, depending whether we are in a high-flow (limit moving upstream) or low-flow period (limit moving downstream). (1) During high-flow periods, lateral exchanges are important for both base and flood components. For the baseflow component both reaches are fed by strongly mineralized lateral inflows. However, for the flood component different results are obtained for each reach, with strongly mineralized inflows for reach 1 and outflows for reach 2. (2) For minor precipitation during low-flow periods, reach 1 reveals similarly mineralized inflows for the base and the flood component. In contrast, different behaviors for the components are obtained for reach 2, with weak inflows for the base and very weak outflows for the flood component. (3) For major precipitation during extremely dry periods, reach 1 presents very low base inflows and important flood outflows, indicating high losses towards the rock matrix in the unsaturated zone.

On the contrary, reach 2 presents very low baseflow inputs of highly mineralized water. The flood component is mainly characterized by lateral outflows followed by very weak inflows. These exchanges during the flood event may be related to conduit-matrix interactions in the saturated zone. The observed constant inflows all along the reaches mainly come from diffuse infiltration for reach 1 and from lateral drainage systems for reach 2. In addition, reach 1 collects mineralized inflows coming from secondary tributaries, related to the presence of several sinkholes in this part of the aquifer. All of the latter shows the non linearity in the hydrogeological behavior of the karst aquifer in response to rainfall events, indicating that the lateral exchanges are not only depending on the amount of precipitation, but also on the initial state of the rock matrix (dry or wet).

Charmoille (2005) studied a rain event and compared water height measured at s1 (Verne), s2 (Fontenotte), and s3 (Fourbanne), between the 24^{th} and 28^{th} of October 2002. The author studied the different time lapses obtained between the rain event peak and the positive trend in the observed water heights, in other terms the time lapse between the maximum rainfall and the start of the water height increase. This is called response time of an aquifer, referring to the time it takes for changes in hydraulic conditions, such as changes in water heights, to propagate through the aquifer and result in observable effects at a particular location. The obtained results are shown in Fig. 2.16, in which we can see time series of water heights measured at Verne (swallow hole), Fontenotte (underground river) and Fourbanne (outlet), as well as the rain rate in bars during the same period. The obtained response time is about 7h 30min for Verne, about 10h for Fontenotte, and about 13h for Fourbanne. These values might change depending on, as mentioned above, the initial state of the aquifer and the intensity of the rainfall, but give us an order of magnitude of the aquifer response to a rainfall event.

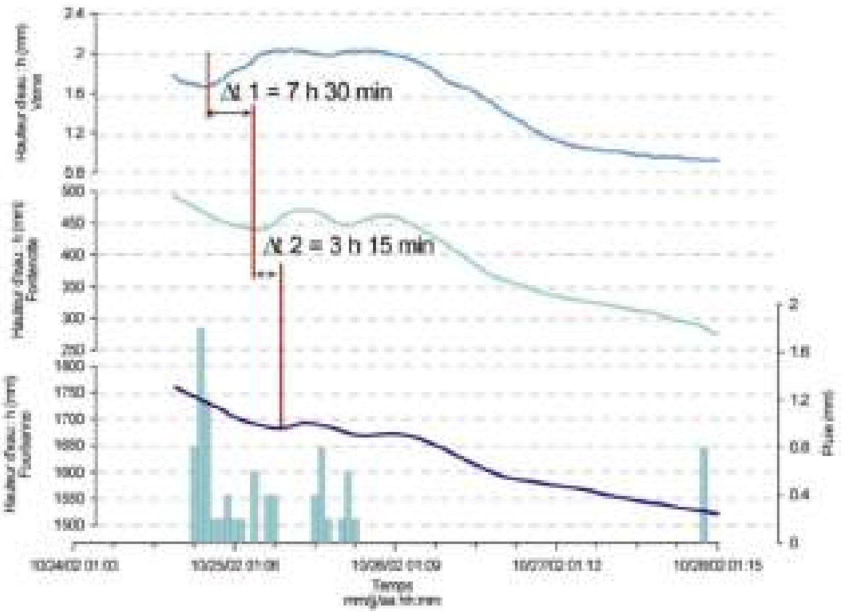
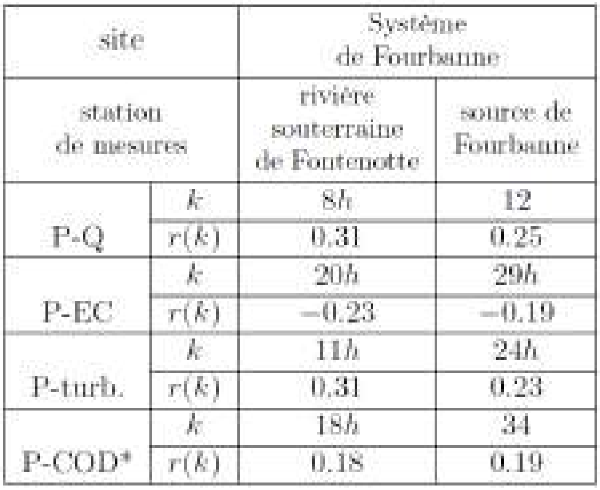


Fig. 2.16 Water height variations at Verne, Fontenotte, and Fourbanne after a rainfall event, between the 24^{th} and 28^{th} of October 2002, as well as time response variation between the 3 locations, from Charmoille (2005).

Cholet (2017) used dye tracing experiments to characterise transport mechanisms along the karst conduit. The detailed analysis of these experiments provides relevant information, particularly for understanding the interactions between transfer pathways and storage phenomena within the network, as well as the transit time lapse (Goldscheider et al., 2008). Injections were done in the swallow hole at Verne and surveyed both in the underground river at Fontenotte (unsaturated zone) and at the Fourbanne system outlet (saturated zone). Transfer time (e.g. time lapse between the first detection of the tracer at two stations) and velocity were computed for the reach 1 (between Verne and Fontenotte) and reach 2 (between Fontenotte and Fourbanne), as well as the restitution rate of the injected dye mass. The restitution rates are lower at the outlet (averaging 62% for all the tests) than at the underground river (averaging 67%), which seems to confirm the loss (or storage) phenomena in the saturated zone of the system. In fact, the saturated zone is a reservoir with a long residence time, thus leading to a strong dilution of the tracer. Tracer can also be absorbed by the rock matrix resulting potential loss. Moreover, it is also possible that the Fourbanne spring is not the unique outlet of the system, additional outlets may exist, covered under the alluvial plain of the Doubs valley. The transit times are dependent on the average flows measured at the Fourbanne source, with a shorter transit time for higher flow levels for both reaches. Beyond, the transit times between both reaches are positively correlated, with higher average (37h) for reach 2 compared to reach 1 (16h). Consequently, higher transit velocities are obtained for reach 1 (average of 0.006 m/s for all the tests) compared to reach 2 (with an average of 0.0047 m/s). This contrast can be related to a higher hydraulic gradient for reach 1 (unsaturated zone) compared to reach 2 (saturated zone). However, all of these observations are trends and the values obtained for each test are different depending on the initial conditions of the aquifer and the rainfall conditions.

Cholet et al. (2017) also studied the reaction times of conductivity, temperature, turbidity, and DOC (dissolved organic carbon) after rainfall events for the underground river at Fontenotte and the Fourbanne spring. They held cross-correlation analysis covering a period of 2.5 months with multiple rain events. The distributions of the later components are highly scattered between different events; however, a slightly lower dispersion is observed at the outlet, suggesting a buffering effect associated with the saturated zone. Correlation between rain and the previous components decreases along the network, illustrating the dispersion of the flood wave within the network, revealing a relatively non-inertial system. The authors also computed correlograms between rain (marked as P) and the different components: flow (Q), electrical conductivity (EC), turbidity (turb), and DOC. The maximum coefficients of correlation (r(k)) as well as the time lapse (k) giving the maximum correlation are given in Fig. 2.17. The hydraulic signal and the chemical signal at the outlet are desynchronized during a flood; the water height and rain correlation reaches a peak much before the correlation between rain and the other factors (EC, turb, and DOC). In fact, an increase in water level is immediately transferred to the source, corresponding mainly to water present in the aquifer since a previous recharge episode, marking the coexistence of water with short and long residence times within the aquifer.



site		Système de Fourbanne	
station de mesures		rivière souterraine de Fontenotte	source de Fourbanne
P-Q	k	$8h$	12
	$r(k)$	0.31	0.25
P-EC	k	$20h$	$29h$
	$r(k)$	−0.23	−0.19
P-turb.	k	$11h$	$24h$
	$r(k)$	0.31	0.23
P-COD*	k	$18h$	34
	$r(k)$	0.18	0.19

Fig. 2.17 The values of k and r(k) from the hourly cross-correlation analyses. The selected k values correspond to the maximum r(k) for the precipitation-discharge, precipitation-DOC (COD in the figure), and precipitation-turbidity cross-correlations, as well as the minimum r(k) for the precipitation-EC cross-correlation, from Cholet (2017).

2.1.5 Soil Water Reserve

Soils constitute the interface between atmospheric processes (rainfall, evapotranspiration) and aquifers. They serve as water storage and play a role in the water balance at the catchment scale. We can distinguish water runoff at the soil surface and water infiltration. The water infiltrating the soil can either reach the underground aquifer by gravity, e.g. flowing through the soil porosity without being stored, or stay stored within the soil because of water retention, e.g. capillarity and adsorption water (Vauthier, 2011). Capillary water consists of two parts: an unusable portion below the wilting point, strongly bound to the soil proprieties, and the available water reserve. This water is held in the macropores and depends on the organic matter content, mineralogy, texture, and gravel content. It is generally categorized into two types of water: capillary water absorbable and non-absorbable by the root system of plants (Baize et al., 1995). This reserve represents the amount of water stored in the soil that can be easily absorbed by plants (Antoni et al., 2011) and is expressed in the height of the water content (mm) in a soil column. It is called the "available water capacity" or "available water content" (AWC).

The following work has been done by two master's students (Yaël Henriet and Erwan Dore) in Water Sciences at the University of Bourgogne-Franche-Comté, in the framework of their master's thesis (Dore, 2022; Henriet, 2022). In order to quantify the spatial distribution of the AWC in the watershed, the latter has been divided into 271 mapping units (MU), which were essentially delineated from topography, geology, and soil cover. Each mapping unit was sampled at least once with a manual soil auger. Given the size of the watershed and the time available for field work, a MU covered around 15 hectares. Each soil auger sampling descended to about 120 cm depth to obtain the parameters necessary for the calculation of the AWC. For each sampling the texture and the percentage of coarse elements was determined for each soil horizon. Texture was estimated using a tactile test, and the percentage of coarse elements were determined

visually by comparison with a reference grid. The soil auger is unable to reach the bedrock for deep soils (more than 120 centimeters), nor does it consider the presence of soil in rock fractures, especially in limestone. Therefore, the AWC values will be associated with minimum values at the sampled soil level. Average depths for soil on limestone are found to be around 50 cm, and on limestone above 100 cm. The water reserve is then computed based on the method presented in Jamagne et al. (1977), taking into account the textures, depth, organic matter content, and coarse element content. The water reserve is calculated using the following formula:
AWC = E x C x (1- CE/100)
With AWC representing the available reserve (mm), E the layer thickness (cm), C the textural coefficient (mm/cm), and CE the percentage of coarse elements in the layer (Ridremont et al., 2011). Texture coefficients were adapted from Jamagne et al. (1977) for the Jura region, where the organic matter content is higher. The resulting water reserve map is shown in Fig. 2.18. The soil thicknesses varies from 0-40 cm to more than 1m. Thin soils are primarily found on limestone slopes. Soils sampled in sinkhole or marls areas are generally thicker. AWC values show strong variations which can be explained by the significant heterogeneity of geologies and topographies at the watershed scale. However, we can deduce a correlation between the AWC and geological formations: high reserves are obtained for marl outcrops and low reserves for limestone, which is coherent with the higher capacity of marl soil in retaining water and the ability of limestone in draining it downstream. In fact, soils on marls have a more fine-grained texture inherited from the marls and are thicker. Soils on limestone are thinner and coarser. The results obtained in agricultural settings present higher ACW than those from forest soils, which are often thinner (Dore, 2022; Henriet, 2022). This demonstrates that land use is tailored to its nature, aiming to optimize its economic exploitation. This water reserve map will be used afterwards in the interpretation of seismic velocity variation estimated on the watershed (see Chap. 4).

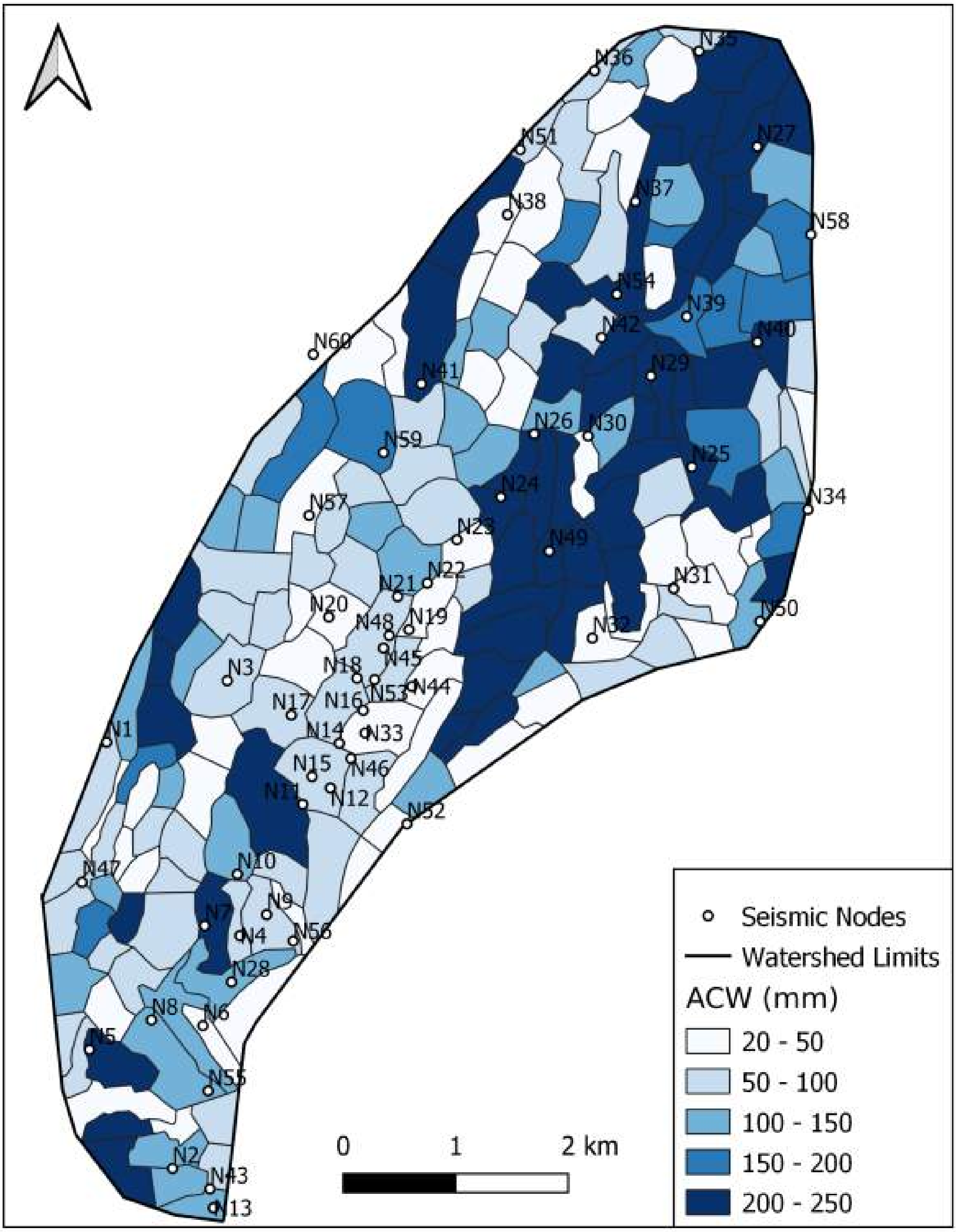


Fig. 2.18 Available water capacity (ACW) map for the Fourbanne watershed, the color scale represents the water volume in mm. The dots indicate the position of the seismic nodes.

2.2 A hydrogeophysical Observatory

2.2.1 Jurrassic Karst Observatory: Long Term Hydrogeological Monitoring

The Jurassic Karst observatory was set up to better understand the degradation of the water quality of the karst rivers in the Jura Mountains over the past 30 years. The primary objective of SNO KARST is to advance scientific knowledge in the functioning of karstic reservoirs. Firstly, it seeks to develop and enhance models that simulate hydrogeochemical flows at the outlets of aquifers. Secondly, it aims to deepen our understanding of the relationships between global and more local (the physicochemical properties of waters) changes during floods (Jourde et al., 2018). The Jurassic Karst Observatory is at the local scale jointly managed by the Chrono-environnement laboratory and the Observatory of Universe Sciences OSU THETA. As mentioned above, the Fourbanne Watershed is one of the sites of the Jurassic Karst observatory and is continuously monitored since 2013.

2.2.2 Juraquake: Long term seismic monitoring

JURAQUAKE is a long term seismological regional network consisting in 7 stations located in eastern France (Fig. 2.19), north of Besançon, and installed in late 2018. This network was established as part of a research project aiming to characterize the micro-seismic activity of the area, and study the potential role of fluid in triggering this activity. The figure illustrates the JURAQUAKE array as well as the recorded earthquakes and quarry blasts through Renass (RESIF - Réseau sismologique et géodesique Français) between 1980 and 2018. Two of the deployed seismometers are located in the Fourbanne watershed: AVEN and FONT. AVEN (Guralp CMG40T 60s-100Hz sensor, connected to a Staneo D3BB-MOB digitizer, 200 Hz sampling frequency) is located in the underground gallery (423 m altitude) at 20 m depth below surface. FONT (Guralp 6TD, 30s-100Hz sensor, 100 Hz sampling frequency) is located at the surface (443 m altitude), at 3 m from the wellhead. For coupling purposes, both sensors are buried 50 cm into the ground. AVEN and FONT are at a slope distance of about 50 m and 60 m from the underground river's channel (Fig. 3.12). These stations are used to continuously monitor the underground river.

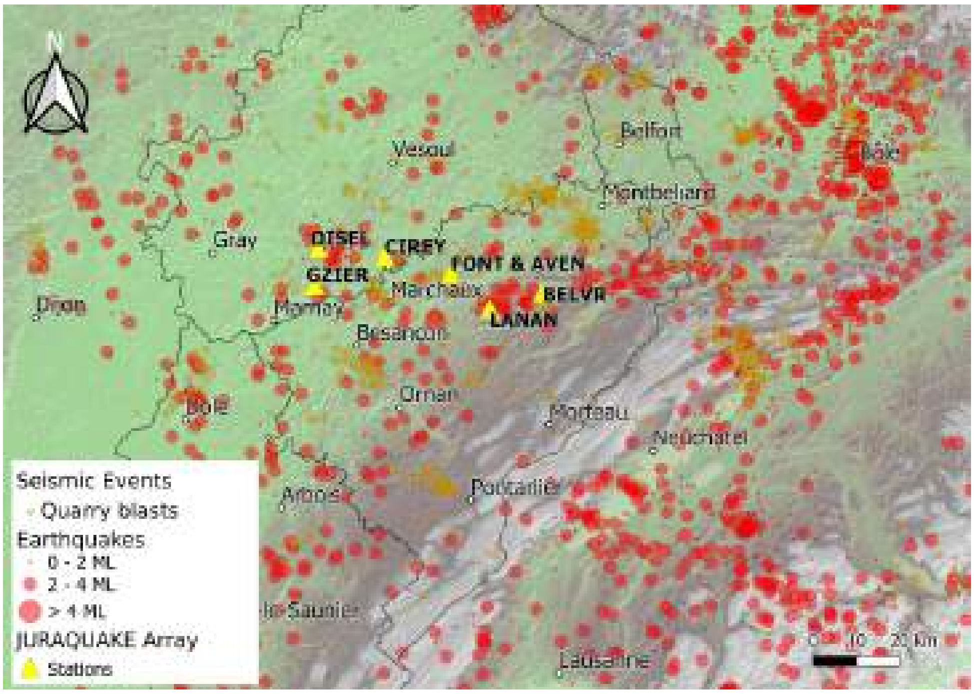


Fig. 2.19 Seismic activity and quarry blasts in the Jura area between 1980 and 2018 (from the RéNaSS catalogue provided by the national network Résif). The yellow triangles indicate the positions of the JURAQUAKE network stations.

2.2.3 FibroKarst: Temperature measurements along underground conduit with fiber optics

The FibroKarst project focuses on understanding karstic systems, particularly addressing the challenge of comprehending recharge and transfer processes within these highly heterogeneous and typically inaccessible environments. It is a collaborative project between Chrono-environnement and EOST/ITES. It consists in employing optical fiber technology for identifying recharge points within a karstic aquifer by continuously measuring water temperature along the conduit, through a method called FO-DTS, fiber optic distributed temperature sensing. In the framework of this project we installed a 1500-meter long optical fiber, of which 800 meters are used for measurements within the underground river, inside the karstic conduit of Fontenotte (Fig. 2.20 and 2.21). Afterwards we did a topographical survey of the deployed fiber. We have also installed 2 temperature probes along the conduit in order to calibrate the temperatures measured by the fiber. Finally, we conducted a high-resolution 3D scan of the cavity with the help of Geotopo, engineering companay specialized in topographical studies, using an instrument called Zeb-Horizon. The objective of this 3D mapping is to be coupled with the results of the optical fiber, to refine the structures controlling the transfers between the matrix and conduit (see section 3.2 of Chapter 3).

Fig. 2.20 Deployment of the optical fiber cable in the underground conduit.

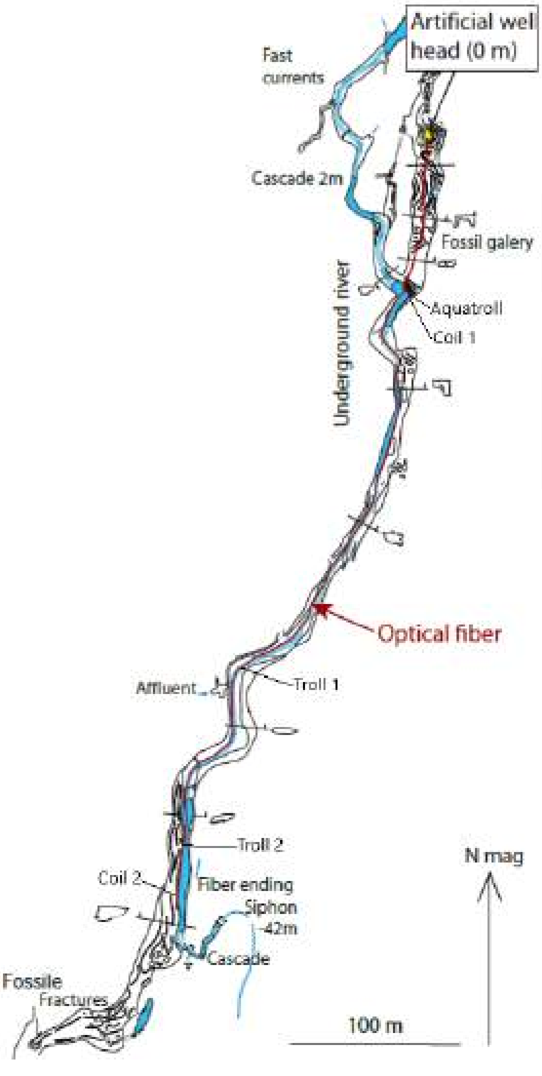


Fig. 2.21 Topographical map of the underground conduit and the instrumentation used in this study. The deployed optical fiber is marked with a red line. Map from by D. Motte (ASDC)

.

2.2.4 Sismeauclim: temporary seismic monitoring using a dense array

2.2.4.1 Presentation

The Sismeauclim project is a Bourgogne Franche-Comté region-funded project (between 2020 and 2023) aiming to combine the fields of seismology, hydrogeology and climatology for the study of karstic reservoirs during floods. It was held in the framework of a collaboration between Chrono-environnement and Biogeosciences laboratories. This project involves locating and quantifying the water transfers, starting from their release as rainfall, up to their conveyance towards the basin's outlet. For this purpose, temporary observation networks of 60 seismometers and rain gauges have been installed for an integrated monitoring of the watershed. Cross-correlation methods have been applied on seismic noise recordings to estimate seismic velocity variations in the medium. The spatial mapping of seismic velocity variations over time may be interpreted in terms of water flow, and soil/matrix saturation, relying on measurements of precipitations falling over the basin, measured through rain gauges, and circulating within the karstic network.

2.2.4.2 The Nodes Network

Prior to the deployment of the network, survey field trips were required to locate the potential positions of each of the instruments. Once located, we went through contacting the mayors of the cities and the owners of the fields to get permissions to install the instruments. From the end of November 2021 til the end of March 2022, 60 seismic nodes were deployed all over the watershed. The positions of the nodes were chosen in order to cover all of the studied area with its different aspects, in other words the zones with different outcrops (limestone and marl) and different altitudes, with a tight mesh around the conduit. The nodes (Fairfiel ZLand 3C) belong to the French national pool of portable seismic instruments SisMob-RESIF (INSU-CNRS). They are compact devices that combine a 3-channel sensor, a datalogger and an internal battery (power supply for 35 days span) in one unit. They include built-in GPS timing for precise synchronization. These sensors record seismic noise at a sampling frequency of 250 Hz, with a gain set at 36 dB for noise amplification.

The deployment of seismic network demanded tool and instrument preparation and field organisation. This was also implied on the maintenance of the instruments. Monthly maintenance is required for the nodes: they shall be uninstalled to recharge the battery and save the data and then deployed again. The latter necessitates three days work: one day for the instruments withdrawal from the field, one day to charge the nodes and retrieve the data and one day to deploy them again (Fig. 2.22). The fieldwork, whether it was for the withdrawal or deployment of the nodes, was done by 3 groups of two persons, combining fellow researchers, engineers, technicians and PhD students. Once the data was collected, it shall be preprocessed and sent to the SisMob-RESIF engineers (Isterre, Grenoble) to be converted to miniseed format and then sent back to us. Accurate GPS locations were measured for all the instruments via GPS TRIMBLE Geo7X, borrowed from MSHE (Maison des Sciences de l'Homme et de l'Environnement) Besançon. The GPS data were processed by Amandine Angeli and Matthieu Thivet (Chrono-environnement/MSHE). A datasheet with the nodes' coordinates and other characteristics concerning their position is presented in the table A.1 in Appendix A.

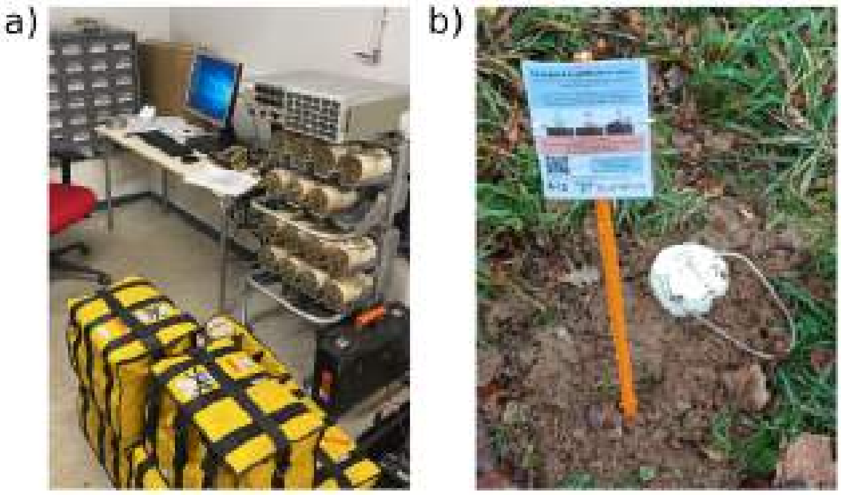


Fig. 2.22 Nodes' maintenance. a) Charging of the nodes' batteries and data collect. b) Node installed on the field, with an informative notice.

2.2.4.3 The Rain Gauges Network

Twenty tipping-bucket rain gauges (Rainwise) were also deployed for a period of a year (from June 2021 to June 2022) to study the spatio-temporal variation of rainfall on the watershed scale (Fig. 2.23). Survey filed trips were also required to locate the potential locations of the rain gauges, and then the concerned land lords were contacted for installation permissions. The rain gauges were installed on 2 m long wooden stakes, far from any shelter or wall, and in the most open space possible. The presence of nearby trees is detrimental to the proper functioning of the rain gauge, as their branches can deflect precipitation. Usually a 30 degree angle circular check is done from the top of the rain gauge. A horizontal plane level should be held for the instrument for better results.

The functioning of these rain gauges is quite simple, every time the bucket tips, a count is transmitted to the data logger and the gauge empties. The data contains the number of tips recorded each 30 minutes, corresponding to 2 ml of water, and knowing this we can find the daily rain rate in mm. However, this type of rain gauge is quite sensitive to debris that can obstruct the tipping of the bucket and thus significantly distort the measurement (leaves and debris accumulating in the collector, insects nesting in the rain gauge, etc.). Two-times per month maintenance is required to obtain high-quality data. This maintenance was usually done by group of two, requiring to clean the instruments, check up on their batteries, change them if necessary and collect the data. The data then should be pre-processed and cleaned, removing field intervention time that correspond to false tip counts.

Fig. 2.23 Maintenance and deployment of a rain gauge.

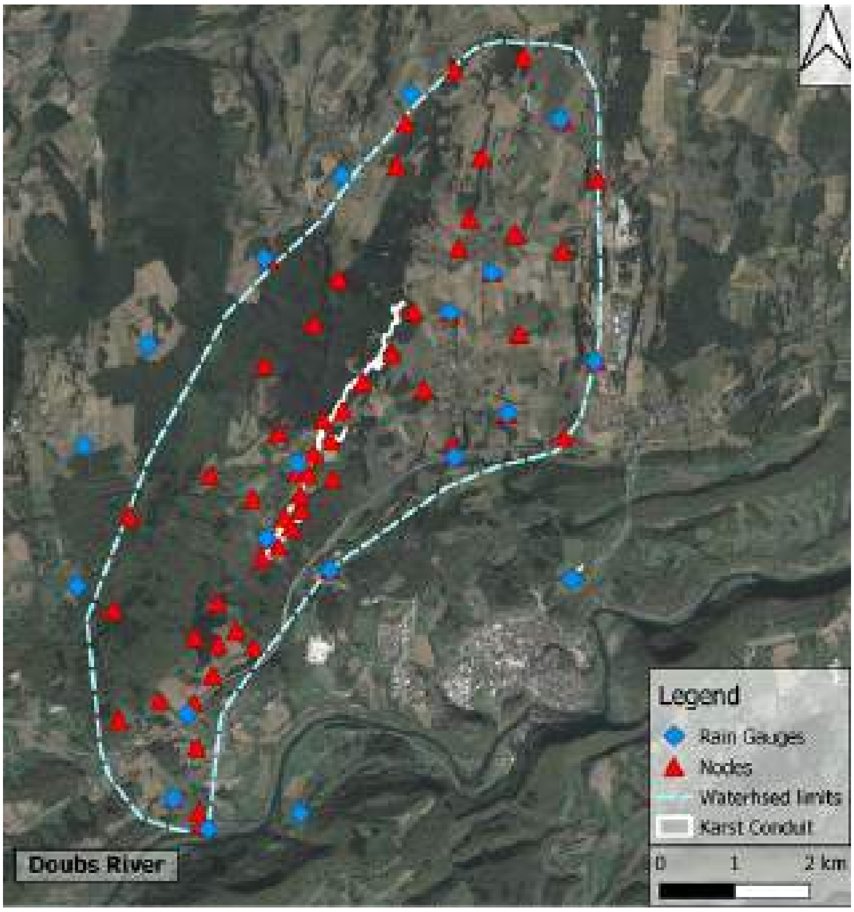


Fig. 2.24 Positions of the seismometers and rain gauges installed for the temporary network (sismeauclim project).

Chapter 3

Remote Monitoring of the Underground River's Hydrodynamics

Keypoints:

- Machine learning based method applied on seismic noise was used to predict underground river water height
- Fiber optic cable was used to monitor temperature changes along underground river.
- Lateral recharge could be detected along the river.

3.1 Machine Learning Prediction of Groundwater Heights from Passive Seismic Wavefield

This section takes up the work published in the journal Geophysical Journal International (GJI) under: Abi Nader, Anthony and Albaric, Julie and Steinmann, Marc and Hibert, Clément and Malet, J.-P and Sue, Christian and Fores, Benjamin and Marchand, A and Gros, M and Celle-Jeanton, Hélène and Pohl, Benjamin and Stefani, V and Boetsch, Anne. (2023). Machine learning prediction of ground water heights from passive seismic wavefield. 234. 1807-1818. 10.1093/gji/ggad160.

3.1.1 Summary

Most of water reservoirs are underground and therefore challenging to monitor. This is particularly the case of karst aquifers which knowledge is mostly based on sparse spatial and temporal observations. In this study, we propose a new approach, based on a supervised machine learning algorithm, the Random Forests, and continuous seismic noise records, that allows the prediction of the underground river water height. The study site is a karst aquifer in the Jura Mountains (France). An underground river is accessible through an artificial shaft and is instrumented by a hydrological probe. The seismic noise generated by the river is recorded by two broadband seismometers, located underground (20 m depth) and at the surface. The algorithm succeeds in predicting water height thanks to signal energy features. Even weak river-induced noise such as recorded at the surface can be detected and used by the algorithm. Its efficiency, expressed by the Nash-Sutcliffe criterion, is above 95% and 53% for data from the underground and surface seismic stations, respectively.

3.1.2 Introduction

Water resource has become an essential environmental and societal issue due to the intensification of its exploitation and its vulnerability to climate change (Andreo et al., 2006; Drew and Hötzl, 1999; Green et al., 2011). Drinking water supply relies mainly on groundwater aquifers, which are generally not directly discernible nor accessible (Chen et al., 2017; McDonnell, 2017). This applies in particular to karst aquifers, which are very heterogeneous in terms of permeability : they are characterized by fast groundwater flows in open conduits (underground rivers) and slow flows in the micro-fractured rock matrix (Ford and Williams, 2007). Since most karst aquifers are inaccessible, their monitoring often relies on punctual observations from piezometers or on spring hydrographs. In order to better understand these systems, it is therefore essential to develop new monitoring approaches, adapted to their heterogeneous geometry and flow dynamics.

The seismic wavefield has proved to provide information about hydrogeological processes (e.g., Larose et al., 2015). Actually, the hydrodynamics of surface rivers have been the subject of several passive seismic studies. For example, the spectral analysis of the ambient seismic noise induced by river flow has allowed to identify sediment transport and deposition within stream segments (Burtin et al., 2008; Schmandt et al., 2013). It has been shown that seismic recordings from geophones installed on the river bank could be used to estimate river discharge (Anthony et al., 2018), water height and bedload transport (Dietze et al., 2019). Regarding groundwater, seismic interferometry methods are effective in detecting water level within the rock matrix, through the measurement of seismic velocity changes in the subsurface (Fores et al., 2018; Vidal et al., 2021; Voisin et al., 2017). Measuring hydrogeological parameters of underground river, which are generally inaccessible, remains however challenging.

Recent advances in research combining machine learning and seismic monitoring have shown that it is possible to identify automatically the sources of seismological events triggered by various geological processes. Actually, the Random Forest algorithm and curated features have been successful in describing landslide micro-seismicity (Provost et al., 2017; Wenner et al., 2021), differentiating between rockfalls and volcano-tectonic earthquakes (Hibert et al., 2017), detecting debris flow events (Chmiel et al., 2021) and establishing seismic lithofacies classification (Kim et al., 2018). The Random Forest algorithm can also be used to predict, in the machine learning term, continuous values. For example the method was applied on laboratory observations to identify hidden signals that precede earthquakes and predict the time remaining before failure (Rouet-Leduc et al., 2017). It was also applied to predict subsurface porosity and capture its spatial variation in reservoirs based on seismic attributes (Zou et al., 2021).

The objective of this study is to propose an innovative approach based on a continuous application of the Random Forest machine learning algorithm on passive seismic wavefield to provide a remote inference of the water height of the underground river. In other terms, we propose to establish a projection of seismic data in a multi-dimensional feature space extracted using a 15-min-long sliding window to the output of 1-D water height values using the mentioned algorithm. This unprecedented application could be the head start in the investigation of other inaccessible water conduits towards a better groundwater estimation and flood forecasting.

3.1.3 Study site and data

The study site is the Fourbanne karst aquifer in the Jura mountains, eastern France (Figure 3.1). It is part of the JURASSIC KARST hydrogeological observatory settled in 2014 (Cholet et al., 2017) and of the french SNO KARST network (Jourde et al., 2018). The local lithology is characterized by Middle Jurassic tabular limestones and shales cross-cut by a series of N-S and NE-SW normal faults, which control the orientation of the underground conduits. The aquifer is primarily fed by allogenic recharge through sinkholes (Cholet et al., 2015). The underground conduit has been explored and mapped over a length of 9 km by speleologists in the unsaturated zone and by cave divers in the saturated zone. The location of the instruments as well as a part of the karst conduit are detailed in Figure 3.1. The seismological data are recorded by two stations of the long-term regional seismic network JURAQUAKE deployed in eastern France since late 2018 and 2019. The station AVEN is located in a fossil gallery at 20 m depth (423 m asl), at the

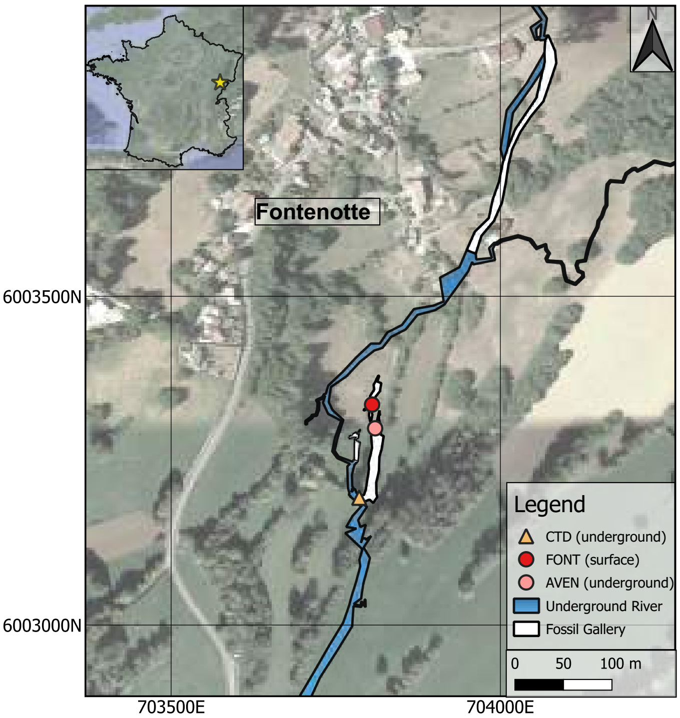


Fig. 3.1 Location map of the site and instrumentation. The underground river and the dry fossil galleries are drawn in blue and white, respectively (topographical data are from D. Motte, ASDC). The red and pink dots show the positions of the seismic stations and the triangle the position of the hydrogeological probe (CTD).

base of a vertical shaft drilled by speleologists (Guralp CMG40T 60s-100Hz sensor, connected to a Staneo D3BB-MOB digitizer). The second station FONT is located at the surface (443 m asl), at 3 m from the wellhead (Guralp 6TD, 30s-100Hz sensor). For coupling purposes both seismometers are dug 50 cm into the cave sediment or surface soil. AVEN and FONT are at a slope distance of about 50 m and 60 m from the underground river's channel. The sampling frequency of these three component seismic stations is 200 Hz for AVEN and 100 Hz for FONT. A hydrological probe (CTD) is installed in the river and records water electrical conductivity, water temperature and water height every 5 minutes.

In this study, we focus on hydrogeological data recorded for two years between September 15^{th} 2019 and September 15^{th} 2021 (Figure 3.2). This period covers two entire hydrogeological cycles, with main rainy seasons in winter and spring. During this period, the CTD recorded a minimal water height of 0.4 m during low water periods, which is measured from the streambed to the water-air interface, and a maximum height of 1.7 m during floods. Seismological data are complete during this period of time at AVEN only. Due to technical problems, there are gaps in the data recorded at FONT and the analysis covers a shorter period of time: between October 27^{th} and December 31^{st} 2019 and between September 15^{th} 2020 and September 15^{th} 2021. Figure 3.2a is a plot of the underground river water height during all of the studied period. Figure 3.2b is a zoom on a flood occurring between November 16^{th} and November 20^{th} 2019. Spectrograms computed from seismological data recorded at FONT and AVEN, during this same

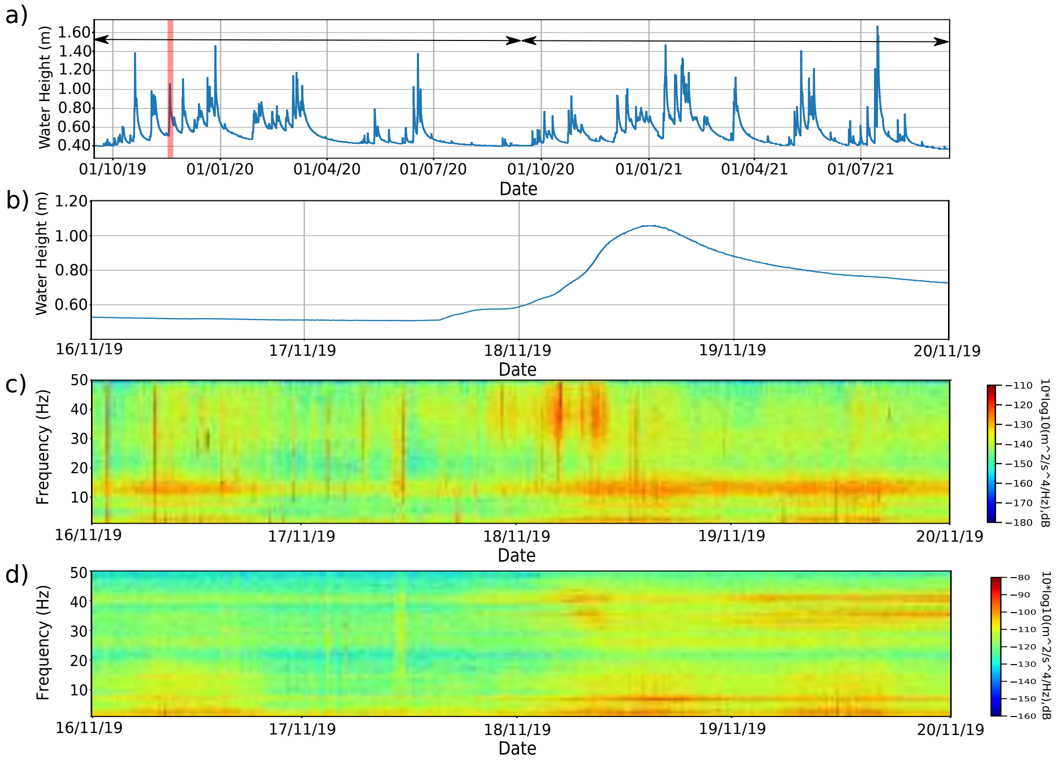


Fig. 3.2 (a) Hydrograph spanning the entire study period in terms of water height. The shaded part is the flood selected for the spectrograms computation. The two arrows indicate the two hydrogeological cycles used for the training and application of the RF algorithm. (b) Hydrograph of the selected flood event in (a), between 16 and 20 November 2019. (c) Spectrogram of the vertical components during the selected flood for signals recorded at FONT and filtered between 1 and 50 Hz. (d) Same as (c) for AVEN.

flood, are presented in Figure 3.2c and Figure 3.2d, respectively. Energy lines between 10 and 20 Hz appearing on both AVEN's and FONT's spectrograms can be related to the anthropogenic activity. Indeed, these energy lines are more marked for FONT than for AVEN due to its location underground, insulated from the surface, thus the river induced noise will appear on its signals' spectrogram with a higher amplitude. Actually, we can notice at this frequency range a day-night variation with more energy during daytime, and less energy during days-off (17^{th} of November is a Sunday). In addition, for the latter frequency range, more energy is manifested on the horizontal components than on the vertical component (Figure B.1 from Appendix B).

The seismic noise induced by the river becomes relatively visible on the spectrograms once the water rises. Due to the position of the seismometer, at the surface, in a field enclosing two horses, and the prominence of noise generated by anthropogenic sources (vehicles on the nearby road, dwellings, agricultural activities, mining) as well as the horses' gaits, the effect of water height variation is hardly detectable on the FONT spectrogram (Figure 3.2c). The noise amplitude increase due to water height increase is clearer on the AVEN spectrogram, which is more isolated from the surface noise (Figure 3.2d). Actually, three main frequency ranges can be associated with water height change (Figure 3.2d): 1-3 Hz, 5-8 Hz and 25-50 Hz. While the low frequency bands (1-8Hz) are visible before the flood, energy at high frequency seems to occur after the flood has started. The seismic noise related to water flow in rivers can be associated to different phenomena (Burtin et al., 2008; Díaz et al., 2014; Gimbert et al., 2014; Schmandt et al., 2013; Tsai et al., 2012). A major source of noise results from the frictional forces produced by the interaction between the turbulent flow and the riverbed. Another one is the bed load particles transport, generally observed at higher frequency.

In order to look more in detail at the noise induced by the flood, we have plotted the noise amplitude against water height with a 5 min time step for FONT (Figure 3.3a-b-c) and AVEN (Figure 3.3d-e-f). Data were filtered at 1-3 Hz, 5-8 Hz and 25-50 Hz, corresponding to the three frequency ranges evoked earlier. Similar trends are obtained for the two stations. While the water height rises from 0.50 m to 0.80 m and decreases below 0.80 m, the variation of the noise versus the variation of water height follow the same path. In addition, at all frequency ranges, between 0.90 m and the flood peak, the noise amplitude as function of water height draws a hysteresis. The latter is generally attributed to bedload transport during increasing flooding (high noise amplitude), and gravel deposit during recession (lower noise amplitude), with lower frequencies corresponding to larger particle movement (Burtin et al., 2008; Díaz et al., 2014; Schmandt et al., 2013). Gimbert et al. (2014) have also shown that turbulence processes in the river could also significantly contribute to the hysteresis curve.

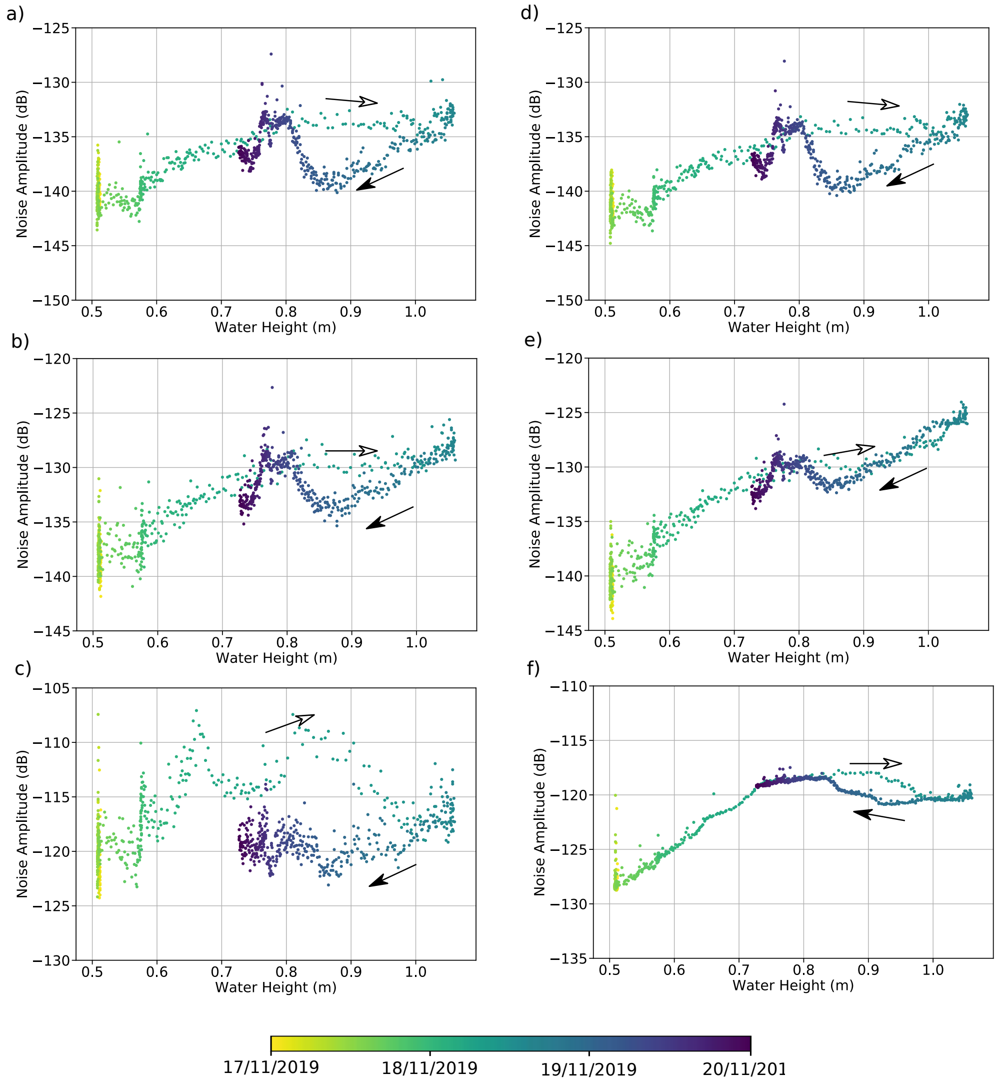


Fig. 3.3 Noise amplitude as function of the water height for for signals recorded at the surface station FONT and filtered between a) 1-3 Hz, b) 5-8 HZ and c) 25-30 Hz. Same for underground station AVEN in d),e) and f). The color bar represents the time scale between 17/11/2019 and 20/11/2019. The white and black arrows correspond respectively to the increase and decrease sections of the water height during the flood.

3.1.4 Feature extraction from seismic data

In this section the objective is to extract characteristic features from the raw seismic data. Firstly, a pre-processing of the seismic data has been performed. The data are decimated to 100 Hz (for AVEN only), and then detrended and filtered between 1 and 50 Hz. The seismic signals are then partitioned using a 15 minutes moving window with an overlap of 50 %, corresponding to a windowing step of 7.5 minutes. Several window lengths were tested. A 15-min-long window was chosen because it requires reasonable CPU time computation and provides sufficient resolution for capturing the beginning of the rise of water during a flood event. Finally, features of the seismic recordings are computed for each window (Table 3.1, see Hibert et al., 2017, for a detailed description of each feature). A total of 72 features are calculated related to the signals' waveform, frequency content, spectral energy, and pseudo-spectrogram. Similar features as Hibert et al. (2017) are used in our study, except for the polarity attributes, with additional frequency bands for the computation of the signal's Kurtosis and energy [1-3 Hz,3-5 Hz, 5-8Hz , 8-10 Hz, 10-15 Hz, 15-20 Hz, 20-25 Hz, 25-30 Hz, 30-35 Hz, 35-40 Hz, 40-45 Hz, 45-50 Hz]. These frequency bands are chosen to cover all of the studied frequency range (1-50 Hz) and target the bands affected by the water height variation obtained during the spectral analysis. These features, as presented in Hibert et al. (2017), are commonly used to identify events or seismic sources within the seismic signals since they are able to cover several aspects of the signals. In the case of the chosen configuration, the computation of features for a station and a year of data takes about 2 weeks of CPU time. The extracted features are used in the algorithm that is explained in the following section.

3.1.5 Method

The Random Forest (RF) algorithm (Breiman, 2001) is a bagging ensemble learning method based on the computation of a large number of decision trees. Each tree in the forest is generated from a random subset of events from the training set and a random subset of features describing the events. The RF algorithm has two modes of application: (1) the classification in which the final result will be a class obtained from the majority of voting, and (2) the regression in which the final result will be a value obtained by averaging the predicted values given by each tree. Increasing the number of trees in the forest helps in the convergence without causing over-fitting but reducing the generalization error (Breiman, 2001), which measures the prediction error of the model over the data set. Because of the random selection of both the training events and features, each decision tree in the Random Forest is unique and the trees are not correlated with each other that helps in reducing the over-fitting, which is one of the advantages of this algorithm. Another advantage is the ability of the algorithm to use a large number of features and assess their importance depending on the attributed case (water height in this study) in the prediction, while going from a multidimensional dataset to a 1D output. The importance of the features helps to better understand the results and provide insights on the link between the seismic signals features and the physics of the phenomena.

Table 3.1 Description of the features used in the algorithm (modified from Hibert et al., 2017). DFT and FFT stand for Discrete Fourier Transform and Fast Fourier Transform respectively.

Feature	Description
Duration	Duration of the signal
RappMaxMean	Ratio of the Max to the Mean of the normalized envelope
RappMaxMedian	Ratio of the Max to the Median of the normalized envelope
AsDec	Ratio of the Ascending to decreasing time of the envelope
KurtoSig	Kurtosis of the Signal
KurtoEnv	Kurtosis of the Envelope
SkewnessSig	Skewness of the Signal
SkewnessEnv	Skewness of the Envelope
CorPeakNumber	Number of peaks in the autocorrelation function
INT1	Energy in the first 1/3 of the autocorrelation function
INT2	Energy in the last 2/3 of the autocorrelation function
INT_RATIO	Ratio of INT1 to INT2
ESi-j	Energy of the seismic signal in the i-j Hz frequency band
Kurtoi-j	Kurtosis of the signal in the i-j Hz frequency band
DistDecAmpEnv	Difference between decreasing coda amplitude and straight line
RatioEnvDur	Ratio between maximum envelope and duration
MeanFFT	Mean FFT
MaxFFT	Max FFT
FmaxFFT	Frequency at Max (FFT)
FCentroid	Frequency of spectrum centroid
Fquart1	Frequency of 1st quartile
Fquart3	Frequency of 3rd quartile
MedianFFT	Median of the normalized FFT spectrum
VarFFT	Variance of the normalized FFT spectrum
NpeakFFT	Number of peaks in the normalized FFT spectrum
MeanPeaksFFT	Mean peaks value for peaks>0.7
E1FFT	Energy in the 1 – NyF/4 Hz (NyF=Nyqusit Frequency) band
E2FFT	Energy in the NyF/4 – NyF/2 Hz band
E3FFT	Energy in the NyF/2 – 3*NyF/4 Hz band
E4FFT	Energy in the 3*NyF/4 – NyF/2 Hz band
gamma1	Spectrum centroid
gamma2	Spectrum gyration radio
gammas	Spectrum centroid width
SpecKurtoMaxEnv	Kurtosis of the envelope of the maximum energy of spectrograms
SpecKurtoMedianEnv	Kurtosis of the envelope of the median energy of spectrograms
Ratioenvspecmaxmean	Ratio of the Max DFT(t) to the Mean DFT(t)
Ratioenvspecmaxmedian	Ratio of the Max DFT(t) to the Median DFT(t)
Distmaxmean	Mean distance bewteen Max DFT(t) Mean DFT(t)
Distmaxmedian	Mean distance bewteen Max DFT Median DFT
Nbrpeakmax	Number of peaks in Max (DFTs(t))
Nbrpeakmean	Number of peaks in Mean (DFTs(t))
Nbrpeakmedian	Number of peaks in Median (DFTs(t))
Rationbrpeakmaxmean	Ratio between the number of peaks in Max (DFTs(t)) and Mean (DFTs(t))
Rationbrpeakmaxmed	Ratio between the number of peaks in Max (DFTs(t)) and Median (DFTs(t))
Nbrpeakfreqcenter	Number of peaks in centroid frequency DFTs(t)
Nbrpeakfreqmax	Number of peaks in Max frequency DFTs(t)
Rationbrfreqpeaks	Ratio between the number of peaks in centroid frequency DFTs(t) and Max frequency DFTs(t)
DISTQ2Q1	Distance Q2 curve to Q1 curve (QX curve = envelope of X quartile of DTFs)
DISTQ3Q2	Distance Q3 curve to Q2 curve
DISTQ3Q1	Distance Q3 curve to Q1 curve

Each decision tree of the RF consists of internal nodes (splits) and terminal nodes (leaves) (Criminisi et al., 2012). The depth of a tree is the number of splits from its root (node 0) to its leaves. It is a measure of the number of splits made by the tree to get a prediction. No limitation was set on the maximum depth in our model: the nodes are expanded until all leaves contain less than 2 samples in the population during the splitting. The deeper the tree, the more splits it has, hence more information will be caught from the data and configured into the model. At each node, the selected feature is used to split the selected subset of data into two separate populations. The best splitting value at each node is found by variance reduction, meaning the value of the feature at the split is the value giving the lowest variance between the predicted (which is the mean value of each obtained population) and the real values, and thus yielding the highest precision. Feature importance is the assignment of a score to features based on their impact on the targeted prediction. The feature giving the lowest variance in the splitting is the feature with the highest importance and is chosen as a root node for the tree. The objective of the feature importance is to assess the influence of each feature on the model decision making in order to interpret the resulted predictions. Another objective can be to select important features for similar applications and gain in processing time.

An example of a tree of the forest resulting from the training of a model with AVEN's data with the maximum depth set at 3 is presented in Figure 3.4. This maximum depth is only used to generate this figure and to be able to visualize the functioning of a tree. A subset of features and data are selected for this tree. The features are sorted according to their importance. At the level of each node, water height is plotted as a function of the feature corresponding to the node. The feature corresponding to the energy of the seismic signal between 40 and 45 Hz (ES40-45), which represents the base 10 logarithm of the integral of the raw seismic signal's envelope filtered between 40 and 45 Hz, is the root of the tree since it is the most important feature. A splitting point for this feature is obtained and the population is divided into two sets accordingly, each set having its own mean water height that gives the lowest variance. Samples having this feature above the resulted threshold will be selected at the right part of the tree, and below the threshold at the left part. The next feature splitting the population will be less important than the preceding feature. This is done at every node until the maximum depth condition is fulfilled. The samples obtained at each node will be satisfying all the above conditions from all the previous nodes of the tree.The final plots represent the water height distribution at different time windows of the remaining samples after the splitting. At the level of each final node is the number of samples obtained for the final split and the prediction of the water height, which is the average value of the water heights of the remaining samples. Given features corresponding to a certain signal, a water height will be obtained at each tree, and having 1000 trees the final prediction will be the mean of all obtained water heights.

Unlike usual applications, the training and testing are here done on continuous signals (water height and seismic) and not on selected events. The process involves independent algorithms for each station (AVEN and FONT). The following steps include training on a certain period of time and testing on another period of time. The choice of these periods of time was controlled by data

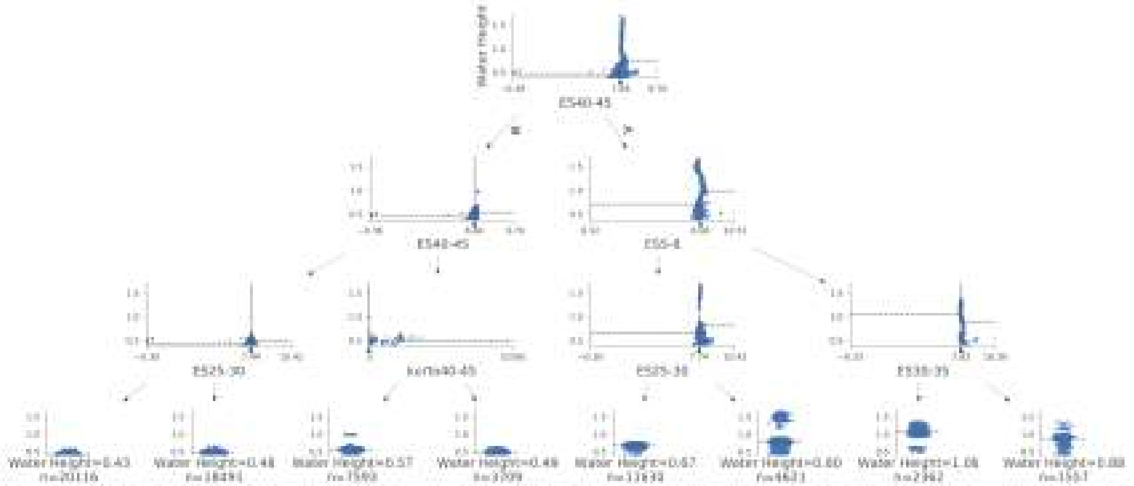


Fig. 3.4 A tree diagram generated from a model using AVEN's data with a maximum depth parameter set at 3.

availability. The training period is the same for AVEN and FONT : from September 15th 2020 to September 15th 2021 (training dataset). The testing period (testing dataset) is from September 15th 2019 to September 15th 2020 for AVEN and only two months for FONT (from October 27th to December 31st 2019). Hence the training dataset counts 70 000 windows (of 15 minutes) for AVEN and FONT, and the testing dataset counts 70 000 windows for AVEN and 13 000 windows for FONT. The choice of training the algorithm on 2020-2021 data and testing it on 2019-2020 instead of the other way round is the lack of data for FONT; this choice allows to have firstly the data of a complete hydrological cycle for the training to cover all potential water heights and secondly a same training period for both stations for results comparison issues. A RF with 1000 trees is then created based on the training dataset by assigning to each window of features the corresponding measured water height. A similar configuration was used as in Provost et al. (2017) and Hibert et al. (2017). The RF model was then applied on the testing dataset which generates an array of water heights. The predicted water heights are then compared with the real values in order to assess the algorithm precision. Finally, to evaluate which features are the most relevant, 10 forests were created and trained, each giving values for the features importance. These values are then averaged over the 10 instances. This number of instances was chosen since it is a reasonable choice in terms of CPU time, knowing that one instance of the algorithm can take several hours since we are using a year of data for the training of a 1000 trees forest.

3.1.6 Results of the regression

The results of the regression analysis performed on data from stations FONT and AVEN are presented in Figure 3.5. The predicted water height was smoothed for both stations using a 10-days moving window to avoid short transitory signals coming from local noise sources. A good fit between observed and predicted values was obtained, as illustrated in Figure 3.5. In order to better assess the quality of the fit we calculated the overall Root Mean Square Error (RMSE)

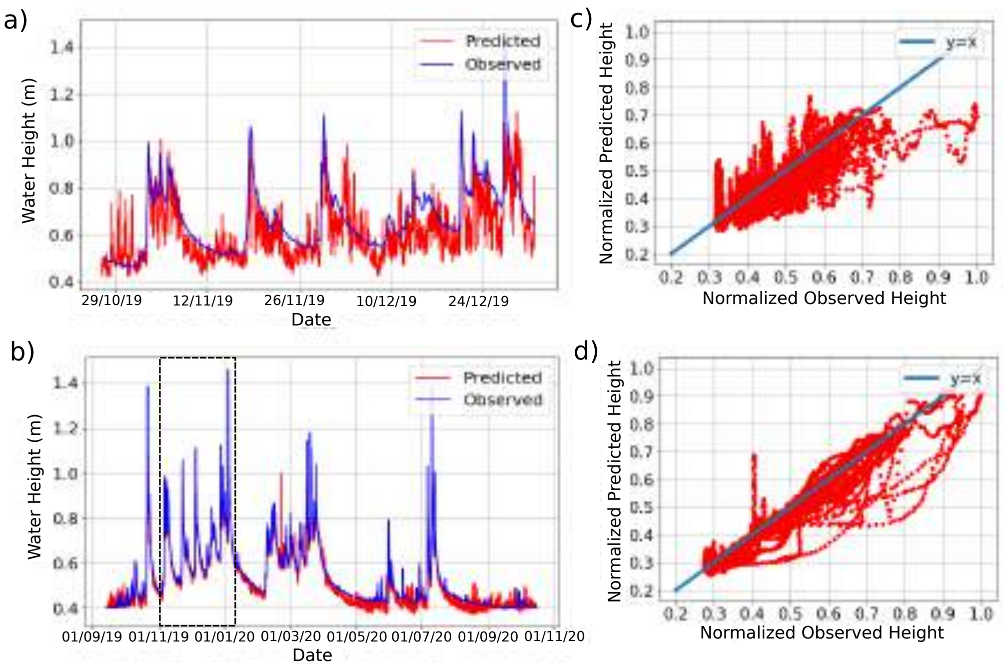


Fig. 3.5 (a) Simulated water height at surface station FONT; the blue line is the water height measured at the CTD; the red line is the predicted water height obtained from the application of the algorithm on the seismological data. (b) Same as (a) for underground station AVEN. The dashed rectangle indicates the period of application in (a). (c) Predicted versus measured water height for simulations carried out at FONT; the blue line represents the 1:1 line.(d) Same as (c) for AVEN.

and the Nash–Sutcliffe efficiency coefficient (NSE) which is commonly used in hydrological models (McCuen et al., 2006) and is given by:

$$NSE = 1 - \frac{\sum_{t=1}^{T} (H_0^t - H_m^t)^2}{\sum_{t=1}^{T} (H_0^t - \bar{H}_0)^2} \tag{3.1}$$

where $\bar{H}_0$ is the mean of observed water heights, and H_m is modeled water height. H_0^t is observed water height at time t.

For FONT, the obtained RMSE is about 0.1 m and the NSE is about 53 %. The prediction shows many outliers or misfits compared to the true water height variation. The misfits are mainly observed during the recession period (e.g. beginning of December 2019; Figures 3.5a, and c). It is most likely due to the position of the station at the surface, at the vicinity of many major noise sources which tends to hide the noise generated by the river. However, the overall shape of the hydrograph is correctly reproduced. For AVEN the quality of the fit is very high, with an RMSE of only 0.03 m and the NSE reaching 95 %. A few outsider peaks can be observed systematically during periods of flood recession : mostly from mid to end of January 2019, from mid March to end of April 2019 and from beginning of July to end of September 2019. For these periods the predicted water heights fall mainly below and only punctually above the observed heights (Figures 3.5b and d). In these cases, the seismic noise generated by the river most likely interferes with other noise sources. Figures 3.5c and d show the predicted values of the water height versus the real measured values on a normalized scale for simulations done with FONT and AVEN, respectively. If the fit between predicted and real values was perfect they would have been aligned along the 1:1 line. We observe that for both stations, most of the points are below the 1:1 line, meaning that the predicted water height is lower than the observed one. This could be due to the difficulty to predict the water heights during low water periods, when the induced noise amplitude is lower. We also computed the Pearson coefficient (R) for both applications to measure the strength and direction of the linear relationship between the predicted and observed values. As expected, the coefficient is high for AVEN (0.98) and lower for FONT (0.70).

3.1.7 Discussion

The features importance resulting from averaging scores over 10 instances of training are presented in Figure 3.6. The most important features retained by the algorithm are waveform and signal energy features for both stations. In the case of FONT, the dominant feature is the Kurtosis between 8 and 20 Hz (a plot showing the variation of the kurtosis feature for signals filtered between 15 and 20 Hz compared with the water height is shown in Figure B.2 of the Appendix B). In the case of AVEN, signal energy is the most important feature, for frequencies between 5 and 8 Hz and at high frequency (above 20 Hz), especially between 40 and 45 Hz (a plot showing the variation of the signal's energy feature between 40 and 45 Hz compared with the water height is shown in Figure B.3 of the Appendix B).

The Kurtosis of the seismic signal is a common feature used in many classification applications which allows detecting natural or anthropogenic seismic events within continuous seismic records (e.g., Baillard et al., 2013; Liang et al., 2008; Ross and Ben-Zion, 2014). This may justify why this feature is the most important for the FONT regression model. The position of this station at the surface near roads and agricultural fields leads to the presence of random punctual peaks in the seismic signals thus causing an increase in the kurtosis. Since we are using 15 minutes windows, multiple few-seconds events can be included leading to a flatter normal distribution and a higher kurtosis value, which can be detrimental to the variation of the kurtosis with the water height. The skewness of the seismic signal is also a common feature used for event detection, especially seismic phase detection (Küperkoch et al., 2010; Ma et al., 2015). Hence, it might also be affected in our case by anthropogenic events. We tested the algorithm without the kurtosis and the skewness features, which allowed improving the results (Figure 3.7).

The predicted hydrograph is smoother, presenting less misfits and similar RMSE (0.14 m) compared with the initial prediction. The NSE is however smaller (13 % against 53 %). Indeed, the amplitude of the floods are not completely recovered. As mentioned above, the Kurtosis feature is usually used for event detection applications, hence its importance in the detection of floods as well as the detection of transitory water height peaks during a flood event. For example for the flood occuring between the 20^{th} and the 26^{th} of December 2019 (Figure 3.7), we can notice that the first peak is correctly reproduced, unlike the following peaks, which are

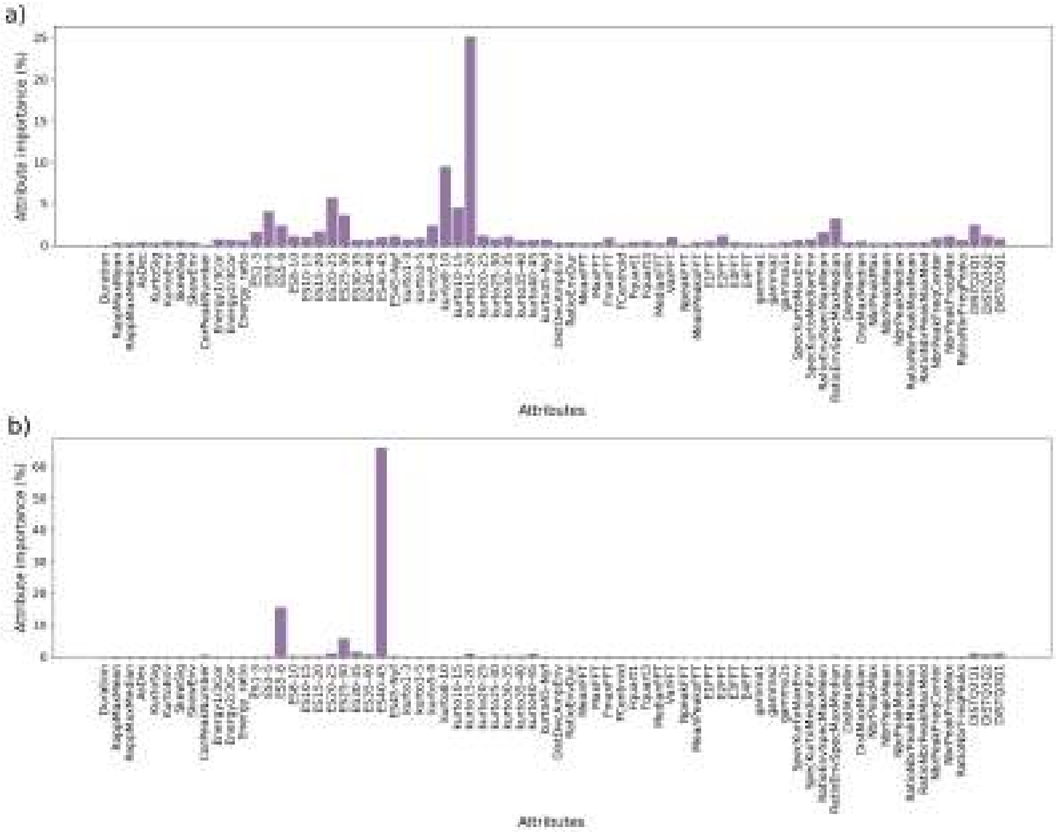


Fig. 3.6 Feature importance obtained by averaging scores over 10 instances of training and testing the RF algorithm for stations a) FONT and b) AVEN.

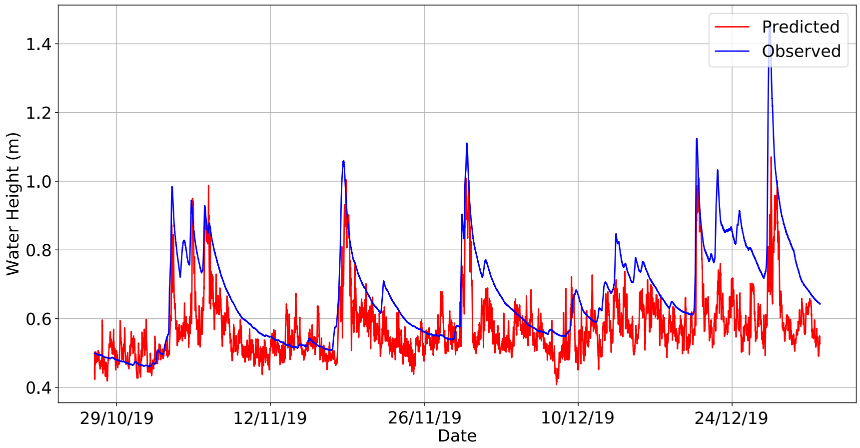


Fig. 3.7 Water heights simulated with FONT's data after removing the kurtosis and skewness features; the blue and red lines correspond to the observed and predicted water height, respectively.

underestimated. Applying the algorithm while removing the kurtosis features at FONT allows to highlight the signal's energy feature, in particular between 1-8 Hz and 20-30 Hz (Figure 3.8).

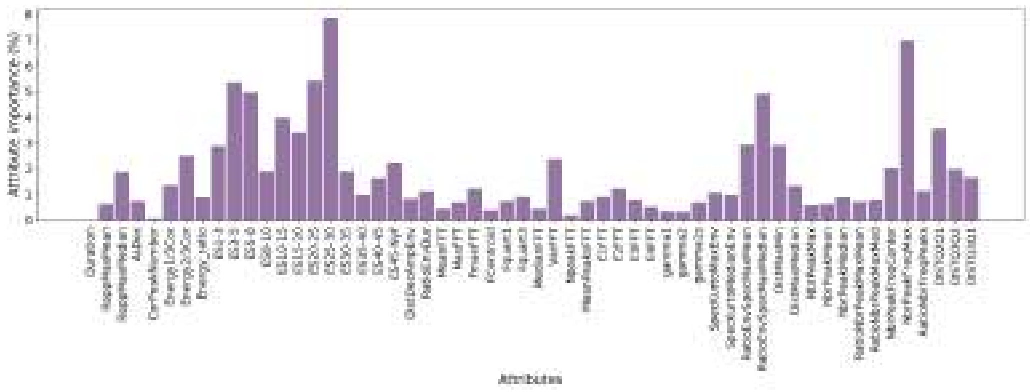


Fig. 3.8 Feature importance obtained by averaging scores over 10 instances of training and testing the RF algorithm for FONT without the kurtosis and skewness features

According to results at the two stations, the main feature of the seismic noise use by the RF algorithm for simulating the water height is the signal energy. At both stations, these features exhibit low and high frequency contents, at 1-8 Hz and above 20 Hz, respectively. As mentioned in section 3.1.3, hydrodynamic processes have a particular spectral signature, which has been described and modeled for surface river (Burtin et al., 2008; Gimbert et al., 2014; Tsai et al., 2012). Observed features are most likely related to these mechanisms : the low frequency band features would correspond to the turbulence, while higher frequency content (above 20 Hz) could be due to bed load transport.

Associated frequency bands are more widely distributed at FONT than at AVEN, which could be explained by their respective location (surface and 20 m depth) affecting their ability to capture river induced signals only. In addition, due to seismic attenuation, a station closer to the underground river, such as AVEN, is able to record higher frequency content, compared to a surface station such as FONT. The channel morphology at Fontenotte is very heterogeneous (the river width can vary in terms of several meters) which makes physical models developed for surface river (Gimbert et al., 2014; Tsai et al., 2012) difficult to apply. However, bed load transport model developed by Tsai et al. (2012) and field observations provided by Burtin et al. (2008) indicate that the smaller is the distance station-river, the higher is the peak frequency. Our results corroborate this, with the second best feature being the seismic energy value in the 25-30 Hz frequency band for the farthest station (FONT - approximate distance of 60 m) and the best feature being the seismic energy fluctuation in the 40 to 45 Hz for the closest and deepest station (AVEN - approximate distance of 50 m). This highlights the strong correlation between the seismic energy in those frequency bands and the dynamics of the river, and hence explains the ability of the RF machine learning model to reproduce the hydrograph from seismic signals.

One limit of the methodology presented here is the necessity of having hydrogeological data for the training. On the other hand, once the training is complete, no access is anymore needed to the underground river to do the maintenance of the hydrological probe and collect its data. In addition, the hydrological probe should be installed a period long enough to cover all potential water heights that might occur. Indeed, the method is not able to predict water heights that it was not trained for, in other terms it cannot extrapolate. Further investigation to unravel this could be to use data at the river's outlet, where the access is generally easy, in order to train the RF and then test it at other positions on the conduit.

Linear regression is a simple and efficient mean allowing to simulate water height as demonstrated by Anthony et al. (2018), who used data from geophones located at the vicinity of the river (1 m from the stream). In this station configuration, the seismic energy is dominated by the noise induced by the river. This is not the case when monitoring an underground river from surface, as observed at FONT where the river-induced signal is most likely too weak and noisy to apply regression analysis. On the other hand, machine learning method can detect the most relevant frequency ranges for performing a successful simulation, even for distant and noisy stations. In fact, multiple features were revealed relevant for the regression, which shows the importance of this method in it's ability to deal with such datasets.

Although the predicted hydrograph fits more correctly the observed values using data from a very high signal to noise ratio sensor closer to the source (such as AVEN), the application of the method using data from the surface station is promising: the prediction fits well the temporal evolution of the real water height permitting the detection of floods as well as roughly estimating the corresponding water height. Better regression for FONT can potentially be obtained by improving the station insulation (e.g. FONT is only buried at 0.5 m depth) which could enhance the signal-to-noise ratio of the recorded signals. Further investigation could be performed to better characterize the effect of the geometry and the dynamics of the river. An approach could be to test the algorithm with data from a denser seismological array installed all along the river as well as at the surface. Finally, the proposed method could be extrapolated to other accessible or inaccessible sites and implemented for real-time applications allowing a continuous monitoring of flood events in order to circumvent the installation of invasive instruments.

3.1.8 Conclusions

We used a Random Forest (RF) algorithm to predict the flow dynamics of an underground karstic river using seismological data. 72 features characterising the seismic signals were computed to train the model and predict the water heights. With an RMSE of 0.03 m obtained for the regression using the data collected from the underground seismological station (AVEN) and of 0.1 m using the data collected from the seismological station at the surface (FONT), the RF model proved to be a reliable method for remote monitoring of the water heights. Feature importance was computed in terms of variance between predicted and observed values. The most important features correspond to the signal energy for AVEN and FONT. They are related to the contribution of the water-riverbed interaction and the bed load transport on the seismic noise content. The results demonstrate the accuracy of the method in predicting underground river water heights, even with weak river-induced and noisy signals (the case for FONT).

3.1.9 Acknowledgments

This project has been funded by the Région Bourgogne-Franche-Comté and OSU THETA. Calculations were performed using HPC resources from DNUM CCUB (Centre de Calcul de l'Université de Bourgogne). We are very thankful to the members of the Association Spéléologique du Doubs Central (ASDC) and J.-P. Simonnet (Chrono-environnement, UFC) for their active support in the field work. We thank D. Motte for his scientific contribution and for giving us access to the topographical data of the Fontenotte cavity. We are also thankful to the land owners and tenants of the instrumented site as well as the municipality actors for their help and availability. We also acknowledge the French Karst National Observatory Service (SNO KARST, https://sokarst.org/) and the OZCAR (Critical Zone Observatories: Research and Application) Research Infrastructure. We are grateful for F. Gimbert C. Stanciu and anonymous reviewer, for their helpful comments that contributed to improve this manuscript.

3.1.10 Data Availability Statement

Data from the long-term regional seismic network JURAQUAKE and the JURASSIC KARST hydrogeological observatory of the SNO KARST network were used in the creation of this manuscript. A description of the data and their access are found in https://www.fdsn.org/networks/detail/5C_2018/ for JURAQUAKE and https://sokarst.org/en/9-observatories-in-france/jurassic-karst-en/ for JURASSIC KARST. Figures were made with Matplotlib version 3.5, available under the Matplotlib license at https://matplotlib.org/. The map was created through QGIS3 available under https://www.qgis.org/fr/site/forusers/download.html. The regression is done using the scikit learn library "RandomForestRegressor". The Python code for the computation of the features is created by C. Hibert and modified by A. Abi Nader for this article and can be found under https://doi.org/10.5281/zenodo.6592237.

3.2 Detecting water recharge mechanisms in a karst reservoir with Fiber Optic Distributed Temperature Sensing

3.2.1 Introduction

Karst aquifers serve as a valuable research platform for understanding groundwater flow, contaminant transport, and for protecting and managing water resources. Understanding the processes of recharge and transfer in these systems is thus of utmost importance. Karst aquifers can be affected by several types of porosity, such as conduits, fractures and matrix, inducing both rapid and slow water movements between potential interconnected reservoirs (Ford and Williams, 2007). The recharge of karst aquifers can be diffuse, with the percolation of precipitations within the epikarst and flowing through fractures and joints until the saturated zone (Williams, 2008). This recharge mechanism is rather slow and continuous. Localized infiltration of surface waters from non-karstified impermeable areas of the catchment can also occur, generally through sinkholes and/or faults (Atkinson, 1990; Ford and Williams, 2007). This so-called allogenic recharge is fast and dependant on the flow rate of the surface stream. The contribution of each recharge mechanism is however difficult to distinguish and quantify. Such assessment would require continuous, long-term and, and as much as possible, large spatial coverage monitoring of different parameters of water flow (e.g., temperature, electrical conductivity, flow rate).

Many geophysical methods have been implemented to investigate and monitor groundwater reservoirs (Goldscheider and Drew, 2014), and more recently Fiber Optic Digital Temperature Sensing (FO-DTS). Actually, this technique has emerged as a cutting-edge technology due to several advantages: it can offer such as high sensitivity, large spatial resolution, continuous monitoring and remote data collection (real-time analysis). For example, FO-DTS has been successful in characterizing stream-aquifer and estuary-aquifer interactions (e.g.in Virginia, Wyoming, and Massachusetts, Day-Lewis et al., 2008). Vertical profiling has been done with FO-DTS to estimate seepage rates in a losing stream (Vogt et al., 2010). This technique was also used for fluid-logging experiments in boreholes as in the Grimsel test site for the disposal of radioactive waste of NAGRA in the Swiss Alps (Hurtig et al., 1994).

In this study, we introduce a pioneering approach in which we use FO-DTS to investigate recharge processes of a karst aquifer. We aim to better understand the exchanges/interactions between its different compartments (conduits, matrix, fractures). To to this, we perform fine monitoring of temperature changes within an underground river and one of its affluent identified in the reservoir. The site is located in Jura, which is part of an in-situ research laboratory (Fourbanne watershed, France, see Chapter 2). An 800 m fiber optics cable was deployed within the underground river and three hydrological probes are used as temperature references for the calibration. The objective is (1) to assess the effect of atmospheric changes (precipitations, air temperature) on the spatio-temporal distribution of the temperature within the reservoir (air cavity, river, affluent) (2) to characterize the the recharge mechanisms in the river and the affluent.

3.2.2 Physical Background of the FO-DTS

3.2.2.1 Functioning

Fiber is normally made of silica (glass) due to its pure qualities and the properties that allows it to carry light internally. The fiber is composed of overlapping layers which are the core, the cladding, and the coating (Fig. 3.9). Light propagates into the core (many individual optical fibers bound together) of the fiber ($\sim 10\mu m$ of diameter). The cladding ($\sim 125\mu m$) acts as a protection and allows the reflection of the light to keep it trapped inside the core (Kawakami and Nishida, 1974). The cladding is usually made of the same material as the core, but with a slightly lower refractive index ($n_{core} > n_{clad}$). The refractive index is a fundamental optical property of a material that describes how much light is bent or refracted when it passes from one medium into another, it is defined as the ratio of the speed of light in vacuum (c = 299792458 m/s), and the phase velocity of light in the medium. Usually the cladding's index is about 1% lower than the core's index. When a light wave moves through the fiber core and reaches the interface between the core and cladding, the distinct properties of these two regions lead to the bending of the light wave back into the core. The coating, or buffer, ($400\mu m$) is made out of one or more coats of plastic material to protect the fiber from the physical environment and mechanical constraints. A layer of strengthening fibers is added for protection. Finally a jacket layer, usually made out of plastic, might be added to give more strength to the fiber. Light pulses should be injected in the fiber with a specific angle in order to avoid beam diffraction and maintain the total internal reflection. This angle, referred to as the acceptance angle, is related to the difference in the indices of refraction of the outside medium (n_{clad}) and the fiber medium (n_{core}) (Lowe et al., 2015). This angle is called the acceptance angle and is given by the following formula:

$$\theta_{acc} = asin(\frac{1}{n_0}\sqrt{n_{core}^2 - n_{clad}^2}) \tag{3.2}$$

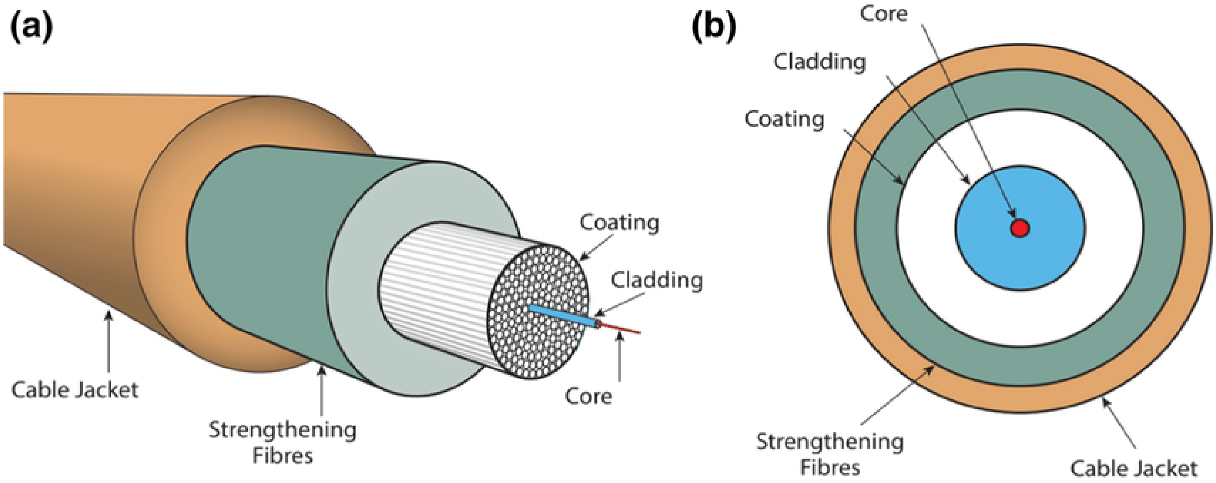


Fig. 3.9 Fiber optic cable structure. a) Sketch of the layers composing the cable and b) cross-section of the cable, from Fenta and Szanyi (2021).

3.2.2.2 Raman Scattering

Even in the purest optical fibers, light is scattered as a result of the disordered (noncrystalline) structure of glass. The three modes of scattering are Rayleigh (elastic), Brioullin and Raman (non elastic). When light strikes matter, it may be reflected at the original energy, or a portion of that light is adsorbed and reemitted at wavelengths just above or below the frequency of the incident light, due to loss or gain from quanta of energy exchanged with electrons. Thermal effects induce lattice oscillations within the solid. Lattice vibration corresponds to the oscillations of atoms in a solid about the equilibrium position. When light falls onto these thermally excited molecular oscillations, an interaction occurs between the light particles (photons) and the electrons of the molecule. This frequency shifted-light, altered by an amount equivalent to the resonance frequency of the lattice oscillation, is referred to as Raman scattering, with the light at frequency below the incident light being referred to as Stokes backscatter, and that above the incident light the anti-Stokes backscatter (Fig. 3.10). Below a critical light intensity, the magnitude of Raman Stokes and anti-stokes scattering is a linear function of the intensity of illumination. The ratio of the magnitudes of the anti-Stokes to Stokes scattered lights eliminates the intensity dependence and provides a quantity that depends exponentially on the fiber temperature (Day-Lewis et al., 2008; Mwakanyamale et al., 2012).

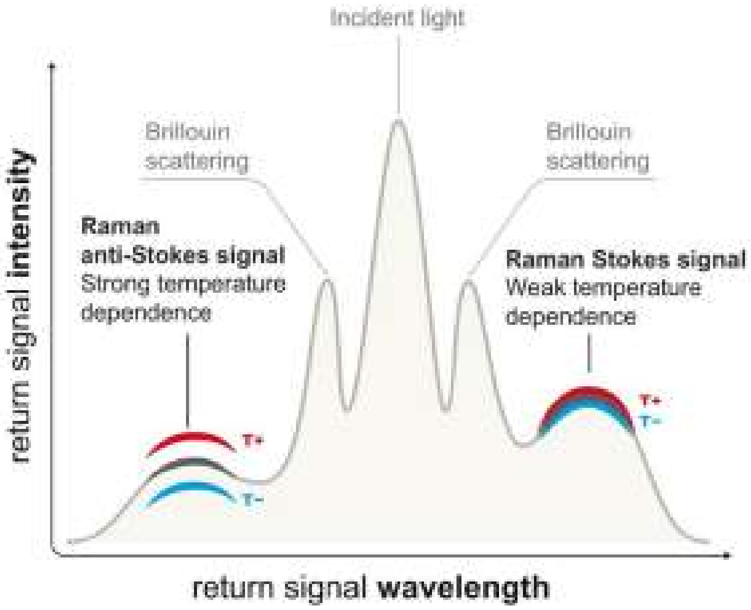


Fig. 3.10 Spectrum of the incident and backscattered signals, from AP Sensing (APSensing, 2018).

The location of the measured backscattered light can be found using the following equation:

$$Location = 2.t_R.V_L \tag{3.3}$$

where t_R is the time taken for the light to come back to the DTS (about $0.1\mu s$), and V_L is the light velocity in fiber. If we take into account the optical index of the silica (n=1.4496 at λ =

$1.064\mu m$), the velocity of the light passing through fiber is given by:

$$V_L = \frac{c}{n} = 206953642.38 \ \ m/s \tag{3.4}$$

Where c is equal to 3.10^8 m/s, and n is dimensionless. Finally, the temperature can be calculated using the following equation (Farahani and Gogolla, 1999):

$$\frac{I(\omega+\Omega)}{I(\omega-\Omega)} = exp\left(\frac{-k\Omega}{k_B T}\right) \tag{3.5}$$

where $I(\omega+\Omega)$ corresponding to intensity of the anti-Stokes signal (in frequency domain), and $I(\omega-\Omega)$ to the Stokes signal. k_B is the Boltzmann constant, T the temperature (K), k the wave number (=$2.\pi.\nu/V_L$ in cm^{-1}) and Ω is the difference in frequency between the Stokes and anti-Stokes signals (Hz).

3.2.2.3 Single-ended Configuration

The configuration for the FO-DTS used in this study is the single-ended configuration: only one end of the fiber is connected to the DTS (Fig. 3.11). Another configuration exists, which is the double-ended configuration. In the latter, the two ends of the fiber are connected to the DTS and measurements are done from both ends. The double ended configuration is usually applied for studies where several kilometers of fiber are deployed, since it yields better precision due to the measurements of attenuation done from the two sides of the fiber. In our case, the single ended configuration implies that the difference in attenuation (between Stokes and anti-Stokes scattering) is generally approximated as constant over the length of the fiber. Nonetheless, this difference in attenuation can be influenced by unintended factors such as sharp bends or connectors joining together two cable portions (Des Tombe et al., 2020).

Fig. 3.11 DTS instrument measurement principle and basic information system setup, from AP sensing .

3.2.3 Studied Site and FO-DTS deployement

The studied site is the underground river flowing into the main conduit of the Fourbanne karst reservoir (see Chapter 2). The conduit is monitored with three hydrological probes (2 simple Troll 100 pressure-temperature probe and an Aquatroll 500 multiparameter probre), measuring water height and water temperature (Fig. 3.12). Precipitations and air temperature are recorded by the meteorological station located at Fontenotte (see Fig. 2.5 in Chapter 2 for the location). A

1500 m long fiber has been deployed in September 2020 between the Fontenotte's artificial well entrance at the surface (North) and mostly the end of the explored unsaturated area in the south (red line on Fig. 3.12). The fiber cable and the DTS were provided by J.P. Malet (EOST-ITES). This challenging installation took 2 days and requested about 15 people. The precise location of the fiber (LAF) was determined based on topographical measurements requiring again several days of field work (teams led by D. Motte (ASDC) and G. Ferhat (EOST-ITES)).

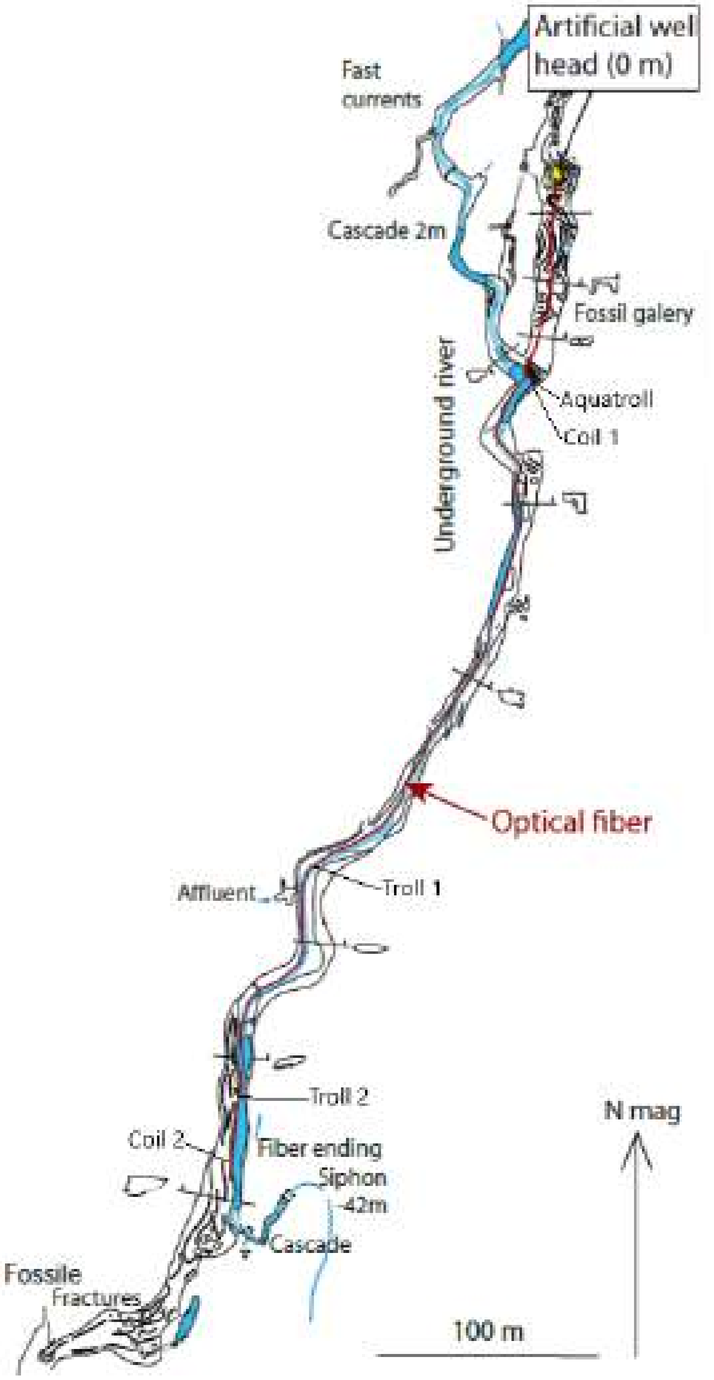


Fig. 3.12 Topographical map of the underground conduit and the instrumentation used in this study. The deployed optical fiber is marked with a red line. Map from by D. Motte (ASDC)

.

A very-high resolution 3D scan of the cavity was also acquired thanks to Geotopo (Gleize, France) with a long-range Zeb-REVO to document the geomorphology of the cavity (Fig. 3.13).

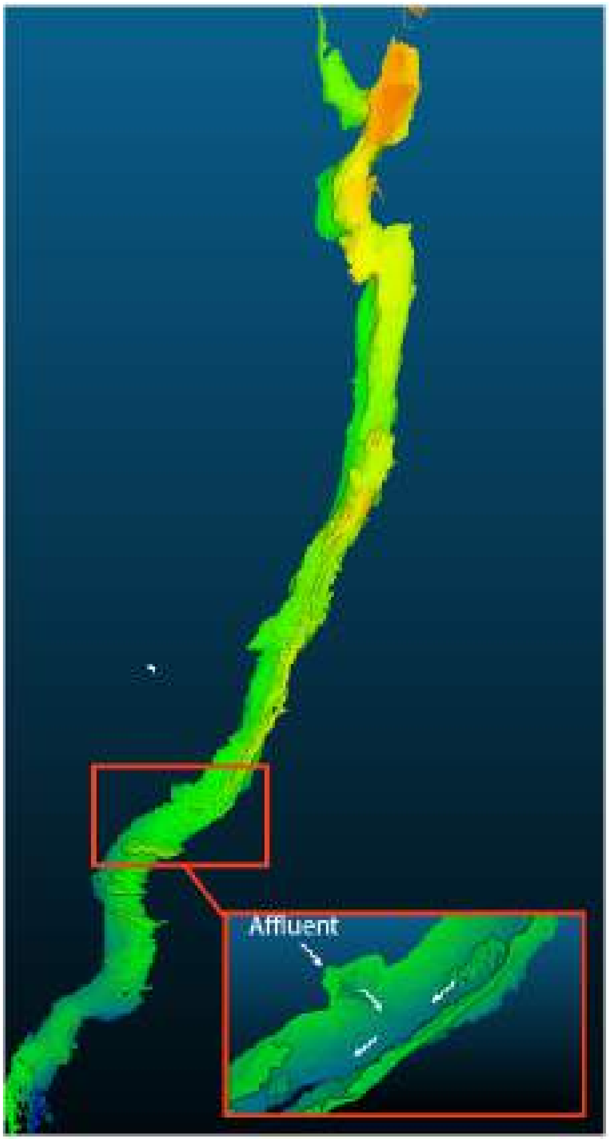


Fig. 3.13 3D scan of the underground conduit (top view) with a zoom on the affluent (cross section). The arrows indicate the flow's direction in the river and from the affluent.

Fig. 3.14 is a sketch of the DTS-cable on site deployment. The DTS is located at the surface and is connected to solar panels in order to supply its battery for continuous recording (see Fig. 3.15). The fiber is connected to the DTS, which is considered as its starting point (0 m length along fiber-LAF). The fiber cable allows to measure the temperature each 5 minutes, with a 50 cm spatial resolution, and to identify temperature variations of the order of 0,01^oC.

In order to locate the positions of the first fiber-water contact in the river, the coils and the affluent along the fiber's lengths, 4 heating bags (hot water bags) were used and deployed at the position of each for 6 minutes, the time needed for the FO-DTS to record 3 traces. Due to the temperature difference between the river and the bags, the fiber will record high temperature peaks in their corresponding positions (see Fig. 3.16).

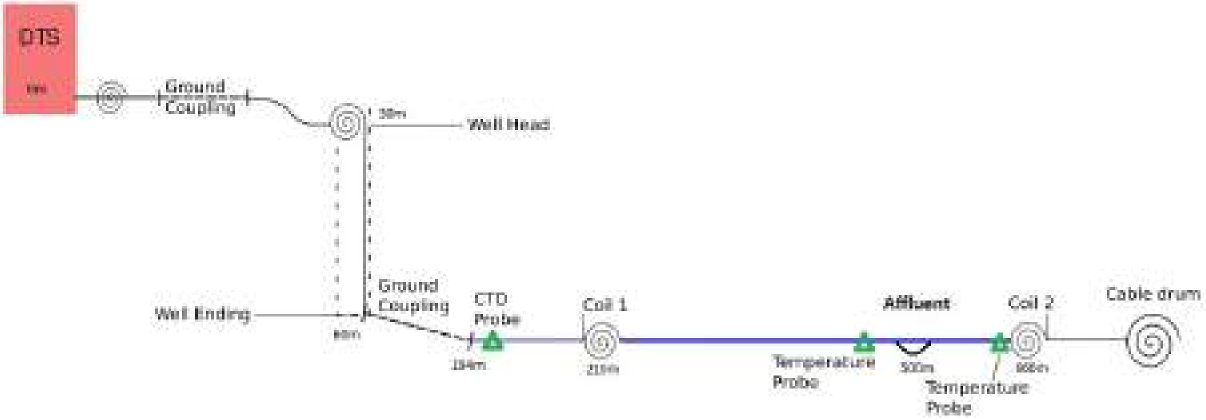


Fig. 3.14 Configuration of the deployed instrumentation (courtesy of Camille Nguyen).

Fig. 3.15 DTS instrument and the solar panels powering its battery.

The fiber is coupled to the ground until reaching the head of the artificial well giving access to the underground conduit. The fiber passes through the drilled well and reaches the river at 194 m LAF, close to the position of a hydrological probe (Aquatroll, 195 m LAF). From there, the fiber lies within the water, ballasted with rocks or attached to the river ground. In two places the fiber has been fixed to the wall, forming coils. The first coil is built up at 220 m, by fixating parallel horizontal lines of fiber on the conduit's wall, which gives at least one measuring point on each line, or each level (see the picture in Fig. 3.17).

The coil permits to measure air temperature inside the cave, but also indirectly measures the water height at the position of the coil, due to the water-air temperature difference. At 492 m the fiber reaches the position of a temperature probe (Troll1 on Fig. 3.12). At ~500 m along the fiber, the cable does a deviation to enter a lateral affluent (Fig. 3.18). Then the cable gets out of

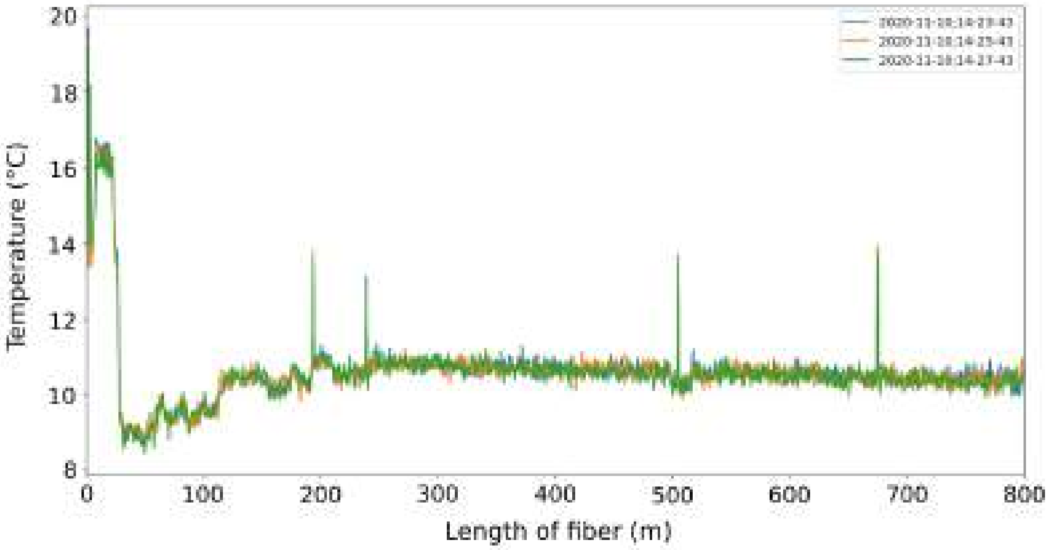


Fig. 3.16 Location of heating points along the fiber, three different traces are plotted corresponding to three different measurements with a two minutes interval (the sampling frequency was changed to 5 minutes afterwards).

Fig. 3.17 Picture of the southern cable coil.

the affluent in a u-turn and into the main conduit again and reaches a second temperature probe at 659 m (Troll2 on Fig. 3.12). Then the fiber reaches the second coil and ends in a cable drum outside the river. The studied period spans from the 15th of March to the 26th of August 2022, when data from all instruments were available.

Fig. 3.18 Picture of fiber deployed within the affluent taken the $14th$ of May 2023.

3.2.4 FO-DTS data Processing

3.2.4.1 Data Conversion

The DTS logger records binary data (big-endian 32-bits format): a trace file is recorded per day corresponding to measurements all along the fiber for the chosen time and spatial steps. In each trace file there are 288 traces (5 minutes time step over a day), each of them consisting of temperature data every 0.5 m along the fiber. Data are then converted from binary to decimal format using a Matlab code from Kusnahadi Susanto (Univ. Padjadjaran, Indonesia/EOST-ITES). The latter results in tabular data containing for each spatial step a recorded temperature, step loss (based on attenuation), and the ratio between the Stokes and anti-Stokes intensities (Hausner and Kobs, 2016).

3.2.4.2 Data Correction

The optical signal traveling through the fiber can experience attenuation and signal loss due to various factors like fiber quality, bends, and other impingements. Correcting for these losses can provide a more accurate representation of the actual temperature distribution along the fiber. The further the position along the fiber, the higher the attenuation, hence the loss in the measurements. Here, we use the method proposed by Hausner et al. (2011) for data correction considering single ended configuration and two reference points. The following calibration scheme is applied at every time step.

Temperature along fiber is given by the following equation:

$$T(z) = \frac{\gamma}{ln\frac{P_S(z)}{P_{aS}(z)} + C - \Delta\alpha z} \tag{3.6}$$

Where $T(z)$ (Kelvin) represents the temperature at the position z (m), γ (Kelvin) is the shift in energy between a photon at the wavelength of the incident laser and the scattered Raman photon (ΔE (J), where energy is calculated with $E = hc/\lambda$, h is the Planck constant: $6,63.10^{-34}$ ($J.s$), c the speed of light in vacuum and λ the wavelength). γ is given by: $\gamma = \Delta E/k$, k being the Boltzmann constant: $1,38.10^{-23}$ ($J.K^{-1}$). C represents a dimensionless parameter specific to the DTS, it encompasses properties of the incident laser and the DTS (incident laser wavelength, back-scattered Raman signals, instrument's photon detector). $\Delta\alpha$ (m^{-1}) is the differential attenuation between the Raman Stokes and anti-Stokes signals (Hausner et al., 2011). The following single-ended calibration methods require a uniform $\Delta\alpha$ along fiber. If this parameter significantly changes, the double-ended configuration is more appropriate. In our case, due to the used length of fiber, $\Delta\alpha$ is fairly uniform.

From eq. (3.6), three parameters should be found for temperature correction. This can be done using three temperature references, which are in our case the hydrological probes, and applying eq. (3.6) at the position of each. However, since the temperatures recorded for the probes in our study are quite similar, corresponding to temperature measured along the river, data correction has been done using two references located at the beginning and at the end of the section lying in the river, the Aquatroll an Troll2, respectively. In this case, two methods are proposed: we should compute $\Delta\alpha$ or fix γ and use two reference points to find the two other missing parameters (Hausner et al., 2011). We present both cases in the following.

The $\Delta\alpha$ can be interpolated from a length of fiber at uniform temperature. Using Beer's Law, that states that the attenuation of light occurs either as a result of distance through substance in absence of other influences, and because $\Delta\alpha = \alpha_{aS} - \alpha_S$, the following relationship was derived between $\Delta\alpha$, z and $\frac{P_S}{P_{aS}}$ (Hausner et al., 2011):

$$e^{-\alpha_{aS} z_2} \frac{P_S(z_2)}{P_{aS}(z_2)} = e^{-\alpha_S z_1} \frac{P_S(z_1)}{P_{aS}(z_1)} \tag{3.7}$$

by detailing, it yields:

$$ln\frac{P_S}{P_{aS}}(z_2) = ln\frac{P_S}{P_{aS}}(z_1) + \Delta\alpha(z_2 - z_1) \tag{3.8}$$

Where the power of the Raman Stokes ($P_S(z)$) and anti-Stokes signals ($P_{aS}(z)$) are used by taking the Neperian logarithm of the ratio (directly measured from the FO-DTS), $\Delta\alpha$ (m^{-1}) is the differential attenuation between the Raman Stokes and anti-Stokes signals and z, the distance (m) (Hausner et al., 2011). We apply eq. (3.8) by using two points along the fiber, by rearranging it as follows, z_1 and z_2 being the corresponding positions of the two chosen points:

$$\Delta\alpha = \frac{ln\frac{P_S(z_1)}{P_{aS}(z_1)} - ln\frac{P_S(z_2)}{P_{aS}(z_2)}}{(z_2 - z_1)} \tag{3.9}$$

It is considerably better to take two points, as far away as possible, in order to have a parameter as reliable as possible. The next step after the computation of $\Delta\alpha$ is to calculate γ and C, using the equation eq. (3.6) and data from the two temperature references. We should note that the correction yields better results if the temperatures recorded at the two references are different. Computing the parameters comes down to resolve the following system of equations:

$$\begin{aligned} T_1(z_1) &= \frac{\gamma}{ln\frac{P_S(z_1)}{P_{aS}(z_1)} + C - \Delta\alpha z_1} \qquad (a) \\ T_1(z_2) &= \frac{\gamma}{ln\frac{P_S(z_2)}{P_{aS}(z_2)} + C - \Delta\alpha z_2} \qquad (b) \end{aligned} \tag{3.10}$$

Having z_1 and z_2 as the position of the two references, $T_{1(z1)}$ and $T_{1(z2)}$ the temperatures recorded at both references, the ratio between Stokes and anti-Stokes being taken from the fiber's data at the same position of the chosen references, and $\Delta\alpha$ calculated previously, we can resolve the latter set of equations and find γ and C. The other method consists in fixing γ and finding $\Delta\alpha$ and C by resolving the two same equations as before (eq. 3.8). γ depends mainly on the material fiber (van de Giesen et al., 2012).

In our case, we set $\gamma = 216.85$ (K), value used for silica, and we used the second method for the data correction.

3.2.4.3 Comparison between corrected and uncorrected data

To compare corrected and uncorrected fiber data, and evaluate the correction scheme, we selected points along the fiber at the vicinity of the hydrological probes(at 195 m, 492 m and at 659 m to compare with the Aquatroll, Troll 1 and Troll 2 respectively). Figure 3.19 illustrates comparison between uncorrected (blue lines), corrected (dashed red lines) and isolated temperature probe (red lines) performed at the Aquatroll (a), Troll 1 (b) and Troll 2 (c). As mentioned before the temperature correction has been done using two references: the Aquatroll and Troll2. The temperature peak recorded on May 15^{th} is an artifact due to human presence linked to maintenance. A first observation is that both corrected and uncorrected time series follow the trends of increase and decrease recorded at each of the hydrological probes after rainfall. We can see in Fig. 3.19a and Fig. 3.19c a perfect fit between each of the respective references and the corrected data. Fig. 3.19 also shows that the fiber tends to underestimate the temperatures: the difference between the uncorrected data and the reference is about 0.25 °C for the Aquatroll and

1 °C for the second Troll. The latter also shows us that the underestimation gets larger the further is the measuring point from the DTS, due to the attenuation properties of the fiber mentioned above. However, in Fig. 3.19b, the corrected data underestimates the reference values by about 0.25 °C. This might be due to the loss in the correction accuracy the further we get from the two reference points, or due to the overestimation of the temperature by the Troll since the difference between both is within the accuracy error in the measurement for the Trolls (which is in the order of ± 0.3 °C). Nevertheless, the corrected time series gives a much better estimation for the Troll time series than the uncorrected temperature. In the following, we use corrected data from the fiber for the analysis of the response of river temperature to precipitation events.

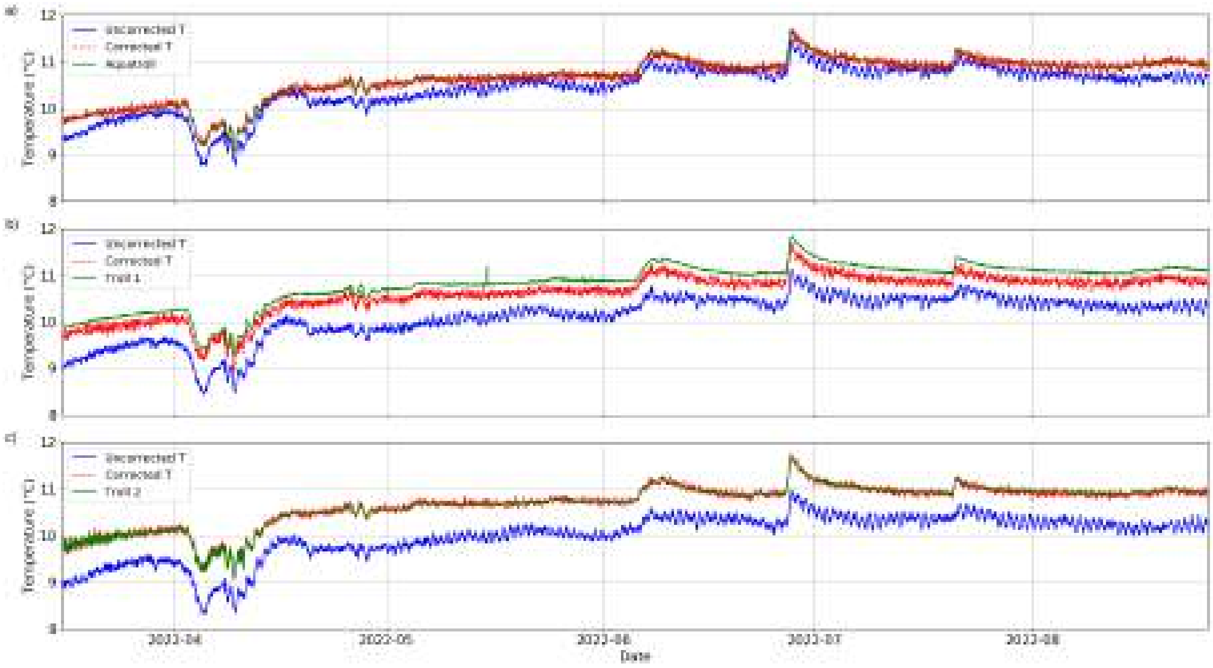


Fig. 3.19 Comparison between corrected and uncorrected fiber temperature data with a) the Aquatroll's temperature data, b) Troll 1 temperature data and c) Troll 1 temperature data, selecting points along fiber at the vicinity of each of the three references (195 m, 492 m and at 659 m).

3.2.5 Temperature analysis within the karst reservoir : Results and Discussion

3.2.5.1 Observations based on hydrogeological and meteorological data

Figure 3.20 shows a plot of the atmospheric and hydrological data recorded between March and end of August 2022. The air temperature data presents daily cycles, corresponding to fluctuations between daytime (high) and nighttime (low) temperatures. The latter also shows a positive trend over these few months, since the studied period starts at the beginning of spring and ends in summer, going from a relatively cold to a relatively hot weather. The total rain amount is about 390 mm during the period. Seven main precipitation events can be pointed out (shaded rectangles on Fig. 3.20, P1 to P7) inducing water level increases, about 8-10 hours later (see Chapter 2;

Cholet, 2017). The water base level tends to decrease during the period which lies from the end of winter in March to warmer season (until the end of summer), from 0.5 m to 0.3 m. In fact, the increase of temperature implies an increase in surface evaporation, thereby limiting infiltration because a larger proportion is rapidly evaporated. Moreover, the change in season implies that rain-bearing systems change in nature, from stratiform in winter to convective in summer (Houze, 1993). Convective precipitation is generally more intense, and of shorter duration, than stratiform precipitation, which is likely to modify the quantities of water reaching the river.

The river's temperature has a positive trend (Fig. 3.20d) which could be related to the global increase of the atmospheric temperature recorded during the studied period (Fig. 3.20a). Rapid temperature changes are also observed within the river during precipitations (e.g. P1 to P7) but aren't directly related to atmospheric temperature changes. Actually, we observe air temperature drop of about 10°C end of May which doesn't affect river temperature. In the same way, air temperature is quite stable during P2 whereas the river presents small but clear change. River temperature variations are concomitant with water level. They are thus controlled by rain water temperature which infiltrates rapidly from the surface.

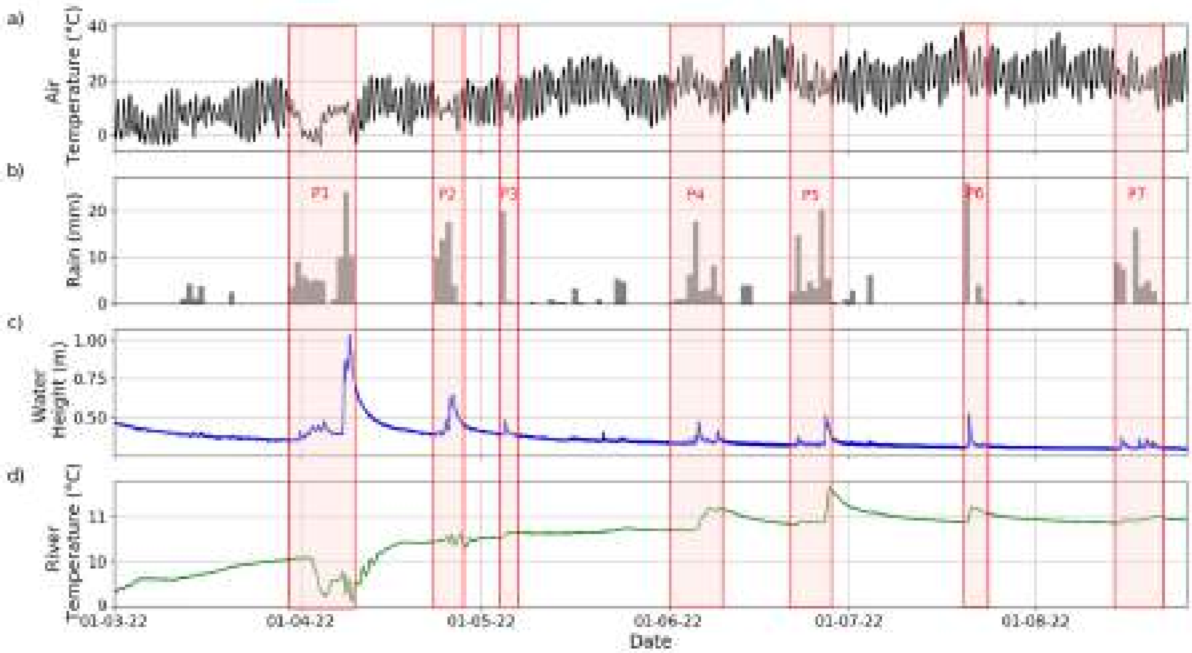


Fig. 3.20 Atmospheric and hydrological data recorded during the studied period. a) Air temperature and b) daily precipitations measured by the Fontenotte weather station. c) Water height and d) water temperature of the underground river measured by the Aquatroll. Seven main precipitation events are highlighted by shaded rectangles.

.

3.2.5.2 Temperature variations along the fiber

3.2.5.2.1 River temperature variations

Figure 3.21 is a spatio-temporal representation of corrected temperatures averaged over 1 h along the fiber during the studied period. Temperature changes more or less homogeneously along the fiber, except for three locations corresponding to the 2 coils (between 209 and 244 m, and between 660 and 681 m) and the affluent (between 498 and 515 m), which are discussed in the following.

The effect of precipitation events on river temperature can be clearly identified all along the fiber, inducing either water temperature decrease or increase. These river temperature changes are almost simultaneous along the monitored section of the river. They depend on precipitation temperature: if the rain is cooler than the river, we observe a decrease in the river temperature (P1) and vice versa (P3 to P7). The variations happen concomitantly with water level increase, typically within 8 to 10 hours after the beginning of rainfall (see Fig. 3.22). This is consistent with fast recharge mechanism, with rainwater infiltrating from surface through faults and/or sinkholes, located upstream, until the river and then to the saturated zone.

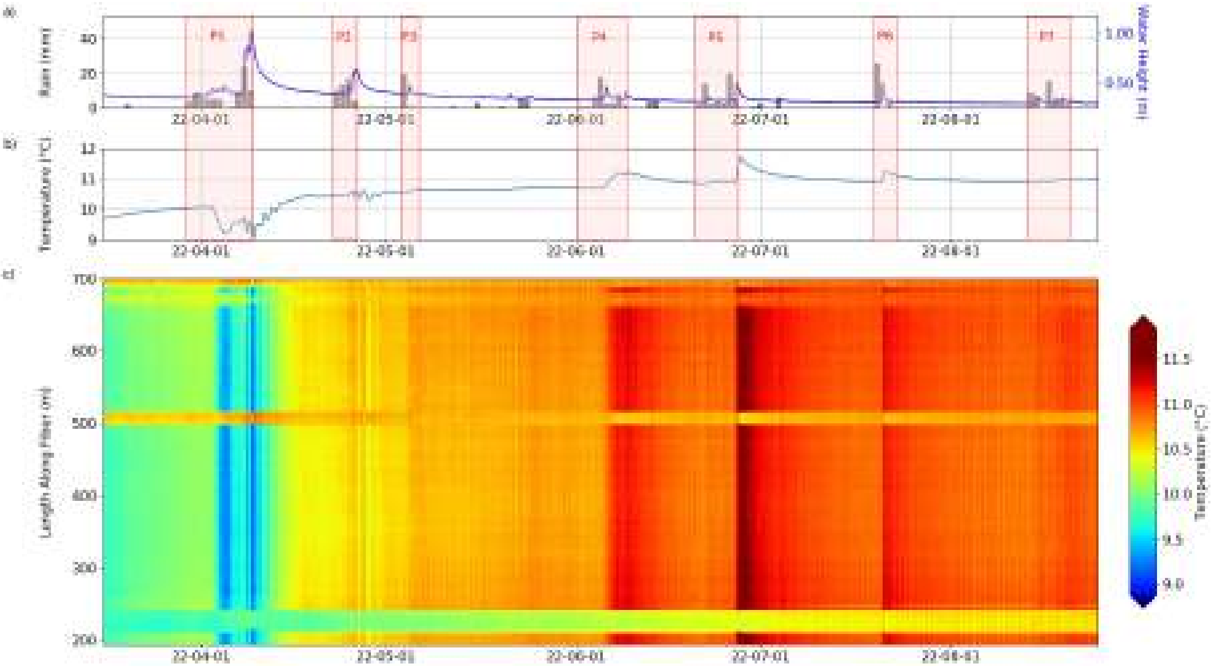


Fig. 3.21 Temperature variations along the fiber, between March and August 2022. a) Daily precipitations and underground river water height. b) Hourly temperature recorded at the Aquatroll. The shaded rectangles correspond to the main precipitation events P1 to P7. c) Hourly temperature along the fiber.

.

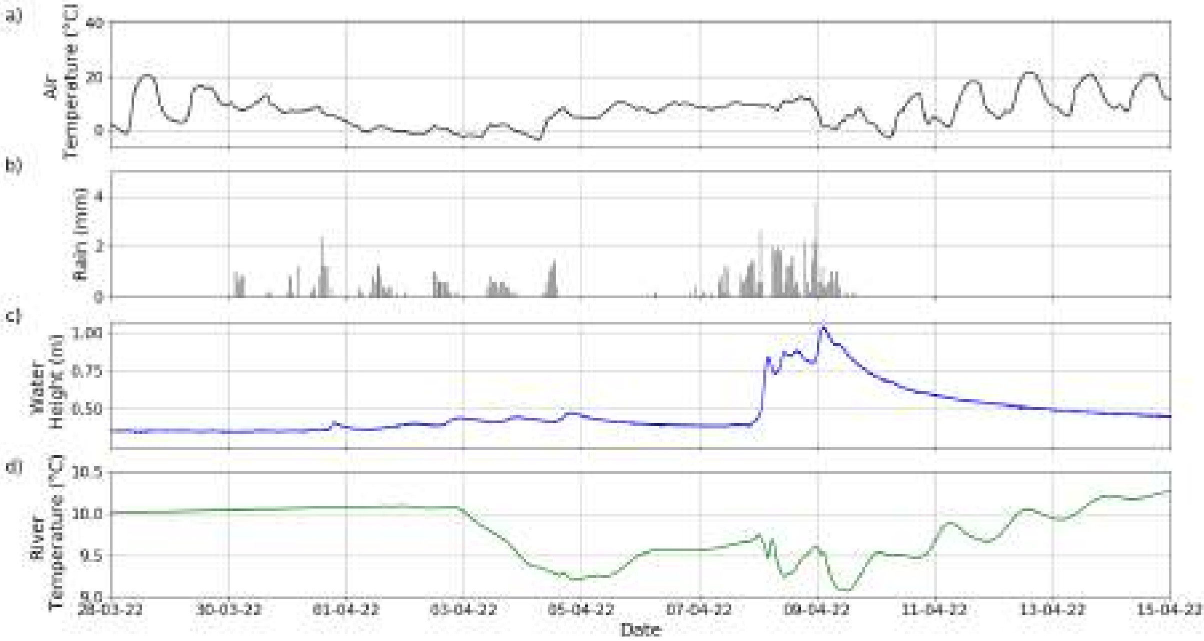


Fig. 3.22 Atmospheric and hydrological data recorded between 28/3/2022 and 15/4/2022 (event P1, see Fig. 3.20). a) Air temperature and b) hourly precipitations measured by the Fontenotte weather station. c) Water height and d) water temperature of the underground river measured by the Aquatroll probe.

3.2.5.2.2 Air temperature variation within the cavity

The cavity air temperature can be measured on the upper part of the coils which are fixed to the cavity wall. We can see in Fig. 3.21 that the temperatures measured at the first coil (around 230 m LAF) are lower than at the second coil (around 675 m LAF), indicating an increase of air temperature with distance from the access shaft. The second coil exhibits temperature variation patterns similar to river water temperature. In order to examine in detail how air temperature varies in cavity, we extracted temperatures time series recorded at specific points along the fiber (Fig. 3.23): at points high up on the wall along the two coils (230 m and 675 m for the first and second coils respectively), and within the river (300 m). We can see there is about 0.8 °C increase between the 1st and the 2nd coil. The same temperature variation patterns are observed at both coils, albeit higher amplitudes for the second coil during precipitation events. As mentioned before, we see the effect of weather warming between March and August : temperatures of the cavity and of the river all increase of about 1 °C. Rapid temperature variations are observed during the largest precipitation events (P1, P2, P4, P5 and P6 on Fig. 3.23). Air temperature at the second coil tends to follow the same variation as river temperature: both temperatures decrease during P1 and increase during the other precipitation events. Theses changes are synchronous with water level change, indicating the direct effect of precipitations, as discussed earlier. On the contrary, temperature at coil 1 is more stable over the studied period, with only weak temperature changes during precipitation events. The largest amplitude temperature change is observed during the first precipitation event (P1): temperature at coil 1 decreases concomitantly with atmospheric temperature (Fig. 3.23b), and before coil 2 and the river. Air temperature variations at coil 1 seems to be affected by atmospheric temperature variations (outside the cavity). This might be due to a draught from the nearby well.

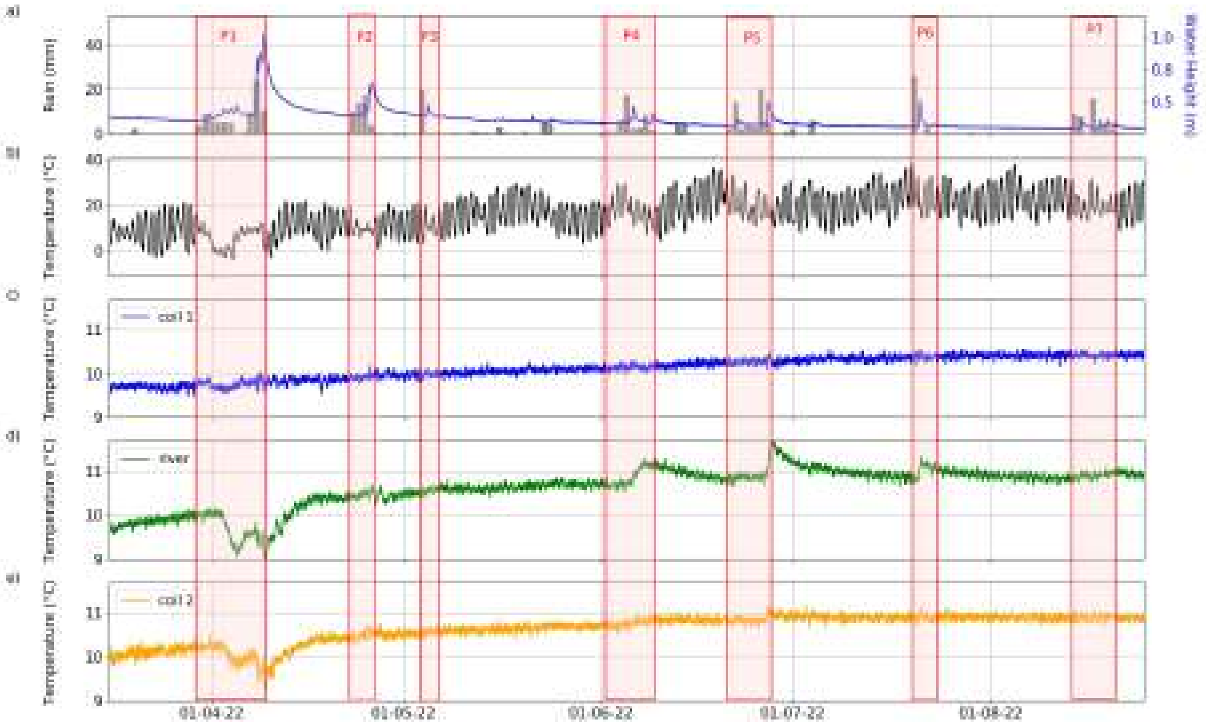


Fig. 3.23 Comparison of air/river/atmospherical temperature variations. a) Daily precipitations and underground river water height. b) Air temperature measured by the meteorological station. Temperature measured along the fiber at c) 230 m (air/wall temperature, coil 1), d) 300 m (river's temperature) and e) 675 m (air/wall temperature, coil 2). The shaded rectangles correspond to the main precipitation events P1 to P7.

3.2.5.2.3 Affluent temperature change

In this section we focus on temperature variations of a small affluent located between ∼495 and 516 m LAF (see Figs. 3.12, 3.18 and 3.24). It was discovered a few years ago by speleologists, who noticed a lateral permanent small stream of water, with variant flow rates depending on the season. In other words, the affluent is continuously supplied with water by diffuse recharge processes. The fiber's cable enters the tributary, does a U-turn and then goes back into the main river. Depending on the flow through the tributary, parts of the fiber might not be completely submerged by affluent water. This can be observed in Fig. 3.24c with horizontal bands depicting temperature alternation of less than 0.5°C between 498 and 516 m LAF. The affluent's temperature is rather constant during the studied period with weak amplitude variations, compared with the river's temperature. Its temperature is higher than the river until May and lower starting the 2nd week of June to the end of the studied period. This observation suggests the water infiltrating at the affluent originates from a separated reservoir. This is discussed in more details in the next section.

During precipitation events, we observe the effect of river overflowing the affluent entrance/exit, in particular during P1: the water is decreasing of about 1°C at 497 m LAF (or

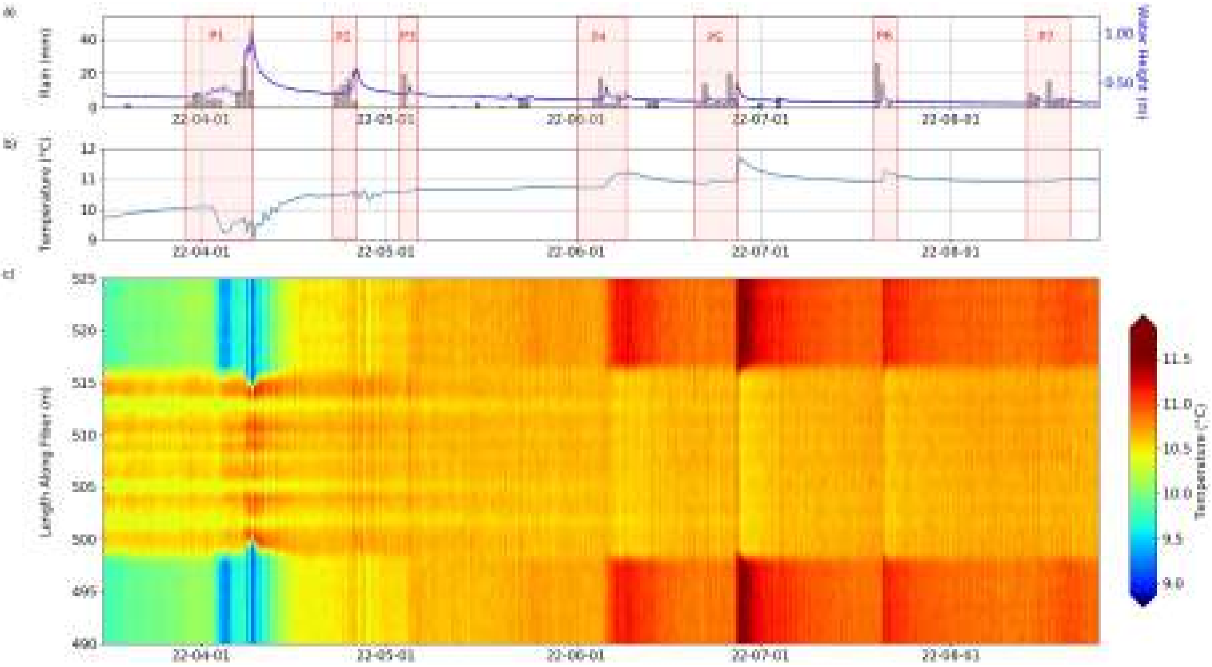


Fig. 3.24 Temperature variations within the affluent, between March 15 and August 26 2022. a) Precipitations and underground river water height. b) Temperature recorded at the Aquatroll averaged over 1 hour. The shaded rectangles correspond to the main precipitation events P1 to P7. c) Temperature variation along the fiber's section in the affluent, averaged over 1 hour. The color bar indicated temperatures going from blue (8.5°C) to red (12°C). Dashed (horizontal) lines showing the beginning of the affluent (at 499 m) and its end (at 516 m).

517 m). The tributary does not seem to have any impact on river temperature, most likely because the river flow is much larger than the affluent flow, so that affluent water is diluted in the river. This is illustrated in Fig. 3.25, which shows temperatures recorded at 3 points located before (497 m), within (509 m) and directly after (517 m) the affluent.

Similarly with the river temperature, the affluent temperature is affected by precipitations, as soon as the water height within the river increases (Figs. 3.24 and 3.25). This is clearly visible during P1, P2, P3, P4, P5 and P6 and consists in a rapid temperature increase, followed by a temperature decrease. During P4, thought, the precipitation induces a temperature decrease.

River and affluent present an almost concomitant response to rain events, more precisely by the water level increase. However, their behavior is not always similar, as they can exhibit opposite temperature variations. Actually, during P1, we observe a temperature increase in the affluent while the temperature decreases in the river after this cold rainfall event. This punctual "warmer/colder" water supply from the affluent seems triggered by water level increase in the whole fracture/conduits network. This suggests that the affluent is fed by a local reservoir from which water is expelled and renewed when water level of the cave river increases. This is discussed in more details in the next section.

Fig. 3.26 shows the recorded temperature in both river and affluent, normalised in respect to the maximum of each, along with the water height recorded at the Aquatroll, during the precipitation event P1 (from the 28^{th} of April to the 15^{th} of March 2022). We can see that the

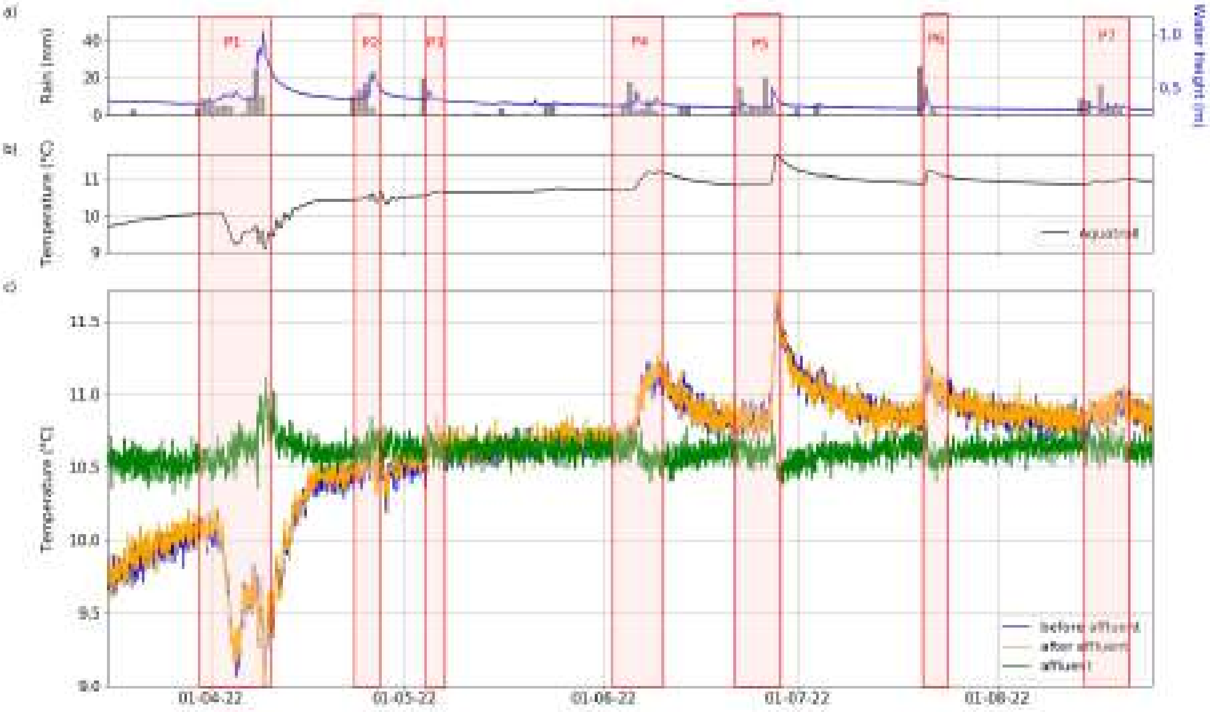


Fig. 3.25 Comparison between temperatures recorded directly before, within and directly after the affluent. a) Daily precipitations and underground river water height. b) Temperature recorded by the Aquatroll, averaged over 1 hour. Temperature measured along the fiber at c) 497 m (before affluent), 509 m (in affluent) and 517 m (after affluent), averaged over 1 hour. The shaded rectangles correspond to the main precipitation events P1 to P7.

affluent water temperature starts increasing when the water height reaches ∼0.45 m, once this threshold is achieved, affluent water temperature variation becomes concomitant with the river's response (in terms of water height and temperature).

Fig. 3.26 Normalised temperatures recorded at the river and the affluent, along with the water height recorded by the Aquatroll probe, during the precipitation event P1.

3.2.5.3 Recharge mechanisms within the unsaturated zone

In order to better understand the origin of water flowing from the affluent, it was equipped with CTD diver probe the 14th of May 2023 for measuring water height, temperature and electrical conductivity with a 15 min sampling rate. The water height is not considered due to the low flow rate in the affluent. We plotted the underground river water height measured by the Aquatroll along with the river and affluent temperature and electrical conductivity in both river and affluent between the 15th of May 2023 and the 26th of August 2023 (Fig. 3.27).

The first precipitation event during the period (29th of June 2023) induces a peak of temperature of +0.1°C and +3°C in the affluent and the river, respectively. The following precipitations are less intense and induce smaller and opposite temperature variations in the affluent (decreases) and the river (increases). Regarding the conductivity, it is decreasing after the 1st precipitation event indicating dilution by rainwater. Actually, electrical conductivity (EC) provides information on the degree of mineralization of the water, which increases with increasing residence time in low permeability fractures and/or the rock matrix (Celle-Jeanton et al., 2003). On the contrary, EC in the affluent increases abruptly after the first precipitation event, indicating expulsion of long residence water to the affluent by piston-effect. The subsequent precipitation events also induce EC increases but with lower amplitudes. These observations come as a support to the hypothesis set above that the affluent is fed by a local reservoir, where water is stored for a long time and thus more mineralized than river or rain water. Thus, during a rain event, the stored water is expelled by pressure transfer (see Fig.3.28, rainy period case), inducing an EC increase in the affluent. The "pushed" water has the temperature of the rock reservoir, which is either higher or lower than the river and rain temperature, inducing a decrease or increase. In addition, the hydrological model proposed by Cholet (2017) shows that in periods of high flow

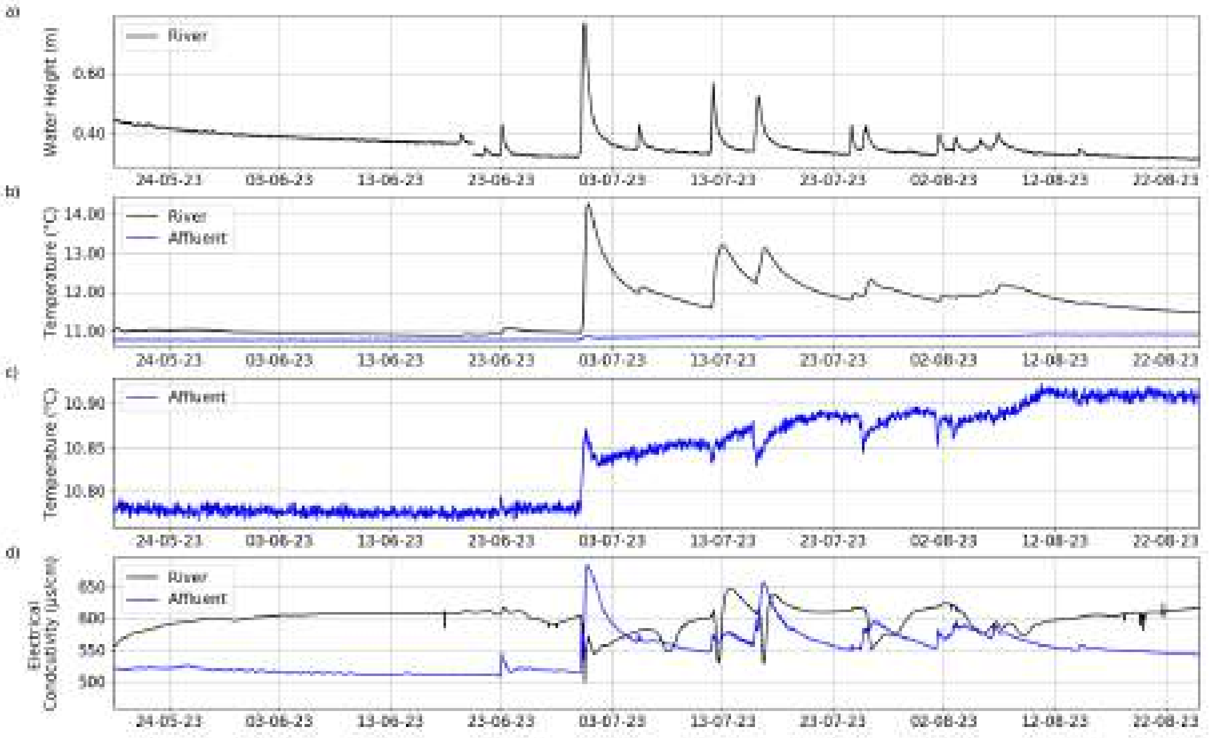


Fig. 3.27 Conductivity and temperature measurements in the affluent and the river. a) Underground river water height measured by the Aquatroll. b) Water temperature of the river and the affluent. c) Water electrical conductivity of the river and the affluent.

and in case of heavy rainfall event, the section of the conduit investigated here (Fontenotte area) is characterized by highly mineralized lateral inflows (see Fig. 2.6 from Chapter 2), which is coherent with our affluent data.

Another interesting observation is that before the 1st rain event (end of June) and after the 9^{th} of August, EC in the affluent is lower than in the river (Fig. 3.27d). This could be related to the hypothesis proposed before, stating that both river and affluent are fed by slow recharge processes but from different origins. The water of the affluent has most likely circulated across a shorter distance within the matrix than the water supplying the river outside rainy periods. The affluent is thus most likely supplied by a local water recharge from the epikarst (see Fig.3.28). This is supported by the change in the reservoir geomorphology in the affluent area as, which can observed on the 3D scan of the conduit (Fig. 3.13). The fractures/faults directions change from N-S/NE-SW north of the affluent to ∼E-W south of it. This change of fractures/faults direction could favor lateral inflow on the river and be responsible for affluent localisation.

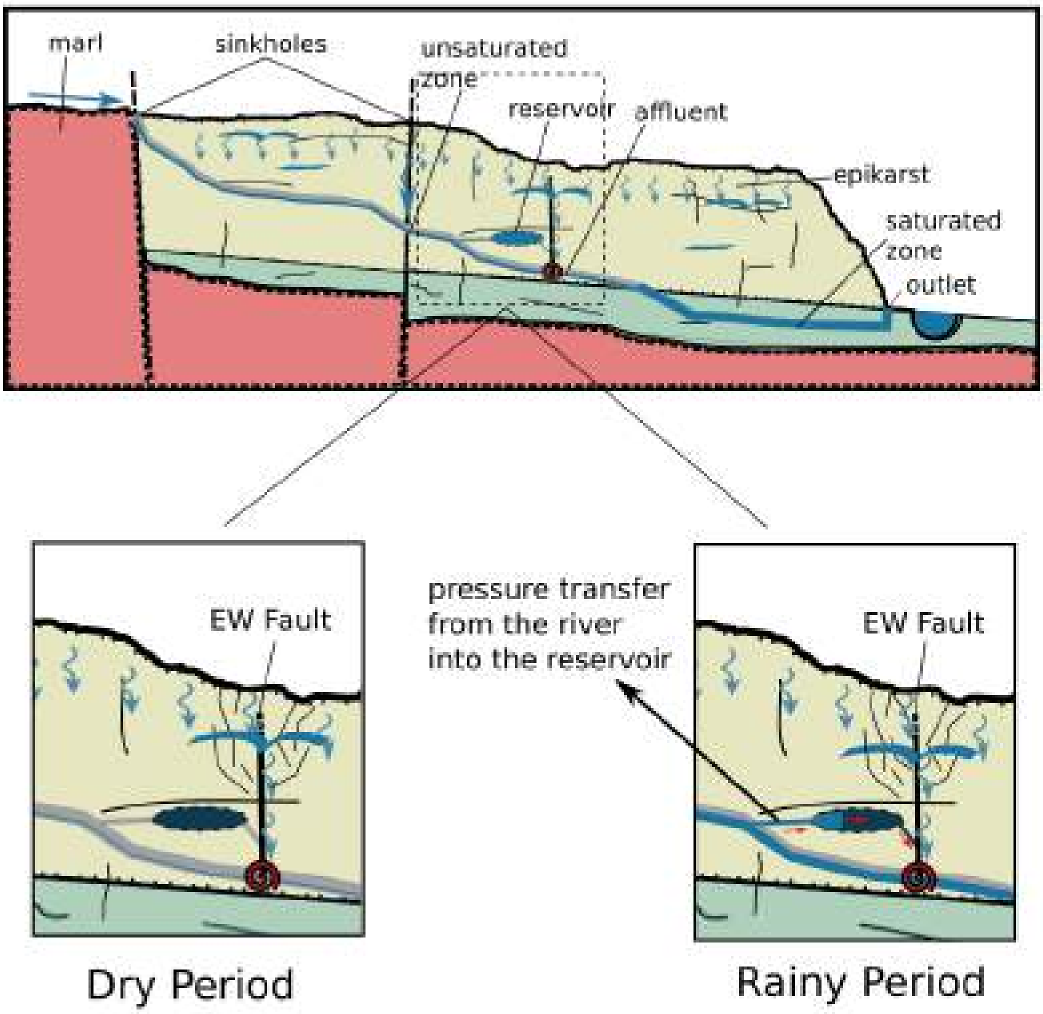


Fig. 3.28 Simplified sketch illustrating the main aspects of the studied karst system: main conduit, saturated and unsaturated zones, impermeable layer, faults and fractures. The studied affluent is represented in a red spiral. Two zooms of the dashed rectangle are presented for two different cases: rainy and dry periods.

3.2.6 Conclusion

We used a Fiber Optic Distributed Temperature Sensing (FO-DTS) to monitor temperature changes in the unsaturated zone of a karst aquifer during a case study carried out in the Fourbanne karst conduit. Air/wall of the cavity and water temperatures were measured and complemented by precipitations, water height and atmospheric temperature. We observe that underground river temperature increases of 2°C with the atmospheric temperature during the 6 months of recording. The air/wall temperature of the cavity is still affected by the atmospheric temperature at 25 m depth, because of a draught from the artificial well. Deeper in the network the air/wall temperature of the conduit is rather controlled by river temperature.

We investigated in detail the response of the river and one of its affluent to atmospheric changes. The analysis was based on the 6 months FO-DTS monitoring period and more recent electrical conductivity measurements. Outside rainy periods, both river and affluent are supplied with water from slow infiltration processes but from different origins. A local, probably epikarstic reservoir is feeding the affluent. In addition, rapid infiltration occurs into the affluent from a reservoir/sink connected to the main river conduit when the water height in the river reaches a threshold of about 0.45 m. The water within this reservoir is pushed out into the affluent by the water river flowing into it during precipitation events.

In this study, we show that FO-DTS is a powerful tool for environmental geophysics applications. We could characterize the recharge mechanisms of a karst aquifer and interactions between its structural components, in particular the main conduit and the surrounding fracture network. Besides, this study illustrates the necessity of continuous and long-term monitoring for such complex medium.

Chapter 4

Seismic Noise Autocorrelation to Track Groundwater flows within the watershed

Keypoints:

- Seismic noise autocorrelation was used to monitor groundwater in karstic media.
- Noise recorded from 60 nodes over 4 month was used to investigate the groundwater circulation on the watershed scale.
- Noise recorded from an underground permanent seismomter was used to monitor the long term groundwater fluctuation.

4.1 Introduction

Studying karst aquifers is vital due to their critical role in water resource management. These complex underground systems require special attention to ensure the sustainable use of freshwater resources, prevent contamination, and minimize risks associated with infrastructure development and natural hazards like sinkholes. Karst aquifers are hence crucial but challenging to study due to their high heterogeneity and inaccessibility.

The world is currently witnessing a paradigm shift in the field of environmental monitoring, with cutting-edge techniques emerging to better understand and manage the Earth's natural resources such as geophysical prospection. Seismic cross-correlation is a powerful technique used in geophysics to monitor changes in the physical properties of soil (Sens-Schönfelder and Wegler, 2006). This method allows the retrieval of the Green's function (or impulse response) between two sensors (Shapiro and Campillo, 2004) and can be used for environmental imaging and monitoring.

For enhanced sensitivity in retrieving velocity variations from cross-correlations, it is more advantageous to examine the temporal variations in the coda of the correlations rather than the direct waves (Poupinet et al., 1984). The coda is used in such application because of its stability in time compared with the direct arrival waves. This method can also be applied in a single-station approach using autocorrelation, which can be computed for same or cross-component signals.

Cross-correlations and autocorrelations methods are more and more used to monitor meteorological and hydrological processes, based on the fact that a water saturated medium will lead to a decrease in seismic velocity. Actually, many studies have shown a clear negative correlation between the groundwater level and the dv/v (e.g. Voisin et al., 2017). Similarly, this negative correlation exists between rainfall and dv/v, potentially with a time lag corresponding to the time of infiltration depending on the local geology (e.g. Fores et al., 2018). Other factors can influence the correlation between the velocity variation and rain, as well as the potential time shift between rain and response. These factors will depend on the physical and mechanical properties of the ground, such as mineralogy, permeability, porosity, which can be irregularly distributed in heterogeneous sites such as karst aquifers. Karst aquifers are characterised by a network of interconnected fissures, fractures and conduits, making them highly heterogeneous and anisotropic. In this study, we will be using seismic noise autocorrelation to have a better understanding of karst's response to precipitations and resulting water infiltration at the scale of a watershed. This study benefits from original data recorded at the Fourbanne watershed (Jura, see Chapter 2). A seismometer located in the underground conduit was used for multi-annual monitoring in order to study long term trends, and a dense seismic array of 60 nodes was used for a 4 months monitoring over the whole watershed.

4.2 Geological setting and instrumentation

4.2.1 Fourbanne karst aquifer

Karst is formed through the process of dissolution within carbonate rocks. Water infiltrates these rocks, carrying carbon dioxide (CO_2), and flows through fractures and weak zones in the rock (Ford and Williams, 2007; Hartmann et al., 2014). As water passes through, it dissolves the rock creating a heterogeneous fractured matrix, thus an underground drainage network. The described phenomenon is denoted as karstification, and depends essentially on the nature of the rock, its morphology, its state of fracturing and climatic conditions. We can distinguish different types of water flows depending on the porosity: fast local flows at the level of sinkholes and slow to fast flows through the fractures and/or rock matrix, depending on the level of porosity of the medium. The studied site is the Fourbanne watershed (Fig. 4.1), a karst aquifer spanning over 30 km² located in the Jura mountains, eastern France. Its delineation has been established based on geological structures and tracer studies conducted in the region (Charmoille, 2005; Cholet, 2017).

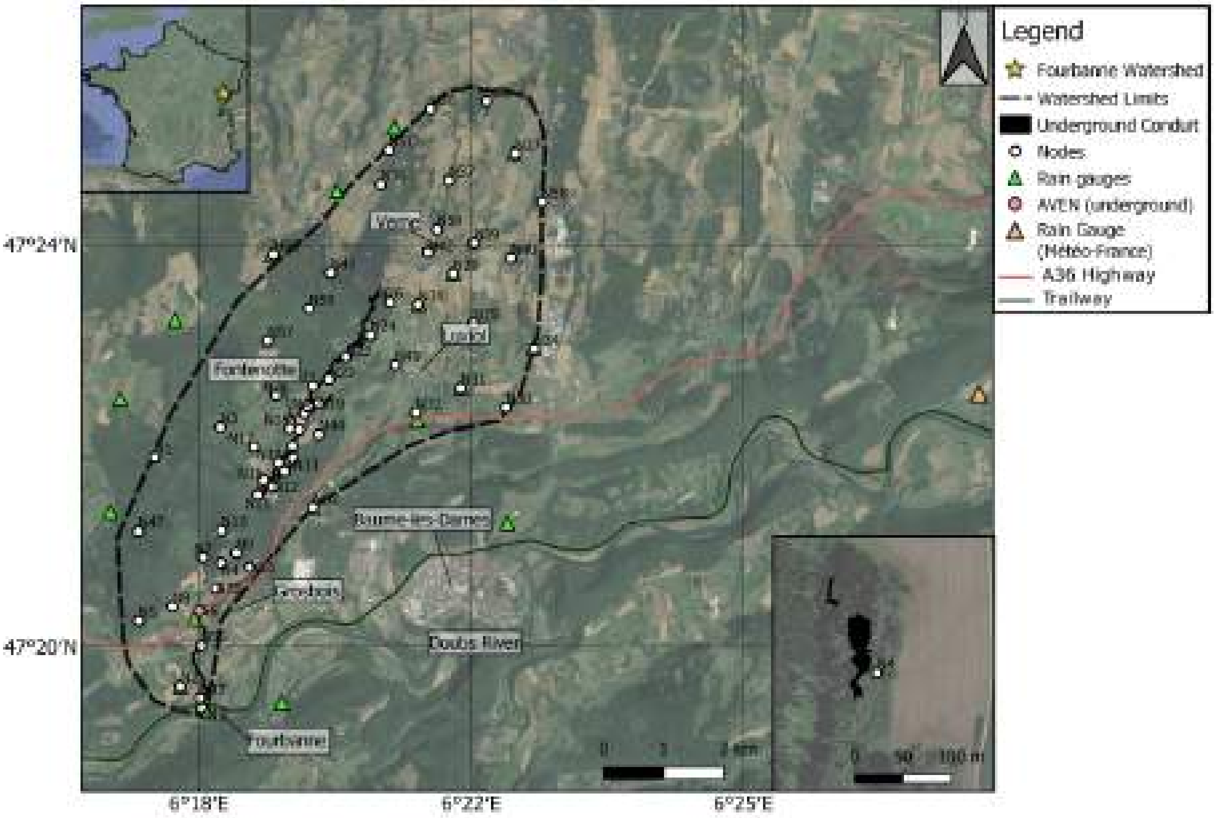


Fig. 4.1 Satellite Google map of the studied area and instrumentation and zoom in around node N04 and two caves found by the speleologists in Grosbois.The topography of the underground conduit has been provided by D. Motte (ASDC) and M. Mendonça Dias.

The Fourbanne catchment basin is located between two major tectonic features of the western section of the Jura Mountains: the Avant-monts anticline overlaps the Haute-Saône Plateau through the Ognon Fault (to the north), and the external zone of the Jura overlaps the Avant-monts at the level of the thrust zone called the "faisceau bisontin" to the south (see Fig. 2.4 from Chapter 2, Charmoille, 2005). It is a component of the JURASSIC KARST hydrogeological observatory, established in 2014 (Cholet et al., 2017), and the French SNO KARST network (Jourde et al., 2018), and hydrogeologically monitored since. It is primarily composed out of Middle Jurassic limestones, intertwined by a network of north-south-oriented faults. These faults bring the lower Jurassic and Triassic marl layers to the forefront (Fig. 4.2). The latter results in zones having limestone outcrops and others with marl outcrops, the first contributing to infiltration through rock matrix and the second to overland flow. The aquifer hosts an underground conduit: about 9 km were explored in the northern part of the conduit, accessible through a 20 m drilled shaft at Fontenotte, and 3 km explored in the southern part of the conduit. The northern section is the unsaturated part of the conduit, and the southern section is saturated : the limit could not be precisely defined yet (Fig. 4.2).

4.2.2 Long-term and temporary networks for seismic, hydrogeological and atmospherical monitoing

The instrumentation used in this study is shown in Figs. 4.1 and 4.2. Long term seismic monitoring has been done using the seismic station AVEN located in the underground gallery at 20 m depth. AVEN is a part of the long-term regional seismic network, JURAQUAKE (see Section 2.2.2 in Chapter 2), deployed in eastern France since late 2018. The station is positioned at a distance of 50 m from the river, buried 50 cm into the sediment of the cave to ensure proper coupling. Additionally, a hydrological probe (CTD) deployed since 2014 measures the underground river water's electrical conductivity, temperature, and height. Temporary experiments were also conducted in 2021 and 2022 in the frame of the *Sismeauclim* project (see Section 2.2.4 in Chapter 2). A dense seismic array composed of 60 nodes from the French national pool of portable seismic instruments SisMob- RESIF (INSU-CNRS) was deployed for 4 months (December 2021 to March 2022) all over the watershed. Initially, we planned to start the experiment in November, in order to catch "dry days" and the first winter rains. However, for logistical reasons, it could not start before the 29th of November. The nodes' positions were chosen in order to cover all of the following conditions: (1) the main geological formations (marls and limestone), (2) the area bordering the explored underground conduit, and (3) the whole watershed surface. Besides, 20 tipping-bucket rain gauges were deployed in summer 2021 in the watershed in order to quantify water volume intakes and detect its potential spatial variability within the watershed.

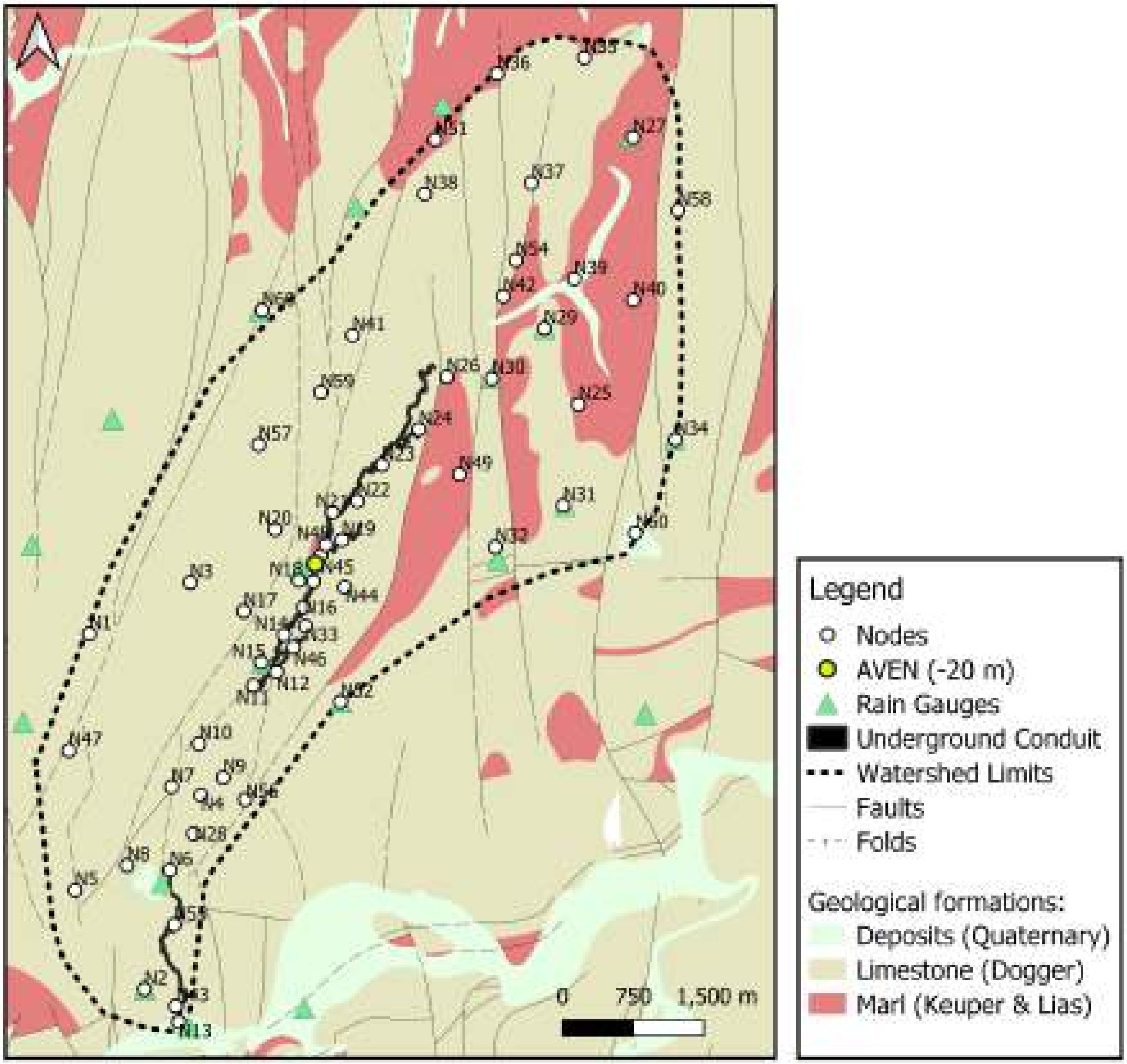


Fig. 4.2 Geological map of the Fourbanne watershed (main formations, faults and folds), showing also its limits (black dotted line), the underground conduit (black line) and the location of the instruments used in this study.

4.3 Data and Method

4.3.1 Hydrological and meteorological data

In this chapter we used the data collected from the CTD in terms of the underground river water height time series. Long term precipitations measurements come from a closeby Météo-france station (Branne, see Fig.4.1). The Branne station is located about 14 km at the east of the watershed's outlet. Cholet (2017) have found out that the Branne station seems to be relevant for analyzing precipitation records in the Fourbanne karstic system: they compared the rain cumulative over a year between the Campbell and Branne station, which gives relatively close results. However, Fontenotte slightly accumulates more rain (around a 50 mm difference for a year of data) as shown in Figure II.30 of Cholet (2017).

Time series obtained from the 20 rain gauges were used for the short term analysis. Due to the highly demanding maintenance of the rain gauges, data might contain gaps or outlier values. In this study, an average value computed from all the rain gauges has been mainly used, resulting in one time series for 1D velocity changes interpretation. Figure 4.3 is a plot of precipitation, collected from Météo-France station between the 1^{st} of January 2018 til the 1^{st} of January 2023 and from the rain gauges between the 1^{st} of October til the 1^{st} of April 2022, and the underground river water height for the same periods. The shown period for the rain gauges presents two months before the seismic nodes acquisition period til the end of it.
The objective of showing hydrological data prior to the seismic acquisition is to showcase the occurrence of rain events which might affect the watershed's saturation, prior to the seismic recordings. Figure 4.3a shows that rain events occur throughout all seasons. The water height records a minimum of 0.3 m and a maximum of 1.75 m (Figure 4.3b). In Figure 4.3c, we can see that after December 1^{st}, three main rain events occurred, inducing water height increases of 0.7 m (see the red frames in Fig 4.3). These periods are from the 2^{nd} til the 11^{th} of December, from the $25^{th}$of December til the 11^{th} of January and from the 28^{th} of January til the 26^{th} of February. We can notice that not all of the rain events generate a peak in the water height: water height peaks depend on whether the rain event follows a mostly dry period (no peaks might be recorded) or a mostly wet period (direct response might be exhibited). After rainfall events, water height can take up to 10 days to get back to its base level, which is the water height prior to the rainfall.

4.3.2 Seismic data

4.3.2.1 Data availability

We used long-term seismic data collected at station AVEN between mid-December 2018 and January 2023 as well as data recorded by the 60-nodes array between December 2021 and March 2022 for a spatio-temporal monitoring of the watershed during a rainy season. The data from AVEN are continuous and do no present any gaps. There is a total of 5-days gap for the data collected from nodes due to the maintenance scheme, these gaps occurred the 4^{th} of January 2022, 1^{st} to 3^{rd} of February 2022 and the 1^{st} of March 2022. In addition, few stations have gaps in data due to technical problems. The data availability for the nodes during the studied period is presented in Fig. 4.4.

4.3.2.2 Noise content and origin

The studied area can be mainly identified as a rural area. It's important to note that rural areas generally have lower levels of anthropogenic noise compared to urban or industrial zones, due to lower number of inhabitants and the type of activities. However, the presence of roads, railways, or industrial activities, can still introduce significant noise contributions. Noise sources within and around the watershed are mainly anthropogenic. They are generated in crossing towns: Baume-les-Dames (5200 inhabitants), Grosbois (250 inhabitants), Fourbanne (175 inhabitants),

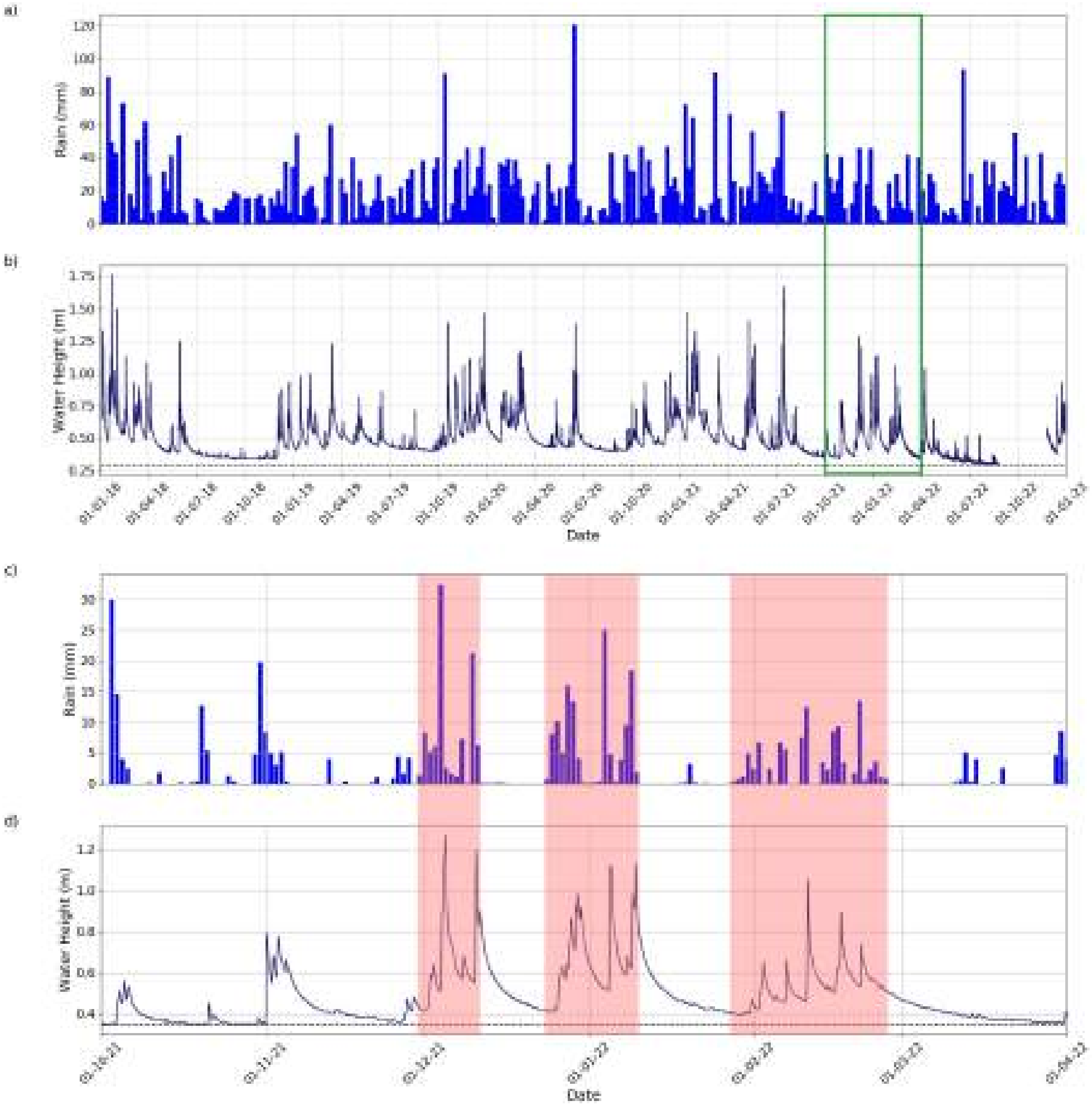


Fig. 4.3 Atmospherical and hydrogeological data used in this chapter. a) Weekly precipitations measured at Branne Météo-France station. b) Water height measured at the Fontenotte CTD. The dashed black line in represents the minimal water height during the considered time period. The green frame indicates the period represented in (c) and (d). c) Daily rainfall obtained by averaging the values recorded on the deployed rain gauges between October 2021 and March 2022. d) Water height measured at the Fontenotte CTD corresponding to the same period as in (c). Pink dashed areas indicates the 3 main rain events occurring during the temporary seismic experiment.

Luxiol (160 inhabitants), Verne (125 inhabitants) and Fontenotte (60 inhabitants) (see Fig. 4.1). The main human activities in these towns are construction, quarrying and agriculture, in terms of both crop and livestock production. Other recurring noise sources are vehicles passing on near departmental roads (D683 passing through the southern section of the watershed, D50 tracing the eastern limit of the watershed, D271 passing through Fontenotte etc...), a highway (A36, passing through the south-eastern and southern sections of the watershed, see Fig. 4.1), and a

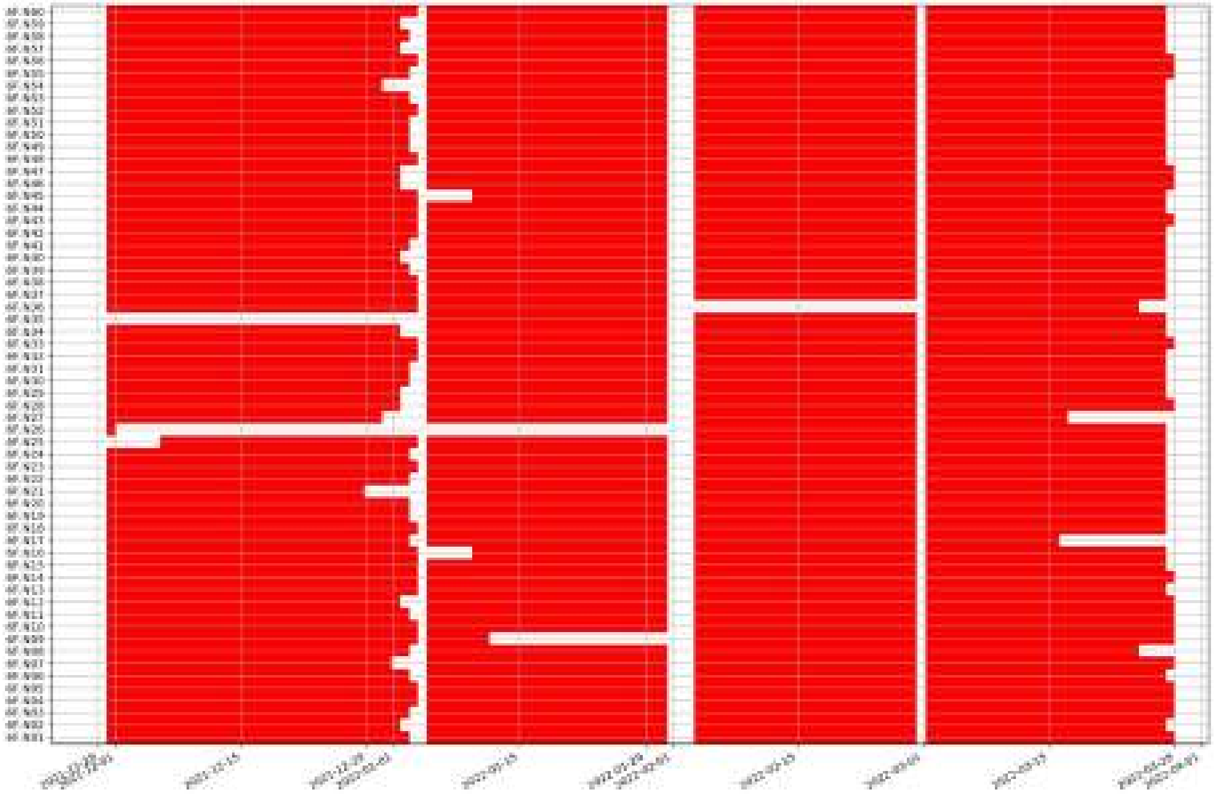


Fig. 4.4 Availability of seismic data for each node of the seismic array during the studied period. The red squares indicate the presence of data, and white squares the lack of data.

railway (passing through Baume-les-Dames, with 3 to 4 trains passing per hour). Water flow in both surface river (Doubs river) and underground river passing through the karst conduit also generates a constant noise, more emphatic during rainy seasons. Additionally, wind turbines located at the northern part of the basin also induce ground vibrations.

An example of seismic signals recorded on the ZZ component by AVEN and nodes 16 and 49 is presented in Fig. 4.5. N16 is located above the conduit unlike N49 (see Fig. 4.1 for stations location). We observe that the noise recorded at AVEN has lower energy than the noise recorded at N49 and N16, since it is located underground, far from anthropogenic noise sources. In this data extract, spanning between the 23^{rd} and the 27^{th} of December, two regional earthquakes occurred on the 24^{th} and 25^{th} of December, SE of Montbeliard (ML 4.1 and 3.8, respectively; arrows on Fig. 4.5). Seismic arrivals related to these events can be easily picked on all stations components indicating a good signal to noise ratio (see Appendix C for more details on these earthquake). During this period, there were also precipitations, which direct effect (raindrops hitting the ground; Dean, 2018) is visible on the seismic traces of the nodes (Fig. 4.5), with an increase of the seismic energy starting the 25th of December, around 4 am. Due to its underground position, AVEN signal is isolated from the direct rain effect.

Noise sources can be investigated in more detail by computing spectrograms Mordret (2014). The latter are presented in Figs. 4.6, 4.7 and 4.8 for AVEN, N16 and N49, respectively. Anthropogenic activity (daily activity versus quiet nights) is visible on all stations, in particular on the nodes. There is also less energy during days-off (December 25^{th} and 26^{th}, which correspond to a

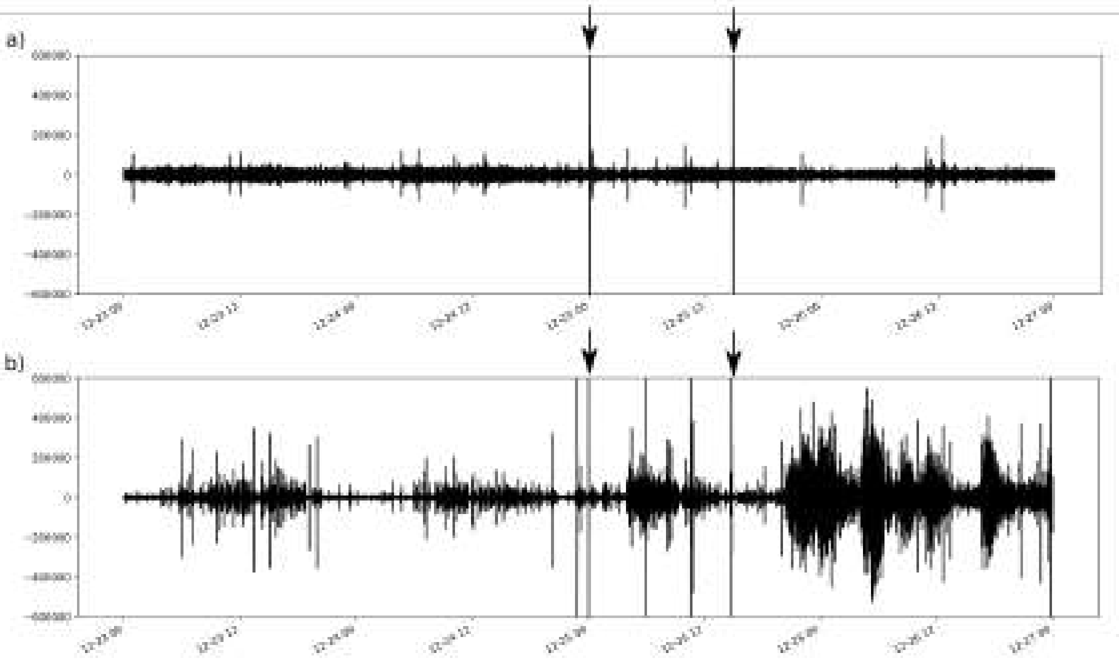


Fig. 4.5 Comparison between hydrological data and seismic signals. a) Hourly rain values between the 23^{rd} and 27^{th} of December 2021. b) Underground river water height corresponding to this period. c) Vertical components of the seismic traces recorded at AVEN d) N16, and e) N49 between the 23^{rd} and the 27^{th} of December 2021. The arrows indicate the earthquakes that occurred on the 25^{th} of December 2021. The signals were detrended, demeaned and filtered between 0.01 and 50 Hz.

Saturday and a Sunday). More energy is manifested on the horizontal components than on the vertical component, which is a common observation for anthropogenic generated noise (Acerra et al., 2002).

A direct effect of precipitations is not visible on AVEN data (Fig. 4.6). However, as described earlier in Chapter 3, an increase in the seismic energy due to the underground water height increase can be pointed out. Actually, once the water height reaches 0.5 m, on the 26^{th} of December at about 4 am, we observe clear noise amplitude increase across a large range of frequencies, and particularly at high frequencies, between 20 and 50 Hz. Regarding N16 and N49 (Figs. 4.7 and 4.8), we observe an increase in the seismic noise amplitude as soon as precipitations are significant (the 12/25/21 4 am). The direct effect of noise can be better distinguished on the vertical component, as horizontal components are more affected by anthropogenic noise. No clear relation with underground water height can be detected on the nodes spectrograms, even on N16 which is above the conduit.

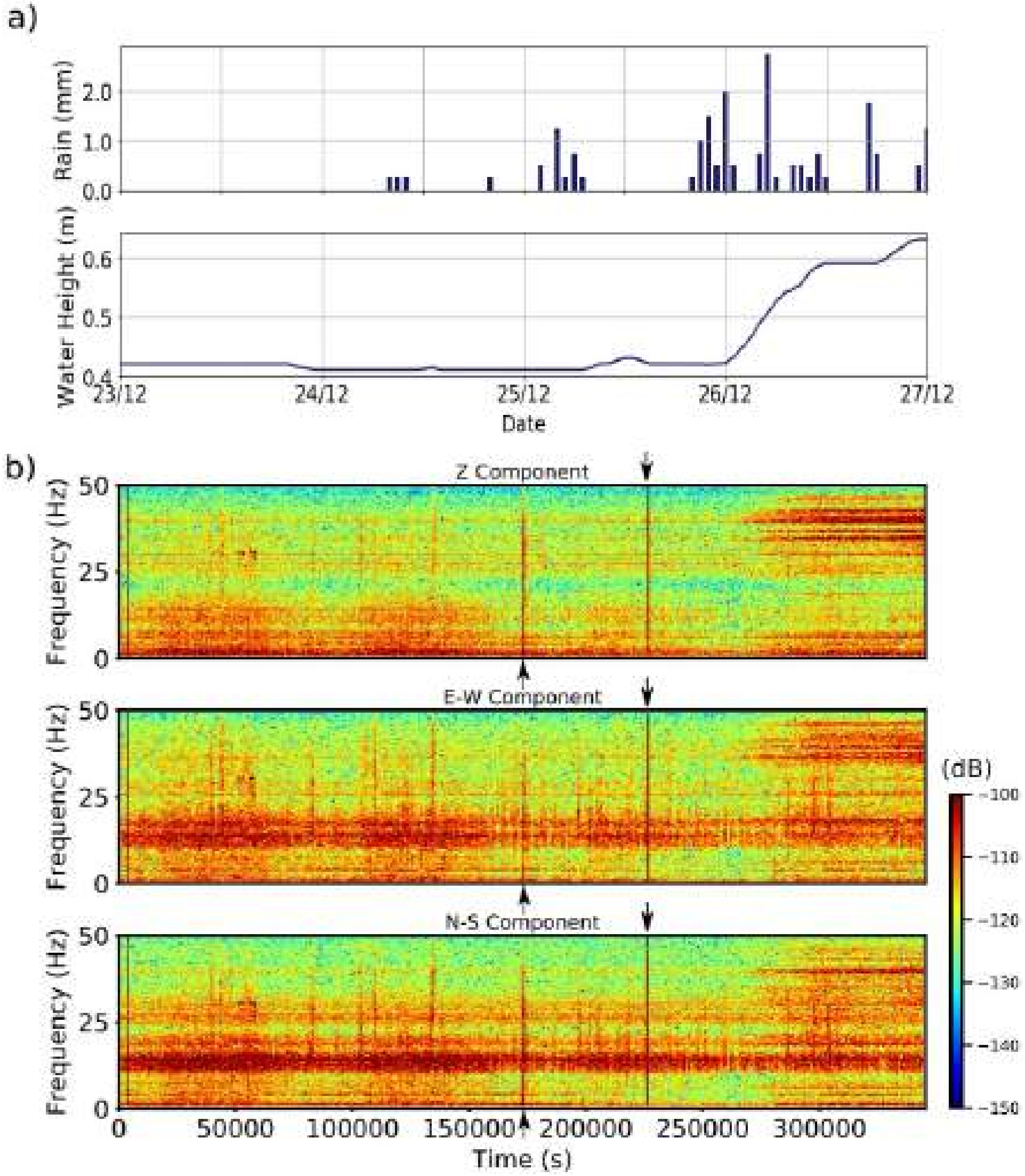


Fig. 4.6 Comparison between hydrological data and seismic spectrograms. a) Hourly rain values between the 23rd and 27th of December 2021 the underground river water height corresponding to this period. b) Spectrogram of the vertical (Z) and horizontal (E-W and N-S) components for data collected from AVEN for the same period as a), filtered between 0.01 and 50 Hz. The two arrows in each spectrogram indicate the two earthquakes that occurred on the 25th of December 2021.

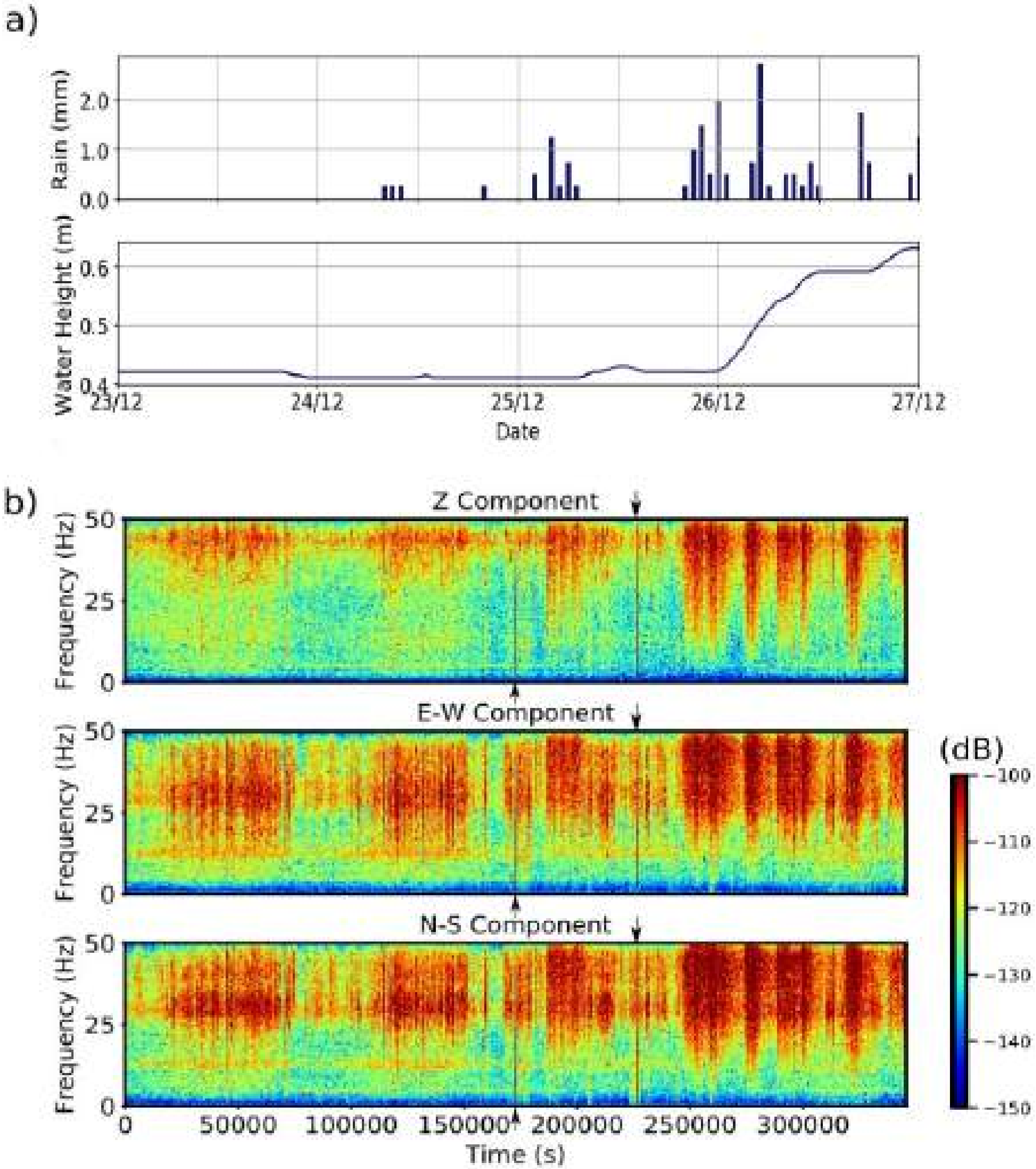


Fig. 4.7 Comparison between hydrologic data and seismic spectrograms. a) Hourly rain values between the $23rd$ and $27th$ of December 2021 and the underground river water height corresponding to this period. b) Spectrogram of the vertical (Z) and horizontal (E-W and N-S) components for data collected from the Node 16 for the same period as a), filtered between 0.01 and 50 Hz. The two arrows in each spectrogram indicate the two earthquakes that occurred on the $25th$ of December 2021.

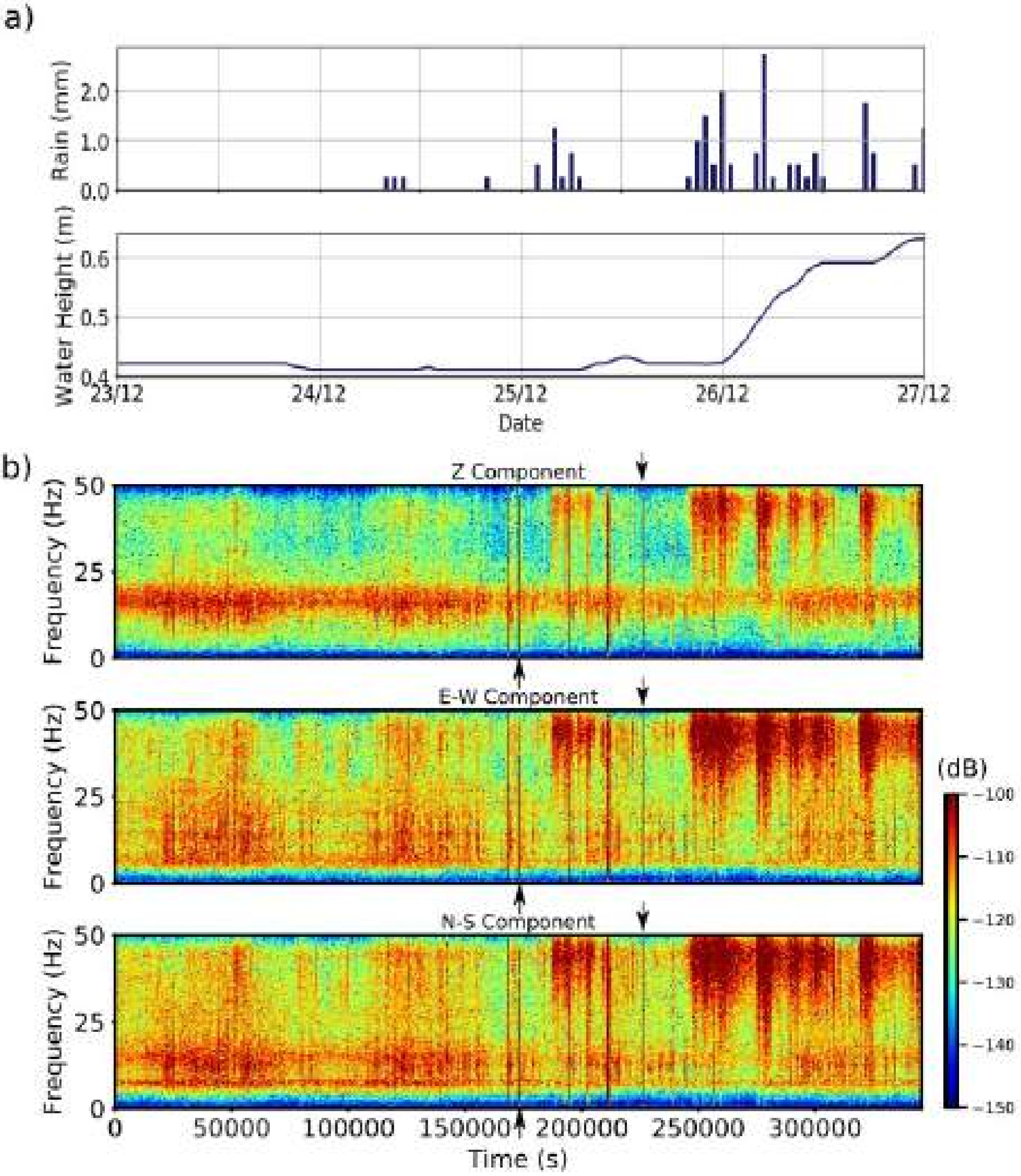


Fig. 4.8 Comparison between hydrologic data and seismic spectrograms. a) Hourly rain values between the 23^{rd} and 27^{th} of December 2021 and the underground river water height corresponding to this period. b) Spectrogram of the vertical (Z) and horizontal (E-W and N-S) components for data collected from the Node 49 for the same period as a), filtered between 0.01 and 50 Hz. The two arrows in each spectrogram indicate the two earthquakes that occurred on the 25^{th} of December 2021.

4.3.3 Noise single-station cross-components correlation

Relative seismic velocity changes (dv/v) were determined from seismic noise correlation, using a methodology similar to that described in the works of Lecocq et al. (2014) and De Plaen et al. (2019). The processing has been done using MSNoise (http://www.msnoise.org/), which is a Python based software used for the analysis of seismic noise. It includes several steps : (1) the preprocessing, (2) the cross-correlation functions computation (CCFs), (4) the reference CCF computation, and (3) the dv/v estimation. Similar workflow, but separated analysis, were applied to the data collected from both the broadband station AVEN and the 60 nodes array.

The scheme starts with signal preprocessing: for each station, all traces containing data for the same day are merged and split to obtain the most complete and continuous chunks possible for each day. The gaps within a day are filled through linear interpolation for gaps smaller then 10 seconds, or else the gaps will be filled with zeros. The one day seismic traces were then resampled to 50 Hz and filtered between 0.01 and 20 Hz. Traces were then partitioned to 30 minutes-windows with 50% overlap. The computation is done for each pair, on the different components to compute and then on the different filters for each 30 minutes-window. Windsorizing was done at 3 times the RMS to remove outlier peaks for more stable signals. Spectral whitening was also applied: the amplitude of the signal is set to 1 for all frequencies, so only the phase of the signal remains (Hobiger et al., 2014).

In this study we used Noise Single-station Cross-components correlation Functions (NSCFs) computed between different components (EN, EZ, NZ) of each station. The cross-correlation analysis, between different station pairs, will be the subject of later works. One of the advantages of the single-station correlation is that the measured velocity changes can be considered as the velocity variation under the station's position. It can also better deal with high frequencies compared to cross-correlation using larger distances pairs, due to high frequency noise attenuation with distance. De Plaen et al. (2016) have shown that dv/v obtained from single-station cross-component correlations are particularly sensitive to saturation changes in the ground. The authors observe a clear decrease in dv/v in response to infiltration of rain. We chose to do cross-components correlation instead of same components correlation since the NSCFs are more stable in time and present a better signal-to-noise ratio De Plaen et al. (2016), compared with same components CCFs.

Calculations were performed using HPC resources from DNUM CCUB (Centre de Calcul de l'Université de Bourgogne). We obtained 21600 daily NSCFs for the 60 nodes array (4 months of data) which represents 4 GB of data. Data processing for one frequency band takes approximately 72 hours. We obtained 8760 daily NSCFs for AVEN (4 years of data) which represents 1 GB of data. Data processing for one frequency band takes about 24 hours.

Then, we computed the reference cross-component autocorrelations functions (RNSCF) as the average of the NSCFs over the studied period. The period selected for the RNSCF depends first on the purpose of the correlation analysis, since the velocity variation is computed relatively to the chosen reference, and second the reference should correspond to a stable period, a condition that might be fulfilled by choosing the stack of the whole recording period as reference. A

RNSCF is obtained for every station, and for each of the components and the chosen frequency filters. An example of the RNSCF for the EN component of the node N06 is given in Fig. 4.9. We can see an asymmetry between the causal and acausal parts of the signal, that corresponds to a non homogeneous distribution of noise sources, with a more important energy coming from a side of a sensor (Froment, 2011). Similar behavior is observed for other components and other nodes.

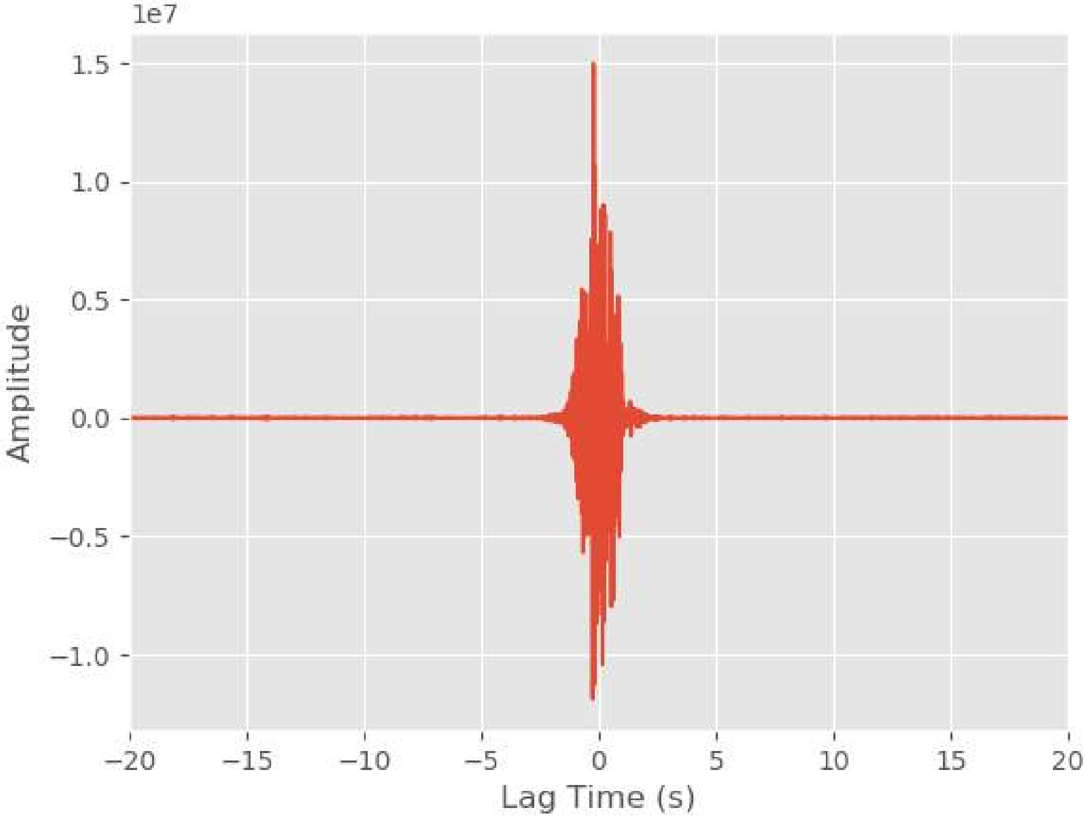


Fig. 4.9 Reference function, corresponding to all days stack, for the EN component cross-correlation for the node N06.

Finally, we estimated daily velocity variations (dv/v) using the "doublets" or Mowing-Window Cross-Spectral (MWCS) method (Poupinet et al., 1984). The MWCS aims at assessing the similarity between the daily correlation function and the reference correlation function within each window. It consists in identifying daily phase shift between the NSCFs and the RNSCF, that corresponds to a time shift in the time domain between selected windows. The minimal coherence was set to 0.7 and the maximum error for the travel time changes computation to 0.1 s. Windows resulting with a coherency less than the minimal coherency, and error larger than the maximum error, are not taken into account in the following. The next step is to use the time windows that fulfill the previous conditions in a linear regression scheme between lags and corresponding time shifts. The linear regression was performed for each day and the daily time shift is thus equal to the slope (dt/t) of the linear regression. The change in the seismic velocity is finally deduced from the travel time variation (Poupinet et al., 1984): -dt/t=dv/v. Negative and positive sides (lags) of the NSCFs were used. The method was applied in the coda of the NSCFs, on time lags between 10 and 50 s, which aren't affected by direct waves. Figure 4.10 is an illustration of 5-days stacked NSCFs obtained for node N06, EN component, between the 7^{th} and the 13^{th} of January 2022. NSCFs exhibit coherent phases in the coda. Coherent phases refer to specific arrivals or wave packets that exhibit consistent waveforms and travel paths and they can provide insights into the structure and properties of the subsurface medium (Froment, 2011).The coda typically consists of scattered waves and reflections that have traveled through complex paths within the Earth's subsurface. It contains valuable information about the subsurface properties and can be analyzed to infer the characteristics of the seismic source and the surrounding medium due to their sensitivity to heterogeneities.

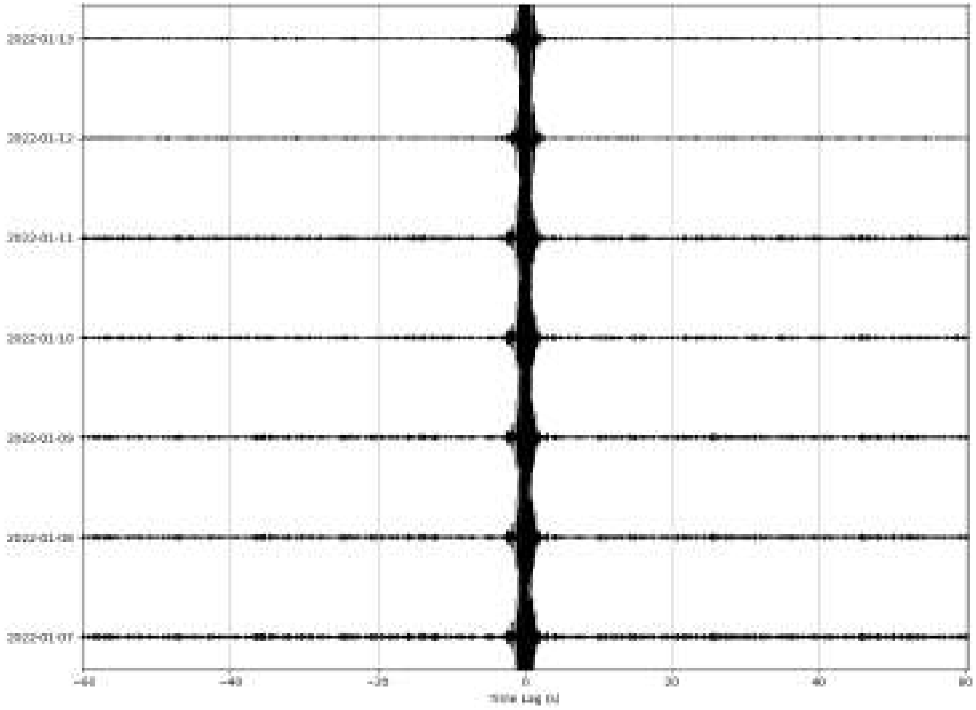


Fig. 4.10 Daily NSCFs for the EN component for the Node N06, between 07/01/22 and 13/01/2022, filtered between 1 and 10 Hz, with a 5 days stack.

We applied the method on stacked and band-passed filtered NSCFs. Regarding the filtering, we performed different tests and finally focused our analysis on 2 different frequency ranges : 2-4 Hz and 8-10 Hz. Stacking NSCFs enables to smooth the data : the stack on a specific date corresponds to the mean cross-component correlation function of the n-days before. The dv/v values can be stacked, which helps in improving coherency.For example, for a 5 days stack, the dv/v affected at day *j* correspond to dv/v stack from day *j-4* to day *j*. We performed 20 days stack for AVEN and and 5 days stacks for the nodes array. Results are described in the following section.

4.4 Velocity variation analysis

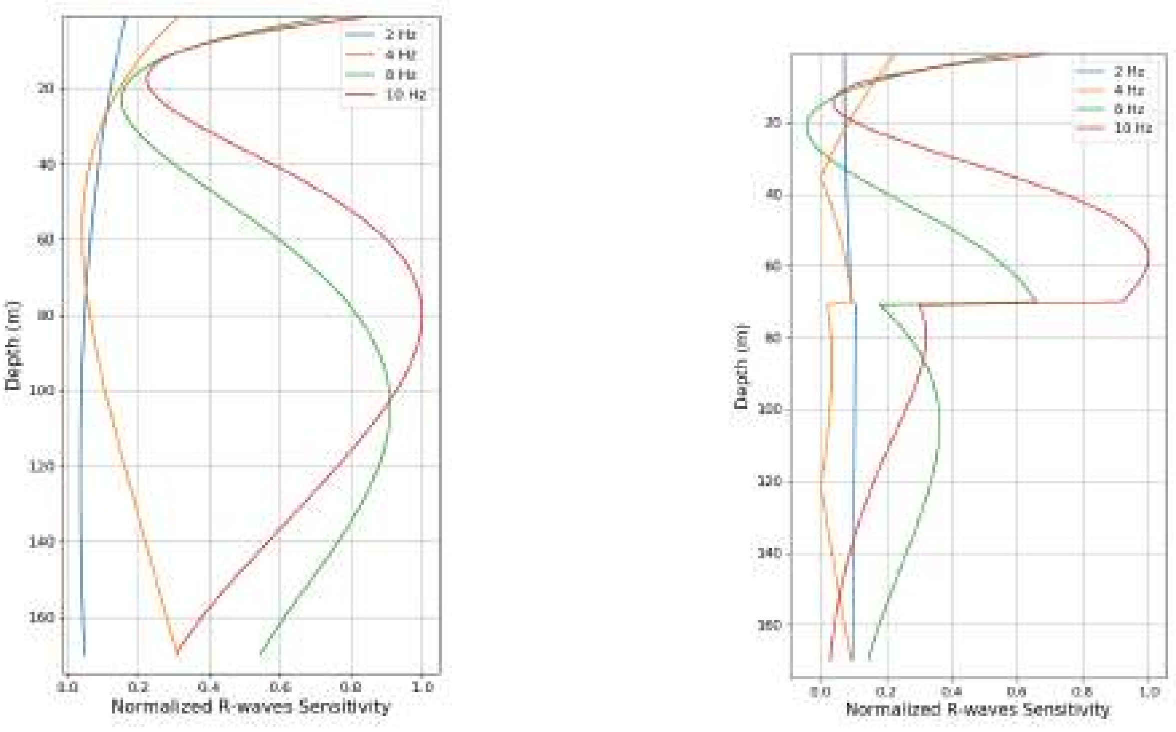
4.4.1 Depth of velocity changes

Rayleigh waves (R-waves) are often inferred to be reconstructed through early codas of the correlation functions due to their predominant energy distribution, as demonstrated in studies such as Mainsant et al. (2012b); Obermann et al. (2013). In order to comprehend variations in seismic wave velocity, it becomes essential to assess R-wave sensitivity to eventually interpret on the depth of the velocity change. The sensitivity is frequency dependent: 2 frequency bands have been investigated, 2-4 Hz and 8-10 Hz, corresponding to the filters used in the single-station correlation (see section 4.3.3). This has been done by considering a one-dimensional medium and using the Fortran based code Surf96 developed by Herrmann (2013). We performed the analysis using two velocity models (MOD1 and MOD2; Table 4.1), which are representative of the geological variability across the watershed (see Section 2.1 in Chapter 2 for the geological description if the site and Fig 4.2). The one layer velocity model MOD1 has been used for the north-eastern part, around Verne, which is mainly composed of impermeable marls (Lias and Trias). MOD2 includes the main geological formations encountered in most of the area : 70 m of fractured and karstified permeable limestones (or Malm) overlying 100 m of lower Jurassic impermeable marls (or Lias) (Charmoille, 2005). The P-wave velocities (Vp) and densities were taken from the Riniken borehole data located in the Jura Mountains (Malz et al., 2015). The S-wave velocities (Vs) were calculated from Vp using a poisson's ratio of 0.25 (average value for limestones, Barton, 2009). The quality factor was set to 20 (average value for calcareous rocks with clay content, Barton, 2009).

The results for the sensitivity kernels for shear wave velocity are plotted in Fig. 4.11. At 2 and 4 Hz, for both models, we observe weak but relatively constant R-waves sensitivity from surface to 170 m. Thus we can expect that potential dv/v variations observed in this frequency range would be related to changes occurring within the Fourbanne aquifer. At 8 and 10 Hz, R-waves sensitivity is relatively high, close to the surface (0 to 10 m depth) for both models. Sensitivity curves present peaks at 80 m and 55 m depth for MOD1 and MOD2, respectively. Thus, the high frequency band most likely allows to investigate velocity variations at shallow surface and at greater depth, within the whole aquifer. In addition, smaller and slow paced velocity variations can be expected for the lower frequency band [2-4 Hz], compared to variations for the higher frequency band [8-10 Hz]. In fact, lower-frequency band waves tend to average out variations over larger volumes of Earth's subsurface, they are characterized by their ability to detect larger-scale variations in subsurface properties, while higher-frequency waves are more sensitive to smaller-scale features.

Table 4.1 Simplified 1D velocity models used in the R-wave sensitivity analysis, characterising the area. MOD1 corresponds to the NE part of the watershed (one layer of marls), and MOD2 includes the aquifer (70 m limestones above 100 m of marls).

lightgray	MOD1				
	Thickness (m)	Vp (km/s)	Vs (km/s)	Density (kg/m^3)	Quality Factor
Marls (Lias)	170	4.5	2.65	2.5	20
lightgray	MOD2				
	Thickness (m)	Vp (km/s)	Vs (km/s)	Density (kg/m^3)	Quality Factor
Limestones (Malm)	70	3	1.75	2.3	20
Marls (Lias)	100	4.5	2.65	2.5	20



(a) MOD1

(b) MOD2

Fig. 4.11 Normalised sensitivity kernel at 2, 4, 8 and 10 Hz, computed for 1D velocity models described in Table 4.1. (a) 170 m of marls (MOD1); (b) 70 m limestone above 100 m of marls. Sensitivity is the partial derivative of R-waves phase velocity with respect to the S-wave velocity.

4.4.2 Long-term velocity variations (AVEN broadband seismic station)

After the processing scheme presented in the section 4.3.3, we have obtained time series of the seismic velocity variation for each of the 3 cross-components EN, EZ and NZ and the two chosen frequency ranges 8-10 Hz and 2-4 Hz. We applied the *LoWeSS* regression algorithm on the time series to smooth the data and reduce unnecessary noisy peaks (Locally Weighted Scatterplot Smoothing; Cleveland, 1979). The algorithm is a supervised learning based method used for non-parametric fitting, which doesn't assume that data follow any specific distribution. An example of the raw and smoothed data for AVEN using the *LoWeSS* method is showcased in Fig. 4.12 (EN cross-component, 2-4 Hz): we can see that the smoothed time series reveals the increasing or decreasing tendencies, while leveling outlier peaks, and without creating a time shift between raw and smoothed data.

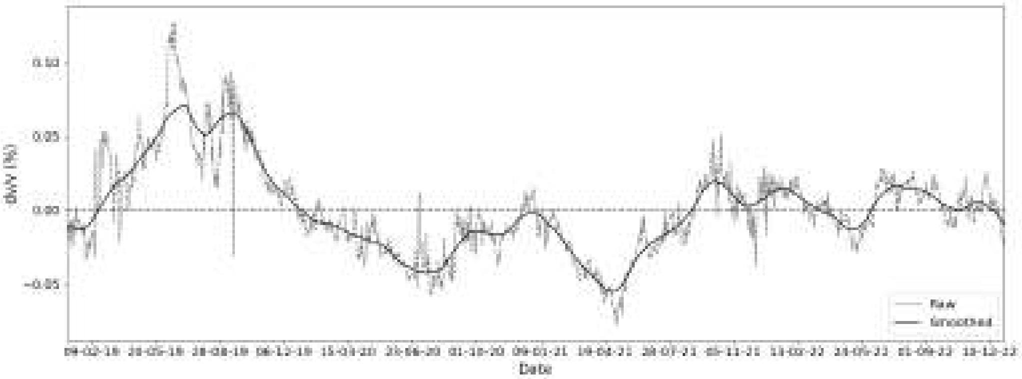


Fig. 4.12 Raw and smoothed velocity variations for AVEN's EN cross-component [2-4 Hz].

The resulting 20 days-stacked 2-4 Hz dv/v time series obtained between September 2019 and December 2022, for the 3 cross-components (EN, EZ, NZ) are shown in Fig. 4.13b,c,d as well as the mean over the cross-components (e). Daily rainfall recorded at Branne station and underground river water height measured at Fontenotte (see Section 2.1.4 in Chapter 2 for details on hydrogeological data) are also plotted in Fig. 4.13a.

We see that it's raining quite regularly over the watershed during the years, and more intensely during winter. In fact, there isn't clear separated dry and rainy seasons, as in the south of France, for example. However, we can point out relatively "dry" time periods, as illustrated in Fig. 4.13 (orange rectangles). During these periods, we observe that 2-4 Hz dv/v for all cross-components tends to increase, while it decreases during more rainy periods. Similar trend is observed for 10-days stack.

Results in the higher frequency range, at 8-10 Hz, are more noisy Fig. 4.14. No clear relationship can be observed between the rain events and the dv/v variations (Fig. 4.14b,c,d). The mean dv/v is flattened out, due to time shifted peaks on the individual cross-components (Fig. 4.14e). It could be due to the location of AVEN station (about 20 m depth), which insulates its from high frequency noise energy.

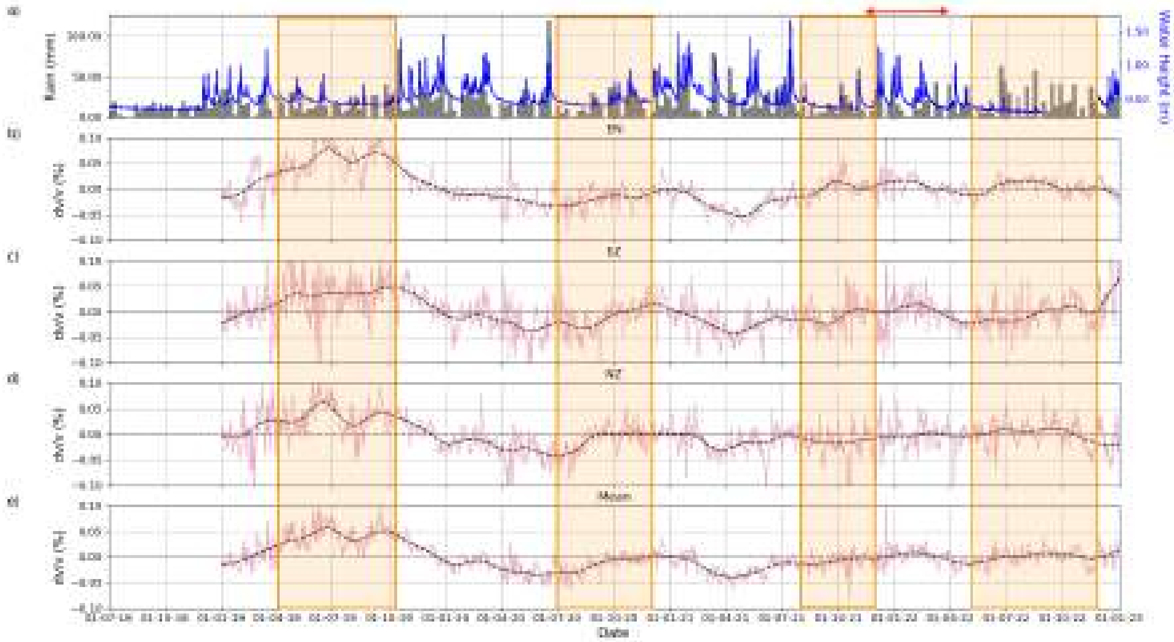


Fig. 4.13 Velocity variations at 2-4 Hz computed at AVEN station between 2019 and 2022. a) Underground water height variation measured at Fontenotte and weekly precipitations measured at Branne. 20-days stacked dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. Black dotted lines correspond to smoothed dv/v. The red arrow indicate the time period of the node array deployment. The shaded sections indicate dry periods.

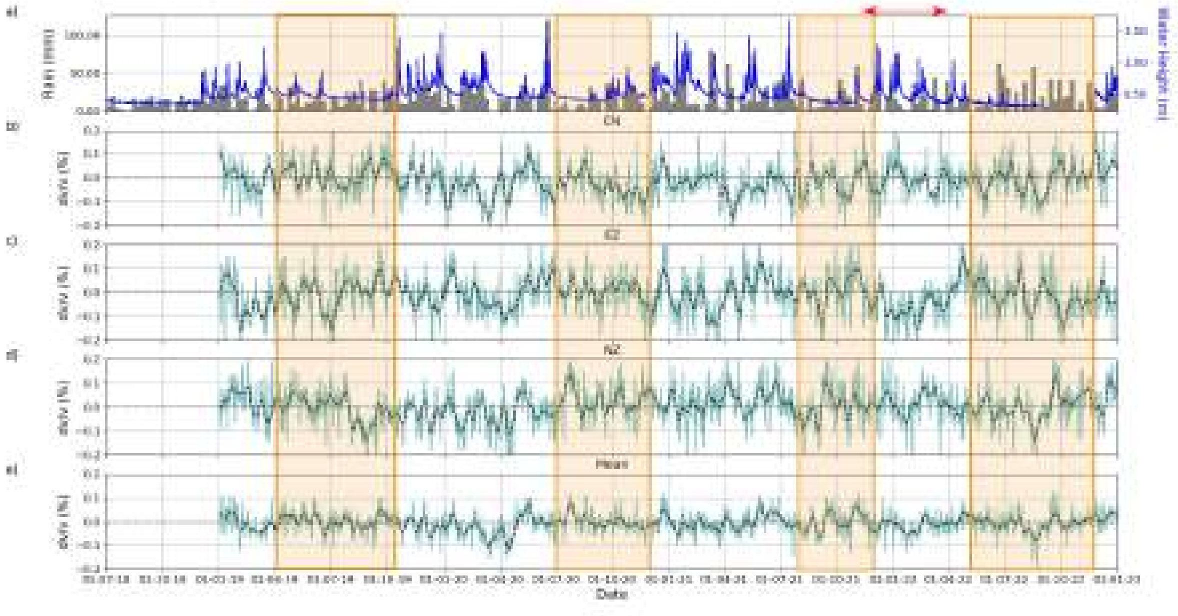


Fig. 4.14 Velocity variations computed at AVEN station between 2019 and 2022 [8-10 Hz]. a) Underground water height variation measured at Fontenotte and weekly precipitations measured at Branne. 20-days stacked dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. Black dotted lines correspond to smoothed dv/v. The red arrow indicate the time period of the node array deployment. The Shaded sections indicate rainy periods.

4.4.3 Four months velocity variations (*Sismeauclim* array)

4.4.3.1 Average velocity variations across the array

After the processing scheme in section 4.3.3, we obtain 4 months dv/v values for each of the nodes' cross-components (EN, EZ and NZ), at 2-4 Hz and 8-10 Hz. To get an overview of the tendencies, we first calculated the mean dv/v over all the nodes, for each cross-component. Averaged 5 days-stacked dv/v computed at 2-4 Hz and 8-10 Hz are plotted in Figs. 4.15 and 4.16b,c,d, respectively. Daily rainfall, averaged over the rain gauges and water height recorded at Fontenotte are also plotted (Fig. 4.15 and 4.16 a). Three main rain periods are observed inducing flood peaks within the underground river (shaded rectangles).

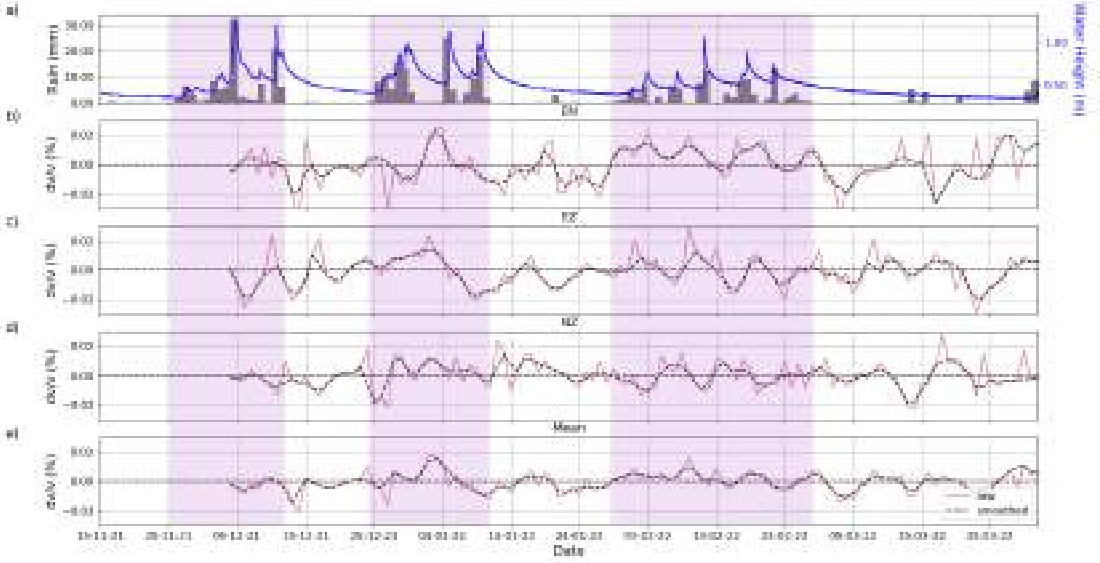


Fig. 4.15 Average dv/v obtained over the node array [2-4 Hz]. a) Daily rainfall averaged over all the rain gauges and underground water height measured at Fontenotte. 5-days stacked dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. Black dotted lines correspond to smoothed dv/v. Shaded rectangles indicate the main rain periods.

Weak dv/v changes are observed on the three cross-components at both frequency ranges (-0.02 to 0.02%). Since the time series are different for each cross-component, we observe a much smoother mean series compared with individual components.

Regarding 2-4 Hz frequency range, we can note that EN and EZ cross-components have comparable behaviors but no clear relationship between dv/v and rainfall or water height can be observed. It could be due to the time scale of the experiment, which lasted 4 months, probably too short to highlight "long-term" behavior as observed at AVEN station. At this frequency range, velocity variations are most likely affected by annual cyclic variations, or rain/dry season variations. They mostly reveal long term variations rather than direct response to rain events. Because of that it might be quite difficult to interpret the velocity variations from the nodes for this frequency range due to the need of long term recording, in the studied period we might be checking the variations within the same cycle (of seasonal variations).

The 8-10 Hz averaged dv/v time series exhibit large variability and the relation between rainfall and averaged dv/v is quite unclear (Fig. 4.16). However, we can see a dv/v decrease during and directly after the first rainfall period right after the deployment of nodes, between 30^{th} November and 14^{th} of December 2021, which marks the end of the flood. After this, for the EN and NZ components, we observe dv/v increase until the second rain event which begins the 24^{th} of December 2021, when dv/v starts to decrease again. The pattern of dv/v increase during dry periods and decrease during wet periods can mostly be recognised for the EN component. However we can also see positive or negative peaks that do not correspond with rainfall occurring time. The latter might be the result of different behaviours or time shifts in variations observed on some of the nodes. The 8-10 Hz frequency band is more sensitive to changes close to the surface, so potentially to rapid changes in the medium right after rain events. In the following sections, we will focus on the 8-10 Hz frequency band and investigate dv/v spatial distribution over the watershed.

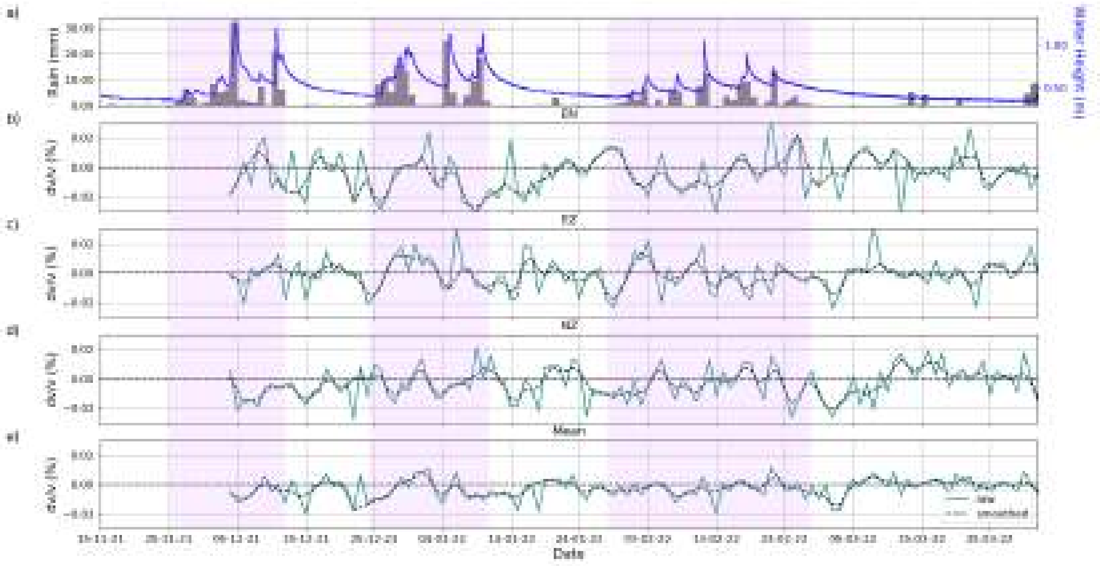


Fig. 4.16 Average dv/v obtained over the node array [8-10 Hz]. a) Daily rainfall averaged over all the rain gauges and underground water height measured at Fontenotte. 5-days stacked dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. Black dotted lines correspond to smoothed dv/v. Shaded rectangles indicate the main rain periods.

4.4.3.2 Average velocity variations within main geological units

In order to assess the impact of the outcrop formation at the vicinity of the nodes, we have looked in more details at 8-10 Hz dv/v measurements under nodes located on marl and limestone outcrops (Fig. 4.17; marl: N25, N40, N49; limestone: N3, N31, N57). The key difference between marl and limestone lies in their water-holding capacity and permeability. Marl generally has a higher water-holding capacity but lower permeability, leading to reduced water flow through the rock. On the other hand, limestone has lower water-holding capacity but higher permeability, facilitating the movement of water through its interconnected pores and fractures (Bell et al., 2001). This is also portrayed in the water reserve map were we can see a higher rate for soil above marls compared to limestones (see Section 2.1.5 in Chapter 2). These differences have implications on the behavior of aquifers and groundwater systems.

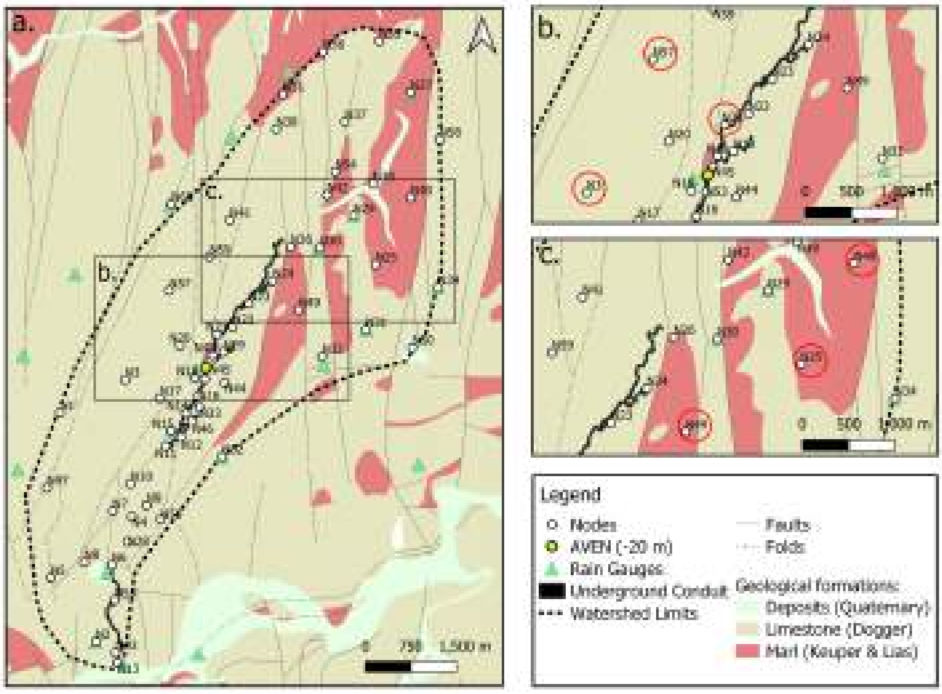


Fig. 4.17 a) Geological map of the Fourbanne watershed (main formations, faults and folds), showing also its limits (black dotted line), the underground conduit (black line) and the seismic array (white dots). The two rectangles indicate the sections zoomed in b and c. b) Zoom in on the nodes selected (in red circles) to analyse the velocity variation on limestone. c) Zoom in on the nodes selected (in red circles) to analyse the velocity variation on marl.

Relative velocity variations obtained from "limestone group" (N3, N31, N57) and "marl group" (N25, N40, N49) are plotted in Figs. 4.18 and 4.19, respectively. Individual dv/v time series are noisy, reflecting spatial variability within each group (see next section). The EN cross-component presents comparable dv/v for both groups, however no clear resemblance is observed for EZ and NZ.

In order to better compare the dv/v response to rainfall in marl and limestone, we have averaged dv/v in each group, as illustrated in Figs. 4.20 and 4.21. For both groups, dv/v time series do not exhibit similar pattern on any of the cross-components during the studied period. As expected from individual dv/v patterns, the mean on the dv/v EN cross-component presents the largest amplitudes compared with other cross-components, since individual dv/v were more coherent in terms of variations. In addition, we observe negative dv/v mainly during or at the end of the "rain periods" for the EN cross-component (shaded rectangles on Figs. 4.20 and 4.21). On

Fig. 4.18 Five-days stacked dv/v measured under 3 nodes located on limestone [8-10 Hz]. a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. dv at nodes N3, N21, N57 for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components.

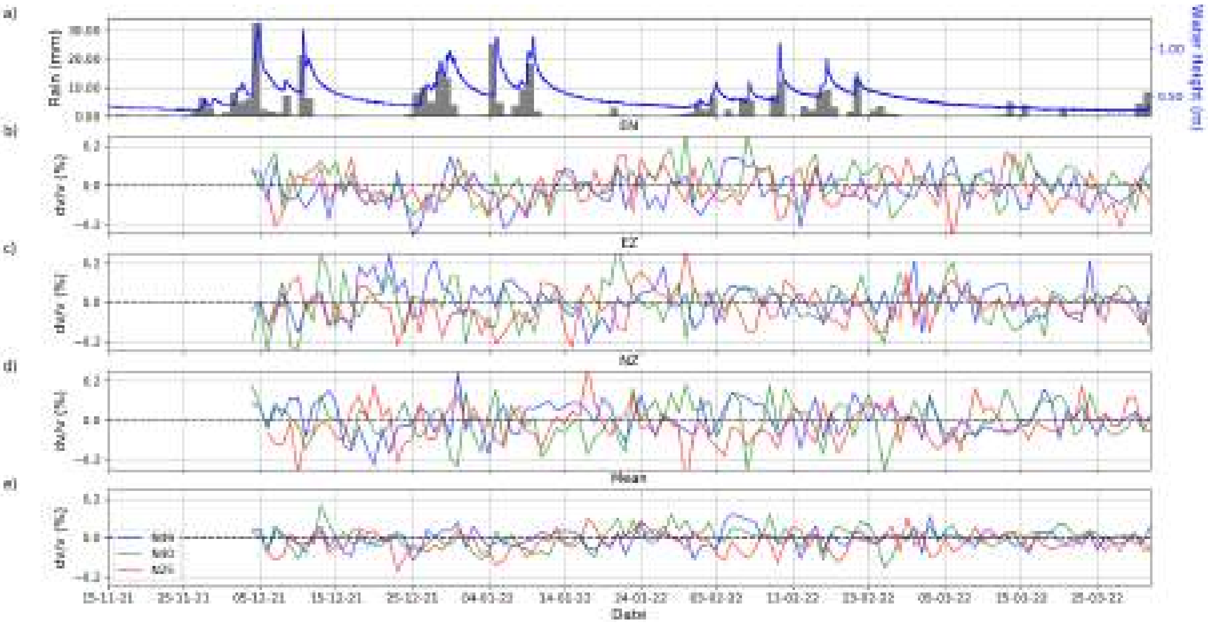


Fig. 4.19 Five-days stacked dv/v measured under 3 nodes located on marl [8-10 Hz]. a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. dv at nodes N25, N40, N49 for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components.

the opposite, positive dv/v are rather observed during "dry periods". This is less clear on the second half of the experiment, from the end of January, which includes less intense but longer rain period. The "marl group" present larger dv/v amplitudes compared with the "limestone group", in particular the negative ones, for all components. This might reflect the larger soil water reserve capacity expected from this material.

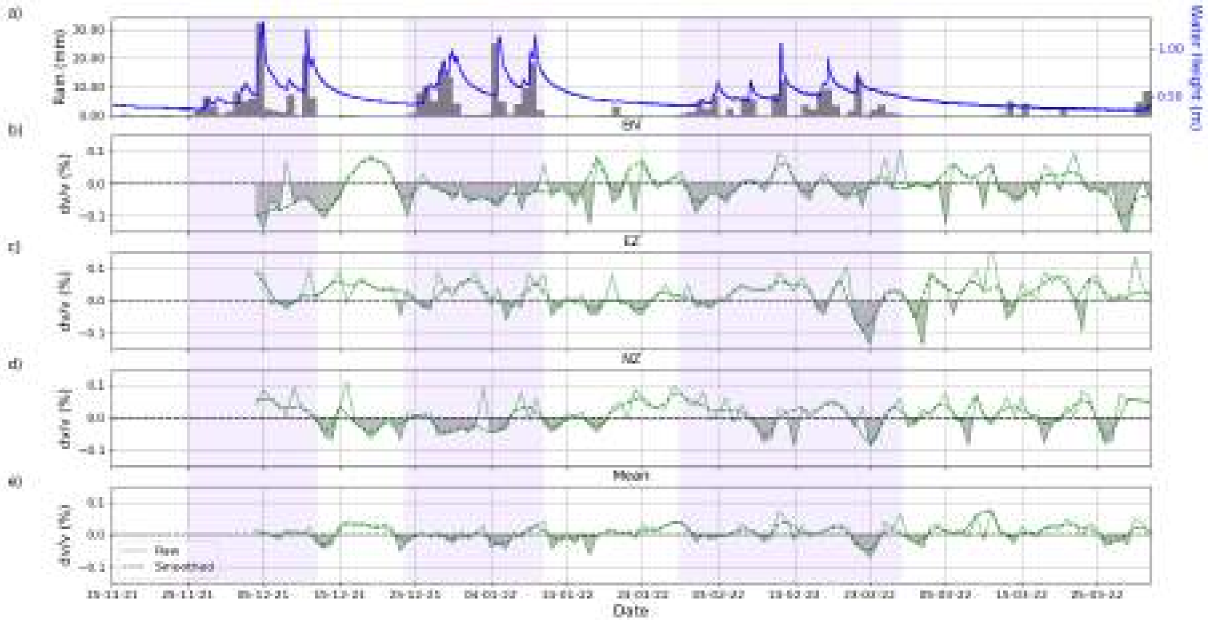


Fig. 4.20 Five-days stacked 8-10 Hz dv/v averaged over 3 nodes installed on limestone (N3, N21, N57). a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. 5-days stacked dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. Dotted lines correspond to smoothed dv/v.

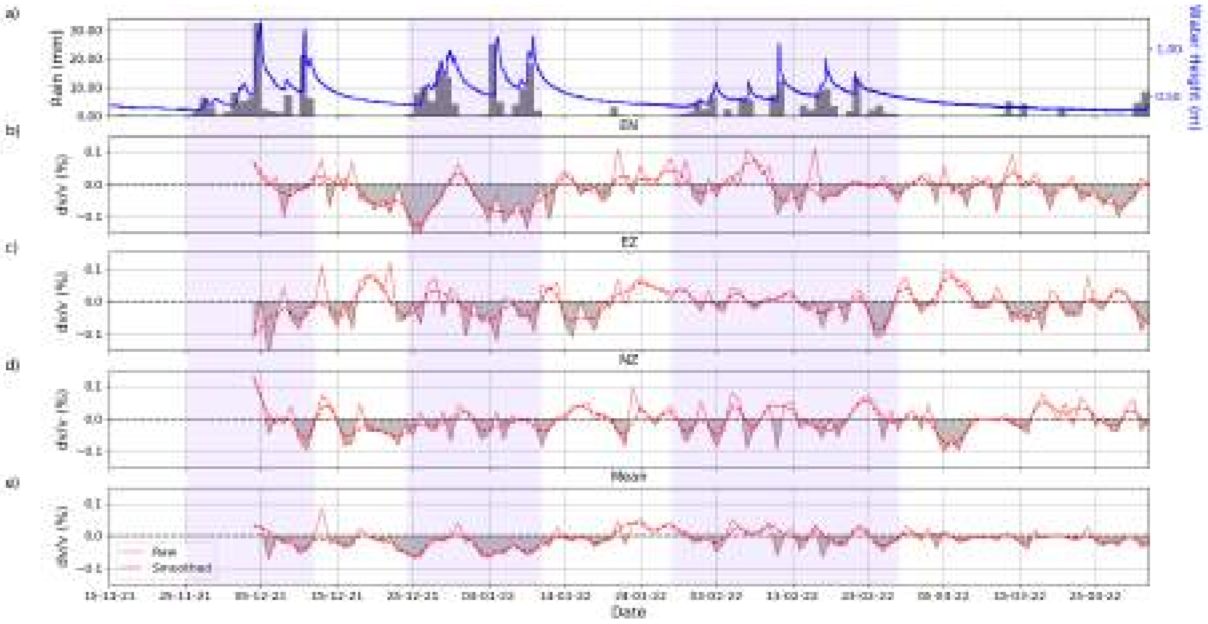


Fig. 4.21 Five-days stacked 8-10 Hz dv/v averaged over 3 nodes installed on marl (N25, N40, N49). a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. 5-days stacked dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. Dotted lines correspond to smoothed dv/v.

4.4.3.3 Spatial distribution of velocity variations during a rain period

Mean of the dv/v stacks for each of the components

To visualise the velocity changes while taking into account the spatial variation on the basin, we chose to investigate the aftermath of the first rain period from 30th November (beginning of experiment) to 14th of December 2021 (end of flood), resulting in a total rain amount of about 100 mm. The mean of the 5-days stacked dv/v over the first rain period is presented in Fig. 4.22. As already observed, the cross-components exhibit different dv/v response. However, globally, we mainly observe negative dv/v in the NE part of the watershed and along the conduit. In Fig. 4.22e, which is the mean over the cross-components, we see that most of negative dv/v are located in the NE (N34, 50, 25, 29, 39) and in the middle of the watershed, from N26 in the north to N4 in the south. There are also positive mean dv/v in the NE, in particular under nodes aligned along a major NS fault (N37, N54, N42, N30, N32) and along the conduit. The amplitude of the mean dv/v is relatively weak in the southern part of the watershed, reflecting opposite behaviours between cross-components : mainly negative for EN and positive for EZ and NZ.

Comparison between mean of the dv/v and the rainfall distribution

The mean velocity variation of the 3 components can be compared with the spatial variation of cumulative precipitations measured at each rain gauge during the selected period (between the 4th and the 13th of December 2021; Fig. 4.23). The rainfall occurred over all of the watershed, but spatial variations can be observed, ranging from 55 mm (least cumulative precipitations) to 85 mm (largest cumulative precipitation). In Fig. 4.23, we see that the NE part of the watershed recorded the highest rain amounts during the studied 10 days period. This is consistent with mean dv/v which tends to decrease in this area (negative slope in Fig. 4.23).

Slopes of the linear regression of the dv/v stacks

We tried another way to investigate spatial distribution tendencies (increase or decrease) of dv/v during a specific period : we computed the slope of the linear regression of the velocity variations over a selected period of time. We recall, the dv/v have been estimated comparing NSCFs with a RNSCF spanning the whole period of experiment (4 months). An example of the approach is illustrated in Fig. 4.24 for N47.

We computed dv/v slope for the 8-10 Hz frequency range during the first rain period for each node and each cross-component as well as for the mean of all of the 3 components (Fig. 4.25). The obtained slope values are between -3 % (corresponding to maximum decrease) and +3 % (corresponding to maximum increase). These values are assigned to a colorbar going from blue (negative slope) to red (positive slope), passing through white (null slope). White dots can indicate either minimal variations during the 10 days for the specific node or succeeding opposite signs peaks, giving a small slope value. Again, we observe different trends between the different cross-components. A positive trend for a certain cross-component can be negative for another cross-component. In fact, the mean cross-component dv/v slope isn't really more explicit (Fig. 4.25 e). Thought, we can note that dv/v slopes are rather negative in the N and NE (N27, N58, N50, N29, N31, N24, N49, N41, N59, N57). Each of the 3 cross-components reveal

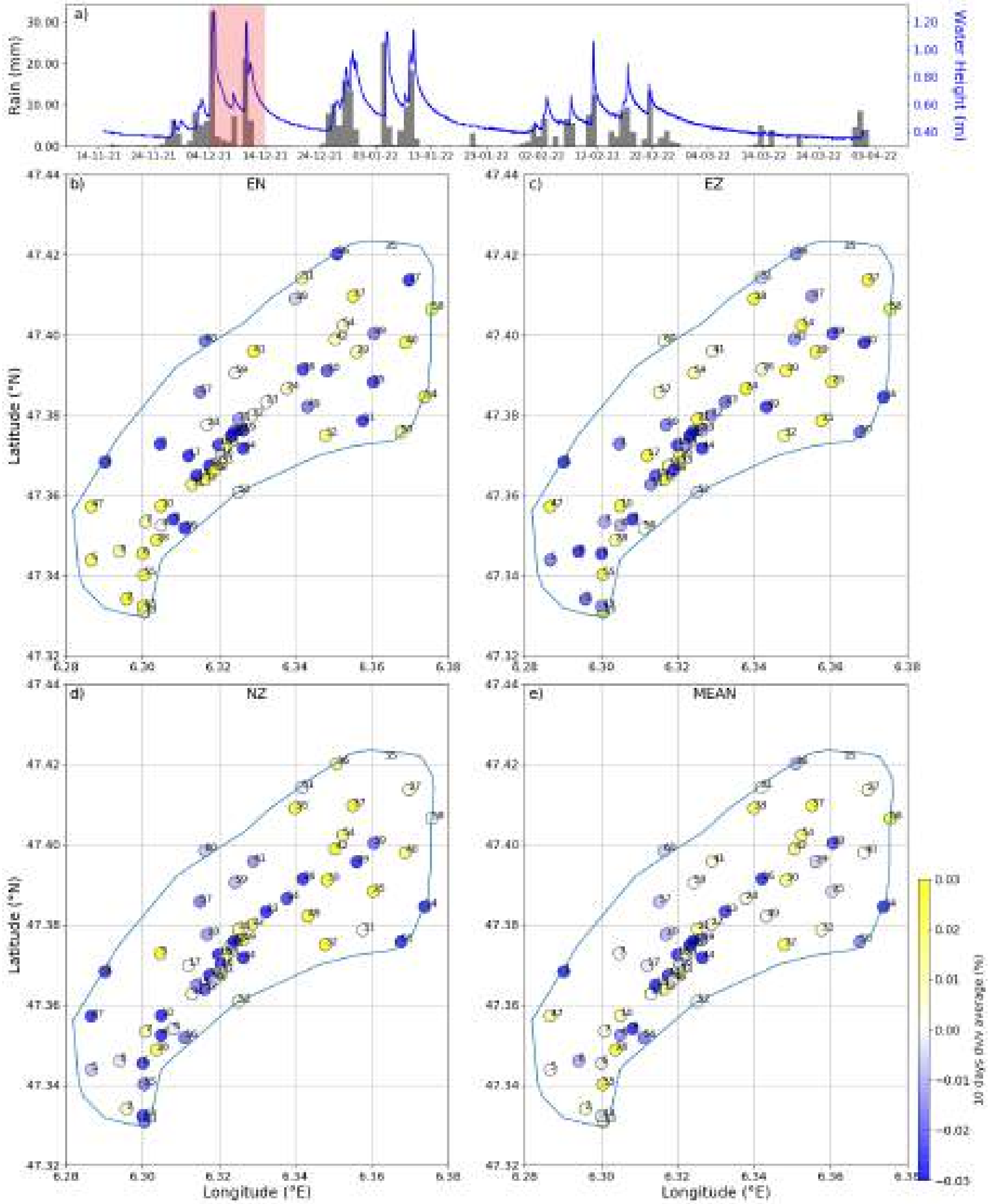


Fig. 4.22 Average of 5-days stacked 8-10 Hz dv/v measured under the nodes array during the first rain period (shaded rectangle). a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. Spatial distribution of averaged dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components.

negative dv/v slopes for some of the nodes along the conduit, but since the trends are different for each cross-component we cannot observe this in the mean dv/v slopes.

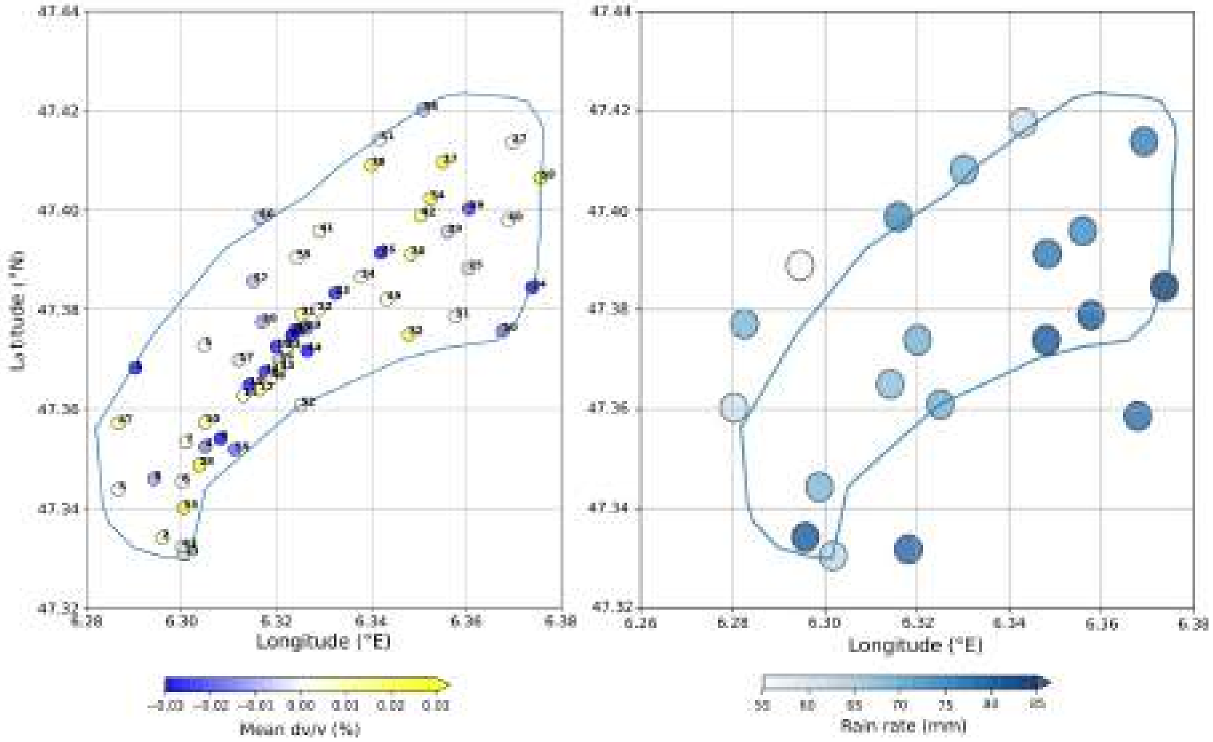


Fig. 4.23 Spatial distribution of averaged dv/v over the 3 cross-components (left) and cumulative rainfall rate (right) between the 4^{th} and the 13^{th} of December 2021.

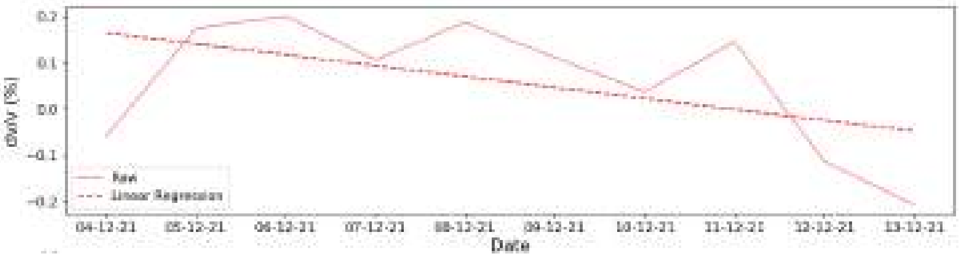


Fig. 4.24 Example of linear regression applied on five-days stacked 8-10 Hz dv/v time series, between the 04/12/2021 and 13/12/2021, measured under node 47. The obtained slope here is -2% (-0.02).

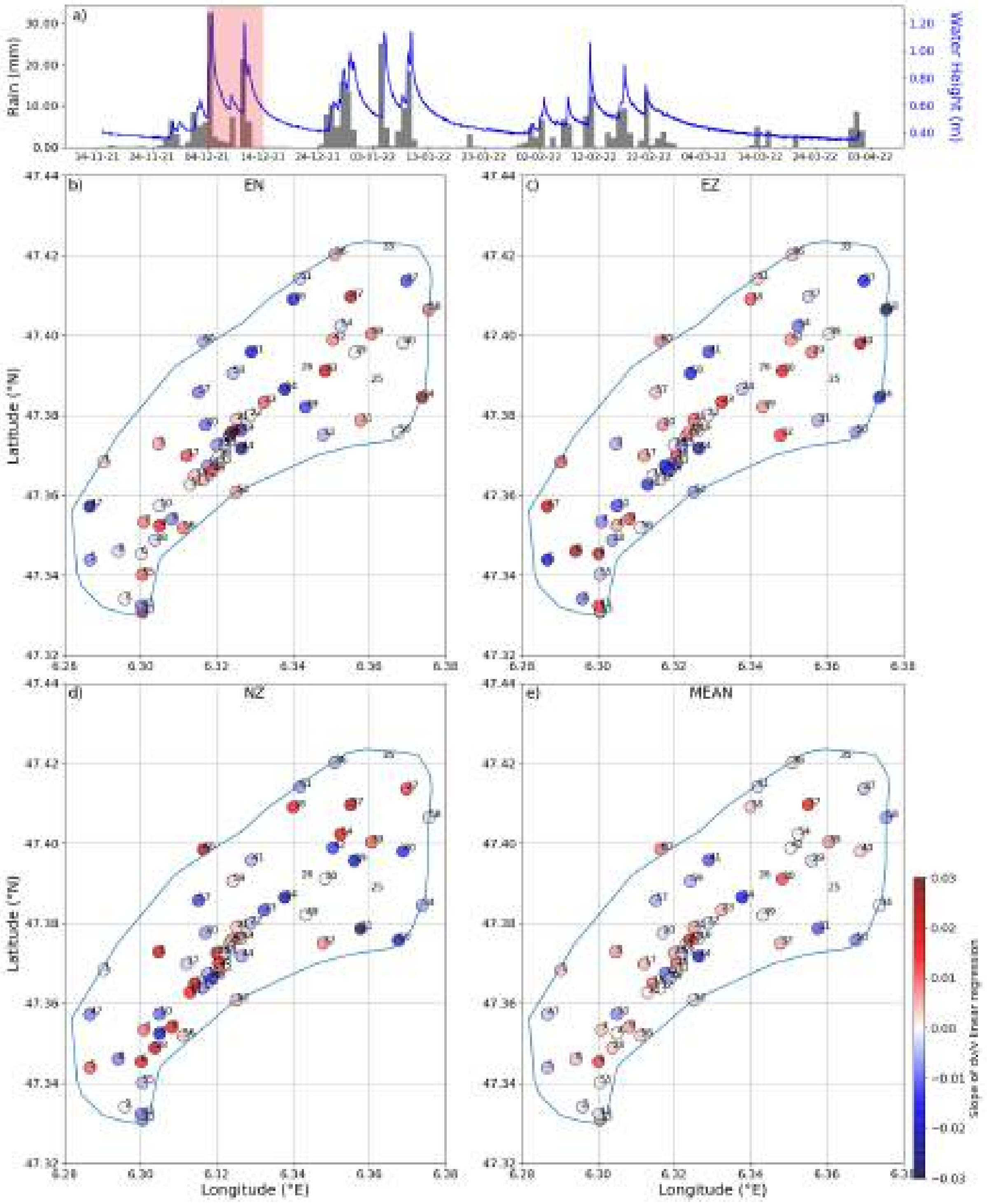


Fig. 4.25 Linear regression slope of five-days stacked 8-10 Hz dv/v measured under the node array during the first rain period (shaded rectangle). a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. Spatial distribution of averaged dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. The colorbar indicates slope values ranging from -0.03 (blue) to +0.03 (red).

Components distribution of the minimum dv/v average

In order to understand the dv/v difference between cross-components, we selected the cross-component yielding the minimum 10 days average (highest negative value) at each node and plotted the result by assigning a color to each cross-component: a gray dot for the minimum value obtained on EN, green for EZ and orange for NZ. This has been done for negative dv/v only, nodes having positive values on the all of the cross-components for the selected period are assigned red dots. The results for the period between the 4^{th} and the 13^{th} of December are plotted in Fig. 4.26. The first observation is that all the nodes (except 5 nodes) have negative values for the averaged dv/v at least on one of their components, indicating a decrease in velocity as a response to the rain period. We can also see that the distribution is done by groups of neighbouring nodes: spatially close nodes tend to have minimal trend within the same component, as observed for example for the nodes N41, N59, N26, N24 and N23. or N08, N06 and N05. None of the components show a dominance with the number of minimal trends compared to others. The nodes having positive dv/v averages are N58, N54, N32, N33 and N28. This is not constant for all rain events: other nodes might have positive values after rainfall, and components at which nodes show negative values might differ. However, considering this specific period, the mentioned nodes might have different behaviors since N54 is located at the vicinity of a sinkhole, N32 and N33 are positioned along NS faults, N58 and N28 are on highly sloped terrains, with slope values about 15% and 11% respectively, which might favor surface flows.

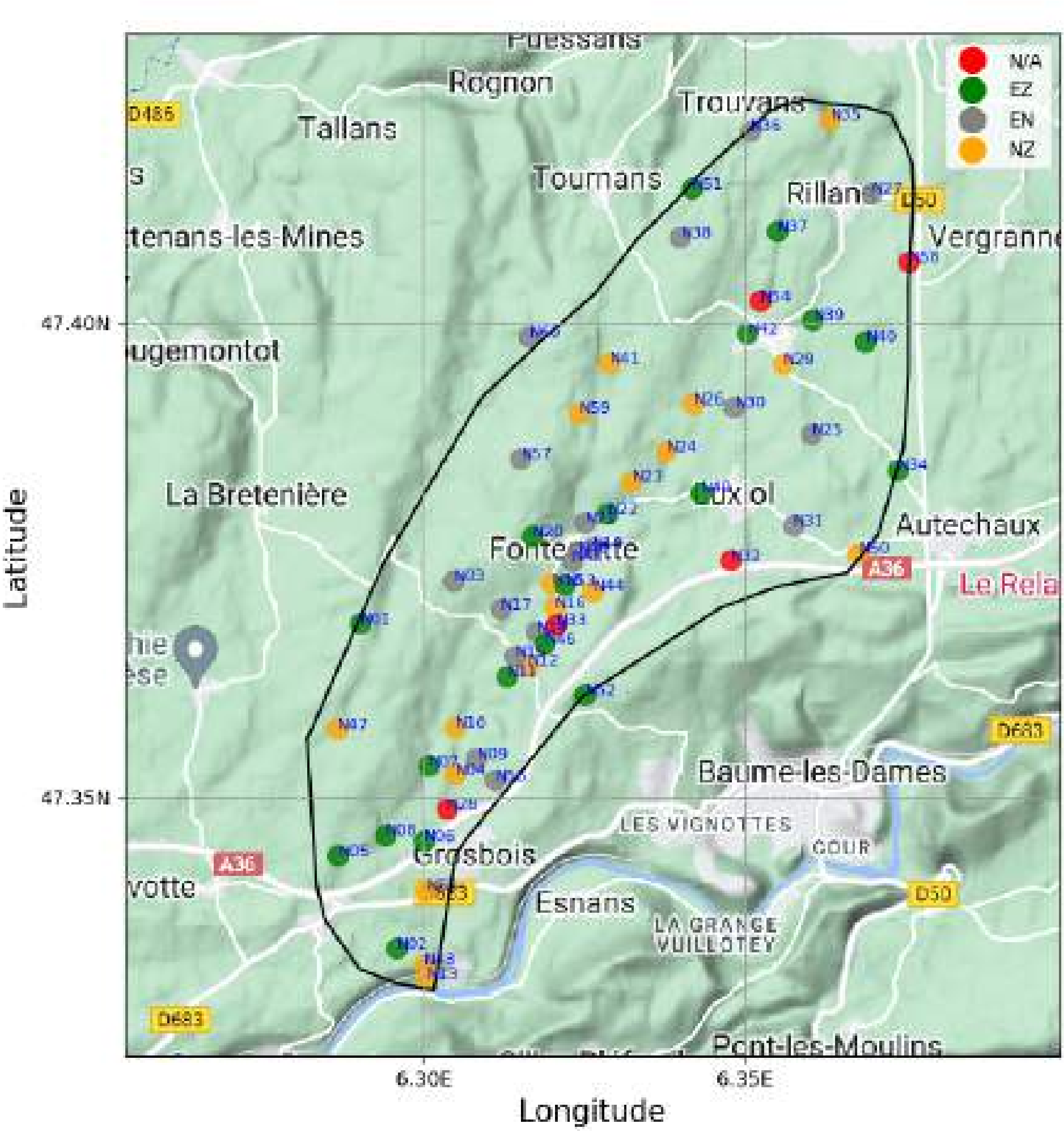


Fig. 4.26 Map with the component corresponding to the highest negative value for the average velocity variations over 10 days, between the 4th and the 13th of December 2021: gray dots correspond to the a minimum yielded on the EN component, green dots for the EZ component and orange dots for the NZ component. The red dots indicate that the average dv/v is positive along all the components.

4.4.3.4 Spatial distribution of velocity variations during a *dry* period

Mean of the dv/v stacks for each of the components

In the same framework as the section above, we decided to visualise the velocity changes while taking into account the spatial variation on the basin during a period without rain. We chose to investigate the period coming directly after the rainy period studied above,between the 14^{th} and the 23^{rd} of December 2021 (shaded rectangle on Fig. 4.28a). These 10 days mark no major rainfall occurring. The mean of the 5-days stacked dv/v over the first *dry* period is plotted in Fig. 4.27. While observing separately the cross-components, we can see that the dv/v distribution is still heterogeneous. Negative dv/v are observed over a larger part of the conduit, extending southward, compared with the rain period.

Slopes of the linear regression of the dv/v stacks

As we did above, we used linear regression to compute the slope of the velocity variations over 10 days for the 8-10 Hz frequency range, for each node and each cross-component (EN, EZ and NZ), as well as for the mean of all of the 3 cross-components. We observe again that different and/or trends are on the different cross-components. Negative trends mostly appear on the southern section of the conduit as long as on the southern end of the watershed and on the western side of the watershed.

Components distribution of the minimum dv/v average

Finally, similarly as for the rainy period, we selected minimal values between the 3 cross-components for the velocity variation 10-days average at each node and plotted the result by assigning a color to each component: gray for the EN component, green for the EZ component and orange for the NZ component. Nodes having positive values on all of the components are shown with red dots. The results for the period between the 14^{th} and the 23^{rd} of December are plotted in Fig. 4.26. As observed for the rainy period, approximately all the nodes (except 4 nodes) have negative values of dv/v average at least on one of their components. The four nodes having positive trends on all their components are N51, N42, N28 (as for the rainy season) and N02. In fact, N42 is located at the vicinity of the Verne sinkhole, which is the largest sinkhole in the watershed and might be draining the water towards the underground river, N51 and N28 are located on highly sloped terrains (slope about 11%) which eases the surface water flows (making the infiltration rather difficult). We can also observe that neighbouring nodes tend to have minimum slopes on the same component. Comparing with the same representation done during the rainy period (Fig. 4.26), we can see that the components are different.

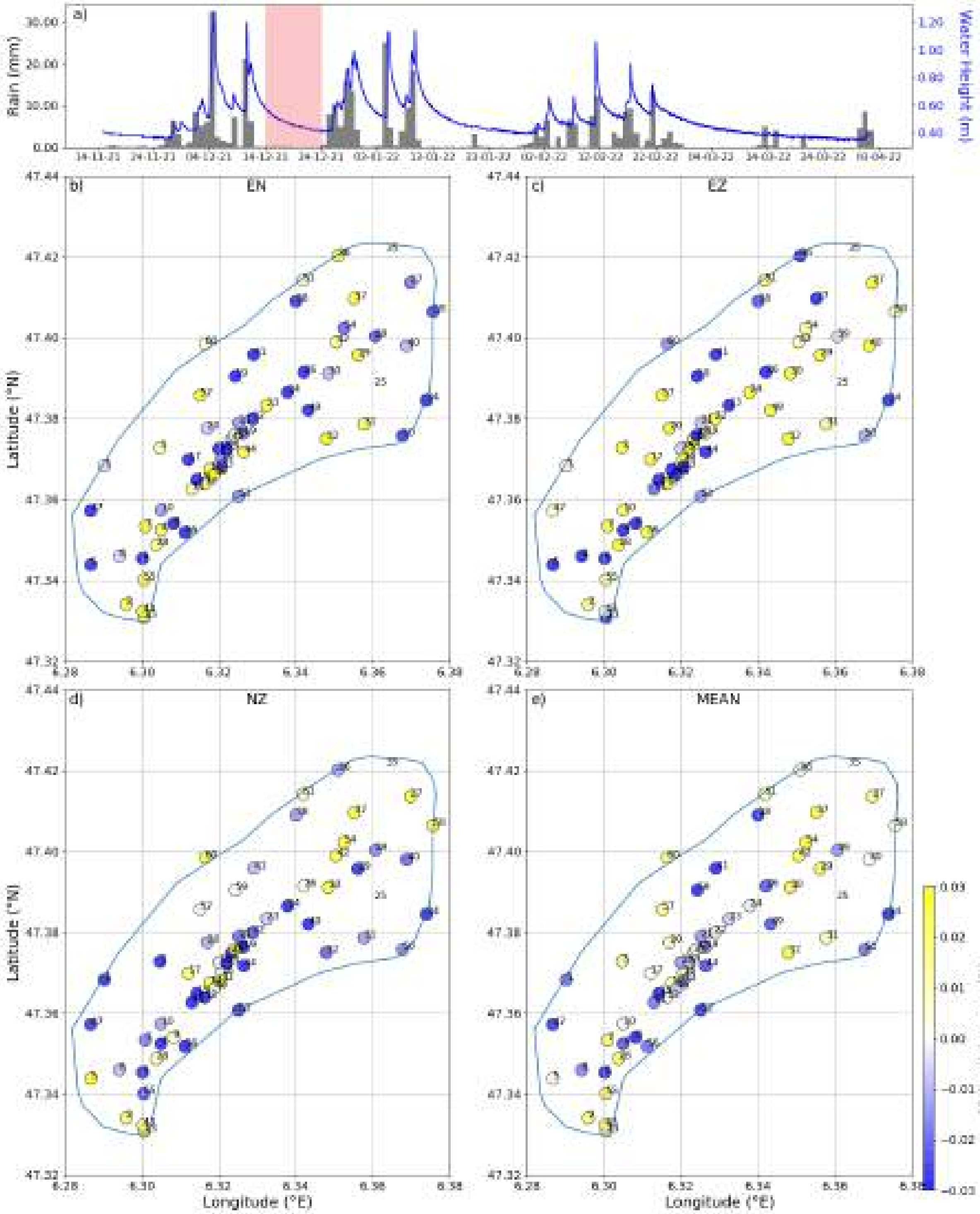


Fig. 4.27 Average of five-days stacked 8-10 Hz dv/v measured under the node array during the first dry period (shaded rectangle), between the 14^{th} and the 23^{rd} of December 2021. a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. Spatial distribution of averaged dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components.

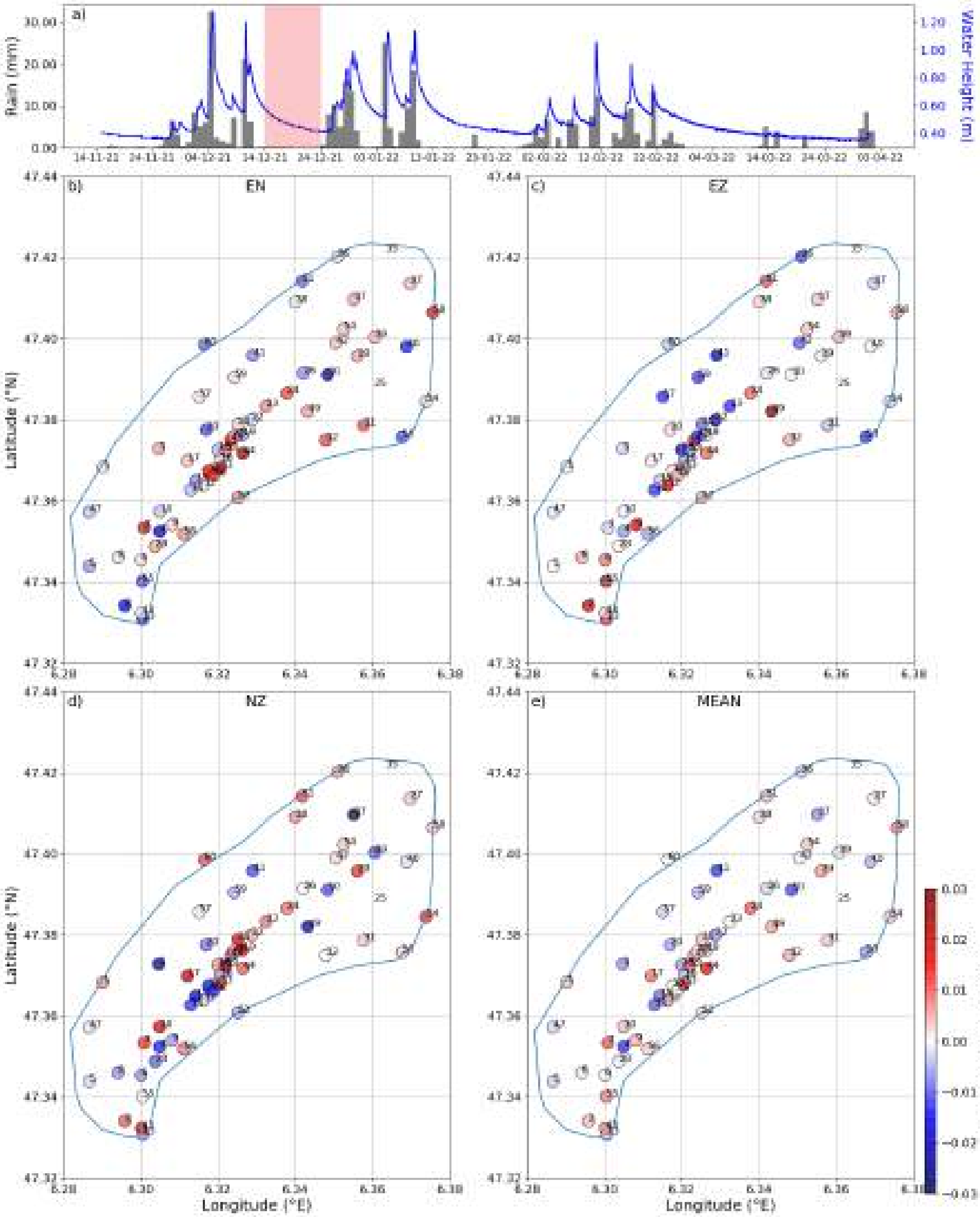


Fig. 4.28 Linear regression slope of five-days stacked 8-10 Hz dv/v measured under the node array during the first dry period (shaded rectangle). a) Daily rainfall (bars) and underground river water height (blue line) during the studied period. Spatial distribution of averaged dv/v for b) EN, c) EZ, d) NZ cross-components and e) the mean of the 3 cross-components. The colorbar indicates slope values ranging from -0.03 (blue) to +0.03 (red).

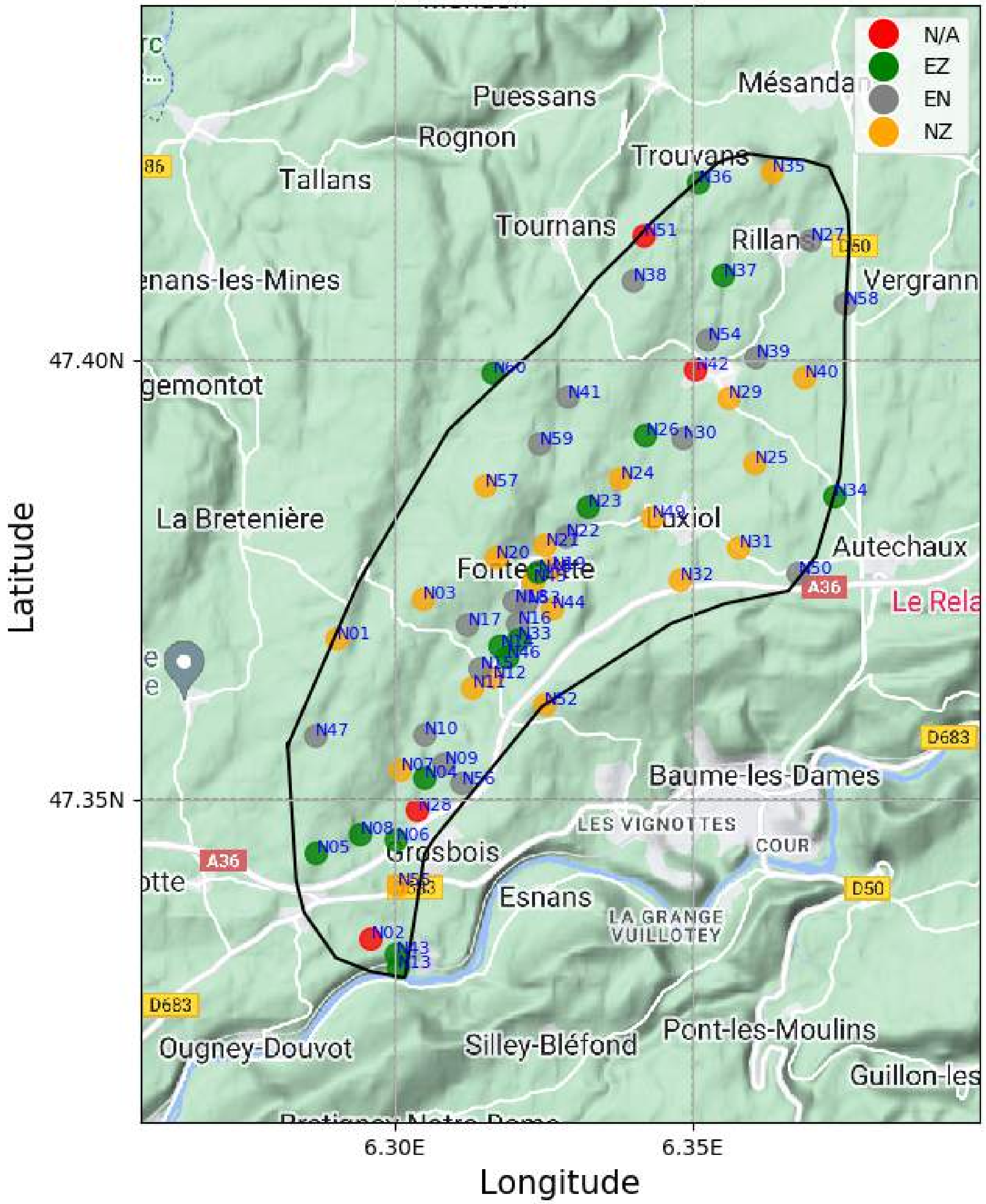


Fig. 4.29 Map with the component corresponding to the highest negative value for the average velocity variations over 10 days, between the 14th and the 23rd of December 2021, between the 3 components EN (gray), EZ (green) and NZ (orange) The red dots indicate that the slope is positive.

4.5 Discussion : Water transfer within the Fourbanne watershed

4.5.1 Saturation changes within highly fractured reservoir

In this section we try to understand the 4 years dv/v variations observed at AVEN station, at 2-4 Hz, characterized by a decrease during and shortly after precipitations. (Fig. 4.13). When precipitations occur, water infiltrates into the ground, inducing changes in its hydro-physical properties that can be detected through seismic measurements (Karato and Wong, 1995). The influence of hydrogeological conditions of the medium on relative seismic velocity has been the subject of several studies (e.g. Fores et al., 2018; Lecocq et al., 2017; Poli et al., 2020; Sens-Schönfelder and Wegler, 2006; Tsai, 2011). Actually, variation in seismic wave velocity may reflect changes in pore pressure, porosity, saturation and density (Voisin et al., 2017). Kim and Lekic (2019) showed that ground water level changes induce correlated changes in P-wave velocity (V_P), but anticorrelated and smaller changes in S-wave velocity (V_S).
The Biot–Gassmann relation can be used to estimate saturation of the medium based on V_S measurements (Gassmann, 1951) :

$$V_S = \sqrt{\frac{\mu}{\rho}} \tag{4.1}$$

where μ is the shear modulus and ρ the bulk rock density given by:

$$\rho(S) = \rho_{\mathrm{min}}(1-\phi) + \rho_{\mathrm{w}}(\phi S) + \rho_{\mathrm{air}}\,\phi(1-S) \tag{4.2}$$

Where ρ_{min}, ρ_w and ρ_{air} are the densities of rock minerals, water, and air, S is the saturation and ϕ is the porosity.

Fores et al. (2018) performed cross-correlations of ambient seismic noise on a karstic site in the Larzac plateau (France). They obtained a velocity variation of 0.2% at 6-8 Hz frequency range with a 6 months delay between the rainy season and the minimum velocity, revealing a slow infiltration rate. They propose the following equation between S and V_s:

$$V_{\mathrm{S}}(S) = V_{\mathrm{S,dry}}\left\{1 - S\left[1 - \sqrt{\frac{\rho_{\mathrm{min}}(1-\phi)}{\rho_{\mathrm{min}}(1-\phi) + \rho_{\mathrm{w}}\phi}}\right]\right\} \tag{4.3}$$

$$\begin{cases} \mu(S) = \mu \\ \rho(S) = \rho_{\mathrm{min}}(1-\phi) + \rho_{\mathrm{w}}(\phi S) + \rho_{\mathrm{air}}\,\phi(1-S) \end{cases}$$

Assuming 100% saturation increase, 10% porosity, $\rho_{\mathrm{min}} = 2800\ kg/m$, Fores et al. (2018) obtained V_S decrease of 2%. Then, they could relate the observed dv/v decrease of 0.4% to saturation changes of about 20%. In our case, considering $\phi = 5\%$ and $\rho_{\mathrm{min}} = 2700\ kg/m$ (unaltered Dogger limestones; J.P. Sizun, Chrono-environnement, personal communication), we obtain a V_S decrease of 2% (Eq.4.3). The maximum velocity decrease of dv/v is about 0.1%, which results in a saturation increase of only 10%. Fores et al. (2018) observed 6 months delay between the rainy season and the minimum velocity, revealing slow infiltration process. In our case though, dv/v decreases while precipitation periods have started, indicating faster infiltration mechanism (AVEN, 2-4Hz). Velocity variation are thus most likely due to water saturation of intensively fractured reservoir(s)/zone(s), within the epikarst.

Regarding the velocity variations obtained using the seismic array, the relation between rainfall and velocity tendencies is less obvious. The spatial distribution of the dv/v with highest negative amplitude correspond to the major rainfall location (see Fig. 4.23) for the studied event. However, the response of the aquifer is still rather heterogeneous, whether while analysing the variations for the nodes' same components, or for a same node on different components. The latter might be explained by the heterogeneity and complexity of the medium, that leads to anisotropy in flows patterns and seismic velocity variations. The effects of anisotropy are more developed in the following.

4.5.2 The effect of soil available water content distribution

From the analysis of the average velocity variation for the mean of the 3 cross-components we can see that zones with marl outcrops (Fig. 4.2), have mostly negative dv/v averages. The next step was to analyse distribution for the mean of the 3 components with the available water content map. As said before, having different behaviors for each of the component, the mean might result with averaging opposites. However, we plotted the mean velocity variation distribution on the water reserve map during both rainy and dry studied periods (Figs. 4.30 and 4.31, respectively). We can notice that the zones exhibiting negative mean velocities correspond mostly with zones having a high water reserve. This is related to the higher water-holding capacity of the soil with marly substratum, compared to the soil with limestone substratum, that will be more porous and highly permeable (see Section 2.1.5 in Chapter 2). Figure 4.11 shows that for the 8-10 Hz frequency range, the R-waves, in addition to being sensitive in the depth of tens of meters, are sensitive to the first few meters in both models. For the nodes located on marl outcrop, they might mostly record variation in the first few meters of soil, that will be more consequent than the variations at depth, explaining the dv/v observed distribution.

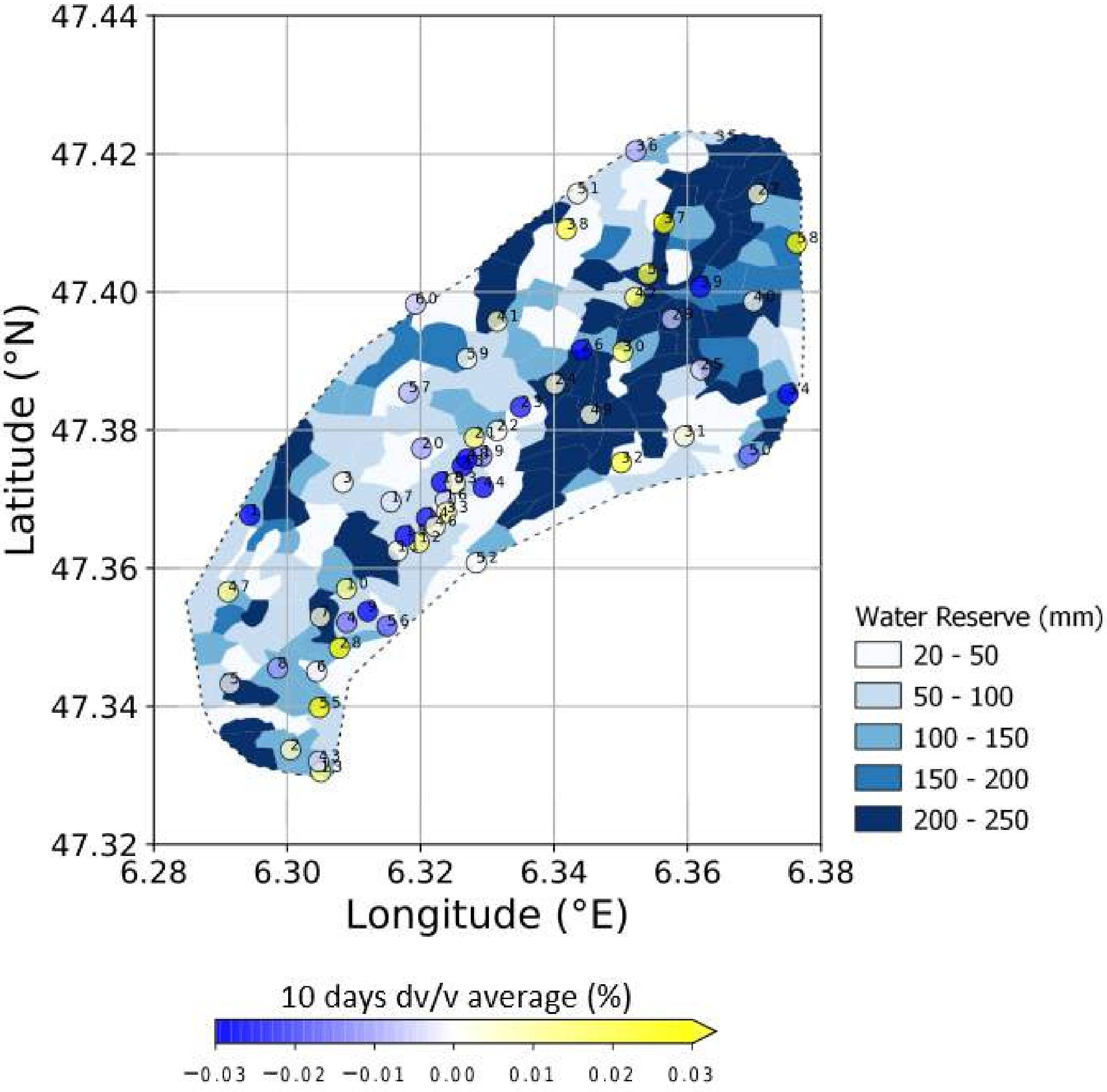


Fig. 4.30 Average of five-days stacked 8-10 Hz dv/v measured under the nodes array during the first rain period for the mean of the 3 cross-components plotted on top of the water reserve. The colorbar indicates average values of dv/v ranging from -0.03 (blue) to +0.03 (yellow). The water reserve values are portrayed using shades of blue for ranges of values, between 20 and 250 mm.

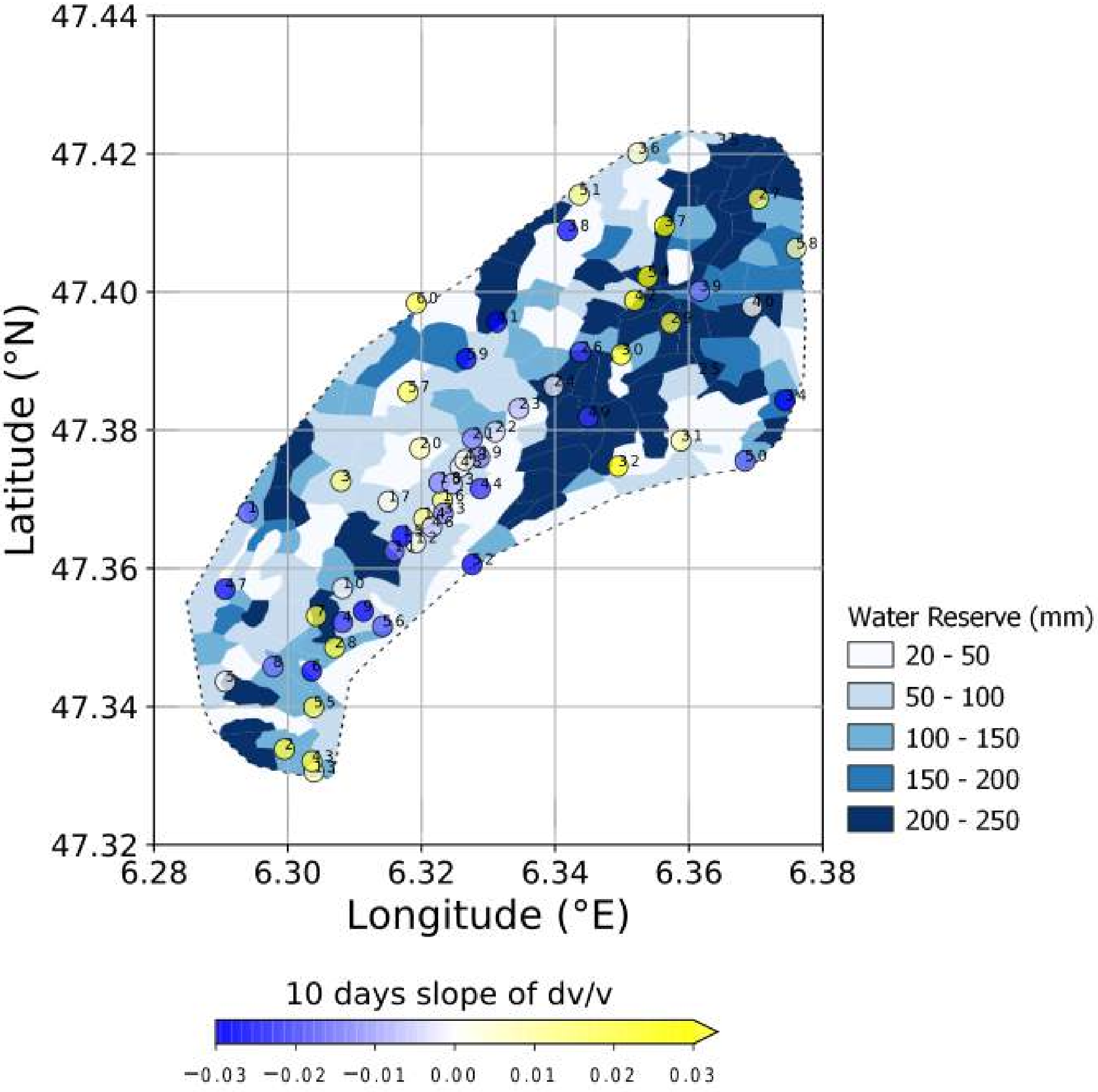


Fig. 4.31 Average of the velocity variations over 10 days, between the 14th and the 23rd of December 2021 for the mean of all components, between 8 and 10 Hz, plotted on top of the water reserve. The colorbar indicates slope values ranging from -0.03 (blue) to +0.03 (yellow). The water reserve values are portrayed using shades of blue for ranges of values, between 20 and 250 mm.

4.5.3 dv/v response in saturated and unsaturated zones

Cholet et al. (2017) have separated the underground conduit into two reaches: reach 1 between Verne and Fontenotte, and reach 2 between Fontenotte and Fourbanne (see Fig. 2.6 from Chapter 2). The reach 1 represents an unsaturated section, and reach 2 is both unsaturated and saturated. The limits separating the unsaturated and saturated sections is not constant: the saturated part is more developed during high flow periods and rainy seasons. In the saturated section, the conduit remains submerged over its entire length with a maximum depth of 25 m (the watershed's outlet is fed by a 25 m "siphon"). In order to have a better understanding of the response of the two different zones to rainfall, we cross-correlated the rain time series with the mean dv/v time series (5 days stack, 8-10 Hz, mean of the 3 cross-components) for all of the recorded period (4 months). The main objective is to find the lag time (in days) corresponding to the highest anticorrelation, as shown in Fig. 4.32, since as mentioned before rainfall and velocity variations are usually anticorrelated. We have done this for all nodes and selected for each the lag and the corresponding correlation value, and plotted the results in Fig. 4.33. This is similar to what has been done in Andajani et al. (2020): they selected lags corresponding to correlation coefficients that are smaller than -20%. In our case since we use the mean of the 3 components, correlation coefficients yield a minimum of -15%, however tests on individual components dv/v time series yield coefficients reaching -30%.

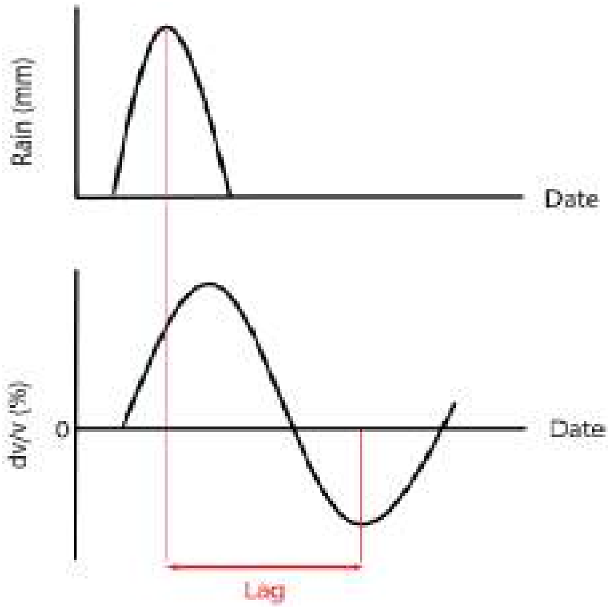


Fig. 4.32 Schematic figure of cross-correlation analysis between seismic velocity changes and the rain time series. The selected lag corresponds to the time shift that give the highest anticorrelation (negative correlation) value.

From Fig 4.33, we can see that we have a maximum negative correlations at short time lags (less than 5 days) on the northern and north-eastern and western parts of the watershed, that correspond mainly to unsaturated zoned. Larger time lags (above one week) are obtained for the nodes along the conduit (specially towards to southern section) as well as the southern part of the watershed which corresponds to a saturated zone. The latter is coherent with what we obtained while analysing the distribution of the average dv/v: during the rainy period we have negative velocities in the N, W and NE parts of the watershed (4.22) and for some nodes along the conduit, during the dry period directly succeeding the rain period, we observe negative velocities mostly along the conduit and in the southern part of the watershed. This means that faster responses are obtained in the unsaturated parts of the aquifer, compared to the saturated part.

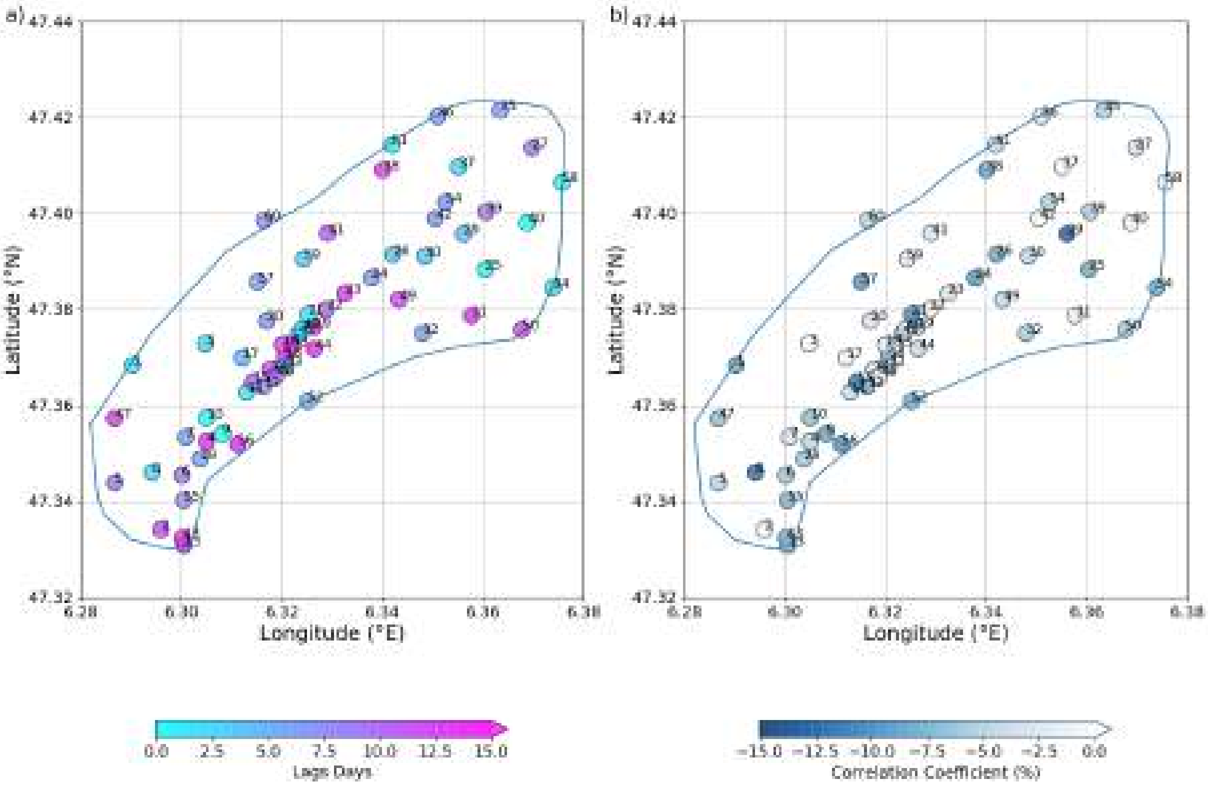


Fig. 4.33 Cross-correlation results between the mean of the 3 components of the dv/v time series (5 days stack, 8-10 Hz) and the rain time series for all of the studied period, in terms of a) Lag days corresponding to the time shift yielding the highest b) negative correlation coefficient for all the nodes.

The slow response in the saturated section was also demonstrated by Cholet et al. (2017) while doing flood monitoring. Figure 4.34 shows the time series of rain, flow rate, electrical conductivity (inverted y axis) and turbidity measured in Fourbanne between the 14th and the 21st of June 2015, to see the outlet's response due to the rainfall event that took place on the 14th of June. Rainfall events yield a decrease in the conductivity, due to the infiltration of less mineralised water in the karst. The lag time between the rainfall event and the negative peak in the electrical conductivity is about 5 days as observed in Fig. 4.34. In fact the flow rate reaches a peak hours after the rain event, this corresponds to water already existing in the saturated part of the karst (due to increase of the electrical conductivity) that is pressured out after rainfall, and the rainfall water reaches the outlet days after the rainfall event. This might be related to the slow response that we observe in the saturated part and could reflect water transfer processes between the unsaturated and saturated part of the conduit.

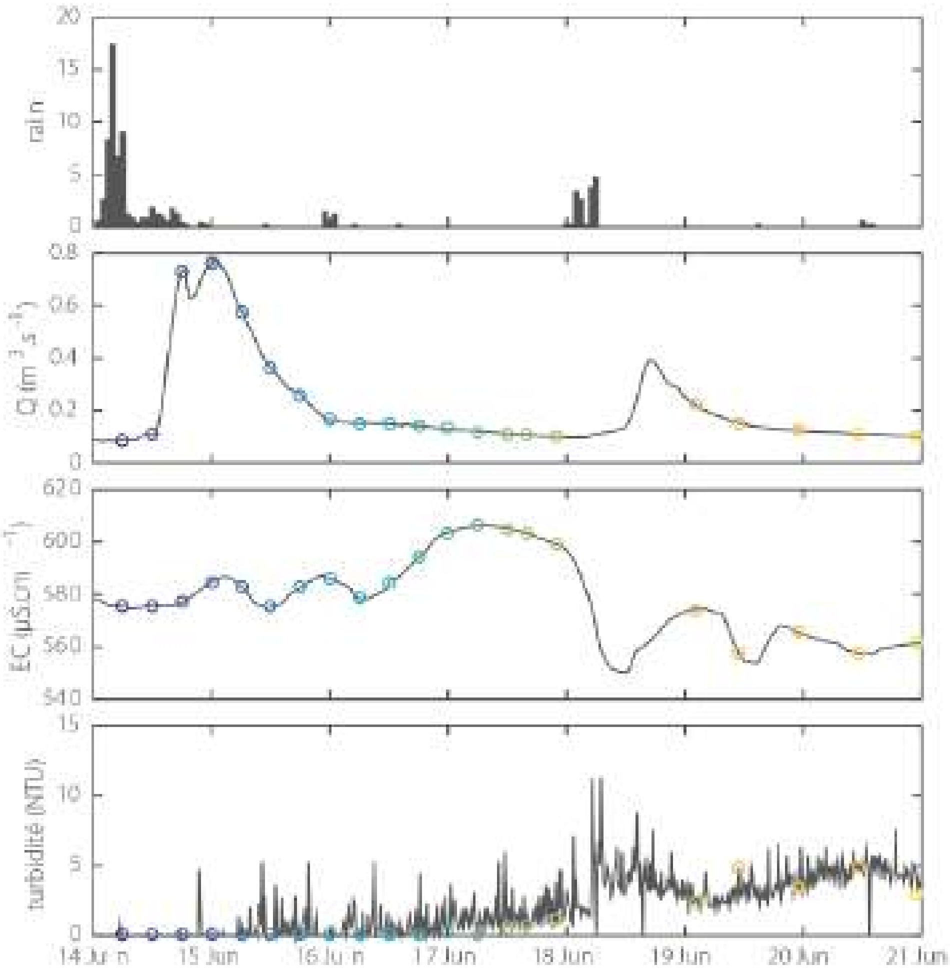


Fig. 4.34 Hydrogeological monitoring of a flood occurring on the 14^{th} of June 2015, showing rain data, flow rate (Q), electrical conductivity (EC) and turbidity measured at the watershed's outlet (Fourbanne), from Cholet et al. (2017).

4.5.4 The effects of anisotropy

The heterogeneity of karst systems makes water flow patterns after rainfall variable and intricate (Bakalowicz, 2005; Einsiedl, 2005). This is influenced by preferred flow paths that may exhibit higher hydraulic conductivity in certain directions compared to others due to fractures, porosity and alteration (Best et al., 2007; Knochenmus and Robinson, 1996; Prasad and Nur, 2005; Winterstein, 1990). Tsai (2009) showed that heterogeneities in the medium can introduce biases on propagation times in noise correlations. In addition, they may indirectly change the source position affecting the results. Having a highly heterogeneous medium, this is mainly what we observe while analysing the 10-days mean dv/v variation: variations between components can be similar but shifted in time (see Fig. 4.16, second rain period, EN and EZ variations), or even completely different (see Fig. 4.25 and Fig. 4.28).

Many previous studies based on cross-component correlations revealed anisotropic results. Machacca-Puma et al. (2019) found an apparent anisotropy of velocity change in single-station cross-components correlation while monitoring the Andesitic Volcano: in their case, the velocity changes are stronger when they are computed between the vertical and the horizontal component compared to the tangential direction with respect to the crater. They related the anisotropy in velocity variation to the anisotropy of the fractures distribution, due to the the predominance of radial faults in the studied structure. Flinders et al. (2020) also encountered anisotropy in seismic velocity variation in single station autocorrelation between the vertical, radial and transverse components, associated with the 2018 lower East Rift Zone (Hawaii). The observed phenomena happened after the eruption and caldera collapse and might be related to along-rift crack opening, giving differences in strain in the along-rift and across-rift directions. Caudron et al. (2021) used long time series to estimate the changes in the shallow subsurface due to volcanic tremor using velocity variation. They observed different values for dv/v using single station cross correlation for each of the components and related it to different waves contributing to the wavefield (dominated by Love waves) or anisotropy (see Fig. 4.35).

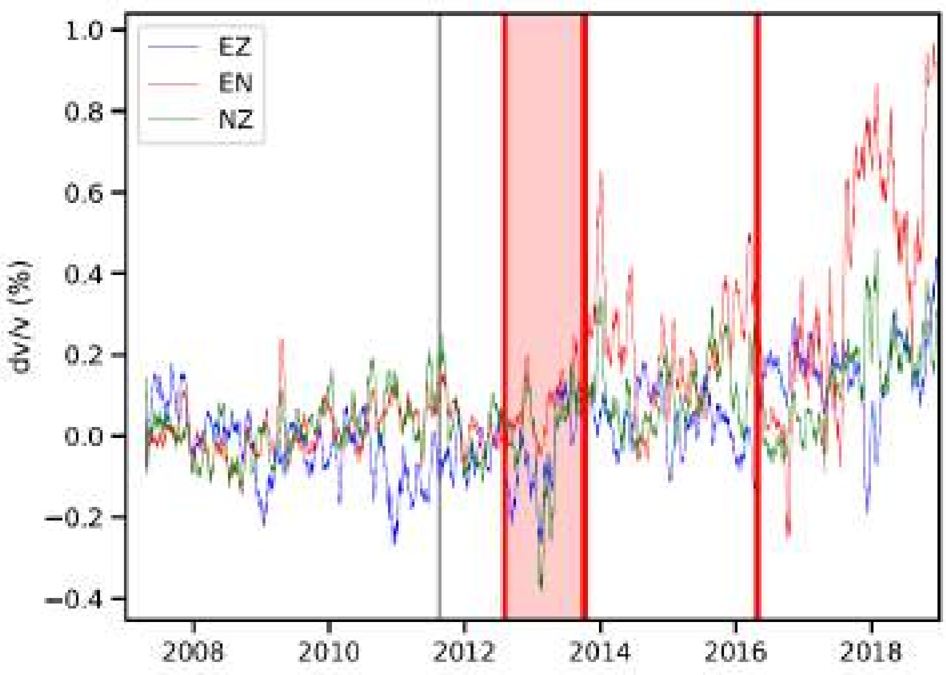


Fig. 4.35 Seismic velocity variations (single station cross-component approach) for the different combinations of components (EN, EZ and ZN), applied for monitoring of the Whakaari/White Island volcano, New Zealand. The 1-day results have been smoothed with a 30-day rolling mean. The red period in 2012– 2013 delineates the magmatic eruption and the last red line mark the 2016 phreatic eruptions (from Caudron et al. (2021) supplementary material).

Kim and Lekic (2019) used 20-year data recorded from a seismic station located in the Gulf Coast Aquifer System of southern Texas and used autocorrelation to study the response of the subsurface to groundwater levels (GWL) in terms of seismic velocity variation. The observed that GWL changes induce correlated changes in V_P, but anticorrelated, much smaller-amplitude changes in V_S. The authors obtained differences in velocity variations between the vertical and horizontal components. They achieved a better correlation between rain and dv/v from the vertical component rather than the horizontal component and related this with the higher sensitivity of V_P to ground water level changes in the case of autocorrelation compared to

V_S Fig. 4.36). The latter might also explain the discrepancies in the dv/v variation between components in our case.

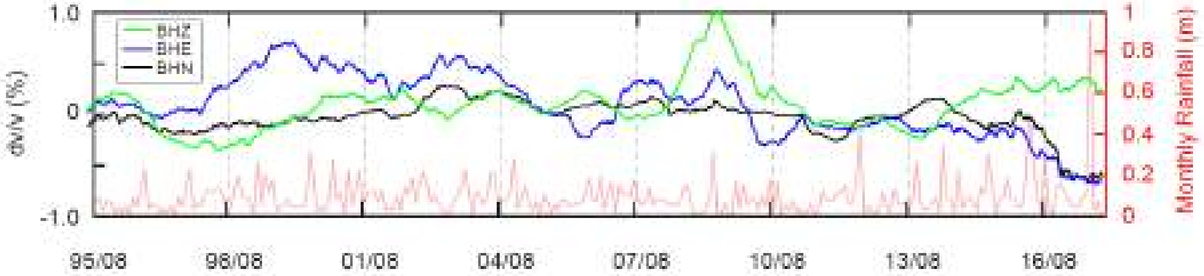


Fig. 4.36 dv/v derived from autocorrelation of BHE, BHN, and BHZ components and monthly rainfall recorded between 1995-2017 in Harris County, TX. Value next to each hurricane symbol (red) indicates its intensity on the Saffir Simpson hurricane wind scale (from Kim and Lekic (2019) supplementary material).

Han et al. (2020) studied the effects of pressure and water saturation on seismic anisotropy partially saturated rocks with aligned fractures. Stating that fractures are usually aligned in a certain preferential direction related to tectonic movements (for example in our case the faults are N-S oriented), which generated anisotropic properties of the rock massif. They found out that compressional wave velocity traveling across fractures are more affected by water saturation than the ones travelling parallel to the fractures. The latter observations might be the reason behind obtaining different velocity variations for different nodes and even different components for the same node, specially that anisotropy is a common issue for karstic media. Anisotropy of the properties of the rock mass will lead to anisotropy in water flow and fractures, hence anisotropy in the velocity variation.

4.5.5 Applicability and limitations

The seismic interferometry by cross-correlation revolutionised the seismology field and widened its applications, specially for monitoring purposes. However this method has its limitations: one of the most important conditions of applicability of the method is having a regular azimuthal distribution of noise sources (Afonin et al., 2019; Wapenaar et al., 2011). It becomes necessary to average over long time intervals due to the presence of heterogeneous and anisotropic distributions of ambient seismic noise sources during short time intervals, as described by Wapenaar et al. (2010). Specifically, in rural regions with low levels of anthropogenic and industrial noise, implementing the method necessitates extended data collection periods due to the scarcity of robust high-frequency wave sources, which exhibit a non-uniform spatial and temporal distribution. Given the results obtained in our study, we can observe that different velocity variations values are obtained for each of the components with sometimes even different trends (velocity can be increasing for a component while decreasing for another). This can be related to noise sources present in the studied area: since it is a rural area the noise generated might not be strong enough at high frequencies to apply the cross correlation method, and in addition they are not homogeneously distributed across the watershed. The main noise source in our case is the A36 highway passing through the east and south parts of the watershed, other noise sources are the departmental roads and small roads in the watershed that are not frequently passed through. Since the neighbouring nodes have mostly the same components that result with the highest negative velocity variation (see Figs. 4.26 and 4.29), this might be related to the fact that the nodes are affected by the same noise source close to their position, thus dominating a certain component.

In such situations, an alternative approach could be to use ballistic waves, which are not scattered by heterogeneities but are generated by localized sources of seismic noise, as suggested by Mulargia (2012). A stacking scheme based on the the global optimization algorithm has been proposed by Afonin et al. (2019) to apply this correlation on scattered waves in areas with low levels of anthropogenic noise. Froment (2011) has also studied the implication of the noise sources on the correlation method and derived a theoretical expression for the error on arrival times in correlations due to a non-isotropic intensity distribution in order to bypass the heterogeneity in noise sources.

Another drawback in the method in our case is the fact that there were no prolonged dry period during the data acquisition period that can be used as a reference on which we can apply the MWCS in order to compare a cross correlation function obtained during a dry period with another obtained during or after rainfall. Figure 4.37 reveals that there is a more or less regular rain distribution all year long in the sector of the studied watershed; all of the months show rather equal rain rate. In addition, we can see in Fig. 4.3 that there were rain events two month prior to the deployment of the seismic array leading to potential residual saturation in the vadose zone from past infiltration which may affect the seismic response to succeeding rain events. Having this, the method may struggle to detect subtle changes in seismic velocity if they are not significant enough to produce noticeable waveform shifts. A solution might be using a moving reference to measure daily velocity variations. In fact, the moving reference method was suggested by James et al. (2017) while monitoring seismic velocity changes in permafrost; the authors obtained large velocity changes between frozen and thawed soils, which masked seasonal cyclic variations, hence they used moving reference to compute MWCS in order to recover the lost seasonal variations. Another solution is computing the velocity changes without a reference, in other terms by calculating the velocity variations between all the pairs of daily cross-correlation functions (Brenguier et al., 2014; Gómez-García et al., 2018; Machacca-Puma et al., 2019). The latter results in a matrix of velocity variations, inversion techniques such as Bayesian least-squares inversion (Tarantola and Valette, 1982) are applied to reconstruct the time series of velocity variation. Computing dv/d without reference is usually applied to avoid the choice of an arbitrary reference cross-correlation, or for the absence of a stable periods to be chosen as reference. In our case, not using a reference can be efficient due to the absence of extended dry periods in the recording.

Fig. 4.37 Distribution of rain and temeprature throughout the year in Baume-les-Dames sector, from climate-data.org.

4.6 Summary of the main results and conclusion

In this study we applied seismic noise cross component correlation to have a localised long term monitoring, using the long term seismic station AVEN, and a 2D short term monitoring, using a 60-nodes array deployed for four months. The objective was to investigate the water saturation in soil and water transfers in the Fourbanne watershed. Two frequency ranges were used, 2-4 Hz and 8-10 Hz, to analyse deep and surface variations respectively. The applied method gave us a view on the complexity and heterogeneity of karst aquifers, sites that are challenging to investigate and understand. The results for AVEN show seasonal variations within the aquifer, related to the change in saturation of epikarstic reservoirs. The results from the nodes reveal different velocity patterns between nodes and even different patterns for the different components of the same node as response to a rain event. A summary with the results for the dv/v monitoring using the seismic array is portrayed in Fig. 4.38. The watershed was divided into 6 sections based on the observations of the dv/v patterns for the nodes. The scheme was done for two different scenarios: first one corresponding to a period with high precipitation rates and the second one corresponding to a dry period directly succeeding the rainy period. The sections were colored based on the dv/v values: white for positive values, light and dark blue for negative values, implying sections with water infiltration/transfer, dark blue indicating higher velocity drops compared to the light blue. The sections' colors have been deduced while analysing the mean dv/v values of the 3 cross-components. Finally, the components mentioned in red indicate the components that affect mostly the mean values, and are chosen while observing dv/v for the individual components. As a final observation, we can conclude:

- During precipitation events, response in term of saturation changes are obtained over all the sections of the watershed, with higher intensity response in the NE and the upper part of the conduit.
- After precipitation event, more water transfers occur in the southern section of the conduit as well as the southern part of the watershed.
- Saturation changes occur during and after precipitations along the conduit, indicating continuous infiltration.

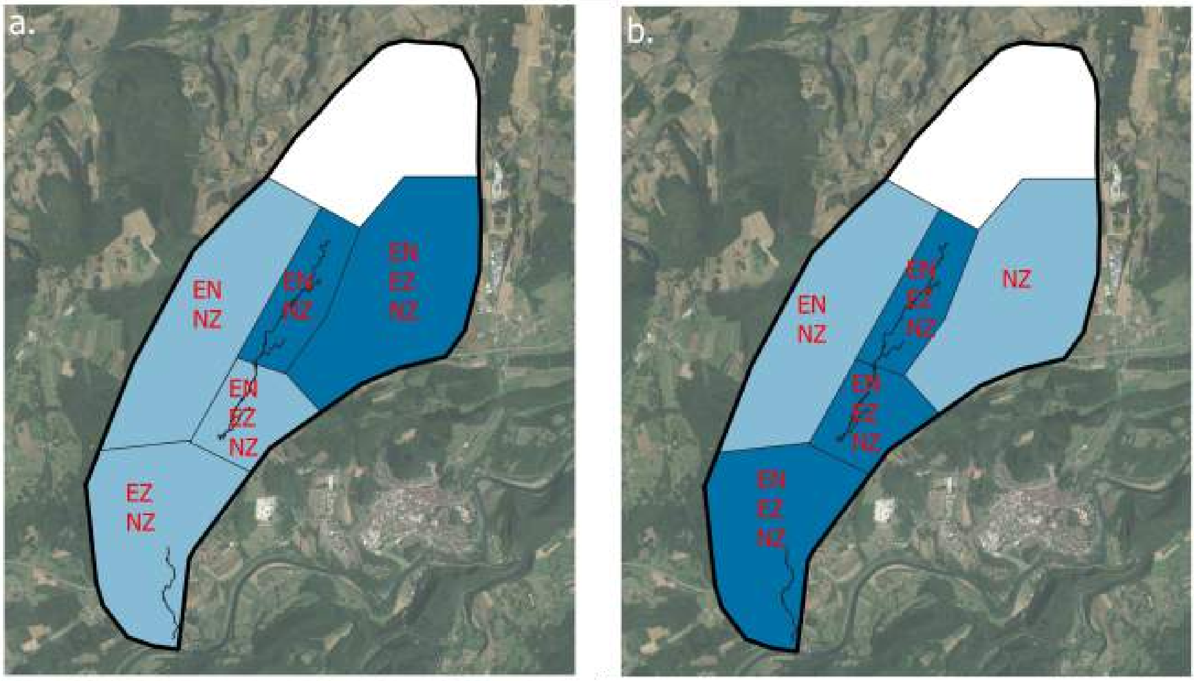


Fig. 4.38 Functional scheme representing the response of the aquifer a) during rainfall event and b) after rainfall event. The watershed was divided into 6 sections. The dark blue color represents sections with important negative dv/v, light blue for low negative dv/v and white for no negative dv/v. The components in red represent the components showing the negative dv/v in each section.

In order to better understand the nodes results inspect concerning the potential anisotropy in water saturation or the irregularity in the noise sources distribution in our studied site, ballistic waves interferometry can be tested in order to circumvent to the problem of low noise levels and irregular distributions. As well, in order to have a better understanding of the hydrogeological behavior of the watershed in response to rain events, a dry period should be included in the acquisition so it can be used as a reference. An alternative can be to apply the same method while using a moving reference or no reference at all. As a second step, cross-correlation method between nodes can be applied with sensitivity kernels in order to analyse water transfers occurring between nodes rather than punctually under each node. And finally, tomography methods can be also applied to image the subsurface and locate conduits and fractures, key elements that might help understand the seismic velocity variations.

Chapter 5

Summary and Conclusions

Throughout this manuscript we have investigated water transfer within a karst aquifer, developed and tested new methods for improving the monitoring of these reservoirs. We focused on a karst system in the Jura mountains, the Fourbanne watershed, where a surface and underground hydrogeogeophysical observatory has been set up. We deployed a dense array of 60 seismic nodes (from the end of November 2021 to the end of March 2022) and 20 rain gauges (from June 2021 to June 2022) to analyse spatio-temporally rainwater infiltration on the watershed scale. In the following, we summarize the results obtained in this study and draw a number of conclusions.

5.1 Hydrodynamic Functioning of the Underground River

This first axis, divided into two parts, aims to better investigate the underground river in terms of recharge and proposes remote methods for continuous monitoring.

In the first part, we initiated a method based on a machine learning algorithm, the Random Forests (RF) algorithm, and continuous recording of seismic noise, to predict the height of the water in the underground river. The RF is an ensemble learning method used in our case for regression, based on computing a large number of decision trees. The trees were generated based on a training set of data which were, in our case, water height and a pack of features computed from the seismic signals, that provide a characterisation of the recorded waveform. The training was done by assigning, for each time window, the water height recorded to the features corresponding to the seismic signal recorded during the same time window. We used data recorded from two broadband seismometers, the first located in the underground conduit at 20 m depth and the other at the surface, as well as water height data recorded from a hydrological probe in the underground river. The application of this method showed the following:

- Combining machine learning and seismic noise can help to retrieve remotely hydrological data.

- Surface seismometer and underground seismometer yielded an efficiency of 53% and 95% in predicting underground river water height.
- Prediction with the seismometer at the surface, even with lower accuracy compared to the seismometer underground, reveals the trends for the water height as well as flood periods.
- This is an unprecedented application for non invasive measurement of underground water height.

In the second part, we used Fiber Optic Distributed Temperature Sensing (FO-DTS) in the unsaturated zone of the karst reservoir. In Fact, 800 m fiber optics cable was deployed within the underground river, measuring temperatures each 5 minutes, with a 50 *cm* spatial resolution. Three hydrological probes were also used as temperature references for the calibration. The following results were obtained:

- River temperature show immediate variation with rain events, with rainwater arriving rapidly through sinkholes and fractures.
- The base temperature in the river increases by 1°C, related to the rising atmospheric temperature while passing from spring to summer.
- A small affluent was monitored showing temperature variations different than the river's and high electrical conductivity peaks in response to precipitations.
- The affluent is continuously supplied by epikarstic recharge outside rain events.
- During precipitations events, water height increases in the main conduit, pushing by piston-effect from a reservoir filled with long residence water and connected to the affluent.

5.2 Saturation change and water transfer in the watershed

In this chapter we used 4 years long data recorded from a broadband seismometer located in the underground conduit and 4 months data recorded from a dense array of 60 nodes to monitor seismic velocity changes (dv/v) in the watershed. We also used data from the hydrological probe measuring the underground river water height, precipitation data from a Meteo-France station in a near village (for long term data) and from 20 rain gauges deployed in the watershed. We used cross component single station correlation method, and analyse velocity variations obtained on the 3 combinations between horizontal and vertical components, and their mean, for two frequency bands 2-4 Hz and 8-10 Hz.

- Long term analysis shows seasonal velocity variations in the low frequency band, with mostly direct dv/v decrease during rainy periods, related to fast infiltration mechanisms in highly fractured media.

- Results of the seismic array show anisotropic velocity changes during and after rainfall, explained by the heterogeneity and complexity of the karst reservoir and thus anisotropy in fractures and water flow.
- Velocity variations reveal faster response to precipitations in the unsaturated part of the watershed compared with the saturated part reflecting water transfer processes.

5.3 Perspectives

In this thesis, we worked with state-of-the-art datasets and techniques to study a karstic reservoir, from the scale of the underground conduit to the scale of the whole watershed. The applied methods helped further discern the complex functioning of the Fourbanne watershed, highly fractured karst aquifer. As mentioned earlier, the main response of the studied system and water infiltration after precipitation is controlled by rapid and localized infiltration of rainwater through fractures and sinkholes. The velocity variations obtained show a heterogeneous system of fractures, with fast changes in saturation after precipitation in unsaturated areas. In addition, velocity variations computed above the explored conduit (in the unsaturated zone) show constant changes in saturation, during and after precipitation, potentially indicating continuous lateral inflows from epikarstic or rock matrix reservoirs. This is also highlighted by the FO-DTS results showing fast water transfer in interconnected structures of the aquifer during precipitations. In addition, the existence of localized reservoirs in the epikarst providing continuous recharge could be inferred. The difficulty remains in being able to locate these reservoirs, and assess their capacity. The first step could be to validate the observations of the velocity variation using other methods such as autocorrelation using ballistic waves or cross-correlation with sensitivity kernels to locate the velocity variations in space and at different depths. The latter would help to better characterize fracture networks. In addition other affluents found by speleologists along the conduit can be monitored using the FO-DTS in order to discern its lateral recharge. Going forward, it would be possible to constrain fractures/matrix/conduits patterns through numerical simulators such as FEFLOW (Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media, Diersch, 2013), that can deal with unsaturated and saturated media. Actually, measured velocity variations could be used to calibrate and constrain the model. Similar work has been done by Fores et al. (2018), using a Hydrus 1D model, and they demonstrated the great potential of dv/v in hydrological model calibration, specially by having multiple stations and working with several frequency ranges for depth sensitivity. The fractures/reservoirs distribution can be modified until reaching convergence between the simulated velocity variations and the ones recorded by the nodes. The combination of ambient seismic noise, fiber optic surveys and hydrological modeling could lead to the 4D imaging of the watershed.

This PhD benefits from a unique collaborative initiative that brings together speleologists, cave divers, farmers, landowners, students, technicians, engineers, and researchers regrouping seismologists, hydrogeologists and climatologists, all working towards a common goal: water resource management. It shows the importance of such collaborations, whether in the complementarity of information that can be provided by each member, or by the technical added value in fieldwork and data processing.

References

Acerra, C., Alguacil, G., Atakan, K., Azzara, R., Bard, P.-Y., Blarel, F., Borges, A., Cara, F., Teves-Costa, P., Duval, A.-M., Guillier, B., Grandison, M., Havskov, J., Ohrnberger, M., Rao, S., Theodoulidis, N., Tvedt, E., Utheim, T., Vidal, S., and Zacharopoulos, S. (2002). *Site Effects Assessment Using Ambient Excitations.*

Afonin, N., Kozlovskaya, E., Nevalainen, J., and Narkilahti, J. (2019). Improving the quality of empirical green's functions, obtained by cross-correlation of high-frequency ambient seismic noise. *Solid Earth*, 10:1621–1634.

Albaric, J., Kühn, D., Ohrnberger, M., Langet, N., Harris, D., Polom, U., Lecomte, I., and Hillers, G. (2021). Seismic Monitoring of Permafrost in Svalbard, Arctic Norway. *Seismological Research Letters*, 92(5):2891–2904.

Andajani, R., Tsuji, T., Snieder, R., and Ikeda, T. (2020). Spatial and temporal influence of rainfall on crustal pore pressure based on seismic velocity monitoring. *Earth Planets Space*, 72.

Andreo, B., Goldscheider, N., Vadillo, I., Vías, J. M., Neukum, C., Sinreich, M., Jiménez, P., Brechenmacher, J., Carrasco, F., Hötzl, H., Perles, M. J., and Zwahlen, F. (2006). Karst groundwater protection: First application of a pan-european approach to vulnerability, hazard and risk mapping in the sierra de líbar (southern spain). *Science of The Total Environment*, 357.

Anthony, R. E., Aster, R. C., Ryan, S., Rathburn, S., and Baker, M. G. (2018). Measuring mountain river discharge using seismographs emplaced within the hyporheic zone. *Journal of Geophysical Research: Earth Surface. 123.*, 123.

Antoni, V., Arrouays, D., Baize, D., Barriuso, E., Bispo, A., Blanca, Y., Boulonne, L., Briand, O., Brossard, M., Cabidoche, Y.-M., Caria, G., Chéry, P., Cluzeau, D., Cousin, I., Couturier, A., Decaëns, T., Denoroy, P., Dequiedt, S., Derrière, N., and Walter, C. (2011). *L'état des sols de France. Gis Sol. 2011. L'état des sols de France. Groupement d'intérêt scientifique sur les sols, 188 p.*

APSensing (2018). Détection de température distribuée (dts).

Atkinson, T. C. (1990). Geomorphology and hydrology of karst terrains by william b. white, oxford university press, oxford and new york, 1988. no of pages:464 isbn 0-19-504444-4. *Earth Surface Processes and Landforms*, 15(1):97–98.

Baillard, C., Crawford, W. C., Ballu, V., Hibert, C., and Mangeney, A. (2013). An Automatic Kurtosis-Based P- and S-Phase Picker Designed for Local Seismic Networks. *Bulletin of the Seismological Society of America*, 104(1):394–409.

Baize, D., Girard, M., Arrouays, D., Aubert, G., Aubrun, A., Aurousseau, P., Beaudou, A., Bornand, M., Boulaine, J., Bresson, L., Brethes, A., Chrétien, J., Favrot, J., Jabiol, B., Jamagne, M., Kaloga, B., Laplace-Dolonde, A., Loyer, J.-Y., Lupascu, G., and Vizier, J.-F. (1995). Référentiel pédologique 1995.

Bakalowicz, M. (1999). *Connaissance et gestion des ressources en eaux souterraines dans les régions karstiques : Agence de l'eau Rhône-Méditerranée-Corse.*, volume 1,13.

Bakalowicz, M. (2005). Karst groundwater: a challenge for new resources. *Hydrogeology Journal*, 13:148–160.

Barton, N. (2009). *Rock Quality, Seismic Velocity, Attenuation and Anisotropy*, volume 15.

Bell, F., Cripps, J., and Culshaw, M. (2001). A review of the engineering behaviour of soils and rocks with respect to groundwater. *Geological Society, London, Engineering Geology Special Publications*, 17:299–306.

Bermejo, L., Ortega, A., Parés, J., Campaña Lozano, I., Bermúdez de Castro, J.-M., Carbonell, E., and Conyers, L. (2020). Karst features interpretation using ground-penetrating radar: A case study from the sierra de atapuerca, spain. *Geomorphology*, 367:107311.

Best, A., Sothcott, J., and McCann, C. (2007). A laboratory study of seismic velocity and attenuation anisotropy in near-surface sedimentary rocks. *Geophysical Prospecting*, 55:609–625.

Bichet, V. and Campy, M. (2008). *Montagnes du Jura, géologie et paysages.*

Binley, A. and Kemna, A. (2005). *DC Resistivity and Induced Polarization Methods*, volume 50, pages 129–156.

Blavoux, B. and Mudry, J. (1986). Influence des pluies estivales sur la qualite des reserves de l'aquifere karstique; role du sol et de l'epikarst dans la concentration des chlorures. *Bulletin de la Société Géologique de France*, II(4):667–674.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N., Briand, X., Emoto, K., and Miyake, H. (2014). Earthquake dynamics. mapping pressurized volcanic fluids from induced crustal seismic velocity drops. *Science (New York, N.Y.)*, 345:80–2.

Burtin, A., Bollinger, L., Vergne, J., Cattin, R., and Nábělek, J. L. (2008). Spectral analysis of seismic noise induced by rivers: A new tool to monitor spatiotemporal changes in stream hydrodynamics. *Journal of Geophysical Research: Solid Earth*, 113(B5).

Caldwell, T., Wolaver, B., Bongiovanni, T., Pierre, J., Robertson, S., Abolt, C., and Scanlon, B. (2020). Spring discharge and thermal regime of a groundwater dependent ecosystem in an arid karst environment. *Journal of Hydrology*, 587:124947.

Caudron, C., Girona, T., Jolly, A., Christenson, B., Savage, M., Carniel, R., Lecocq, T., Kennedy, B., Lokmer, I., Yates, A., Hamling, I., Park, I., Kilgour, G., and Mazot, A. (2021). A quest for unrest in multiparameter observations at whakaari/white island volcano, new zealand 2007–2018. *Earth, Planets and Space*, 73.

Celle-Jeanton, H., Emblanch, C., Mudry, J.-N., and Charmoille, A. (2003). Contribution of time tracers (mg2+, toc, 13ctdic, no3) to understand the role of the unsaturated zone: A case study—karst aquifers in the doubs valley, eastern france. *Geophysical Research Letters - GEOPHYS RES LETT*, 30.

Champollion, C., Deville, S., Chery, J., Doerflinger, E., Le Moigne, N., Bayer, R., Vernant, P., and Mazzilli, N. (2018). Estimating epikarst water storage by time-lapse surface-to-depth gravity measurements. *Hydrology and Earth System Sciences*, 22:3825–3839.

Charmoille, A. (2005). TraÇage hydrochimique des interactions hydrauliques et mecaniques entre les volumes permeables et peu permeables au sein des aquiferes fractures carbonates aquifère karstique de fourbanne (avant-pays jurassien, 25) laboratoire naturel de coaraze (alpes méridionales, 06).

Chen, Z., Auler, A., Bakalowicz, M., Drew, D., Griger, F., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Stevanovic, Z., Veni, G., and Goldscheider, N. (2017). The world karst aquifer mapping project: concept, mapping procedure and map of europe. *Hydrogeology Journal*, 25.

Cheng, Q.-B., Chen, X., Tao, M., and Binley, A. (2019). Characterization of karst structures using quasi-3d electrical resistivity tomography. *Environmental Earth Sciences*, 78.

Chmiel, M., Walter, F., Wenner, M., Zhang, Z., McArdell, B. W., and Hibert, C. (2021). Machine learning improves debris flow warning. *Geophysical Research Letters*, 48(3).

Cholet, C. (2017). *Fonctionnement hydrogéologique et processus de transport dans les aquifères karstiques du Massif du Jura*. PhD thesis.

Cholet, C., Charlier, J.-B., Moussa, R., Steinmann, M., and Denimal, S. (2017). Assessing lateral flows and solute transport during floods in a conduit-flow-dominated karst system using the inverse problem for the advection–diffusion equation. *Hydrology and Earth System Sciences*, 21(7):3635–3653.

Cholet, C., Steinmann, M., Charlier, J.-B., and Denimal, S. (2015). Comparative study of the physicochemical response of two karst systems during contrasting flood events in the french jura mountains. *Environmental Earth Sciences*, 1:1–9.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836.

Criminisi, A., Shotton, J., and Konukoglu, E. (2012). *Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning*. Number 2-3. NOW Publishers, foundations and trends® in computer graphics and vision: vol. 7: no 2-3, pp 81-227 edition.

Day-Lewis, F., Johnson, C., Dawson, C., Nelms, D., Eddy-Miller, C., Wheeler, J., Riverton, W., and Harvey, C. (2008). Fiber-optic distributed temperature sensing: A new tool for assessment and monitoring of hydrologic processes.

De Plaen, R. S. M., Cannata, A., Cannavo', F., Caudron, C., Lecocq, T., and Francis, O. (2019). Temporal changes of seismic velocity caused by volcanic activity at mt. etna revealed by the autocorrelation of ambient seismic noise. *Frontiers in Earth Science*, 6.

De Plaen, R. S. M., Lecocq, T., Caudron, C., Ferrazzini, V., and Francis, O. (2016). Single-station monitoring of volcanoes using seismic ambient noise. *Geophysical Research Letters*, 43(16):8511–8518.

Dean, T. (2018). The seismic signature of rain. *ASEG Extended Abstracts*, 2018:1.

Des Tombe, B., Schilperoort, B., and Bakker, M. (2020). Estimation of temperature and associated uncertainty from fiber-optic raman-spectrum distributed temperature sensing. *Sensors*, 20:2235.

Diersch, H.-J. (2013). *FEFLOW—Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media*.

Dietze, M., Lagarde, S., Halfi, E., Laronne, J. B., and Turowski, J. M. (2019). Joint sensing of bedload flux and water depth by seismic data inversion. *Water Resources Research*, 55(11):9892–9904.

Doolittle, J., Jenkinson, B., Hopkins, D., Ulmer, M., and Tuttle, W. (2006). Hydropedological investigations with ground-penetrating radar (gpr): Estimating water-table depths and local ground-water flow pattern in areas of coarse-textured soils. *Geoderma*, 131:317–329.

Dore, E. (2022). Cartographie de la réserve utile en eau des sols et cartographie des risques sur le bassin versant karstique de fourbanne. Master's thesis.

Drew, D. and Hötzl, H. (1999). *Karst hydrogeology and human activities: impacts, consequences and implications*. A.A. Balkema.

Díaz, J., Ruíz, M., Crescentini, L., Amoruso, A., and Gallart, J. (2014). Seismic monitoring of an Alpine mountain river. *Journal of Geophysical Research: Solid Earth*, 119(4):3276–3289.

Dörfliger, N., Crochet, P., Guerin, R., Jozja, N., Marsaud, B., Mondain, P. H., Muet, P., and V., P. (2010). Guide méthodologique. les outils de l'hydrogéologie karstique pour la caractérisation de la structure et du fonctionnement des systèmes karstiques et l'évaluation de leur ressource. *BRGM*.

Einsiedl, F. (2005). Flow system dynamics and water storage of a fissured-porous karst aquifer characterized by artificial and environmental tracers. *Journal of Hydrology*, 312(1):312–321.

Farahani, M. and Gogolla, T. (1999). Spontaneous raman scattering in optical fibers with modulated probe light for distributed temperature raman remote sensing. *Journal of Lightwave Technology*, 17(8):1379–1391.

Fenta, M. and Szanyi, J. (2021). Fibre optic methods of prospecting: A comprehensive and modern branch of geophysics. *Surveys in Geophysics*.

Flinders, A., Caudron, C., Johanson, I., Taira, T., Shiro, B., and Haney, M. (2020). Seismic velocity variations associated with the 2018 lower east rift zone eruption of kīlauea, hawaii. *Bulletin of Volcanology*, 82.

Ford, D. and Williams, P. D. (2007). *Karst hydrogeology and geomorphology*. John Wiley & Sons.

Fores, B., Champollion, C., Le Moigne, N., Bayer, R., and Chery, J. (2017). Assessing the precision of the igrav superconducting gravimeter for hydrological models and karstic hydrological process identification. *Geophysical Journal International*, 208:269–280.

Fores, B., Champollion, C., Mainsant, G., Albaric, J., and Fort, A. (2018). Monitoring saturation changes with ambient seismic noise and gravimetry in a karst environment. *Vadose Zone Journal*, 17(1):170163.

Froment, B. (2011). *Utilisation du bruit sismique ambiant dans le suivi temporel de structures geologiques*. Theses, Universite de Grenoble.

Gaillardet, J., Braud, I., Hankard, F., Anquetin, S., Bour, O., Dörfliger, N., de Dreuzy, J.-R., Galle, S., Galy, C., Gogo, S., Gourcy, L., Habets, F., Laggoun, F., Longuevergne, L., Borgne, T., Naaim-Bouvet, F., Nord, G., Simonneaux, V., Six, D., and Zitouna-Chebbi, R. (2018). Ozcar: The french network of critical zone observatories. *Vadose Zone Journal*, 17.

Gassmann, F. (1951). Elastic waves through a packing of spheres. *GEOPHYSICS*, 16(4):673–685.

Gimbert, F., Tsai, V. C., Amundson, J. M., Bartholomaus, T. C., and Walter, J. I. (2016). Subseasonal changes observed in subglacial channel pressure, size, and sediment transport. *Geophysical Research Letters*, 43(8):3786–3794.

Gimbert, F., Tsai, V. C., and Lamb, M. P. (2014). A physical model for seismic noise generation by turbulent flow in rivers. *Journal of Geophysical Research (Earth Surface)*, 119(10).

Goldscheider, N. and Drew, D. (2014). *Methods in karst hydrogeology: IAH: International Contributions to Hydrogeology, 26.* Crc Press.

Goldscheider, N., Meiman, J., Pronk, M., and Smart, C. (2008). Tracer test in karst hydrogeology and speleology. *International Journal of Speleology*, 37:27–40.

Green, T. R., Taniguchi, M., Kooi, H., Gurdak, J. J., Allen, D. M., Hiscock, K. M., Treidel, H., and Aureli, A. (2011). Beneath the surface of global change: Impacts of climate change on groundwater. *Journal of Hydrology*, 405(3):532–560.

Grimes, K. (2006). Syngenetic karst in australia: A review. *Helictite*, 39:27–38.

Gómez-García, C., Brenguier, F., Boué, P., Shapiro, N., Droznin, D., Droznina, S., Senyukov, S., and Gordeev, E. (2018). Retrieving robust noise-based seismic velocity changes from sparse data sets: synthetic tests and application to klyuchevskoy volcanic group (kamchatka). *Geophysical Journal International*, 214:1218–1236.

Han, T., Gurevich, B., Fu, L.-Y., Qi, Q., Wei, J., and Chen, X. (2020). Combined effects of pressure and water saturation on the seismic anisotropy in artificial porous sandstone with aligned fractures. *Journal of Geophysical Research: Solid Earth*, 125(1):e2019JB019091.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., and Weiler, M. (2014). Karst water resources in a changing world: Review of hydrological modeling approaches. *Reviews of Geophysics*.

Hausner, M. and Kobs, S. (2016). Identifying and correcting step losses in single-ended fiber-optic distributed temperature sensing data. *Journal of Sensors*, in press.

Hausner, M., Suárez, F., Glander, K., van de Giesen, N., Selker, J., and Tyler, S. (2011). Calibrating single-ended fiber-optic raman spectra distributed temperature sensing data. *Sensors (Basel, Switzerland)*, 11:10859–79.

Henriet, Y. (2022). Cartographie du réservoir utile en eau des sols agricole d'un bassin versant karstique (fourbanne, doubs). application à l'évaluation des transferts. Master's thesis.

Herrmann, R. B. (2013). Computer Programs in Seismology: An Evolving Tool for Instruction and Research. *Seismological Research Letters*, 84(6):1081–1088.

Hibert, C., Provost, F., Malet, J.-P., Maggi, A., Stumpf, A., and Ferrazzini, V. (2017). Automatic identification of rockfalls and volcano-tectonic earthquakes at the Piton de la Fournaise volcano using a Random Forest algorithm. *Journal of Volcanology and Geothermal Research*, 340:130–142.

Hobiger, M., Wegler, U., Shiomi, K., and Nakahara, H. (2014). Single-station cross-correlation analysis of ambient seismic noise: Application to stations in the surroundings of the 2008 iwate-miyagi nairiku earthquake. *Geophysical Journal International*, 198.

Houze, R. (1993). Cloud dynamics. *International Geophysics Series*, 53.

Hurtig, E., Großwig, S., Jobmann, M., Kühn, K., and Marschall, P. (1994). Fibre-optic temperature measurements in shallow boreholes: experimental application for fluid logging. *Geothermics*, 23(4):355–364.

Jamagne, M., Begon, J. C., and Hardy, R. (1977). Soil mapping, a vital element in the management and conservation of rural resources. *Pedologie*, 27(1):9–43.

James, S. R., Knox, H. A., Abbott, R. E., and Screaton, E. J. (2017). Improved moving window cross-spectral analysis for resolving large temporal seismic velocity changes in permafrost. *Geophysical Research Letters*, 44(9):4018–4026.

Jourde, H., Massei, N., Mazzilli, N., Binet, S., Batiot-Guilhe, C., Labat, D., Steinmann, M., Bailly-Comte, V., Seidel, J., Arfib, B., Charlier, J., Guinot, V., Jardani, A., Fournier, M., Aliouache, M., Babic, M., Bertrand, C., Brunet, P., Boyer, J., Bricquet, J., Camboulive, T., Carrière, S., Celle-Jeanton, H., Chalikakis, K., Chen, N., Cholet, C., Clauzon, V., Soglio, L. D., Danquigny, C., Défargue, C., Denimal, S., Emblanch, C., Hernandez, F., Gillon, M., Gutierrez, A., Sanchez, L. H., Hery, M., Houillon, N., Johannet, A., Jouves, J., Jozja, N., Ladouche, B., Leonardi, V., Lorette, G., Loup, C., Marchand, P., de Montety, V., Muller, R., Ollivier, C., Sivelle, V., Lastennet, R., Lecoq, N., Maréchal, J. C., Perotin, L., Perrin, J., Petre, M., Peyraube, N., Pistre, S., Plagnes, V., Probst, A., Probst, J., Simler, R., Stefani, V., Valdes-Lao, D., Viseur, S., and Wang, X. (2018). Sno karst: A french network of observatories for the multidisciplinary study of critical zone processes in karst watersheds and aquifers. *Vadose Zone Journal*, 17.

Karato, S.-i. and Wong, T.-f. (1995). Rock deformation: Ductile and brittle. *Reviews of Geophysics*, 33(S1):451–457.

Kawakami, S. and Nishida, S. (1974). Characteristics of a doubly clad optical fiber with a low-index inner cladding. *IEEE Journal of Quantum Electronics*, 10(12):879–887.

Kim, D. and Lekic, V. (2019). Groundwater variations from autocorrelation and receiver functions. *Geophysical Research Letters*, 46.

Kim, Y., Hardisty, R., Torres Parada, E., and Marfurt, K. (2018). Seismic-facies classification using random forest algorithm. pages 2161–2165.

Knochenmus, L. and Robinson, J. (1996). Descriptions of anisotropy and heterogeneity and their effect on ground-water flow and areas of contribution to public supply wells in a karst carbonate aquifer system. *US Geological Survey Water Supply Paper*, 2475:1–43.

Küperkoch, L., Meier, T., Lee, J., Friederich, W., and Group, E. W. (2010). Automated determination of P-phase arrival times at regional and local distances using higher order statistics. *Geophysical Journal International*, 181(2):1159–1170.

Larose, E., Carrière, S., Voisin, C., Bottelin, P., Baillet, L., Gueguen, P., Walter, F., Jongmans, D., Guillier, B., Garambois, S., Gimbert, F., and Massey, C. (2015). Environmental seismology: What can we learn on earth surface processes with ambient noise? *Journal of Applied Geophysics*, 116.

Lecocq, T., Caudron, C., and Brenguier, F. (2014). Msnoise, a python package for monitoring seismic velocity changes using ambient seismic noise. *Seismological Research Letters*, 85:715–726.

Lecocq, T., Longuevergne, L., Pedersen, H. A., Brenguier, F., and Stammler, K. (2017). Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modeling. *Scientific Reports*, 7(1).

Liang, Z., Wei, J., Zhao, J., Liu, H., Li, B., Shen, J., and Zheng, C. (2008). The statistical meaning of kurtosis and its new application to identification of persons based on seismic signals. *Sensors (Basel, Switzerland)*, 8(8):5106–5119.

Loke, M. (2001). Electrical imaging surveys for environmental and engineering studies. *A practical guide to 2D and 3D surveys*.

Loke, M. (2012). *2-D and 3-D Electrical Imaging Surveys*.

Lowe, M., Donaldson, N., Spiro, A., Gosselin, C., and Nadeau, M. (2015). Fiber optics in medicine module. *MedEdPORTAL*, 11:10292.

Ma, M., Wang, S., Yuan, S., Wang, J., and Wang, T. (2015). The comparison of skewness and kurtosis criteria for wavelet phase estimation.

Machacca-Puma, R., Lesage, P., Larose, E., Lacroix, P., and Anccasi-Figueroa, R. M. (2019). Detection of pre-eruptive seismic velocity variations at an andesitic volcano using ambient noise correlation on 3-component stations: Ubinas volcano, peru, 2014. *Journal of Volcanology and Geothermal Research*, 381:83–100.

Mainsant, G., Larose, E., Brönnimann, C., Jongmans, D., Michoud, C., and Jaboyedoff, M. (2012a). Ambient seismic noise monitoring of a clay landslide: Toward failure prediction. *Journal of Geophysical Research (Earth Surface)*, 117:1030–.

Mainsant, G., Larose, E., Brönnimann, C., Jongmans, D., Michoud, C., and Jaboyedoff, M. (2012b). Ambient seismic noise monitoring of a clay landslide: Toward failure prediction. *Journal of Geophysical Research: Earth Surface*, 117(F1).

Malz, A., Madritsch, H., and Kley, J. (2015). Improving 2d seismic interpretation in challenging settings by integration of restoration techniques: A case study from the jura fold-and-thrust belt (switzerland). *Interpretation*, 3(4):SAA37–SAA58.

Maréchal, J.-C. and Rouillard, J. (2020). Groundwater in France: Resources, Use and Management Issues. In *Sustainable Groundwater Management: A Comparative Analysis of French and Australian Policies and Implications to Other Countries*, Global Issues in Water Policy, pages 17–45. Springer International Publishing, Cham.

McCuen, R. H., Knight, Z., and Cutter, A. G. (2006). Evaluation of the nash–sutcliffe efficiency index. *Journal of Hydrologic Engineering*, 11(6):597–602. Publisher: American Society of Civil Engineers.

McDonnell, J. J. (2017). Beyond the water balance. 10(6):396–396.

Mordret, A. (2014). *Utilisation des correlations de bruit sismique dans un contexte de réseau industriel dense : application au réseau de Valhall*. PhD thesis. Thèse de doctorat dirigée par Shapiro, Nikolai M. Géophysique Institut de physique du globe (Paris ; 1921-....) 2014.

Muchaidze, I. (2008). Imaging in karst terrain using electrical resistivity tomography. Master's thesis.

Mulargia, F. (2012). The seismic noise wavefield is not diffuse. *The Journal of the Acoustical Society of America*, 131(4):2853–2858.

Mwakanyamale, K., Slater, L., Day-Lewis, F., Elwaseif, M., and Johnson, C. (2012). Spatially variable stage-driven groundwater-surface water interaction inferred from time-frequency analysis of distributed temperature sensing data. *Geophysical Research Letters*, 39(6).

Obermann, A., Planès, T., Larose, E., and Campillo, M. (2013). Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise. *Journal of Geophysical Research: Solid Earth*, 118(12):6285–6294.

Poli, P., Marguin, V., Wang, Q., D'Agostino, N., and Johnson, P. (2020). Seasonal and coseismic velocity variation in the region of l'aquila from single station measurements and implications for crustal rheology. *Journal of Geophysical Research: Solid Earth*, 125(7):e2019JB019316.

Poupinet, G., Ellsworth, W. L., and Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the calaveras fault, california. *Journal of Geophysical Research: Solid Earth*, 89(B7):5719–5731.

Prasad, M. and Nur, A. (2005). Velocity and attenuation anisotropy in reservoir rocks. pages 1652–1655.

Provost, F., Hibert, C., and Malet, J.-P. (2017). Automatic classification of endogenous landslide seismicity using the random forest supervised classifier. *Geophysical Research Letters*, 44(1):113–120.

Ridremont, F., Lejeune, P., and Claessens, H. (2011). Méthode pragmatique d'évaluation de la réserve en eau des stations forestières et cartographie à l'échelle régionale (Wallonie, Belgique). *BASE*.

Ross, Z. E. and Ben-Zion, Y. (2014). Automatic picking of direct P, S seismic phases and fault zone head waves. *Geophysical Journal International*, 199(1):368–381.

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., and Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, 44(18):9276–9282.

Schmandt, B., Aster, R. C., Scherler, D., Tsai, V. C., and Karlstrom, K. (2013). Multiple fluvial processes detected by riverside seismic and infrasound monitoring of a controlled flood in the grand canyon. *Geophysical Research Letters*, 40(18):4858–4863.

Sens-Schönfelder, C. and Wegler, U. (2006). Passive image interferometry and seasonal variations of seismic velocities at merapi volcano, indonesia. *Geophysical Research Letters*, 33:L21302.

Shapiro, N. M. and Campillo, M. (2004). Emergence of broadband rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, 31(7).

Simon, N., Bour, O., Lavenant, N., Porel, G., Nauleau, B., Pouladi, B., Longuevergne, L., and Crave, A. (2021). Numerical and experimental validation of the applicability of active-dts experiments to estimate thermal conductivity and groundwater flux in porous media. *Water Resources Research*, 57.

Sommaruga, A. (1996). *Geology of the central Jura and the molasse basin: new insight into an evaporite-based foreland fold and thrust belt*. PhD thesis.

Tarantola, A. and Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2):219–232.

Torrese, P. (2019). Investigating karst aquifers: Using pseudo 3-d electrical resistivity tomography to identify major karst features. *Journal of Hydrology*, 580:124257.

Tsai, V., Minchew, B., Lamb, M., and Ampuero, J. P. (2012). A physical model for seismic noise generation from sediment transport in rivers. *Geophysical Research Letters*, 39.

Tsai, V. C. (2009). On establishing the accuracy of noise tomography travel-time measurements in a realistic medium. *Geophysical Journal International*, 178(3):1555–1564.

Tsai, V. C. (2011). A model for seasonal changes in gps positions and seismic wave speeds due to thermoelastic and hydrologic variations. *Journal of Geophysical Research: Solid Earth*, 116(B4).

Ukil, A., Braendle, H., and Krippner, P. (2012). Distributed temperature sensing: Review of technology and applications. *IEEE Sensors Journal - IEEE SENS J*, 12:885–892.

van de Giesen, N., Steele-Dunne, S., Jansen, J., Hoes, O., Hausner, M., Tyler, S., and Selker, J. (2012). Double-ended calibration of fiber-optic raman spectra distributed temperature sensing data. *Sensors (Basel, Switzerland)*, 12:5471–85.

Vauthier, Q. (2011). Estimation des réserves utiles des sols de Seine-et-Marne. Master's thesis, UHP - Université Henri Poincaré.

Vidal, C., Zaccarelli, L., Pintori, F., Bragato, P., and Serpelloni, E. (2021). Hydrological effects on seismic-noise monitoring in karstic media. *Geophysical Research Letters*, 48.

Vogelgesang, J., Holt, N., Schilling, K., Gannon, M., and Tassier-Surine, S. (2019). Using high-resolution electrical resistivity to estimate hydraulic conductivity and improve characterization of alluvial aquifers. *Journal of Hydrology*, 580:123992.

Vogt, T., Schneider, P., Hahn-Woernle, L., and Cirpka, O. (2010). Estimation of seepage rates in a losing stream by means of fiber-optic high-resolution temperature profiling. *Journal of Hydrology*, 380:154–164.

Voisin, C., Guzmán, M. A. R., Réfloch, A., Taruselli, M., and Garambois, S. (2017). Groundwater Monitoring with Passive Seismic Interferometry. *J Water Resour Prot*, 9(12):1414–1427.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. (2010). Tutorial on seismic interferometry. part i: Basic principles and applications. *Geophysics*, 75:75A195–75209.

Wapenaar, K., van der Neut, J., Ruigrok, E., Draganov, D., Hunziker, J., Slob, E., Thorbecke, J., and Snieder, R. (2011). Seismic interferometry by crosscorrelation and by multidimensional deconvolution: A systematic comparison. *Geophysical Journal International*, 12:5758.

Wenner, M., Hibert, C., van Herwijnen, A., Meier, L., and Walter, F. (2021). Near-real-time automated classification of seismic signals of slope failures with continuous random forests. *Natural Hazards and Earth System Sciences*, 21(1).

Williams, P. (2008). The role of the epikarst in karst and cave hydrogeology: A review. *International Journal of Speleology*, 37.

Winterstein, D. F. (1990). Velocity anisotropy terminology for geophysicists. *GEOPHYSICS*, 55(8):1070–1088.

Zou, C., Zhao, L., Xu, M., Chen, Y., and Geng, J. (2021). Porosity prediction with uncertainty quantification from multiple seismic attributes using random forest. *Journal of Geophysical Research: Solid Earth*, 126(7).

Appendix A

Nodes data-sheet

Table A.1 is a listing of the position of the nodes, their altitude, the slope (%), the slope's aspect and the water reserve of their corresponding position. Altitude, slope and aspect are obtained via a terrain analysis, conducted through QGIS, based on the Digital Elevation Models (DEMs) of the watershed, at a 1 m resolution. Values indicated for each node result from averaging continuous values calculated at a 5 m radius around the node. Slope informs about how steep the terrain is, aspect is the compass direction of the slope. An aspect of 0 (or 360) means that the slope is North-facing, 90 East-facing, 180 South-facing, and 270 West-facing.

Table A.1 The deployed seismic nodes with their corresponding coordinates (WGS 84), the slope (%), the slope's aspect (in degrees, relative to the North direction) and the water reserve of their corresponding position.

Node	Latitude	Longitude	Altitude (m)	Slope (%)	Aspect (degrees)	ACW (mm)
N1	47.3683	6.29024	430.1	7.6	262.9	90
N2	47.3341	6.29591	340.1	0.3	301.8	119
N3	47.3728	6.30473	416.6	2.2	214.0	63
N4	47.3524	6.30497	329.1	4.6	239.9	92
N5	47.3438	6.28673	363.9	3.6	319.3	210
N6	47.3454	6.30017	329.7	4.6	244.9	39
N7	47.3533	6.30088	341.3	4.2	124.9	214
N8	47.346	6.29416	333.8	15.6	56.4	138
N9	47.354	6.30822	341.8	2.4	278.1	63
N10	47.3573	6.30497	331.7	1.2	178.1	117
N11	47.3627	6.31299	350.5	4.4	225.0	78
N12	47.3639	6.31635	376.4	6.9	269.1	78
N13	47.3308	6.30052	261.5	2.1	225.0	121
N14	47.3674	6.3176	374.8	7.2	290.6	61
N15	47.3649	6.31423	363.2	2.5	170.2	78
N16	47.37	6.32056	380.9	7.8	298.0	37
N17	47.3698	6.3121	380.7	3.4	220.3	100
N18	47.3726	6.32	377.6	7.4	53.7	61
N19	47.3763	6.32636	378.9	10.5	309.3	50
N20	47.3775	6.317	419.6	7.2	104.0	30
N21	47.3789	6.32518	364.0	7.7	182.9	85
N22	47.3799	6.32873	372.6	6.4	209.8	111
N23	47.3833	6.33243	415.8	5.3	117.9	44
N24	47.3865	6.33783	373.7	7.5	159.9	225
N25	47.3883	6.36044	402.1	4.7	84.4	225
N26	47.3914	6.34205	376.5	0.5	22.5	232
N27	47.4136	6.36974	393.4	2.2	269.7	227
N28	47.3488	6.30375	330.4	10.7	216.7	138
N29	47.3957	6.35608	402.0	3.2	291.8	235
N30	47.3911	6.34838	376.7	3.2	100.6	35
N31	47.3786	6.35769	424.7	3.4	257.8	35
N32	47.375	6.34789	426.5	5.3	312.0	33
N33	47.3682	6.32069	398.2	7.3	284.3	37
N34	47.3845	6.37388	418.2	4.7	271.0	156
N35	47.4214	6.36335	431.0	3.6	21.0	223
N36	47.4202	6.35097	378.9	13.4	305.8	63
N37	47.4096	6.35512	413.7	13.7	103.5	234
N38	47.409	6.34002	436.8	3.7	217.8	48
N39	47.4003	6.3606	381.2	0.0	233.4	176
N40	47.398	6.36879	390.6	2.5	223.8	227
N41	47.3958	6.32903	395.9	1.4	303.0	235
N42	47.3989	6.35047	379.2	1.0	193.8	58
N43	47.3323	6.3002	331.8	14.0	147.1	121
N44	47.3717	6.32644	415.4	8.4	298.0	39
N45	47.3749	6.32328	355.9	8.0	262.5	85
N46	47.3662	6.31889	387.7	7.8	287.6	78
N47	47.3572	6.28667	407.8	2.3	156.1	71
N48	47.3759	6.324	363.8	7.1	280.1	85
N49	47.382	6.34324	382.2	2.4	100.6	224
N50	47.3757	6.36768	437.9	4.5	61.6	130
N51	47.4142	6.34182	388.9	10.5	324.5	63
N52	47.3608	6.32516	434.1	9.6	178.2	33
N53	47.3724	6.32207	373.7	9.2	312.4	85
N54	47.4023	6.35248	386.9	10.5	150.0	234
N55	47.3401	6.30047	334.9	13.3	31.0	138
N56	47.3518	6.31121	343.9	4.5	234.2	63
N57	47.3857	6.31517	460.5	7.5	290.4	38
N58	47.4064	6.37562	414.7	15.4	277.1	170
N59	47.3905	6.32425	424.4	6.8	80.7	174
N60	47.3985	6.31645	412.0	4.3	179.6	96.5

Appendix B

Appendix for publication in Chapter 3

B.1 Spectrograms of the horizontal components of the seismic signals

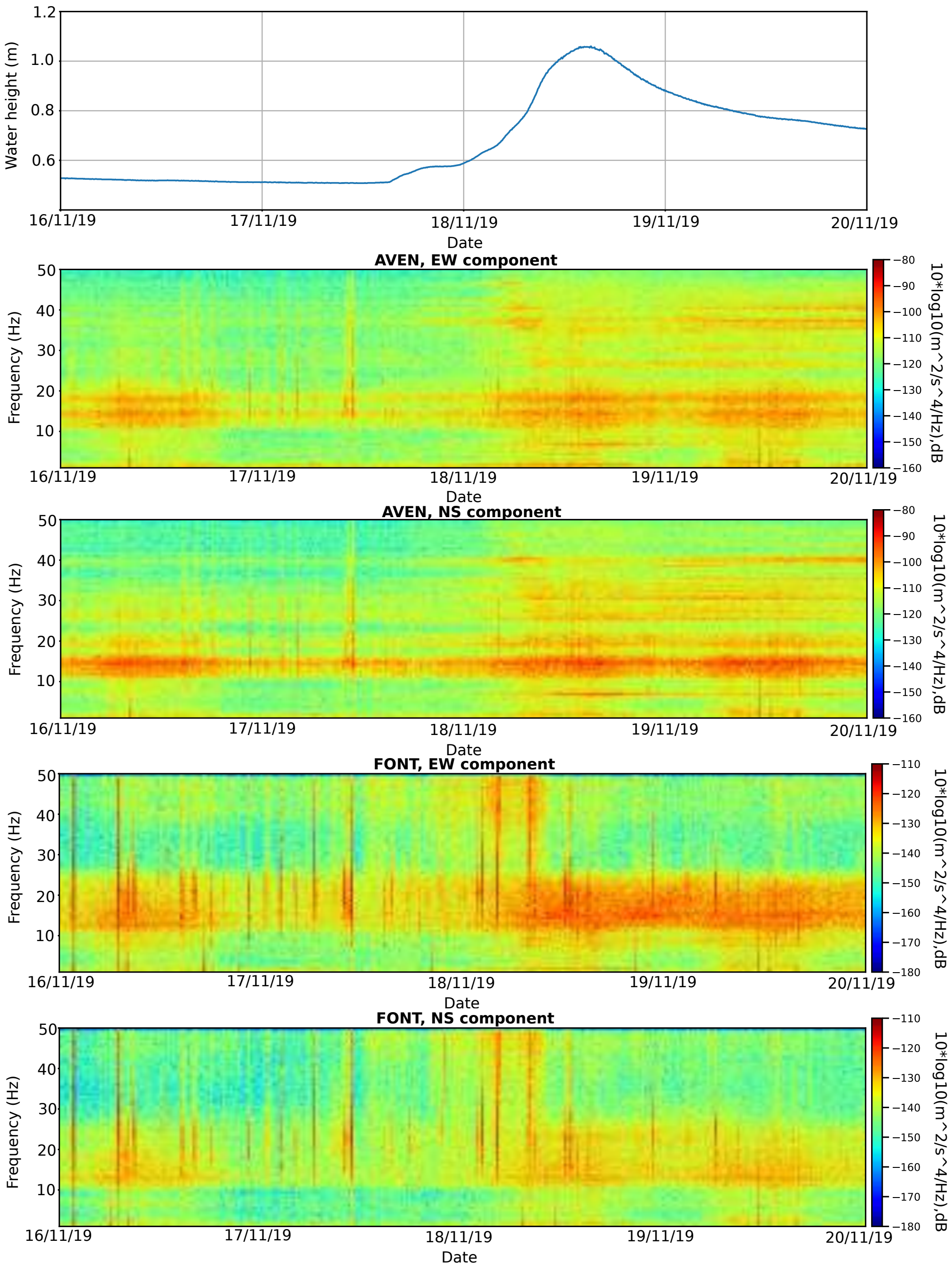


Fig. B.1 Water height for the flood recorded between 16/11/19 and 20/11/19 along with the spectrograms of the horizontal components during the selected flood for signals recorded at AVEN and FONT and filtered between 1 and 50 Hz. EW corresponds to the East-West component, NS to the North-South component.

B.2 Temporal evolution of features

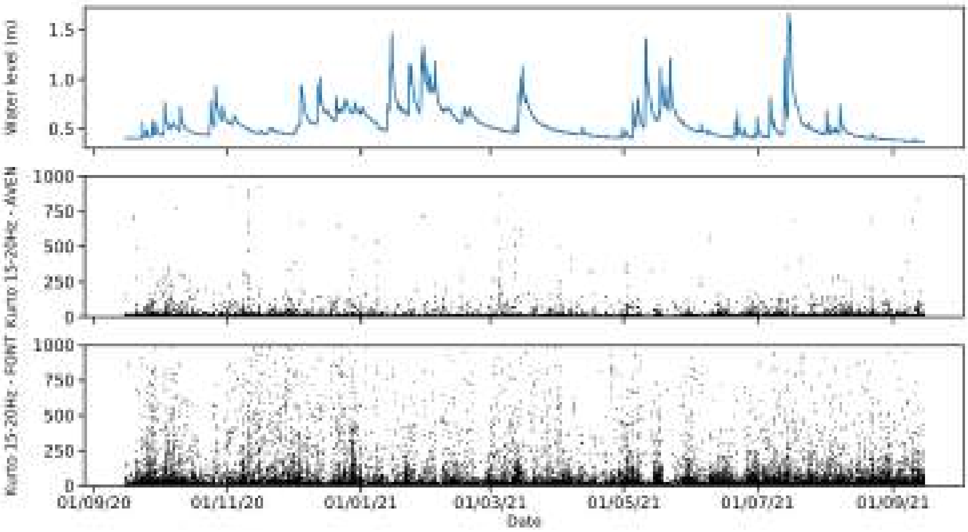


Fig. B.2 Water height recorded for the hydrological cycle used for the training and the signal's kurtosis feature between 15 and 20 Hz for the signals recorded at AVEN and FONT during the same cycle.

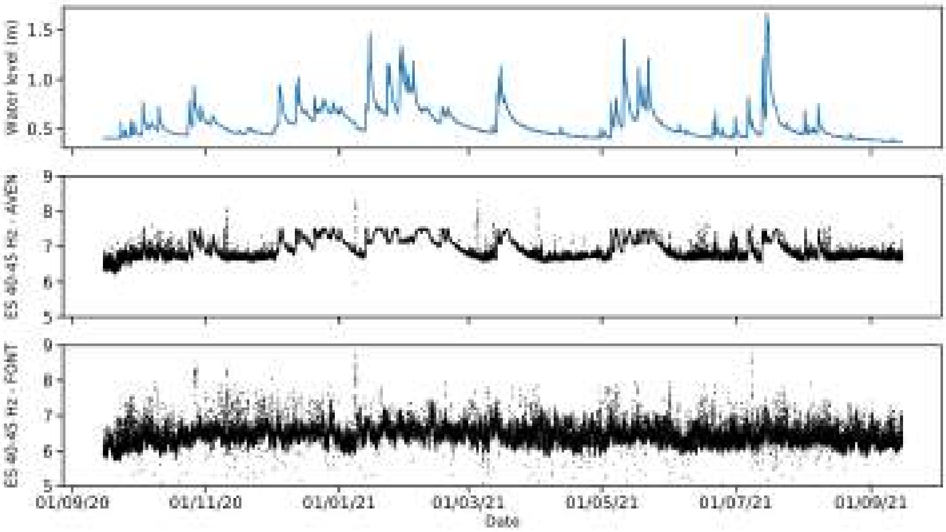


Fig. B.3 Water height recorded for the hydrological cycle used for the training and the signal's energy feature between 40 and 45 Hz for the signals recorded at AVEN and FONT during the same cycle.

Appendix C

The Montbéliard earthquakes

Two regional earthquakes were recorded by the installed network: they occurred on the $24^t h$ and $25^t h$ of December, SE of Montbeliard (ML 4.1 and 3.8, respectively; see Fig. C.2).The first one is a M_L=4.1, occurred at 24/12/21 23h59, 20 km (France) , and the second one is a M_L=3.8, occurred at 25/12/21 14h49, 18 km SE of Montbéliard (France). The vertical components of the seismic signals recorded for these earthquakes at node 49 are shown in Fig. C.1.

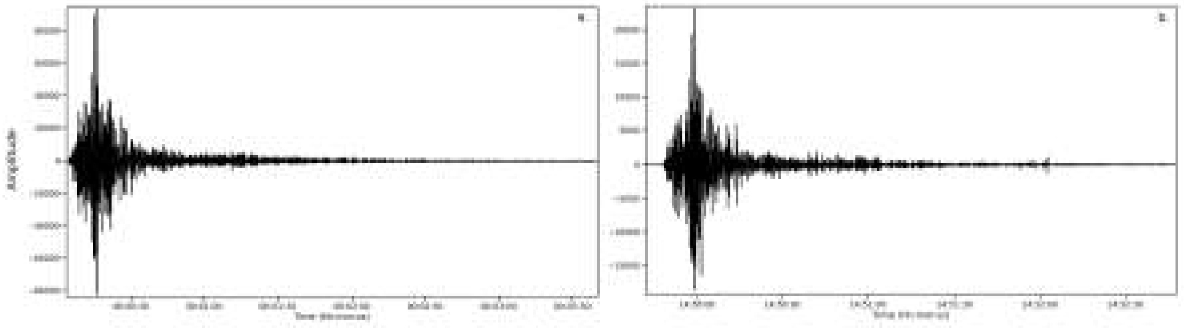


Fig. C.1 Seismic signals recorded at the node 49 for a) earthquake M_L=4.1, 24/12/21 23h59, 20 km SE of Montbéliard (France) and b) earthquake M_L=3.8, 25/12/21 14h49, 18 km SE of Montbéliard (France).

The next step was to visualize the seismic waves arrivals on the watershed scale. The second earthquake (M_L=3.8, 25/12/21 14h49) was chosen. Its epicenter is about 50 km far from the Fourbanne watershed (Fig. C.2). We chose the vertical components of the signals, that were decimated to 50 Hz, in other terms we obtain 50 recordings (frames) per second for the 60 nodes. Two different frames are select and shown in Fig. C.3: the first one shows the first arrivals at the nodes located at the north east of the watershed and the second one shows that the earthquake hit all the nodes, with red and blue dots for positive and negative amplitudes respectively. This is coherent with the position of the epicenter at the north east of the watershed, thus the first arrivals of the earthquake signals at the nodes that are the nearest to the epicenter. The latter shows the importance of the Juraquake network, putting emphasis on the recurrence of earthquakes in the region.

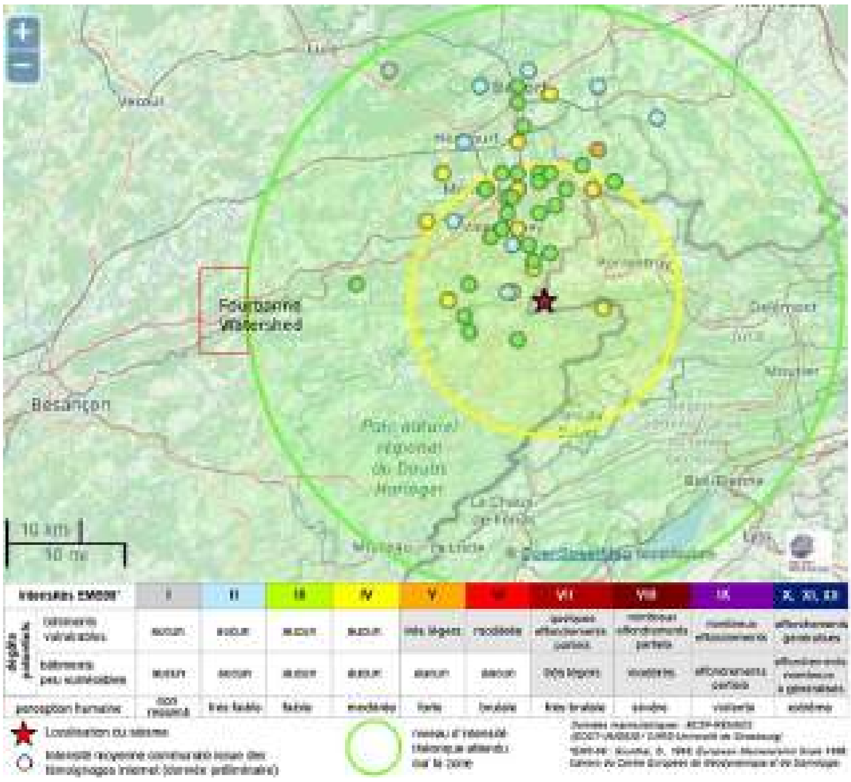


Fig. C.2 Epicenter of the M_L 3.8 earthquake (24/12/21 23h59, 20 km SE of Montbeliard (France). The colored dots show the felt intensity at different locations using the EMS98 scale. The red rectangle shows the location of the Fourbanne Watershed. (Source: BCSF RENASS).

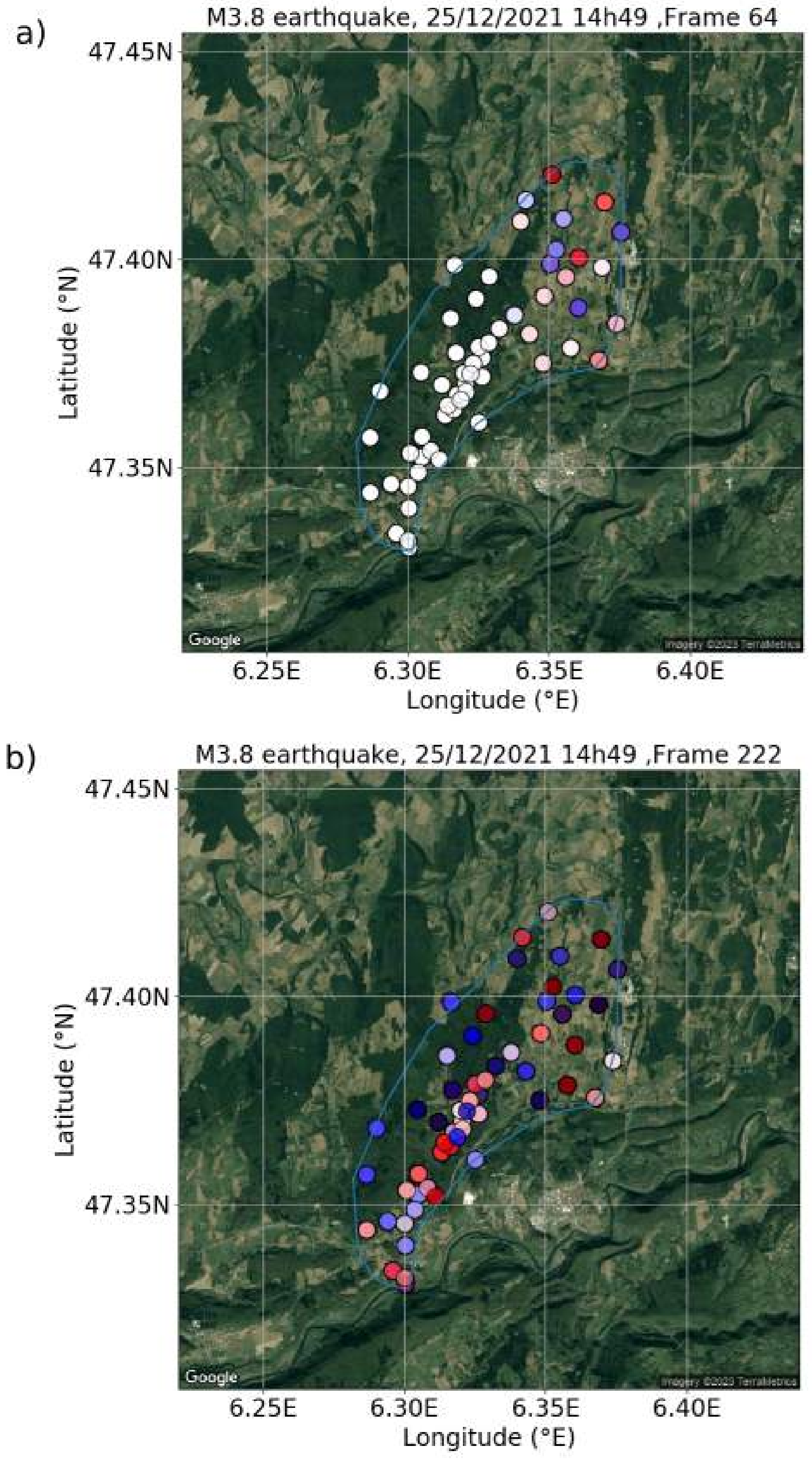


Fig. C.3 Arrivals of the earthquake (M_L 3.8, 24/12/21 23h59) signals at two different time lapses, red and blue dots correspond to positive and negative amplitudes respectively, normalized relatively to the maximum and minimum amplitudes.

Appendix D

Résumé Long

D.1 Introduction

Selon Ford and Williams (2007), environ 25% de la population mondiale dépend des eaux souterraines extraites des aquifères karstiques. En France, ils occupent 35% de la surface du territoire, et contribuent à 40% de l'approvisionnement total en eau potable (Chen et al., 2017). Plus précisément, dans la région de Franche-Comté, 80% des prises d'eau pour l'alimentation en eau potable sont prélevées dans des aquifères karstiques (Bakalowicz, 1999).

Le karst est une formation géologique que l'on trouve dans les roches carbonatées, telles que le calcaire et la dolomie. L'eau s'infiltre, enrichie en dioxyde de carbone (CO2), et s'écoule à travers les fractures et les zones de faiblesse de la roche, dissolvant progressivement les roches en formant des voies d'écoulement préférentielles et des réseaux de drainage souterrains (see Fig. D.1) (Dörfliger et al., 2010; Ford and Williams, 2007). Cette transformation se produit sur une courte durée à l'échelle géologique : de quelques milliers à quelques dizaines de milliers d'années. En créant des vides de tailles différentes au sein du massif, les systèmes karstiques deviennent des milieux très hétérogènes, plus ou moins perméables.

Une grande variabilité spatiale peut être rencontrée dans la subsurface des aquifères karstiques, en raison des multiples porosités à l'échelle et des différentes perméabilités du milieu. Cette dernière entraîne une variabilité spatio-temporelle des processus hydrologiques en cours. La surveillance hydrométrique et d'autres méthodes conventionnelles fournissent des observations ponctuelles et des indications globales sur la réponse hydrologique de l'aquifère, sans expliquer sa structure et la distribution de ses hétérogénéités dans l'espace. Les méthodes hydrogéophysiques viennent en complément de la surveillance hydrométrique et fournissent des moyens non invasifs d'obtenir des informations sur la structure souterraine des environnements karstiques. Ces méthodes peuvent également aider à identifier les voies de recharge potentielles et à mieux comprendre la dynamique des zones saturées et non saturées et l'interaction entre les deux.

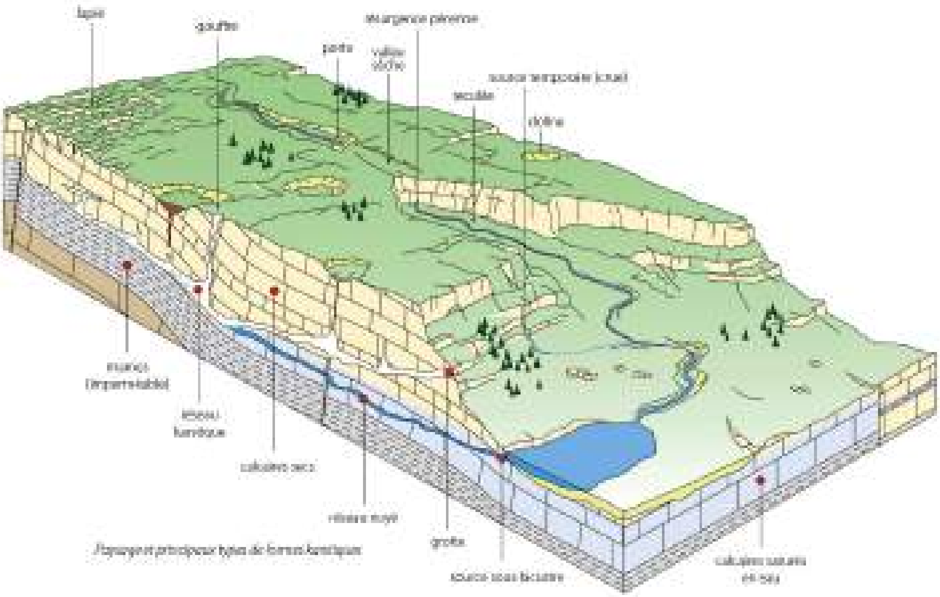


Fig. D.1 Shéma conceptuel d'un aquifère karstique dans le Massif du Jura (Bichet and Campy, 2008).

D.2 Le bassin versant de Fourbanne

Le site d'étude est l'aquifère karstique de Fourbanne dans le Massif du Jura, est de la France. Il fait partie de l'observatoire hydrogéologique JURASSIC KARST installé en 2014 (Cholet et al., 2017) et du réseau français SNO KARST (Jourde et al., 2018). La lithologie locale est caractérisée par des calcaires et des marnes du Jurassique moyen recoupés par une série de failles normales N-S et NE-SW, qui contrôlent l'orientation des conduits souterrains (see Fig. D.2). L'aquifère est principalement alimenté par une recharge allogène à travers des pertes (Cholet et al., 2015). Le conduit souterrain a été exploré et cartographié sur une longueur de 9 km par des spéléologues dans la zone non saturée et une longueur de 2 km par des plongeurs dans la zone saturée. L'aquifère est instrumenté par des sondes hydrologiques, deux sismomètres large bande et d'une fibre optique (pour les mesures de température), pour un suivi continu du réservoir souterrain. Dans le cadre de ce projet de thèse (SISMEAUCLIM), nous avons déployé 60 sismomètres (de fin novembre 2021 à fin mars 2022) et 20 pluviomètres (de juin 2021 à juin 2022) pour analyser la reponse du bassin versant en termes d'infiltration et de transfert d'eau souterraine.

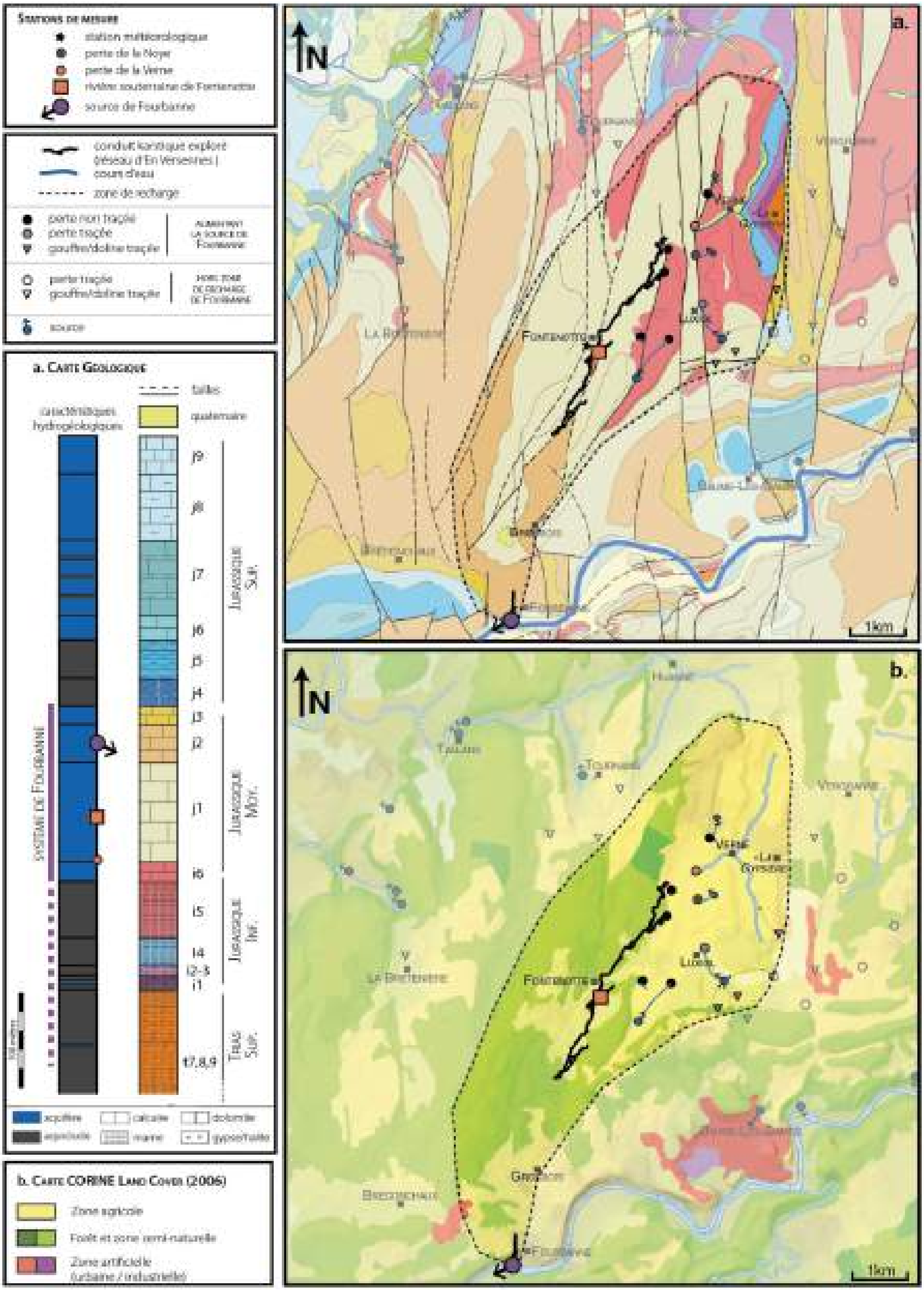


Fig. D.2 Description de la zone de recharge du système karstique de Fourbanne avec (a) la carte géologique et (b) la carte de l’occupation des sols (Cholet et al., 2017).

D.3 Hydrodynamique de la rivière souterraine

D.3.1 Prédiction de la hauteur d'eau de la rivière souterraine en utilisant une méthode d'intelligence artificielle et des enregistrements continus de bruit sismique

Introduction

Nous présentons une méthode basée sur un algorithme d'apprentissage artificiel, l'algorithme Random Forests (RF ou forêts aléatoires), et l'enregistrement continu du bruit sismique, pour prédire la hauteur de l'eau dans la rivière souterraine. Le RF est une méthode d'apprentissage ensembliste qu'on utilise dans le bute de la régression, basée sur le calcul d'un grand nombre d'arbres de décision. L'objectif de cette étude est de proposer une approche innovante permettant de fournir une inférence à distance de la hauteur d'eau de la rivière souterraine. En d'autres termes, nous proposons d'établir une projection des données sismiques dans un espace de caractéristiques multidimensionnelles extraites à l'aide d'une fenêtre coulissante de 15 minutes vers la sortie de valeurs de hauteur d'eau en 1-D à l'aide de l'algorithme mentionné.

Matériels et méthodes

On a utilisé les données des deux sismomètres permanents AVEN et FONT, situés dzans la galerie souterraine à 20 m de profondeur et à la surface respetivement, et de la sonde hydrologique CTD (voir Fig. D.3). Les signaux sismiques brutes subissent un pre-traitement (decimation, filtrage, elimination des tendances linéaires) et puis sont decoupés en fenëtres de 15 min, avec un recouvrement de 50%. Après cela, les features sont calculées pour chaque fenêtre (voir Hibert et al., 2017, pour une description détaillée de chaque caractéristique). Au total, 72 features sont calculés, liés à la forme d'onde, le contenu fréquentiel, l'énergie spectrale et le pseudo-spectrogramme des signaux. Des caractéristiques similaires à celles de Hibert et al. (2017) sont utilisées dans notre étude, à l'exception des attributs de polarité, avec des bandes de fréquences supplémentaires, choisies pour couvrir l'ensemble de la gamme de fréquences étudiée (1-50 Hz). Ces features, sont couramment utilisées pour identifier des événements ou des sources dans les signaux sismiques, car elles sont capables de couvrir plusieurs aspects des signaux. L'apprentissage se fait en attribuant, pour chaque fenêtre temporelle, la hauteur d'eau aux features correspondant au signal sismique enregistré au cours de la même fenêtre temporelle. Un algorithme indépendant est réalisé pour chacun des sismomètres. Chaque algorithme a était entraîné sur une année de données hydrologiques et sismiques et puis testé sur des données d'une année pour AVEN et 3 mois pour FONT.

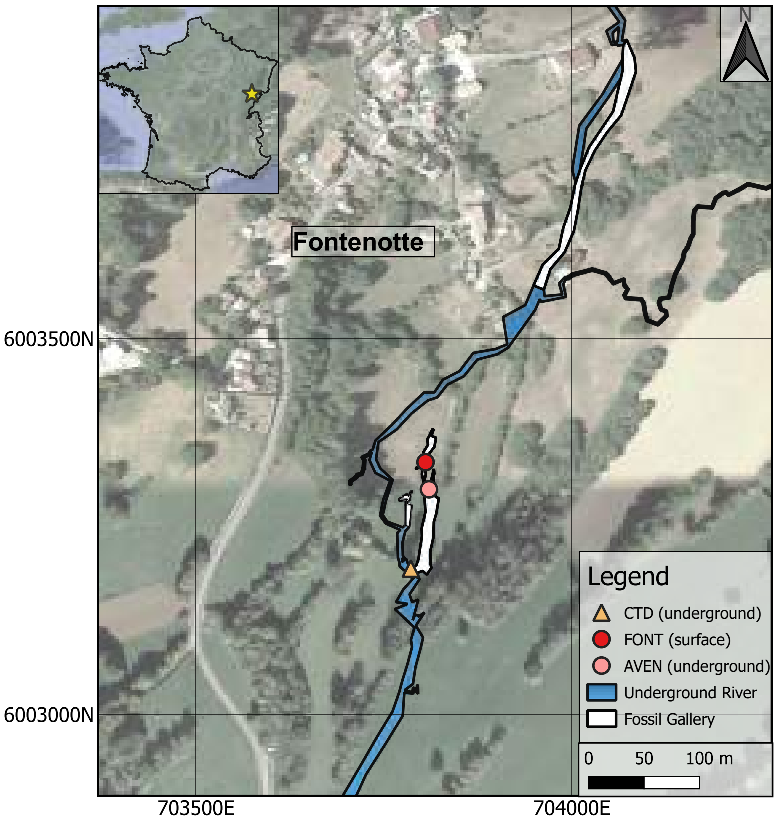


Fig. D.3 Carte de localisation du site et de l'instrumentation. La rivière souterraine et les galeries fossiles sont dessinées respectivement en bleu et en blanc (les données topographiques proviennent de D. Motte, ASDC). Les points rouges et roses indiquent les positions des stations sismiques et le triangle la position de la sonde hydrogéologique (CTD).

Résultats et conclusion

Les résultats de la prédiction se trouvent dans la Figure D.4. On a utilisé le critère de Nash-Sutcliffe pour calculer l'efficacité des méthodes, ce qui donne une valeur de 95% pour AVEN et de 53% pour FONT. Nous avons également calculé l'importance de chaque feature dans la prédiction, les plus importantes étant les features associées à l'énergie du signal pour les deux sismomètres, en raison de l'interaction rivière-lit et des mécanismes de transport de charges. La RF est une méthode robuste pour l'estimation de la hauteur d'eau des rivières en utilisant le bruit sismique, même dans le cas d'une rivière souterraine qui génère de faibles signaux enregistrés à la surface. La prédiction avec FONT, révèle les tendances de la hauteur d'eau ainsi que les périodes de crue, même si le signal sismique induit par la rivière est faible à la surface. Cette application pourrait servir de point de départ à l'étude d'autres conduits d'eau inaccessibles en vue d'une meilleure estimation des eaux souterraines et d'une meilleure prévision des inondations.

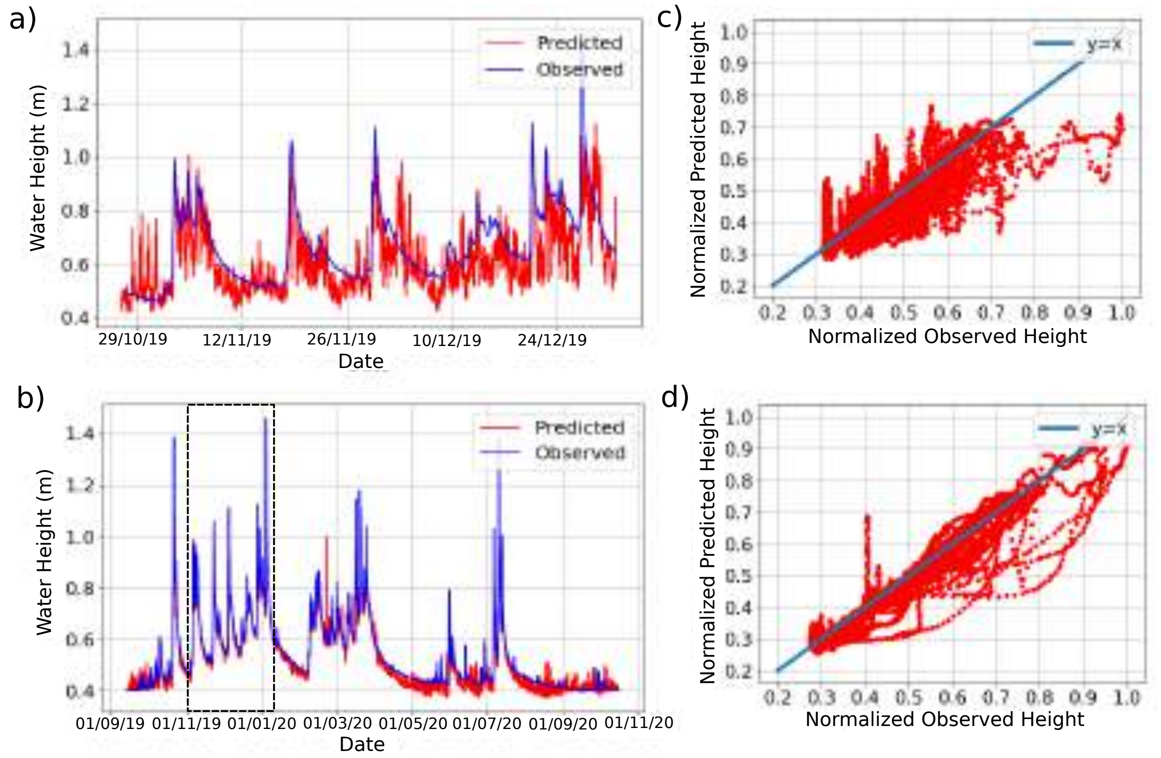


Fig. D.4 (a) Hauteur d'eau simulée avec le sismomètre FONT ; la ligne bleue est la hauteur d'eau mesurée avec la CTD ; la ligne rouge est la hauteur d'eau prédite obtenue par l'application de l'algorithme sur les données sismologiques. (b) Identique à (a) pour le sismomètre souterrain AVEN. Le rectangle en pointillé indique la période d'application en (a). (c) Hauteur d'eau prédite par rapport à la hauteur d'eau mesurée pour les simulations effectuées avec FONT ; la ligne bleue représente la ligne 1:1. (d) Identique à (c) pour AVEN.

D.3.2 Monitoring de la rivière souterraine à travers une fibre optique: Fiber Optic Distributed Temperature Sensing (FO-DTS)

Introduction

Dans cette étude, nous présentons une approche novatrice dans laquelle nous utilisons la méthode FO-DTS (Fiber Optic Distributed Temperature Sensing) pour étudier les processus de recharge d'un aquifère karstique, l'aquifère de Fourbanne. Nous visons à mieux comprendre les échanges/interactions entre ses différents compartiments (conduits, matrice, fractures). Pour ce faire, nous effectuons un suivi des changements de température au sein d'une rivière souterraine et d'un de ses affluents identifié dans le réservoir. L'objectif est (1) d'évaluer l'effet des changements atmosphériques (précipitations, température de l'air) sur la distribution spatio-temporelle de la température à l'intérieur du réservoir (cavité aérienne, rivière, affluent) (2) de caractériser les mécanismes de recharge dans la rivière et l'affluent.

Matériels et méthodes

Un câble à fibre optique de 800 m a été déployé dans la rivière souterraine, mesurant les températures toutes les 5 minutes, avec une résolution spatiale de 50 *cm*, sous le principe de FO-DTS (Fiber optic Distributed Temperature Sensing). Trois sondes hydrologiques (une Aquatroll 500 multiparamètres et deux Troll 100) sont également utilisées comme références de température pour la correction de la température de la fibre optique et éviter les pertes et attenuations tout au long de la fibre. Figure D.5 montre l'instrumentation utilisé sur le tracé du conduit souterrain.

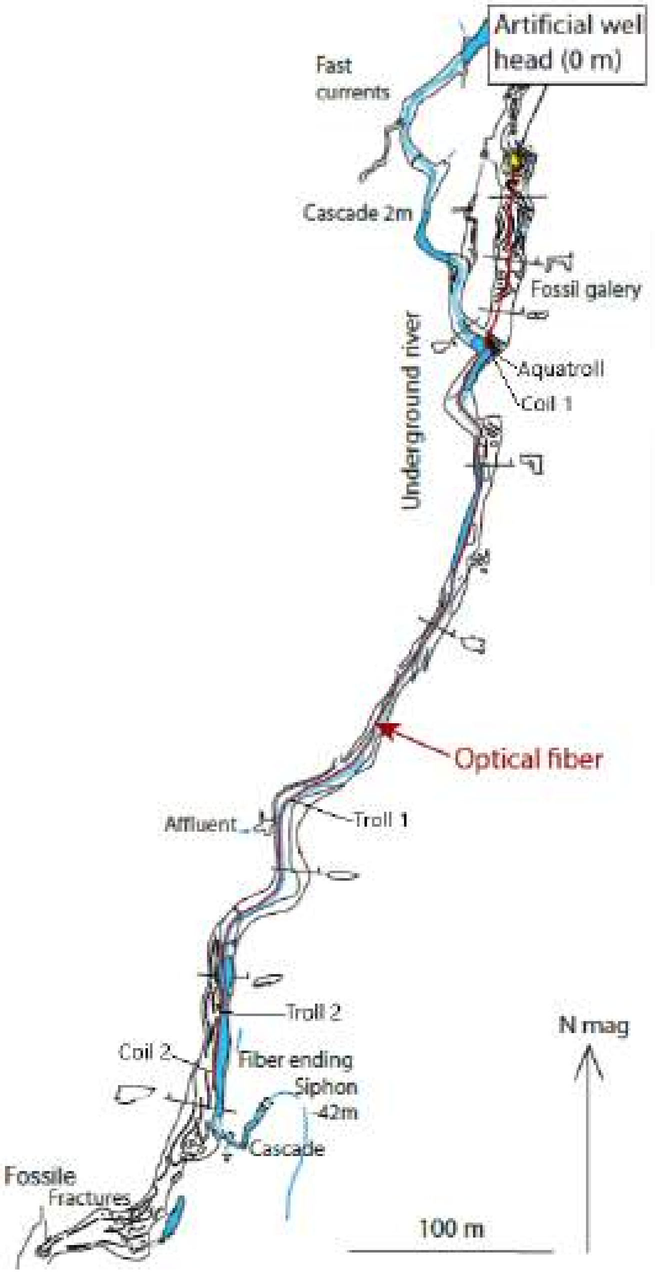


Fig. D.5 Carte topographique du conduit souterrain et de l'instrumentation utilisée dans cette étude. La fibre optique déployée est marquée par une ligne rouge. Carte établie par D. Motte (ASDC)

.

La configuration du FO-DTS utilisée dans cette étude est la configuration à une seule extrémité : une seule extrémité de la fibre est connectée au numériseur. La configuration de l'installation est schématisé dans la Figure. D.6. La fibre est couplée au sol jusqu'à la tête du puits artificiel donnant accès au conduit souterrain. La fibre traverse le puits foré et atteint la rivière à 194 m LAF, près de la position d'une sonde hydrologique (Aquatroll, 195 m LAF). A partir de là, la fibre se trouve dans l'eau, lestée par des rochers ou attachée au sol de la rivière. En deux endroits, la fibre a été fixée au mur, formant des serpentins. Le premier serpentin est construit à 220 m, en fixant des lignes horizontales parallèles de fibre sur le mur permettant de mesurer la température de l'air à l'intérieur de la grotte, mais aussi de mesurer indirectement la hauteur de l'eau à la position du serpentin, en raison de la différence de température entre l'eau et l'air. A 492 m, la fibre atteint la position d'une sonde de température . A ∼500 m le long de la fibre, le câble dévie pour entrer dans un affluent latéral. Ensuite, le câble sort de l'affluent en faisant un demi-tour et retourne dans le conduit principal pour atteindre une deuxième sonde de température à 659 m. La fibre atteint ensuite le deuxième serpentin et se termine dans un enrouleur de câble à l'extérieur de la rivière.

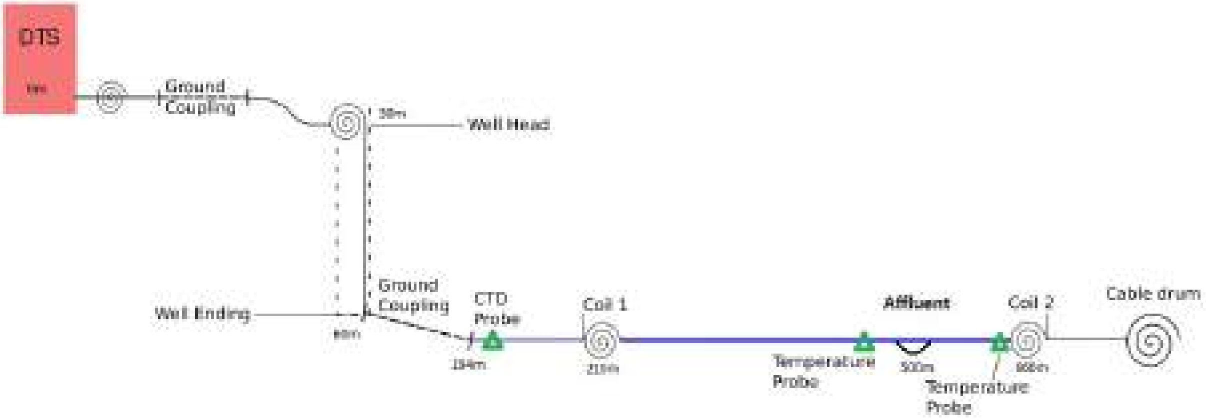


Fig. D.6 Disposition de l'instrumentation utilisée (d'après C. Nguyen).

Le DTS enregistre des données binaires. Les données sont ensuite converties du format binaire au format décimal, aboutissant à des données tabulaires contenant pour chaque pas spatial une température enregistrée, une perte de pas (basée sur l'atténuation), et le rapport entre les intensités Stokes et anti-Stokes. Le signal optique traversant la fibre peut subir une atténuation et une perte de signal en raison de divers facteurs tels que la qualité de la fibre, les courbures et d'autres impacts. La correction de ces pertes peut fournir une représentation plus précise de la distribution réelle de la température le long de la fibre. Plus la position le long de la fibre est éloignée, plus l'atténuation est élevée, d'où une perte dans les mesures. Nous utilisons la méthode proposée par Hausner et al. (2011) pour la correction des données en considérant une configuration à extrémité unique et deux points de référence, qui sont l'Aquatroll et la deyuxièmre Troll. Le schéma d'étalonnage est appliqué à chaque pas de temps.

Résultats et conclusion

Figure D.7 est une représentation spatio-temporelle des températures corrigées moyennées sur 1 h le long de la fibre pendant la période étudiée. La température évolue de manière plus ou moins homogène le long de la fibre, à l'exception de trois endroits correspondant aux 2 serpentins et à l'affluent. L'effet des précipitations sur la température de la rivière peut être clairement identifié tout au long de la fibre, induisant une diminution ou une augmentation de la température de l'eau. Ces changements de température de la rivière sont presque simultanés le long de la section étudiée. Ils dépendent de la température des précipitations : si la pluie est plus froide que la rivière, on observe une baisse de la température de la rivière et vice versa. Les variations se produisent en même temps que l'augmentation du niveau de l'eau, généralement dans les 8 à 10 heures suivant le début des précipitations. Ceci est cohérent avec le mécanisme de recharge rapide, l'eau de pluie s'infiltrant de la surface à travers les failles et/ou les dolines, situées en amont, jusqu'à la rivière et ensuite jusqu'à la zone saturée.

Les températures de la cavité et de la rivière augmentent toutes d'environ 1 °C, cette augmentation peut être liée à l'augmentation de la température atmosphérique durant la période étudiée (passage du printemps en été). En ce qui concerne la température de l'air de la cavité, elle est affectée par la température atmosphérique près du puit et par les variations de la température de l'eau plus profondément dans le conduit.

Concernant l'affluent, des variations de température différentes de celles de la rivière sont enregistrés. La température de l'affluent est plutôt constante pendant la période étudiée avec des variations de faible amplitude, comparée à la température de la rivière. Sa température est plus élevée que celle de la rivière jusqu'en mai et plus basse à partir de la 2^e semaine de juin jusqu'à la fin de la période étudiée. Cette observation suggère que l'eau qui s'infiltre à l'affluent provient d'un réservoir séparé. L'affluent ne semble pas avoir d'impact sur la température de la rivière, probablement parce que le débit de la rivière est beaucoup plus important que celui de l'affluent, de sorte que l'eau de l'affluent est diluée dans la rivière. La rivière et l'affluent présentent une réponse presque concomitante aux événements pluvieux, plus précisément par l'augmentation du niveau d'eau. Cependant, leur comportement n'est pas toujours similaire, car ils peuvent présenter des variations de température opposées. En effet, lors du premier événement de pluie, nous observons une augmentation de la température dans l'affluent alors que la température diminue dans la rivière après cet épisode de pluie froide. Cet apport ponctuel d'eau "plus chaude/plus froide" par l'affluent semble déclenché par l'augmentation du niveau d'eau dans l'ensemble du réseau de fractures/conduits. Cela suggère que l'affluent est alimenté par un réservoir local dont l'eau est expulsée et renouvelée lorsque le niveau d'eau de la rivière de la grotte augmente. Ce point est discuté plus en détail dans la section suivante.

Afin de mieux comprendre l'origine de l'eau s'écoulant de l'affluent, il a été équipé d'une sonde CTD le 14^{th} de mai 2023 pour mesurer la hauteur de l'eau, la température et la conductivité électrique (CE). Nous avons tracé la hauteur d'eau de la rivière souterraine mesurée par l'Aquatroll ainsi que la température et la conductivité électrique de la rivière et de l'affluent (Fig. D.5). La CE dans l'affluent augmente brusquement après la première précipitation, ce qui indique l'expulsion de l'eau à longue durée de séjour vers l'affluent par effet de piston. Ces

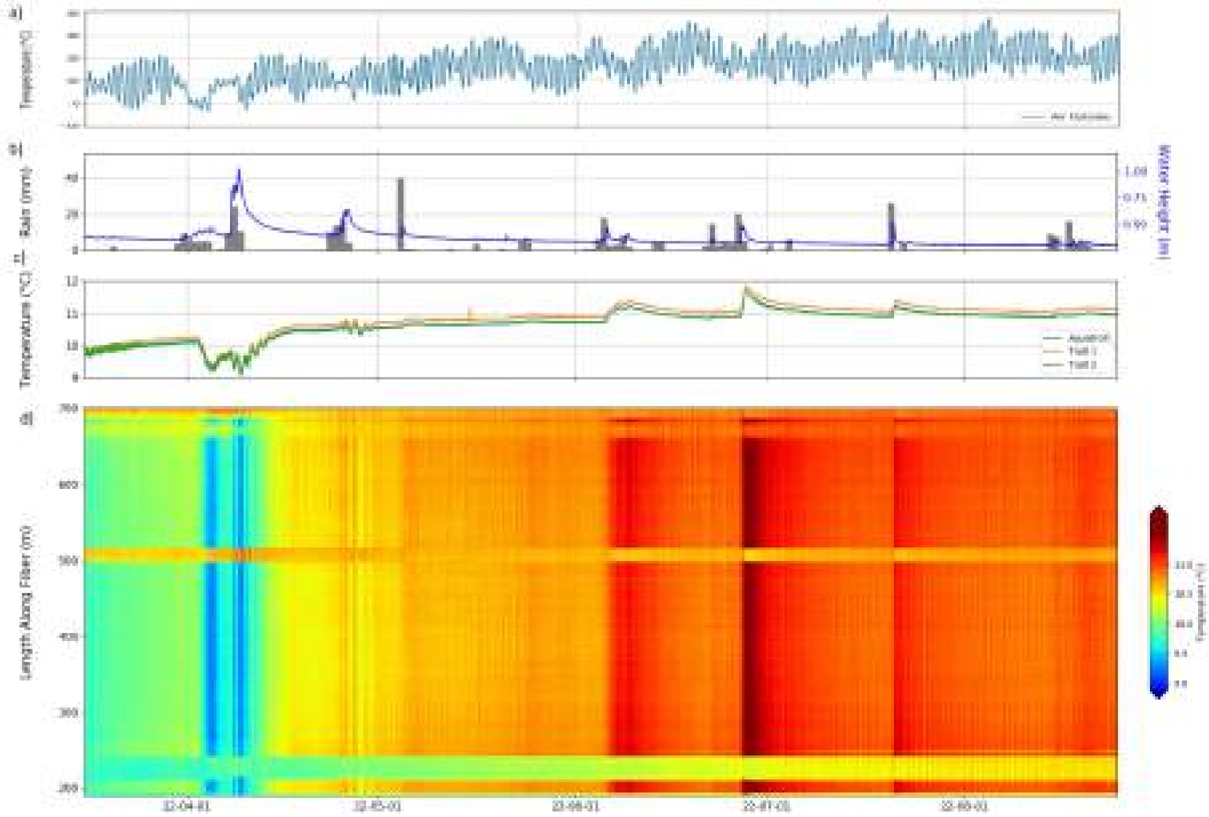


Fig. D.7 Variations de température le long de la fibre, entre mars et août 2022. a) Température atmosphérique b) Précipitations journalières et hauteur d'eau de la rivière souterraine. c) Température horaire enregistrée à l'Aquatroll. d) Température horaire le long de la fibre.

observations viennent étayer l'hypothèse émise plus haut selon laquelle l'affluent est alimenté par un réservoir local, où l'eau est stockée pendant longtemps et donc plus minéralisée que l'eau de rivière ou l'eau de pluie. Ainsi, lors d'un événement pluvieux, l'eau stockée est expulsée par transfert de pression. Hors évenement pluvieux, l'affluent est continuellement alimenté par les recharges épikarstiques. Figure D.9 est un schéma conceptuel qui illustre l'affluent ainsi que les reservoirs qui le rechargent hors et durant les périodes de pluie.

Cette étude montre l'efficacité de la FO-DTS dans la surveillance de l'environnement, ainsi que l'importance des études continues pour caractériser les mécanismes de recharge dans des systèmes aussi complexes.

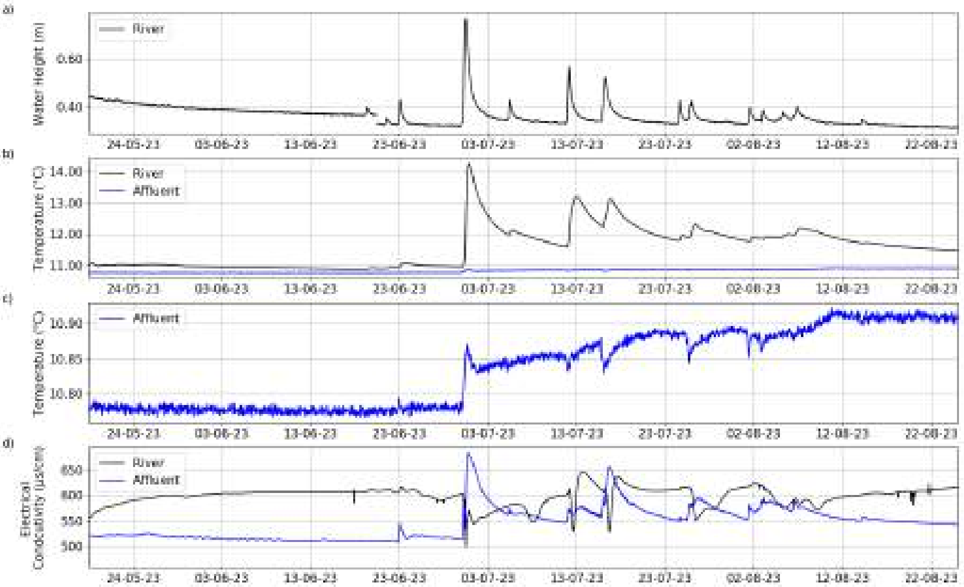


Fig. D.8 Mesures de conductivité et de température dans l'affluent et la rivière. a) Hauteur de l'eau de la rivière souterraine mesurée par l'Aquatroll. b) Température de l'eau de la rivière et de l'affluent. c) Conductivité électrique de l'eau de la rivière et de l'affluent.

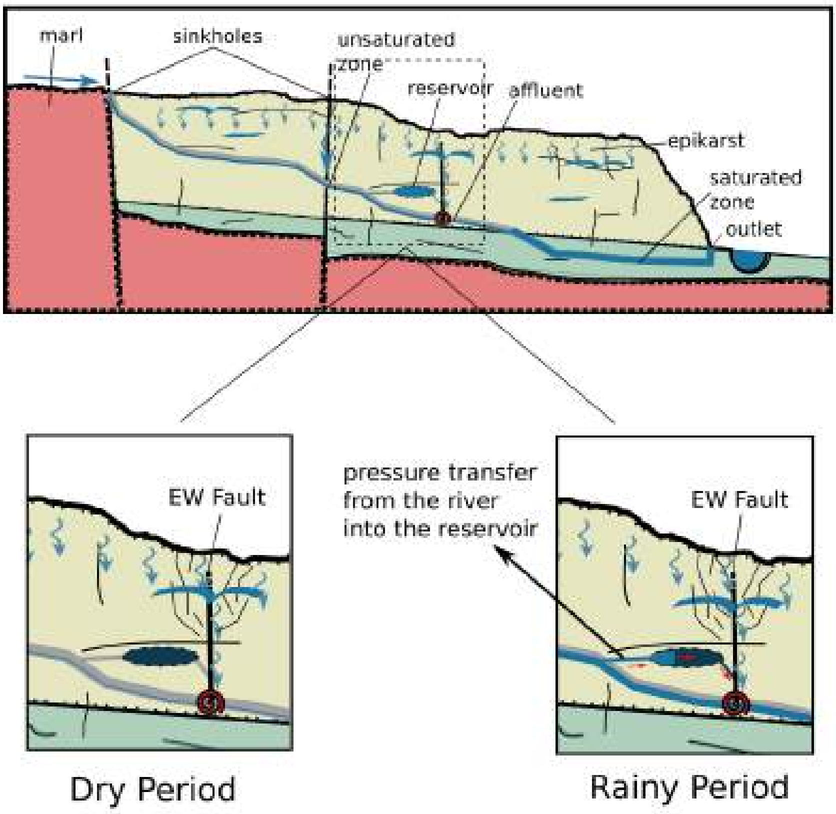


Fig. D.9 Schéma conceptuel illustrant les recharges potentiels de l'affluent.

D.4 Autocorrélation du bruit sismique pour le suivi des eaux souterraines dans les réservoirs karstiques

Introduction

La méthode de cross-corrélation sismique est une technique puissante utilisée en géophysique pour le suivi des changements dans les propriétés physiques du sol (Sens-Schönfelder and Wegler, 2006). Cette méthode permet d'extraire la fonction de Green (ou réponse impulsionnelle) entre deux capteurs (Shapiro and Campillo, 2004) et peut être utilisée pour l'imagerie et le monitoring de l'environnement. Pour améliorer la sensibilité de l'extraction des variations de vitesse à partir des cross-corrélations, il est plus avantageux d'examiner les variations temporelles dans la coda des corrélations plutôt que dans les ondes directes (Poupinet et al., 1984). La coda est utilisée dans ce type d'application en raison de sa stabilité dans le temps par rapport aux ondes directes d'arrivée. Cette méthode peut également être appliquée à une seule station en utilisant l'autocorrélation, qui peut être calculée pour des signaux de même composante ou des cross-composantes. Les méthodes de corrélations sismiques sont de plus en plus utilisées pour surveiller les processus météorologiques et hydrologiques, en se basant sur le fait qu'un milieu saturé d'eau entraîne une diminution de la vitesse sismique. Cette corrélation négative existe entre les précipitations et la dv/v, potentiellement avec un décalage correspondant au temps d'infiltration en fonction de la géologie locale. D'autres facteurs peuvent influencer la corrélation entre la variation de vitesse et la pluie, ainsi que le décalage temporel potentiel entre la pluie et la réponse. Ces facteurs dépendent des propriétés physiques et mécaniques du sol, telles que la minéralogie, la perméabilité et la porosité, qui peuvent être réparties de manière irrégulière dans des sites hétérogènes tels que les aquifères karstiques. Dans cette étude, nous utilisons l'autocorrélation du bruit sismique pour mieux comprendre la réponse du karst aux précipitations et à l'infiltration d'eau qui en résulte à l'échelle d'un bassin versant.

Matériels et méthodes

Un sismomètre situé dans le conduit souterrain a été utilisé pour un suivi pluriannuel (de debut 2019 jusqu'à fin 2022) afin d'étudier les tendances à long terme, et un réseau sismique dense de 60 nodes a été utilisé pour un suivi de 4 mois (de début décembre 2021 jusqu'à fin mars 2022) sur l'ensemble du bassin versant. Pour l'analyse long terme, les dv/v obtenues sont comparés avec les données de pluie enregistrés à la station Météo-France à Branne et la heuteur d'eau de la rivière souterraine. Pour l'analyse avec les nodes, les dv/v obtenues sont comparés avec les données de pluie enregistrés par un réseau de 20 pluviomètres deployés dans le cadre de ce projet. Les positions des instrumentations utilisées sont présentées sur la carte satellite du bassin versant dans la Figure D.10.

La zone étudiée peut être principalement identifiée comme une zone rurale. Il est important de noter que les zones rurales présentent généralement des niveaux de bruit anthropique plus faibles que les zones urbaines ou industrielles, en raison du nombre moins élevé d'habitants et

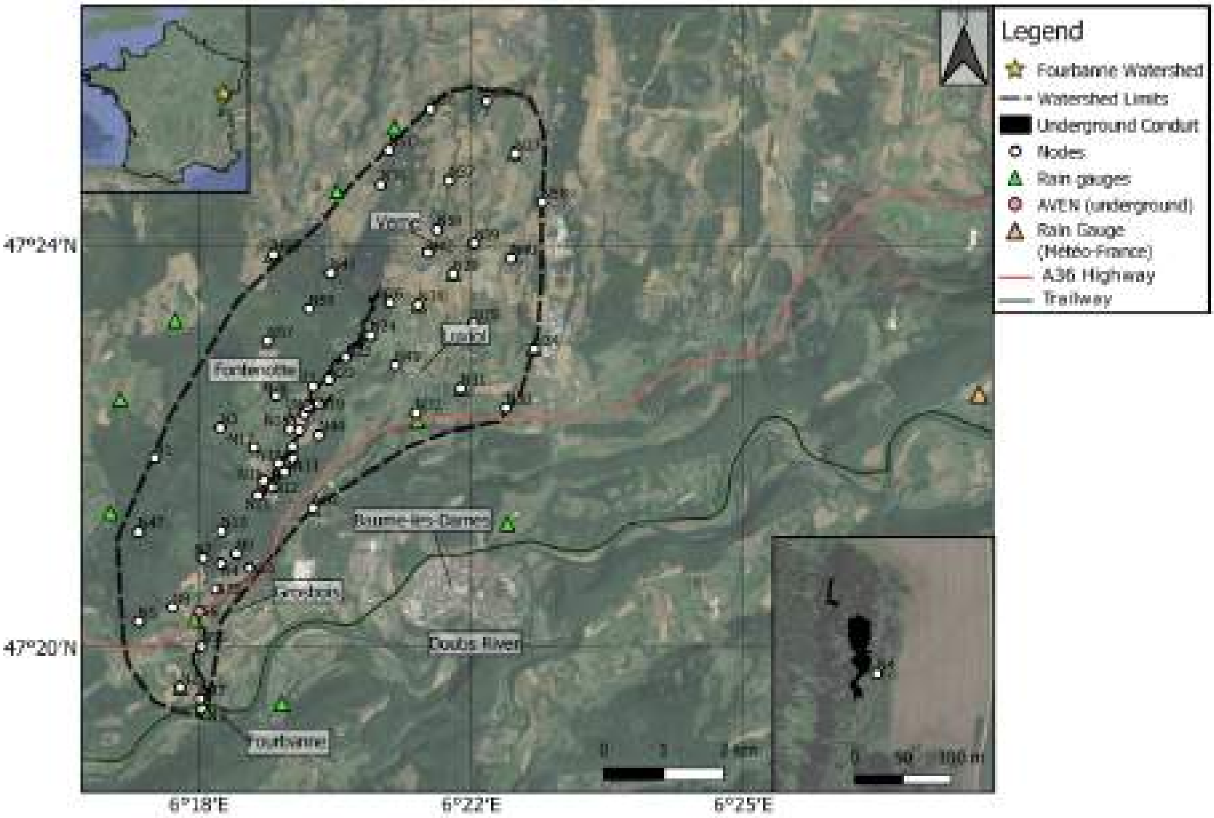


Fig. D.10 Carte satellite Google de la zone étudiée et de l'instrumentation et zoom autour du node N04 et des deux grottes trouvées par les spéléologues à Grosbois. La topographie du conduit souterrain a été fournie par D. Motte (ASDC) et M. Mendonça Dias.

du type d'activités. Cependant, la présence de routes, de voies ferrées ou d'activités industrielles peut encore générer des contributions sonores significatives. Les sources de bruit à l'intérieur et autour du bassin versant sont principalement anthropogéniques. Elles sont générées dans les villages. Les principales activités humaines dans ces villes sont la construction, l'exploitation de carrières, l'agriculture, et l'élevage. D'autres sources de bruit récurrentes sont le passage de véhicules sur des routes départementales, une autoroute, et une voie ferrée (passant par Baume-les-Dames, avec 3 à 4 trains passant par heure). L'écoulement de l'eau dans la rivière de surface (le Doubs) et dans la rivière souterraine qui traverse le conduit karstique génère également un bruit constant, plus accentué pendant les saisons des pluies. En outre, les éoliennes situées dans la partie nord du bassin induisent également des vibrations du sol.

Les changements relatifs de vitesse sismique (dv/v) ont été déterminés à partir de la corrélation du bruit sismique, en utilisant une méthodologie similaire à celle décrite dans les travaux de Lecocq et al. (2014) et De Plaen et al. (2019). Le traitement a été effectué à l'aide de MSNoise (http://www.msnoise.org/), un logiciel basé sur Python utilisé pour l'analyse du bruit sismique. Il comprend plusieurs étapes : (1) le prétraitement, (2) le calcul des fonctions de corrélation (CCF), (4) le calcul de la CCF de référence (moyenne des CCF sur la période étudiée), et (3) l'estimation dv/v. Un processus similaire, mais une analyse séparée, ont été appliqués aux données collectées à partir de la station AVEN et du réseau de 60 nodes. Dans cette étude, nous avons utilisé les fonctions de corrélation des composantes croisées (NSCF) calculées entre les différentes composantes (EN, EZ, NZ) de chaque station, appliqué sur la coda des NSCF. Nous avons choisi d'effectuer une corrélation entre les différentes composantes plutôt qu'une corrélation entre les mêmes composantes, car les NSCF sont plus stables dans le temps et présentent un meilleur rapport signal/bruit De Plaen et al. (2016), par rapport aux CCF des mêmes composantes. Nous avons appliqué la méthode à des NSCFs empilés et filtrés par passe-bande. En ce qui concerne le filtrage, nous avons effectué différents tests et avons finalement concentré notre analyse sur deux gammes de fréquences différentes : 2-4 Hz et 8-10 Hz.

Résultats du monitoring long terme avec AVEN

Les réusltats de dv/v entre 2 et 4 Hz pour AVEN, stackés sur 20 jours, sont montrés dans la Figure D.11, avec la heuteur d'eau de la rivière souterraine et les données de pluie. On constate qu'il pleut assez régulièrement sur le bassin versant au cours des années, et plus intensément en hiver. En fait, il n'y a pas de séparation nette entre saison sèche et saison des pluies, comme dans le sud de la France, par exemple. Cependant, on peut mettre en évidence des périodes relativement "sèches". Pendant ces périodes, nous observons que les dv/v de 2-4 Hz pour toutes les composantes croisées ont tendance à augmenter, alors qu'ils diminuent pendant les périodes plus pluvieuses. En fait, lorsque des précipitations se produisent, l'eau s'infiltre dans le sol, induisant des changements dans ses propriétés hydro-physiques qui peuvent être détectés par des mesures sismiques (Karato and Wong, 1995). L'influence des conditions hydrogéologiques du milieu sur la vitesse sismique relative a fait l'objet de plusieurs études (e.g. Fores et al., 2018; Lecocq et al., 2017; Poli et al., 2020; Sens-Schönfelder and Wegler, 2006; Tsai, 2011). La variation de la vitesse des ondes sismiques peut refléter des changements dans la pression interstitielle, la porosité, la saturation et la densité (Voisin et al., 2017). (Kim and Lekic, 2019) a montré que les variations du niveau de la nappe phréatique induisent des changements corrélés de la vitesse des ondes P (V_P), mais des changements anticorrélés et plus faibles de la vitesse des ondes S (V_S). La relation de Biot-Gassmann peut être utilisée pour estimer la saturation du milieu sur la base des mesures de V_S (Gassmann, 1951) :

$$V_S = \sqrt{\frac{\mu}{\rho}} \tag{D.1}$$

Une augmentation de la saturation entraîne une augmentation de la densité et une diminution de la rigidité, ce qui se traduit par une diminution de V_S. Fores et al. (2018) a observé un délai de 6 mois entre la saison des pluies et la vitesse minimale, révélant un processus d'infiltration lent. Dans notre cas cependant, dv/v diminue alors que les périodes de précipitations ont commencé, ce qui indique un mécanisme d'infiltration plus rapide. Les variations de vitesse sont donc très probablement dues à la saturation en eau de réservoirs/zones intensément fracturés, au sein de l'épikarst.

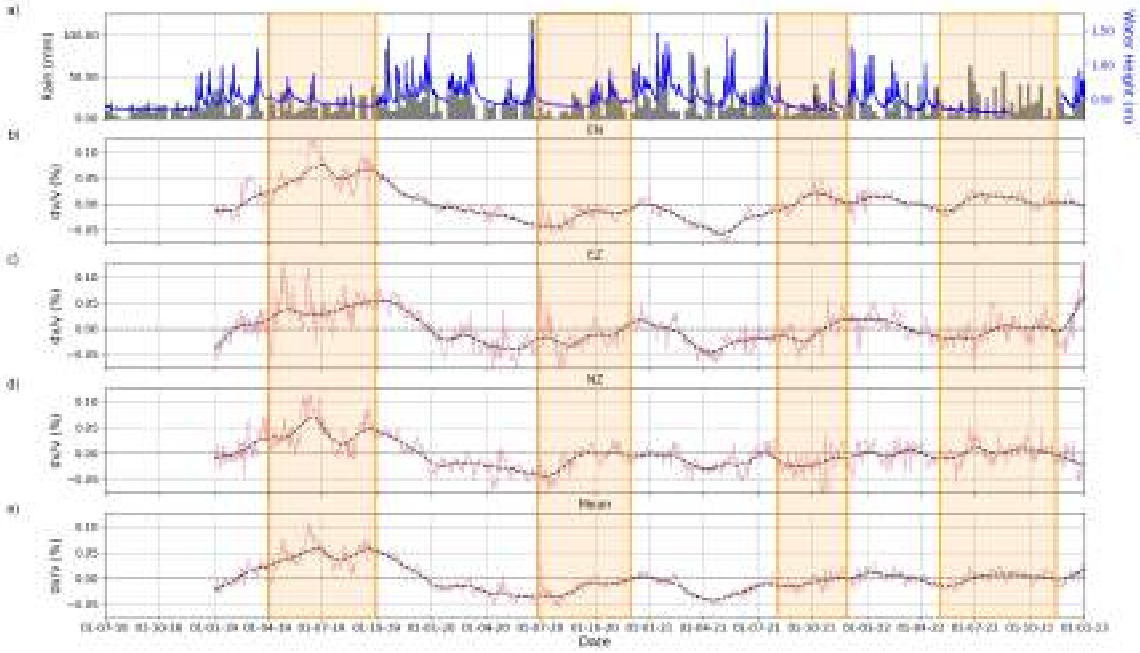


Fig. D.11 Variations de vitesse à 2-4 Hz calculées à la station AVEN entre 2019 et 2022. a) Variation de la hauteur d'eau souterraine mesurée à Fontenotte et précipitations hebdomadaires mesurées à Branne. Dv/v empilés sur 20 jours pour b) EN, c) EZ, d) NZ cross-components et e) la moyenne des 3 cross-components. Les lignes pointillées noires correspondent aux dv/v lissées. Les sections ombrées indiquent les périodes sèches.

Résultats du monitoring 2D avec les 60 nodes

Pour visualiser les changements de vitesse tout en tenant compte de la variation spatiale sur le bassin, nous avons choisi d'étudier les conséquences de la première période de pluie du 30^{th} novembre (début de l'expérience) au 14^{th} décembre 2021 (fin de la crue), résultant en une quantité totale de pluie d'environ 100 mm. La moyenne des dv/v stackés sur 5 jours (filtrés entre 8 et 10 Hz) au cours de la première période de pluie est présentée dans la Fig. D.12. On peut voir que les composantes présentent des réponses dv/v différentes. Cependant, globalement, nous observons principalement des dv/v négatifs dans la partie NE du bassin versant et le long du conduit. Dans la Fig. D.12(e), qui est la moyenne sur les composantes, nous voyons que la plupart des dv/v négatifs sont situés dans la partie NE et au milieu du bassin versant, de N26 au nord à N4 au sud. Il y a également des dv/v moyens positifs dans le NE, en particulier sous les nodes alignés le long d'une faille majeure NS et le long du conduit. L'amplitude de la dv/v moyenne est relativement faible dans la partie sud du bassin versant, reflétant des comportements opposés entre les composantes : principalement négative pour EN et positive pour EZ et NZ.

L'analyse de la variation de la vitesse moyenne pour la moyenne des 3 composantes croisées montre que les zones avec des affleurements de marne (Fig. D.2), ont des moyennes dv/v majoritairement négatives. En effet, les sols sur marnes ont une texture plus fine héritée des marnes et sont plus profonds. Les sols sur calcaire sont moins profonds et plus grossiers. Cela peut être concrétiser en fonction des reserves utiles en eau. Les reserves utiles en eau sont donc plus élevés dans les sols ou il y a un affleurement marneux par rapport aux sols ou il y a un

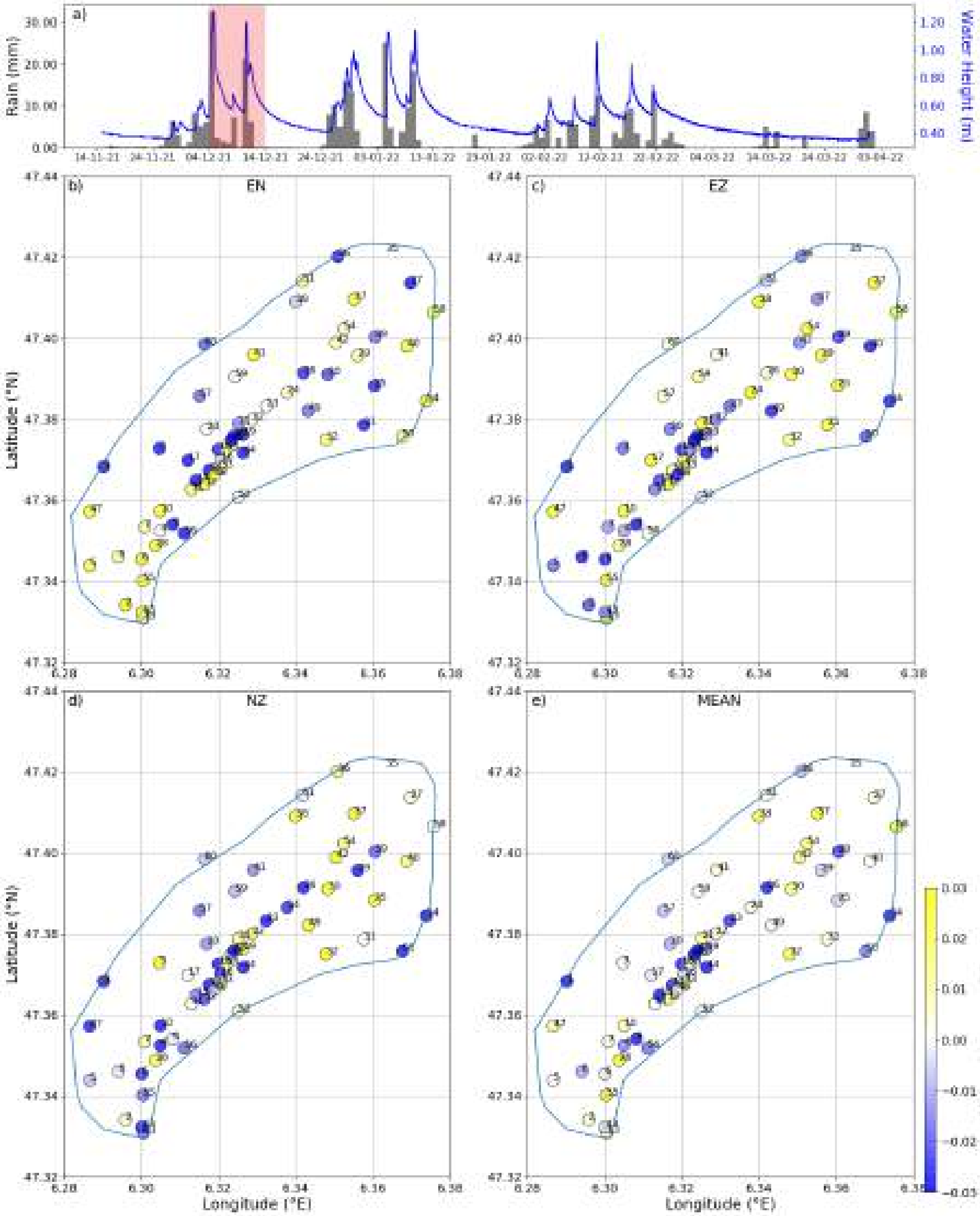


Fig. D.12 Moyenne des dv/v 8-10 Hz stackés sur 5 jours mesurés sous le réseau de nodes pendant la première période de pluie (rectangle ombré). a) Pluies journalières (barres) et hauteur d'eau de la rivière souterraine (ligne bleue) pendant la période étudiée. Distribution spatiale des dv/v moyens pour b) EN, c) EZ, d) NZ et e) la moyenne des 3 composantes.

affleurement calcaire, lié à la plus grande capacité des marnes à retenir l'eau et la capacité des calcaires à la drainer vers l'aval. Nous pouvons remarquer que les zones présentant des vitesses moyennes négatives correspondent principalement à des zones ayant une réserve d'eau élevée.

En fait, la gamme de fréquences 8-10 Hz, les ondes R, en plus d'être sensibles à une profondeur de plusieurs dizaines de mètres, sont sensibles aux premiers mètres dans les deux modèles. Pour les nodes situés sur des affleurements marneux, il est possible qu'ils enregistrent principalement des variations dans les premiers mètres du sol, qui seront plus conséquentes que les variations en profondeur, ce qui explique la distribution dv/v observée.

Variations des vitesses sismiques dans les zones saturée et non saturée

Afin de mieux comprendre la réponse des deux zones différentes aux précipitations, nous avons établi une cross-corrélation entre la série temporelle des précipitations et la série temporelle moyenne de dv/v (stack de 5 jours, 8-10 Hz, moyenne des 3 composantes) pour l'ensemble de la période enregistrée (4 mois). L'objectif principal est de trouver le temps de décalage (en jours) correspondant à l'anticorrélation la plus élevée, puisque, comme nous l'avons déjà mentionné, les variations des précipitations et de la vitesse sont généralement anticorrélées. Nous avons procédé ainsi pour tous les nodes et sélectionné pour chacun d'eux le décalage et la valeur de corrélation correspondante, et nous avons représenté les résultats sur la Fig. D.13. Ceci est similaire à ce qui a été fait dans Andajani et al. (2020) : ils ont sélectionné des décalages correspondant à des coefficients de corrélation qui sont inférieurs à -20%. Dans notre cas, puisque nous utilisons la moyenne des 3 composantes, les coefficients de corrélation donnent un minimum de -15%, cependant les tests sur les composantes individuelles dv/v séries temporelles donnent des coefficients atteignant -30%.

D'après la Fig D.13, nous pouvons voir que nous avons un maximum de corrélations négatives à des décalages temporels courts (moins de 5 jours) sur les parties nord et nord-est et ouest du bassin versant, qui correspondent principalement à des zones non saturées. Des décalages plus importants (supérieurs à une semaine) sont obtenus pour les nodes le long du conduit (surtout vers la section sud) ainsi que pour la partie sud du bassin versant qui correspond à une zone saturée. Ce dernier point est cohérent avec ce que nous avons obtenu en analysant la distribution du dv/v moyen : pendant la période des pluies, nous avons des vitesses négatives dans les parties N, W et NE du bassin versant et pour certains nodes le long du conduit, pendant la période sèche succédant directement à la période des pluies, nous observons des vitesses négatives principalement le long du conduit et dans la partie sud du bassin versant. Cela signifie que des réponses plus rapides sont obtenues dans les parties non saturées de l'aquifère, par rapport à la partie saturée. Cela pourrait refléter des processus de transfert d'eau entre les parties non saturées et saturées du conduit.

Effets de l'anisotropie

L'hétérogénéité des systèmes karstiques rend les régimes d'écoulement de l'eau après les précipitations variables et complexes (Bakalowicz, 2005; Einsiedl, 2005). Ce phénomène est influencé par des voies d'écoulement privilégiées qui peuvent présenter une conductivité hydraulique plus élevée dans certaines directions que dans d'autres en raison des fractures, de la porosité et de l'altération (Best et al., 2007; Knochenmus and Robinson, 1996; Prasad and Nur, 2005; Winterstein, 1990). (Tsai, 2009) a montré que les hétérogénéités dans le milieu peuvent

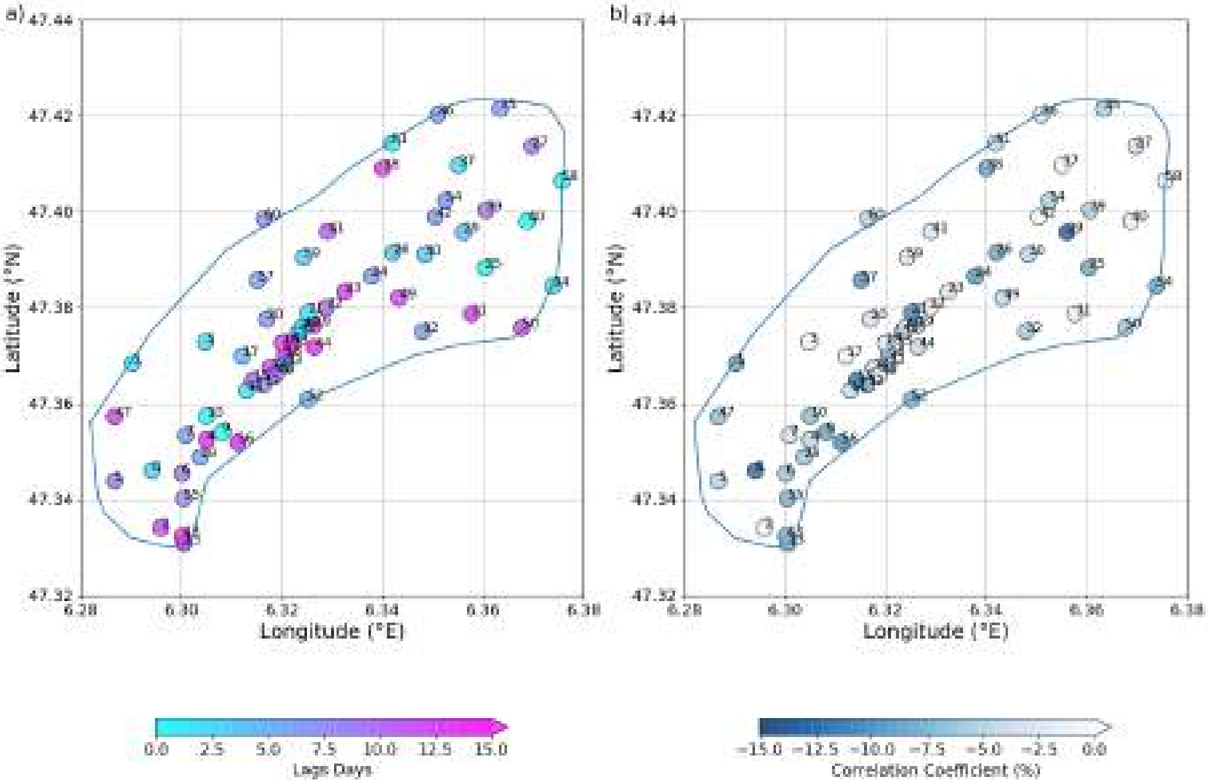


Fig. D.13 Résultats de la corrélation entre la moyenne des 3 composantes de la série temporelle dv/v (stack de 5 jours, 8-10 Hz) et la série temporelle de pluie pour toute la période étudiée, en termes de a) jours de décalage (lags) correspondant au décalage temporel produisant b) le coefficient de corrélation le plus élevé négatif pour tous les nodes.

introduire des biais sur les temps de propagation dans les corrélations de bruit. En outre, elles peuvent indirectement modifier la position de la source, ce qui affecte les résultats. Dans un milieu très hétérogène, c'est principalement ce que nous observons en analysant la variation moyenne de dv/v sur 10 jours : les variations entre les composantes peuvent être similaires mais décalées dans le temps, ou même complètement différentes.

Han et al. (2020) a étudié les effets de la pression et de la saturation en eau sur l'anisotropie sismique des roches partiellement saturées avec des fractures alignées. Ils ont indiqué que les fractures sont généralement alignées dans une certaine direction préférentielle liée aux mouvements tectoniques (par exemple, dans notre cas, les failles sont orientées N-S), ce qui a généré des propriétés anisotropes du massif rocheux. Ils ont découvert que la vitesse des ondes de compression traversant les fractures est plus affectée par la saturation en eau que celles voyageant parallèlement aux fractures. Ces dernières observations pourraient être à l'origine de l'obtention de variations de vitesse différentes pour différents nodes et même de composantes différentes pour un même nœud, d'autant plus que l'anisotropie est un problème courant pour les milieux karstiques. L'anisotropie des propriétés de la masse rocheuse entraînera une anisotropie dans l'écoulement de l'eau et dans les fractures, d'où une anisotropie dans la variation de la vitesse.

Limitations de la méthode

L'interférométrie sismique par crosscorrélation a révolutionné le domaine de la sismologie et élargi ses applications, notamment à des fins de surveillance. Cependant, cette méthode a ses limites : l'une des principales conditions d'applicabilité de la méthode est une distribution azimutale régulière des sources de bruit (Afonin et al., 2019; Wapenaar et al., 2011). Il devient nécessaire d'établir une moyenne sur de longs intervalles de temps en raison de la présence de distributions hétérogènes et anisotropes des sources de bruit sismique ambiant pendant de courts intervalles de temps, comme le décrit (Wapenaar et al., 2010). Plus précisément, dans les régions rurales où les niveaux de bruit anthropique et industriel sont faibles, la mise en œuvre de la méthode nécessite des périodes de collecte de données prolongées en raison de la rareté des sources d'ondes robustes à haute fréquence, qui présentent une distribution spatiale et temporelle non uniforme. Au vu des résultats obtenus dans notre étude, nous pouvons observer que différentes valeurs de variations de vitesse sont obtenues pour chacune des composantes avec parfois même des tendances différentes (la vitesse peut augmenter pour une composante alors qu'elle diminue pour une autre). Cela peut être lié aux sources de bruit présentes dans la zone étudiée : étant donné qu'il s'agit d'une zone rurale, les bruits générés peuvent ne pas être suffisamment forts à haute fréquence pour appliquer la méthode de corrélation et, en outre, ils ne sont pas répartis de manière homogène dans le bassin versant. La principale source de bruit dans notre cas est l'autoroute A36 qui traverse les parties est et sud du bassin versant, les autres sources de bruit sont les routes départementales et les petites routes du bassin versant qui ne sont pas fréquemment traversées.

Une autre limite de la méthode dans notre cas est le fait qu'il n'y a pas eu de période sèche prolongée pendant la période d'acquisition des données qui puisse servir de référence sur laquelle nous pouvons appliquer le MWCS afin de comparer une fonction de corrélation obtenue pendant une période sèche avec une autre obtenue pendant ou après des précipitations. En outre, il y a eu des événements pluvieux deux mois avant le déploiement du réseau sismique, ce qui conduit à une saturation résiduelle potentielle dans la zone vadose due à l'infiltration passée qui peut affecter la réponse sismique aux événements pluvieux ultérieurs.

Synthèse

Dans cette étude, nous avons appliqué la méthode d'autocorrélation du bruit sismique pour obtenir un monitoring localisé à long terme, en utilisant la station sismique AVEN, et un suivi 2D à court terme, en utilisant un réseau de 60 nodes déployés pendant quatre mois. L'objectif était d'étudier la saturation en eau du sol et les transferts d'eau dans le bassin versant de la Fourbanne. La méthode appliquée nous a donné un aperçu de la complexité et de l'hétérogénéité des aquifères karstiques, des sites qui sont difficiles à étudier et à comprendre. Les résultats pour AVEN montrent des variations saisonnières dans l'aquifère, liées au changement de saturation des réservoirs épikarstiques. Les résultats des nodes révèlent des modèles de vitesse différents d'un node à l'autre et même des modèles différents pour les différentes composantes d'un même node en réponse à un événement pluvieux. Un résumé des résultats de la surveillance dv/v à l'aide du réseau sismique est présenté dans la Fig. D.14. Le bassin versant a été divisé en 6 sections sur la base des observations des modèles dv/v pour les nodes. Le schéma a été réalisé pour deux scénarios différents : le premier correspondant à une période avec des taux de précipitations élevés et le second correspondant à une période sèche succédant directement à la période des pluies. Les sections ont été colorées en fonction des valeurs dv/v : blanc pour les valeurs positives, bleu clair et bleu foncé pour les valeurs négatives, impliquant des sections avec infiltration/transfert d'eau, le bleu foncé indiquant des chutes de vitesse plus importantes que le bleu clair. Les couleurs des sections ont été déduites en analysant les valeurs dv/v moyennes des 3 composantes transversales. Enfin, les composants mentionnés en rouge indiquent les composants qui affectent principalement les valeurs moyennes, et sont choisis en observant les dv/v pour les composants individuels. En guise d'observation finale, nous pouvons conclure :

- Pendant les précipitations, la réponse est obtenue dans toutes les sections du bassin versant, avec une réponse plus intense dans le nord-est et la partie supérieure du conduit.
- Après un épisode de précipitations, davantage de transferts d'eau se produisent dans la section sud du conduit ainsi que dans la partie sud du bassin versant.
- changements de saturation se produisent pendant et après les précipitations le long du conduit, ce qui indique une infiltration continue.

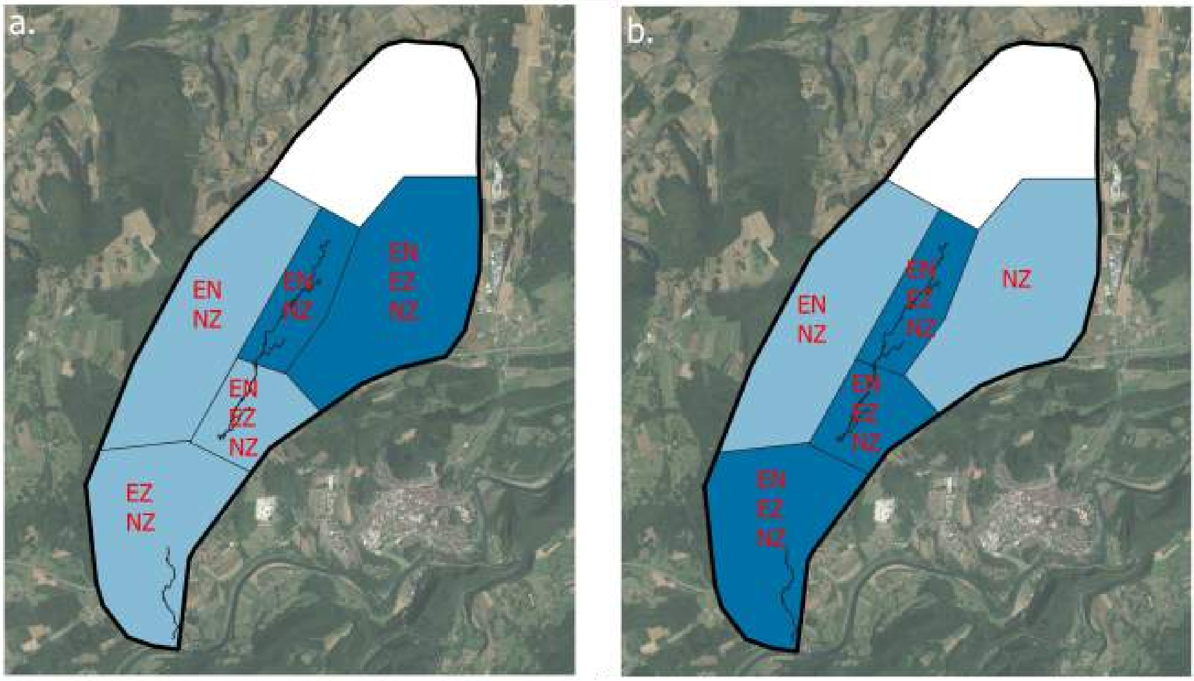


Fig. D.14 Schéma fonctionnel représentant la réponse de l'aquifère a) pendant les précipitations et b) après les précipitations. Le bassin versant a été divisé en 6 sections. La couleur bleu foncé représente les sections avec un dv/v négatif important, le bleu clair pour un dv/v négatif faible et le blanc pour un dv/v négligeable. Les composantes en rouge représentent les composantes montrant le dv/v négatif dans chaque section.

D.5 Conclusion

Dans cette thèse, nous avons travaillé avec des ensembles de données et des techniques de pointe pour étudier un réservoir karstique, en partant de l'échelle du conduit souterrain et en allant jusqu'à l'échelle de tout le bassin versant.La méthode d'apprentissage automatique Random Forest appliquée aux signaux sismiques permet de prédire à distance la hauteur d'eau des rivières souterraines, et donc de surveiller en permanence les crues sans devoir installer des instrumentations invasives. L'utilisation de la fibre optique permet le monitoring continu de la temperature de la rivière souterraine, avec une haute resolution spatio-temporelle. Elle peut être utilisée pour détecter et analyser les affluents afin de comprendre les mécanismes de recharge de la rivière. Cette méthode nous a montré la présence de réservoirs potentiels alimentant des affluents par des failles perpendiculaires à la direction principale du conduit, fournissant des apports latéraux permanents. La présence de fractures remplies d'eau peut également être détectée grâce à la méthode d'autocorrélation de bruit sismique, qui a donné des résultats anisotropes potentiellement liés à différentes directions de fractures dans l'épikarst. Enfin, ce travail montre que la combinaison de plusieurs méthodes hydrogéophysiques complémentaires peut aider à mieux discerner des systèmes aussi hétérogènes et complexes.