UQAC
Université du Québec
à Chicoutimi

REGIONAL GROUNDWATER LEVELS IN CRYSTALLINE AQUIFERS: STRUCTURAL DOMAINS, GROUNDWATER LEVEL MONITORING, AND FACTORS CONTROLLING THE RESPONSE TIME AND VARIABILITY

Par

Attoumane Abi

Thèse présentée à l'Université du Québec à Chicoutimi dans le cadre d'un programme en collaboration avec l'Université du Québec à Montréal en vue de l'obtention du grade de Philosophiæ Doctor (Ph. D.) en sciences de la Terre et de l'atmosphère

Québec, Canada

ABSTRACT

This thesis aims to determine the degree to which fracture networks control the response time and fluctuation of groundwater levels in regional crystalline aquifers in comparison to topography, sediment deposits, precipitation and snowmelt. In this respect, the compartmentalization of the crystalline aquifer into structural domains is necessary, in order to take into account the heterogeneity of the crystalline aquifer in relation to the different fracture networks existing in the rock mass. Field investigations were conducted in the Lanaudiere region, Quebec, Canada, where the underlying crystalline rock outcrops in several locations, allowing access to outcrops for fracture sampling. In addition, four unequipped boreholes drilled into the crystalline rock were available for fracture sampling.

Typically, fracture sampling involves the collection of multiple fracture samples, which involve numerous fracture clusters. Grouping fracture samples into structural domains is generally useful for geologists, hydrogeologists, and geomechanicians as a region of fractured rocks is subdivided into sub-regions with similar behavior in terms of their hydromechanical properties. One of the commonly used methods to group fracture samples into structural domains is Mahtab and Yegulalp's method, considering the orientation of fracture clusters and ignoring several fracture parameters, such as fracture spacing, aperture, and persistence, that are important for fluid circulation in the rock mass. In this thesis, we proposed a new cluster-based similarity method that considers cluster orientation as well as the aperture, persistence and spacing. In addition, a method for compartmentalizing a given study area into structural domains using Voronoi diagrams has also been proposed. The proposed method is more suitable than the previous method for applications in hydrogeology and rock mechanics, especially for regional studies of fluid flow in the rock mass.

The study of response time and variability of groundwater levels requires a groundwater level monitoring network. The inclusion of private boreholes in these monitoring networks can provide a cost-effective means of obtaining a larger data set; however, the use of these boreholes is limited by the fact that frequent pumping, in these boreholes, generates outliers in the recorded time series. In this thesis, a slope criterion is applied to identify and remove outliers from groundwater level time series from exploited private boreholes. Nevertheless, the removal of outliers creates a missing value problem, which biases the subsequent time series analysis. Thus, 14 imputation methods were used to replace missing values. The proposed approach is applied to groundwater level time series from a monitoring network of 20 boreholes in the Lanaudiere region, Quebec, Canada. The slope criterion is shown to be very effective in identifying outliers in exploited private boreholes. Among the characteristics of the missing value pattern, the gap size and gap position in the time series are the most important parameters that affect the performance of the imputation methods. Among the imputation methods tested, linear and Stineman interpolations, and Kalman filtering were the most effective. This thesis demonstrates that privately operated boreholes can be used for groundwater monitoring by removing outliers and imputing missing values.

At local and regional scales, groundwater level is controlled by several factors. The most commonly studied factors are climatic, geologic and geomorphologic controls on groundwater level variability and response time, and in many cases only one controlling factor is considered in the analysis. However, many other factors can affect groundwater level variability and response time, such as the sediment deposit properties and fracture network characteristics in crystalline aquifers. In this study, a more inclusive approach is

used to consider climatic, geomorphological, and fracture network parameters as potential controlling factors. A total of 18 parameters were analyzed for interrelationships as each controlling factor is described by several parameters. The study analyzed a two-year record of groundwater levels in 20 boreholes, drilled into the crystalline rock of the Canadian Shield in the Lanaudière region, Québec, Canada. Factors associated to geomorpgology and fracture network are related to groundwater level variability and its response time. Of the various parameters analyzed in each control factor, sediment thickness and local slope of the geomorphological factor, as well as average persistence and equivalent hydraulic conductivity of the fracture network factor, are most closely related to groundwater level variability and response time. However, further studies are needed to elucidate the physical processes behind certain interrelationships between fracture network parameters and groundwater level variability parameters.

RÉSUMÉ

Cette thèse a pour but de déterminer le degré auquel les réseaux de fractures contrôlent le temps de réponse et la fluctuation du niveau des eaux souterraines dans les aquifères cristallins régionaux par rapport à la topographie, aux dépôts de sédiments, aux précipitations et à la fonte des neiges. À cet égard, la compartimentation de l'aquifère cristallin en domaines structuraux est nécessaire, afin de prendre en compte l'hétérogénéité de l'aquifère cristallin par rapport aux différents réseaux de fractures existants dans le massif rocheux. Des investigations de terrain ont été menées dans la région de Lanaudière, Québec, Canada, où la roche cristalline sous-jacente affleure à plusieurs endroits, permettant un accès aux affleurements pour l'échantillonnage des fractures. De plus, quatre forages non équipés, forés dans la roche cristalline, étaient disponibles pour l'échantillonnage des fractures.

Habituellement, l'échantillonnage de fractures comprend la collecte de plusieurs échantillons de fractures, qui impliquent de nombreux groupes de fractures. Le regroupement des échantillons de fractures en domaines structuraux est généralement utile pour les géologues, les hydrogéologues et les géomécaniciens dans la mesure où une région de roches fracturées est subdivisée en sous-régions ayant un comportement similaire en termes de propriétés hydromécaniques. L'une des méthodes couramment utilisées pour regrouper les échantillons de fractures en domaines structuraux est celle de Mahtab and Yegulalp, considérant l'orientation des clusters de fractures et ignorant plusieurs paramètres de fractures, tels que l'espacement, l'ouverture et la persistance des fractures, qui sont importants pour la circulation des fluides dans le massif rocheux. Dans cette thèse, nous avons proposé une nouvelle méthode de similarité basée sur les clusters qui considère l'orientation des clusters ainsi que l'ouverture, la persistance et l'espacement des clusters. En outre, une méthode pour la compartimentation d'une zone d'étude donnée en domaines structuraux au moyen de diagrammes de Voronoï a également été proposée. La méthode proposée est plus adaptée que la méthode précédente pour des applications en hydrogéologie et en mécanique des roches, notamment pour les études régionales de la circulation des fluides dans la masse rocheuse.

L'étude du temps de réponse et de la variabilité du niveau des eaux souterraines nécessite un réseau de surveillance du niveau des eaux souterraines. L'inclusion de forages privés dans ces réseaux de surveillance peut fournir un moyen peu coûteux d'obtenir un ensemble plus large de données ; cependant, l'utilisation de ces forages est limitée par le fait que le pompage fréquent de ces forages génère des valeurs aberrantes dans les séries temporelles enregistrées. Dans cette thèse, un critère de pente est appliqué pour identifier et éliminer les valeurs aberrantes des séries temporelles du niveau des eaux souterraines provenant de forages privés exploités. Néanmoins, l'élimination des valeurs aberrantes crée un problème de valeurs manquantes, ce qui biaise l'analyse ultérieure des séries temporelles. Ainsi, 14 méthodes d'imputation ont été utilisées pour remplacer les valeurs manquantes. L'approche proposée est appliquée aux séries temporelles du niveau des eaux souterraines provenant d'un réseau de surveillance de 20 forages dans la région de Lanaudière, Québec, Canada. Le critère de pente s'avère très efficace pour identifier les valeurs aberrantes dans les forages privés exploités. Parmi les caractéristiques du modèle de valeurs manquantes, la taille et la position des lacunes dans la série temporelle sont les paramètres les plus importants qui affectent les performances des méthodes d'imputation. Parmi les méthodes d'imputation testées, les interpolations linéaires et de Stineman, ainsi que le filtrage de Kalman ont été les plus efficaces. La présente thèse démontre que les

forages privés exploités peuvent être utilisés pour la surveillance des eaux souterraines en éliminant les valeurs aberrantes et en imputant les valeurs manquantes.

À l'échelle locale et régionale, le niveau des eaux souterraines est contrôlé par plusieurs facteurs. Les facteurs les plus couramment étudiés sont les contrôles climatiques, géologiques et géomorphologiques sur la variabilité du niveau des eaux souterraines et le temps de réponse, et dans de nombreux cas, un seul facteur de contrôle est pris en compte dans l'analyse. Cependant, de nombreux autres facteurs peuvent affecter la variabilité du niveau des eaux souterraines et le temps de réponse, tels que les propriétés des dépôts de sédiments et les caractéristiques du réseau de fractures dans les aquifères cristallins. Dans cette étude, une approche plus globale est utilisée pour considérer les paramètres climatiques, géomorphologiques et du réseau de fractures comme des facteurs de contrôle potentiels. Au total, 18 paramètres ont été analysés pour déterminer les interrelations, sachant que chaque facteur de contrôle est décrit par plusieurs paramètres. L'étude a analysé un jeu de données de deux ans sur les niveaux d'eau souterraine dans 20 forages réalisés dans la roche cristalline du Bouclier canadien dans la région de Lanaudière, au Québec, Canada Les facteurs liés à la géomorphologie et au réseau de fractures sont liés à la variabilité du niveau des eaux souterraines et à son temps de réponse. Parmi les divers paramètres analysés dans chaque facteur de contrôle, l'épaisseur des sédiments et la pente locale du facteur géomorphologique, ainsi que la persistance moyenne et la conductivité hydraulique équivalente du facteur réseau de fractures, sont les plus étroitement liés à la variabilité du niveau des eaux souterraines et à son temps de réponse. Toutefois, des études complémentaires sont nécessaires pour élucider les processus physiques à l'origine de certaines interrelations entre les paramètres du réseau de fractures et les paramètres de variabilité du niveau des eaux souterraines.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ACKNOWLEDGEMENTS

Since the day I have landed in Canada, back to March 2019, I have been in good hands with a research team that is more than remarkable. I would like to express my infinite gratitude to my supervisors for their invaluable guidance and advice during this Ph.D. journey. Pr. Julien Walter, Pr. Romain Chesnaux and Pr. Ali Saeidi you have put in place a friendly working environment in which people are happy to be part of, as you care about your students and you are always available. Thanks for your time and sharing your knowledge. At the end I am happy to say that not only you have been my mentors but also my friends. Julien Walter, thanks for all those times passed in the field collecting data and telling stories about the geology of Canada. I could feel the passion in you. Romain Chesnaux, thanks for inviting us many times to collect mushrooms in the wild. I have enjoyed it and I hope we will do it again and again, as far as I am still in Canada. It is also quite strange that a Comorian learned how to fish in Canada, not in Comoros. I would like to thank all the members of the research group R2Eau, especially Professors Alain Rouleau and Pierre Cousineau, Dr. Anouck Ferroud, Dr. Alireza Shahbazi, Laura-Pier Perron Desmeules, Othniel Glodie Dominique Ngombe, Chaima Miled. I would also like to thank my colleagues, closest friends and housemates in Canada, Daouda Meité and Adoubi Vincent De Paul Adombi (a.k.a. l'homme de prépa, le bosseur). I would also like to thank the research team of the PACES-Lanaudiere for their assistance during fieldwork, friendship and support, especially Dr. Anouck Ferroud, Melanie Lambert and David Noel.

I would also like to thank my family for their support and encouragement, my father Boina Issa Attoumane, my mother Madi Mlatamou Fatima and brothers Attoumane Fahad and Attoumane Tahar. I would like to thank my wife, Issoufi Zidine Tharouoi, for her support, love and sacrifices. You have been wonderful, as always. I dedicate this thesis to you and our son Ayham.

I could not end this section without acknowledging the financial support of the PACES-Lanaudiere project without whom this Ph.D. project won't be possible. Thus, my gratitude goes to the PACES-Lanaudière stakeholders, the Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC) of Quebec, the Université du Québec à Chicoutimi and the five county municipalities of the Lanaudière région (Joliette, Autray, Matawinie, Montcalm and Assomption). I also acknowledge financial contributions from my supervisors through their personal research funds and contributions from the University of Quebec at Chicoutimi through their scholarship program FUQAC.

I would like to thank some friends from around the world who directly or indirectly supported me in this journey. Dr. Dembélé Moktar (Burkina Faso), Paulo Sérgio Lourenço Saveca (Mozambique), Xueyi Zhu (China), Dr. Tibor Stigter (Netherlands), Kamil Abdallah Amarillis (Comoros), Chakira Abacar (Comoros), Florence Gaju Kagabika (Rwanda) and many others.

CHAPTER 1

INTRODUCTION

1.1 Crystalline aquifer

1.1.1 Occurrence and geological settings

A saturated geological formation that can transmit sufficient water in economically exploitable quantities is called an aquifer and the water it contains is called groundwater (Fetter 2000, Şen 2015). Groundwater is an important water resource in regions where the scarcity of surface water resources has led to an early exploitation of those underground water reservoir. In those regions, groundwater resources are considered, in many cases, the only sustainable water resources for water supply and irrigation (Fornés et al. 2005a, Sharp 2014). For instance, more than 50 percent of the world's municipalities use groundwater for water supply, especially in rural areas with scattered population (Fornés et al. 2005a, Raposo et al. 2012a).

Among the different types of aquifers, crystalline aquifers represent a significant part, since shield rocks cover a total land area of about 20% of the earth and are particularly present in Canada, the United States of America and Scandinavian countries, where crystalline rocks form vast areas of Precambrian shields (Gustafson and Krásný 1994, Singhal and Gupta 2010a). Several authors (Krásný 2002, Singhal and Gupta 2010a, Anaba Onana et al. 2017, Maurice et al. 2019) have reported the presence of crystalline rocks in all continents, and their ages vary between 2.4 and 3.6 Ga in Africa (Wright 1992), 2.6 and 3.1 in Canada (Card 1990) and 0.2 Ga to 3 Ga in the USA (DiPietro 2013). Commonly, crystalline rocks include igneous and metamorphic rocks, and are frequently designated as hard rocks (Davis and Turk 1964, Singhal and Gupta 2010a).

The crystalline rock mass is composed of the rock matrix and discontinuities (Jing and Stephansson 2007a). The rock matrix refers to the intact rock and discontinuities is a broad term which stands for faults, joints, dykes, fractures and other geological features that interrupt the continuity or cohesion of the rock mass (Jing and Stephansson 2007a, Singhal and Gupta 2010b, Gudmundsson 2011). The voids contained in the rock matrix are referred to as primary porosity whereas the voids in fractures are known as secondary porosity (Roques et al. 2014). Fracture is defined as any break or crack in the rock mass and depending on the deformation mode, a fracture may be considered as a fault or joint (Berkowitz 2002, Gudmundsson 2011). In fact, joint is generated by a tensile stress which tends to pull apart the rock in opposite direction, while fault is generated by a shear stress which break the rock in a “sliding” movement along a plane (Bonnet et al. 2001a, Gudmundsson 2011). It should be noted, however, that in most papers, researchers use the term fracture as a synonym for joint, while faults are explicitly referred to as faults (Gustafson and Krásný 1994, DesRoches et al. 2014, Roques et al. 2016a). Equally important, depending on the observation scale, even the intact rock may be considered discontinuous as it is composed by more than a single mineral at the micro scale (Hatzor et al. 2017). This study is addressing the mesoscale (a few centimeters) and macroscale (outcrop scale and regional scale).

The potential of crystalline rock aquifers in providing water has been neglected for several years, due to economic and technical reasons (Sharp 2014). It was not until very late in the 20th century that groundwater studies of hard rock aquifers began, in different parts of the world, in response to practical problems such as water supply for human consumption and irrigation, nuclear waste disposal, and seepage in underground infrastructure such as tunnels (Sharp 2014). Singhal and Gupta (2010a) reported that it was the International Symposium on Fractured Rocks, held in 1967 in Dubrovnik-Yugoslavia, which put forward systematic groundwater studies on hard rocks. From the early stage,

some theoretical groundwater related concerns emerged as hydrogeologists realized that the assumptions of homogeneous and isotropic porous media are not systematically applicable to crystalline aquifers (Krásný and Sharp 2007). In fact, hydrogeologists were not considering the role of fracture networks in groundwater flow in crystalline aquifers (Krásný and Sharp 2007). However, this began to change by the end of the 20th century, when the importance of fractured systems and the heterogeneity of crystalline aquifers gained acceptance (Krásný and Sharp 2007). Nowadays, the importance of fracture networks is well recognized in channeling and controlling the groundwater flow in fractured rocks (Lei et al. 2017, Berre et al. 2018, Wenli et al. 2019a).

1.1.2 Fractures and hydraulic properties

The underground rock mass is always fractured to some extent, mainly due to tectonic activities (Hencher 2015, Berre et al. 2018). Moreover, fractures gradually develop throughout the geological history to form fracture networks (Hencher 2015). Rock fractures may start up from the rock formation and evolve throughout the time as the rock mass is subject to stress changes (Hencher 2015). Several authors report that fractures begin to mature at the microscopic level, where, micro-fractures are created, then linked together to develop a macroscopic fracture (Hencher 2015, Fossen 2016).

The permeability is a fluid independent parameter, which expresses the ease that a medium has to let fluid pass through it, while the hydraulic conductivity is a fluid dependent parameter which rates the ability of a formation to transfer water (Singhal and Gupta 2010d, Gudmundsson 2011). Within the crystalline rock mass, the rock matrix permeability is several orders of magnitude lower than the fracture permeability (Gustafson and Krásný 1994, Tsang and Stephansson 1996, Berre et al. 2018, Zhu 2019). As a result, groundwater flow in crystalline rock is commonly assumed to take place only in the fractures (Scesi and

Gattinoni 2010a, Zhang 2013, Zhu 2019). It can be noted that the rock matrix permeability and fracture permeability are commonly referred to as first order permeability and second order permeability, respectively (Scesi and Gattinoni 2010a, Berre et al. 2018). The dependence of the rock mass permeability and hydraulic conductivity on fracture parameters is well documented in the literature, both in hydrogeology (Scesi and Gattinoni 2010a, Sharp 2014, Stoll et al. 2019) and in rock mechanics (Gudmundsson 2011, Hencher 2015). The most obvious parameter is the aperture, which is well illustrated in equation 1.1 that expresses the hydraulic conductivity of a single fracture (K_f) as a function of the aperture, in laminar conditions (Scesi and Gattinoni 2010a, Zhang 2013).

$$K_f = k\frac{g}{\nu} = \frac{g a_h^2}{12\nu} \qquad \text{(Eq. 1.1)}$$

where: k is the rock mass permeability, g is the gravitational acceleration, v is the dynamic viscosity of water, a_h is the hydraulic aperture. In theoretical cases of smooth fracture walls, with no roughness, the hydraulic aperture is equal to the mechanical aperture, which is measured on the rock mass (Scesi and Gattinoni 2010a, Zhang 2013). However, most of the time, the fracture walls are rough and have asperities (Figure 1.1) that reduce the effective aperture available for water circulation in the fracture (Witherspoon et al. 1980, Scesi and Gattinoni 2010a, Zhang 2013). In this regards, the hydraulic aperture (a_h) is linked to the mechanic aperture (a) by a correction factor C as follows $a_h^2 = a^2/C$ (Zhang 2013). Consequently, the hydraulic conductivity also depends on the fracture walls roughness and equation 1 can be rewritten as follows (Zhang 2013):

$$K_f = \frac{g a^2}{12 C \nu} \qquad \text{(Eq. 1.2)}$$

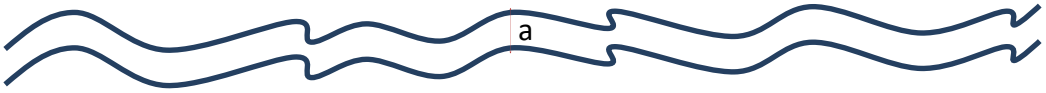


Figure 1.1. Illustration of the mechanical aperture and the asperities in a rough fracture

The hydraulic conductivity of the rock mass is also influenced by the fracture spacing, as it impacts the number of fractures in the rock mass (Scesi and Gattinoni 2010a, Zhang 2013). In fact, the rock mass usually contains several fractures that are mostly organized in fracture sets (Priest 1993, Scesi and Gattinoni 2010a, Schultz 2019). A fracture set is a group of fractures that share similar characteristics, and most of the time, it is the orientation that is used as a grouping factor to create those fracture sets (Priest 1993, Tokhmechi et al. 2011, Liu et al. 2020). Since each fracture in a fracture set i will participate in the hydraulic conductivity of the rock mass, the hydraulic conductivity of a fracture set K_i is given by (Zhang 2013):

$$K_i = \frac{g(a_{ave})^3}{12\nu C_{ave} s_{ave}} \qquad \text{(Eq. 1.3)}$$

where a_{ave} is the average aperture of the fracture set, C_{ave} is the average correction factor of the fracture set, and $s_{ave}$is the average spacing of the fractures within the fracture set. However, the rock mass commonly contains more than just one fracture set in the real world (Schultz 2019). As such, the hydraulic conductivity of the rock mass (K_N), having n fracture sets, in the direction N is given by (Shahbazi et al. 2020):

$$K_N = \sum_{i=1}^{n} K_i \cos^2 \alpha_i \qquad \text{(Eq. 1.4)}$$

where K_i is the hydraulic conductivity of the fracture set *i* as provided in equation 1.3, and α_i is the angle between the fracture set *i* and the direction N, as illustrated in Figure 1.2.

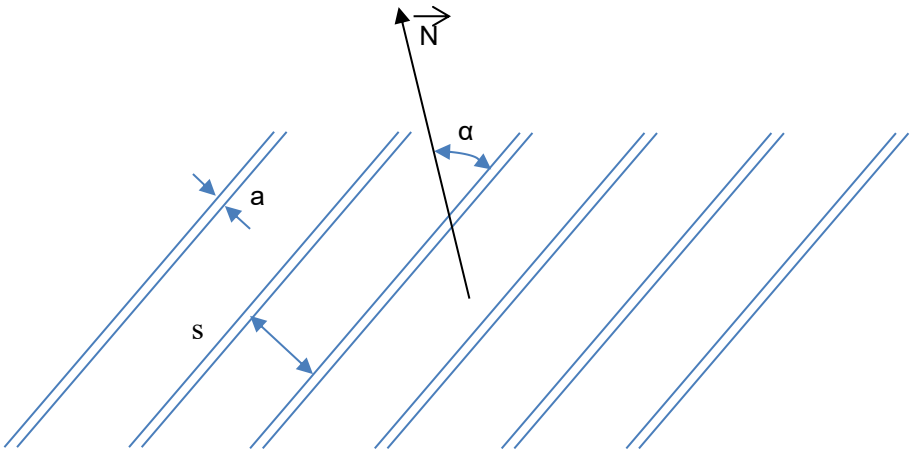


Figure 1.2. Illustration of a fracture set with an average aperture a, an average spacing s, angle α between the fracture set and the direction N toward which the hydraulic conductivity KN is computed.

Tectonic activities and the stress field regulate the degree of fracturing and thus the distribution of hydraulic conductivities in the rock mass (Krásný 2002, Scesi and Gattinoni 2010a). Furthermore, the hydraulic conductivities in the rock mass can vary with scale (Nastev et al. 2004), and such variability is accentuated in areas that have been intensely affected by tectonic activities (Krásný 2002). However, the anisotropy of the hydraulic conductivity in a given rock mass has been shown to be less pronounced with an increasing number of fracture sets with diverse attitudes (spatial orientation) (Gudmundsson 2011).

It is clear from the previous paragraphs that different fracture networks, composed of different fracture sets, will result in different hydraulic conductivities, as described in Equation 1.4. Furthermore, these different fracture networks can occur in the same study area investigated for hydrogeological studies, which then results in a heterogeneous and anisotropic hydraulic property of the rock mass (Kulatilake et al. 1990, Nastev et al. 2004, Scesi and Gattinoni 2010a). This raises questions about how to distinguish between the different fracture networks and how the fracture networks' boundaries are defined? However, in order to address these questions, the concept of structural domains discussed in the next section needs to be introduced.

1.1.3 Structural domains and groundwater storage

As crystalline rocks with dissimilar fracture networks have different hydrogeological properties, it is advisable to distinguish the numerous fracture networks existing in a study area for hydrogeological studies. For that to be done, geologists and hydrogeologist have used for decades the concept of structural domains, which compartmentalizes a given study area into subareas of homogeneous characteristics (Turner et al. 1963, Kulatilake et al. 1990). Traditionally, geologists and hydrogeologists used major structural elements like faults and main rock alterations for the delineation of structural domains (Turner et al. 1963, Priest 1993, Andersson et al. 2013, Follin and Hartley 2014). However, the inconvenience of taking into considerations only the major structural elements in the domain definition is that the information from the fractures is not taken into consideration, while they play an important role in the circulation of fluid in fractured rocks (Kulatilake et al. 1990, Andersson et al. 2013, Follin and Hartley 2014). To overcome this problem, geotechnicians and rock mechanicians have created methods, mainly used in civil engineering projects such as dams, tunnels and mining, that use fracture parameters to define structural domains (Miller 1983, Mahtab and Yegulalp 1984, Zhou et al. 2019, Guo et al. 2020). In fact, fractures not only affect the circulation of fluid in the rock mass, they also impact the mechanical properties of the rock mass (Kulatilake et al. 1990, Zhou et al. 2019, Li et al. 2021b). Despite the many fracture-based structural domain methods that have been developed (Martin and Tannant 2004, Zhang et al. 2016b, Song et al. 2018, Zhou et al. 2018, Li et al. 2021b) assessing the similarity between fracture samples or rock zones, not much has been objectively done to address the compartmentalization of crystalline rocks into structural domains and their delineations. Compartmentalization allows considering the heterogeneity

of fracture networks existing in a crystalline rock, since the compartments (structural domain) have different fracture characteristics of one another.

As presented in the previous section 1.1.1, an aquifer can be considered as a reservoir with a storage capacity, consisting of the voids contained in the geological formation, in which groundwater is stored. These voids are called porosity in hydrogeological studies and in the case of a crystalline aquifer, the porosity of the aquifer can be assimilated to the porosity of the fractures, since it is the porosity that effectively participates in the circulation of water in the rock mass (Singhal and Gupta 2010a). Instead of the storage (S), it is the change in groundwater storage (∆S) that is commonly used in hydrogeology, especially for performing groundwater balance and recharge assessment (Dewandel et al. 2010, Mizan et al. 2019b). The change in groundwater storage is mathematically given by the following expression (Dewandel et al. 2010):

$$\Delta S = S_y \cdot \Delta H \qquad \text{(Eq. 1.5)}$$

where ∆S is the change in groundwater storage over a period of time, Sy is the specific yield, and ∆H is the change in groundwater levels. The specific yield represents the portion of the porosity that can be emptied as the groundwater level decreases in the aquifer (Şen 2015). In unconfined condition, the specific yield is also defined as the volume of water that can be removed by gravity from a saturated rock or soil relative to the volume of that rock or soil (Fetter 2000, Dewandel et al. 2010, Şen 2015). Therefore, the specific yield in crystalline aquifers is an intrinsic property that is linked to the fracture characteristics of the rock mass (Dewandel et al. 2010, Delottier et al. 2018). As such, the specific yield is different from one structural domain to another, since fracture parameters are different between structural domains. It can be deduced from equation 1.5 that the aquifer storage varies over time due to the variation of the groundwater level throughout the time. Equation 1.5 also implies that the specific yield does not change over time in a given location, which is why the specific yield is a constant in the equation of change in groundwater storage.

1.2 Problem statement

As demonstrated in the previous sections, there is no doubt about the role of the fracture network in groundwater circulation and storage in crystalline aquifers. However, the role of fracture networks in the response time and variability of groundwater levels is yet to be studied, especially in comparison to the many other factors influencing the response time and variability of groundwater levels. In fact, other factors such as rainfall pattern and sediment deposits can influence the infiltration of water through the unsaturated zone to the aquifer, which in turn influences the response time and variability of the groundwater level (Gleeson et al. 2009a, Cai and Ofterdinger 2016). The fraction of water from precipitation or surface water bodies that effectively infiltrates and reaches the water table is defined as recharge (Jie et al. 2011, Chesnaux 2013, Huet et al. 2016). This link between recharge and groundwater level rise has resulted in factors influencing recharge also being deemed to influence groundwater level variation. For example, the precipitation pattern was reported to influence both recharge (Dickinson et al. 2004, Cai and Ofterdinger 2016, Kotchoni et al. 2019), on the one hand, and groundwater level (Chen et al. 2004, Miles and Novakowski 2016, Lorenzo-Lacruz et al. 2017, Dong et al. 2022), on the other hand. Similarly, the topography is reported to influence both recharge (Gleeson and Manning 2008, Singhal and Gupta 2010e) and groundwater level (Detty and K. J. McGuire 2010, Rinderer et al. 2016). Numerous other parameters were also reported to influence the recharge and groundwater level such as the land cover (de Vries and Simmers 2002, Singhal and Gupta 2010e), the evapotranspiration (Chen et al. 2002, Okkonen and Kløve 2010), the thickness and hydraulic conductivity of the sediment deposits (Gleeson et al. 2009a, Wright and Novakowski 2019a). It is worth mentioning that in some scenarios, earthquakes (Fetter 2000) and tides (Liu et al. 2018) also affect the groundwater level fluctuation.

In the numerous studies of factors controlling the groundwater level fluctuation and response time in crystalline aquifers, the majority investigate the influence of a single factor, mainly the rainfall (Lee and Lee 2000, Salve et al. 2012, Padilla et al. 2014, Cai and Ofterdinger 2016). However, few studies have combined the analysis of two factors such as topographic and rainfall controls on groundwater level fluctuation and response time (Forster and Smith 1988, Rinderer et al. 2014, 2016). Furthermore, most studies on factors controlling the groundwater level fluctuation and response time in crystalline aquifers ignore the fracture characteristics, while fractures may play an important role in the form of preferential flow that can generate rapid response of the groundwater level. In this regard, there is the need for more inclusive studies on factors controlling the groundwater level fluctuation and response time in crystalline aquifers, especially at the regional scale.

Studies of the factors controlling groundwater levels, groundwater storage and recharge are important for understanding the recharge mechanisms and the fluctuation of groundwater levels over time. Such studies allow for better management of the groundwater resource for its exploitation and protection (Gleeson et al. 2009a). For instance, knowing that the rainfall influences the recharge, Kotchoni et al. (2019) developed a regression model to assess recharge as a function of rainfall. Similarly, Okkonen and Kløve (2010), constructed a regression model that involves rainfall, snowmelt, and evapotranspiration to predict groundwater levels. In terms of groundwater resource protection perspective, Gleeson et al. (2009), reported for example that crystalline aquifers with thin sediment deposits, of less than two meters, are at higher risk of pollution because the groundwater level response time can be extremely rapid (i.e., extremely fast recharge) compared to crystalline aquifers with thick sediment thickness.

1.3 Hypothesis

A hypothesis is defined as an uncertain answer to a research question that can be tested for its veracity or falsity (Pruzan 2016). In this regard and given the different research questions highlighted in the previous sections, several hypotheses are formulated as follows:

- The regional crystalline rock can be compartmentalized into zones of homogeneous fracture characteristics,
- The noises generated by the frequent pumping in the groundwater level time series recorded from exploited private boreholes can be efficiently removed in the recorded data,
- Fracture characteristics influence the response time and groundwater level fluctuation in crystalline aquifers,
- The topography (or landforms) and sediment characteristics control the groundwater level fluctuation and response time in crystalline aquifers.

1.4 Research objectives

The main objective of this thesis is to determine the extent to which fracture networks control the response time and fluctuation of groundwater levels in regional crystalline aquifers in comparison with the topography, sediment deposits, rainfall and snowmelt pattern. In order to address that global research objective, the following specific research objectives have been formulated:

- Compartmentalize crystalline rocks into structural domains of homogeneous fracture networks that consider the important fracture parameters involved in the calculation of the hydraulic conductivity in crystalline rocks;

- Optimize the use of exploited private boreholes in the regional monitoring of the groundwater levels within the monitoring network;
- Determine the influence of the fracture network and fracture parameters on groundwater level fluctuation and response time in crystalline aquifers;
- Determine the influence of the topography (landform), sediment deposits, rainfall and snowmelt pattern on groundwater level fluctuation and response time in regional crystalline aquifers.

1.5 Research methodology

For the research objectives to be investigated, an experimental field site was implemented in the Lanaudiere region, in the province of Quebec, Canada. The chosen study area is bounded by the Sainte-Julienne Fault in the NW, the Maskinongé River in the north, the Saint-Cuthbert Fault in the SE and the Bayonne River in the SW (Figure 1.3). The Lanaudiere experimental site, shown in Figure 1.3, was well suited for the study of the present research objective, since the underlying crystalline rock is outcropping in several locations, allowing access to the outcrops for fracture sampling. Also, many unequipped boreholes, drilled in the crystalline aquifers were available for fracture sampling using acoustic and optical borehole imaging. In order to compartmentalize the study area into structural domains, the similarity between the fracture samples is commonly assessed. Fracture based structural domain methods commonly use stereonets of fracture orientations to assess the statistical similarity between fracture samples (Kulatilake et al. 1990, Martin and Tannant 2004, Phi 2016). In this regards, a stereonets of fracture orientations is generated for each fracture sample, on which the pairwise assessment of statistical similarity between the fracture samples is performed (Miller 1983, Mahtab and Yegulalp 1984, Martin and Tannant 2004). For fracture samples with orientations arranged in clusters on the

stereonet, Mahtab and Yegulalp (1984) proposed a method to evaluate the pairwise similarity of fracture samples. However, the method proposed by Mahtab and Yegulalp (1984) is only based on fracture orientation, whereas the hydraulic conductivity of fractured rocks also depends on other fracture parameters such as spacing and aperture, as presented in Section 1.1.2. In this thesis, a method that considers fracture orientation, spacing, aperture, and persistence is proposed for pairwise fracture sample similarity analysis, as well as an objective method to compartmentalize the study area into structural domains.

Furthermore, the local population depends on groundwater for its water supply, and most households have drilled a borehole to access water. There were 986 enumerated boreholes in the study area of 465 km2. With such a number of boreholes, a monitoring network of 20 boreholes was implemented, composed of 14 private boreholes that were subjected to frequent pumping (pumped) and six normal piezometers that were not exploited (unpumped). This hybrid implementation of a groundwater monitoring network was tested in this thesis because it takes advantage of the many existing boreholes in the study area, avoids the numerous invasive boreholes that would be required to implement 20 monitoring boreholes, and also avoids the inconvenience of filling or sealing these boreholes at the end of the research project in order to prevent any risk of subsequent contamination. While the use of private boreholes optimizes the groundwater monitoring network in reducing implementation cost, it comes with the challenges of the noises generated by the frequent pumping that can be considered as outliers. In this regards, one of the main contributions of this thesis is the development of a methodology to identify and remove outliers in groundwater level time series and apply imputation methods to replace the missing values in the groundwater level time series.

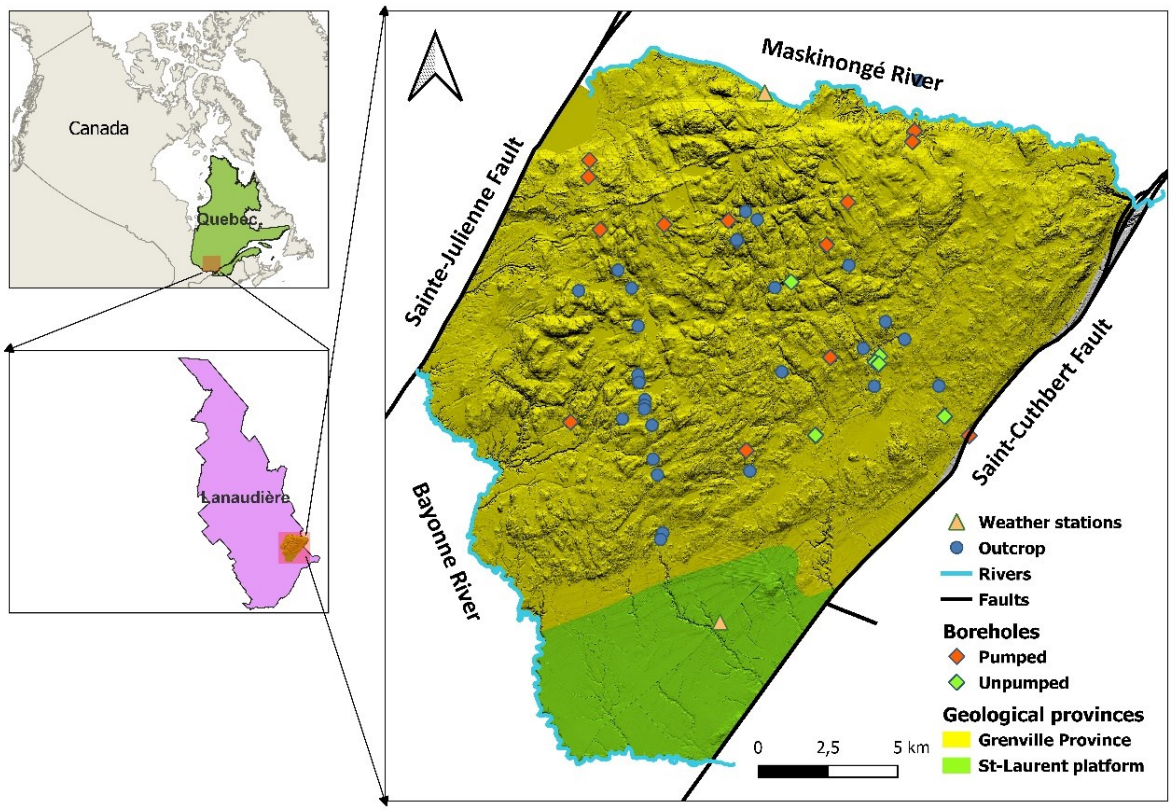


Figure 1.3. Study area in the Lanaudiere region with the investigated outcrops, the borehole monitoring network and the weather stations.

The compartmentalization of the study area into structural domains and the groundwater level time series collected from the previously described monitoring network were necessary to determine the level of influence of topography, sediment deposits, precipitation, and fracture characteristics on the response time and variability of groundwater level in crystalline aquifers. The analysis of the level of influence was performed using the nonparametric Spearman correlation, which allows identifying linear or non-linear association between two variables. The response time was obtained using a cross-correlation analysis between groundwater level and precipitation (rainfall and snowmelt).

1.6 Thesis outline and co-authorship

This thesis is organized in manuscripts as chapters 2 to 4 were written as independent manuscripts that are published or submitted to publication in peer-reviewed journals. The present thesis is written by the first author of these independent manuscripts, Mr. Attoumane Abi, under the supervision and critical review of the co-authors Pr. Julien Walter, Pr. Romain Chesnaux and Pr. Ali Saeidi. Chapter 2 is published in the peer review journal “Quarterly Journal of Engineering Geology and Hydrogeology”, whereas chapter 3 and chapter 4 are under review in the peer review journal “Hydrogeology Journal”, at the time of writing this document. The different chapters are organized as follows:

The present chapter 1, the introduction of this thesis, provides an overview of the state of the art about crystalline aquifers with regards to their occurrence, geological settings, the fractures and their role in the hydraulic properties of crystalline rocks, structural domains and the groundwater storage in crystalline aquifers. The introduction also provides the problem statement along with the hypothesis, research objectives and methodology of this thesis.

The second chapter presents a methodology developed under the scope of this thesis for the compartmentalization of crystalline rocks into structural domains. The proposed methodology is cluster based and considers the fracture orientation, spacing, aperture and persistence. The issue of boundaries between structural domains is also addressed in chapter 2, especially for regional studies. This methodology addresses the first research objective that consists of subdividing the crystalline rock into compartments of statistically homogeneous fracture networks. These fracture networks are then utilized in subsequent study, in chapter four, to determine the influence of the fracture network and fracture parameters on groundwater level fluctuation and response time in crystalline aquifers.

Groundwater level time series are indispensable for addressing the main research objective of this thesis and the third chapter is about data acquisition and treatment of groundwater level time series. More precisely, the third chapter proposed a methodology for implementing a groundwater level monitoring network composed of piezometers and exploited private boreholes in rural or relatively urbanized areas. Implementing such monitoring network is made possible since the proposed methodology processes the noises generated in the exploited private boreholes by removing the outliers and imputing the missing values in groundwater level time series. To the knowledge of the author, it is the first time that such an approach was proposed to implement a groundwater level monitoring network.

The fourth chapter takes advantage of the results obtained in chapters 2 and 3 to determine the extent to which fracture networks control the response time and fluctuation of groundwater levels in regional crystalline aquifers in comparison with the topography, sediment deposits, rainfall and snowmelt pattern (Figure 1.4). In fact, the methodology developed in Chapter 2 was used to characterize the different fracture networks in which each borehole is located in the crystalline rock, whereas the groundwater level time series needed to analyze the aquifer response times and groundwater level variability were obtained from the methodology proposed in Chapter 3. While most of the time, studies of factors controlling aquifer response time and groundwater level variability in crystalline aquifers consider rainfall and snowmelt patterns or sediment deposition, and omit fracture network parameters, a more comprehensive approach was undertaken in Chapter 3, considering the fracture network, topography, sediment deposits, rainfall, and snowmelt patterns as potential controlling factors.

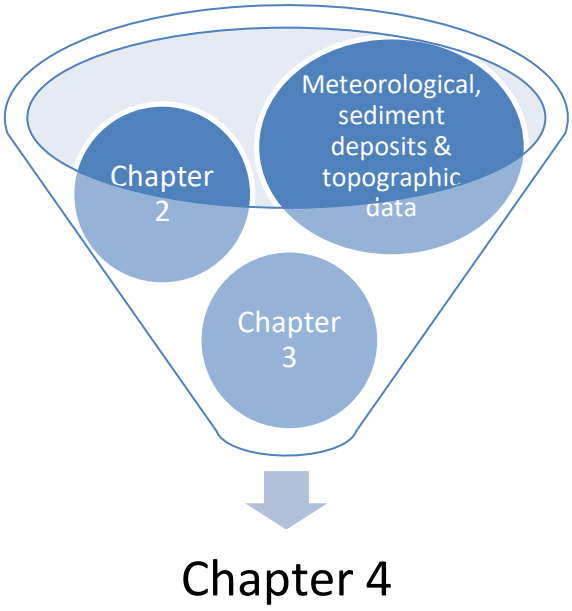


Figure 1.4. Link between the independent chapters of this thesis, highlighting that the data generated in chapter 2 and chapter 3, along with meteorological and topographic data were required in chapter 4.

The fifth and last chapter discusses the global findings of this thesis, provide recommendations, research perspectives, and conclusions.

1.7 References

Anaba Onana, A.B., Ndam Ngoupayou, J.R., and Mvondo Ondoa, J. 2017. Analysis of crystalline bedrock aquifer productivity: Case of central region in Cameroon. Groundwater for Sustainable Development, 5: 66–74. doi:10.1016/j.gsd.2017.05.003.

Andersson, J., Skagius, K., Winberg, A., Lindborg, T., and Ström, A. 2013. Site-descriptive modelling for a final repository for spent nuclear fuel in Sweden. Environmental Earth Sciences, 69: 1045–1060. doi:10.1007/s12665-013-2226-1.

Berkowitz, B. 2002. Characterizing flow and transport in fractured geological media: A review. Advances in Water Resources, 25: 861–884. doi:10.1016/S0309-1708(02)00042-8.

Berre, I., Doster, F., and Keilegavlen, E. 2018. Flow in Fractured Porous Media: A Review of Conceptual Models and Discretization Approaches. Transport in Porous Media,. doi:10.1007/s11242-018-1171-6.

Bonnet, E., Bour, O., Odling, N.E., Davy, P., Main, I., Cowie, P., and Berkowitz, B. 2001. Scaling of fracture systems in geological media. Reviews of Geophysics, 39: 347–383. doi:10.1029/1999RG000074.

Cai, Z., and Ofterdinger, U. 2016. Analysis of groundwater-level response to rainfall and estimation of annual recharge in fractured hard rock aquifers, NW Ireland. Journal of Hydrology, 535: 71–84. doi:10.1016/j.jhydrol.2016.01.066.

Card, K.D. 1990. A review of the Superior Province of the Canadian Shield, a product of Archean accretion. Precambrian Research, 48: 99–156. doi:10.1016/0301-9268(90)90059-Y.

Chen, Z., Grasby, S.E., and Osadetz, K.G. 2002. Predicting average annual groundwater levels from climatic variables: an empirical model. Journal of Hydrology, 260: 102–117. doi:10.1016/S0022-1694(01)00606-0.

Chen, Z., Grasby, S.E., and Osadetz, K.G. 2004. Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada. Journal of Hydrology, 290: 43–62. doi:10.1016/j.jhydrol.2003.11.029.

Chesnaux, R. 2013. Regional recharge assessment in the crystalline bedrock aquifer of the Kenogami Uplands, Canada. Hydrological Sciences Journal, 58: 421–436. doi:10.1080/02626667.2012.754100.

Davis, S.N., and Turk, L.J. 1964. Optimum Depth of Wells in Crystalline Rocks. Groundwater, 2: 6–11. doi:10.1111/j.1745-6584.1964.tb01750.x.

Delottier, H., Pryet, A., Lemieux, J.M., and Dupuy, A. 2018. Estimating groundwater recharge uncertainty from joint application of an aquifer test and the water-Table fluctuation method. Hydrogeology Journal, 26: 2495–2505. doi:10.1007/s10040-018-1790-6.

DesRoches, A., Danielescu, S., and Butler, K. 2014. Structural controls on groundwater flow in a fractured bedrock aquifer underlying an agricultural region of northwestern New Brunswick, Canada. Hydrogeology Journal, 22: 1067–1086. doi:10.1007/s10040-014-1134-0.

Detty, J.M. and K. J. McGuire. 2010. Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment. Hydrological Processes, 24: 2222–2236. doi:10.1002/hyp.7656.

Dewandel, B., Perrin, J., Ahmed, S., Aulong, S., Hrkal, Z., Lachassagne, P., Samad, M., and Massuel, S. 2010. Development of a tool for managing groundwater resources in semi-arid hard rock regions: application to a rural watershed in South India. Hydrological Processes, 24: 2784–2797. doi:10.1002/hyp.7696.

Dickinson, J.E., Hanson, R.T., Ferré, T.P.A., and Leake, S.A. 2004. Inferring time-varying recharge from inverse analysis of long-term water levels. Water Resources Research, 40. doi:10.1029/2003WR002650.

DiPietro, J.A. 2013. Chapter 13 - Crystalline Deformation Belts. In Landscape Evolution in the United States. Edited by J.A. DiPietro. Elsevier, Boston. pp. 197–217. doi:10.1016/B978-0-12-397799-1.00013-0.

Dong, L., Guo, Y., Tang, W., Xu, W., and Fan, Z. 2022. Statistical Evaluation of the Influences of Precipitation and River Level Fluctuations on Groundwater in Yoshino River Basin, Japan. Water, 14: 625. Multidisciplinary Digital Publishing Institute. doi:10.3390/w14040625.

Fetter, C.W., Jr. 2000. Applied Hydrogeology. In 4 edition. Pearson, Upper Saddle River, N.J.

Follin, S., and Hartley, L. 2014. Approaches to confirmatory testing of a groundwater flow model for sparsely fractured crystalline rock, exemplified by data from the proposed high-level nuclear waste repository site at Forsmark, Sweden. Hydrogeology Journal, 22: 333–349. doi:10.1007/s10040-013-1079-8.

Fornés, J.M., Hera, Á. la, and Llamas, M.R. 2005. The silent revolution in groundwater intensive use and its influence in Spain. Water Policy, 7: 253–268. doi:10.2166/wp.2005.0016.

Forster, C., and Smith, L. 1988. Groundwater flow systems in mountainous terrain: 2. Controlling factors. Water Resources Research, 24: 1011–1023. doi:10.1029/WR024i007p01011.

Fossen, H. 2016. Structural Geology. In 2nd edition. Cambridge University Press.

Gleeson, T., and Manning, A.H. 2008. Regional groundwater flow in mountainous terrain: Three-dimensional simulations of topographic and hydrogeologic controls. Water Resources Research, 44. doi:10.1029/2008WR006848.

Gleeson, T., Novakowski, K., and Kurt Kyser, T. 2009. Extremely rapid and localized recharge to a fractured rock aquifer. Journal of Hydrology, 376: 496–509. doi:10.1016/j.jhydrol.2009.07.056.

Gudmundsson, A. 2011. Rock Fractures in Geological Processes. In 1 edition. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511975684.

Guo, L., Wu, L., Zhang, J., Liao, M., and Ji, Y. 2020. Identification of homogeneous region boundaries of fractured rock masses in candidate sites for Chinese HLW repository. Bulletin of Engineering Geology and the Environment, 79: 4221–4243. doi:10.1007/s10064-020-01837-4.

Gustafson, G., and Krásný, J. 1994. Crystalline Rock Aquifers: Their Occurrence, Use And Importance. Applied Hydrogeology, 2: 64–75. doi:10.1007/s100400050051.

Hatzor, Y.H., Ma, G., Shi, G., Ma, G., and Shi, G. 2017. Discontinuous Deformation Analysis in Rock Mechanics Practice. CRC Press. doi:10.4324/9781315687032.

Hencher, S. 2015. Practical Rock Mechanics. CRC Press. doi:10.1201/b18923.

Huet, M., Chesnaux, R., Boucher, M.-A., and Poirier, C. 2016. Comparing various approaches for assessing groundwater recharge at a regional scale in the Canadian Shield. Hydrological Sciences Journal, 61: 2267–2283. doi:10.1080/02626667.2015.1106544.

Jie, Z., van Heyden, J., Bendel, D., and Barthel, R. 2011. Combination of soil-water balance models and water-Table fluctuation methods for evaluation and improvement of

groundwater recharge calculations. Hydrogeology Journal, 19: 1487–1502. doi:10.1007/s10040-011-0772-8.

Jing, L., and Stephansson, O. 2007. 1 - Introduction. In Developments in Geotechnical Engineering. Edited by L. Jing and O. Stephansson. Elsevier. pp. 1–21. doi:10.1016/S0165-1250(07)85001-2.

Kotchoni, D.O.V., Vouillamoz, J.-M., Lawson, F.M.A., Adjomayi, P., Boukari, M., and Taylor, R.G. 2019. Relationships between rainfall and groundwater recharge in seasonally humid Benin: a comparative analysis of long-term hydrographs in sedimentary and crystalline aquifers. Hydrogeology Journal, 27: 447–457. doi:10.1007/s10040-018-1806-2.

Krásný, J. 2002. Quantitative hardrock hydrogeology in a regional scale. Norges geologiske undersøkelse, Bulletin 439: 7–14.

Krásný, J., and Sharp, J.M. 2007. Groundwater in Fractured Rocks. In 1st edition. CRC Press, London.

Kulatilake, P.H.S.W., Wathugala, D.N., Poulton, M., and Stephansson, O. 1990. Analysis of structural homogeneity of rock masses. Engineering Geology, 29: 195–211. doi:10.1016/0013-7952(90)90050-B.

Lee, J.-Y., and Lee, K.-K. 2000. Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system. Journal of Hydrology, 229: 190–201. doi:10.1016/S0022-1694(00)00158-X.

Lei, Q., Latham, J.-P., and Tsang, C.-F. 2017. The use of discrete fracture networks for modelling coupled geomechanical and hydrological behaviour of fractured rocks. Computers and Geotechnics, 85: 151–176. doi:10.1016/j.compgeo.2016.12.024.

Li, L., Li, Y., and Chen, J. 2021. Evaluation of the structural similarity of fractured rock masses based on multiple fracture parameters. Bulletin of Engineering Geology and the Environment, 80: 2189–2198. doi:10.1007/s10064-020-02063-8.

Liu, C.-Y., Chia, Y., Chuang, P.-Y., Chiu, Y.-C., and Tseng, T.-L. 2018. Impacts of hydrogeological characteristics on groundwater-level changes induced by earthquakes. Hydrogeology Journal, 26: 451–465. doi:10.1007/s10040-017-1684-z.

Liu, T., Zheng, J., and Deng, J. 2020. A new iteration clustering method for rock discontinuity sets considering discontinuity trace lengths and orientations. Bulletin of Engineering Geology and the Environment,. doi:10.1007/s10064-020-01921-9.

Lorenzo-Lacruz, J., Garcia, C., and Morán-Tejeda, E. 2017. Groundwater level responses to precipitation variability in Mediterranean insular aquifers. Journal of Hydrology, 552: 516–531. doi:10.1016/j.jhydrol.2017.07.011.

Mahtab, M.A., and Yegulalp, T.M. 1984. Similarity Test For Grouping Orientation Data In Rock Mechanics. American Rock Mechanics Association. Available from https://www.onepetro.org/conference-paper/ARMA-84-0495 [accessed 20 October 2020].

Martin, M.W., and Tannant, D.D. 2004. A technique for identifying structural domain boundaries at the EKATI Diamond Mine. Engineering Geology, 74: 247–264. doi:10.1016/j.enggeo.2004.04.001.

Maurice, L., Taylor, R.G., Tindimugaya, C., MacDonald, A.M., Johnson, P., Kaponda, A., Owor, M., Sanga, H., Bonsor, H.C., Darling, W.G., and Gooddy, D. 2019. Characteristics of high-intensity groundwater abstractions from weathered crystalline bedrock aquifers in East Africa. Hydrogeology Journal, 27: 459–474. doi:10.1007/s10040-018-1836-9.

Miles, O.W., and Novakowski, K.S. 2016. Large water-Table response to rainfall in a shallow bedrock aquifer having minimal overburden cover. Journal of Hydrology, 541: 1316–1328. doi:10.1016/j.jhydrol.2016.08.034.

Miller, S.M. 1983. A statistical method to evaluate homogeneity of structural populations. Journal of the International Association for Mathematical Geology, 15: 317–328. doi:10.1007/BF01036073.

Mizan, S.A., Dewandel, B., Selles, A., Ahmed, S., and Caballero, Y. 2019. A simple groundwater balance tool to evaluate the three-dimensional specific yield and the two-dimensional recharge: application to a deeply weathered crystalline aquifer in southern India. Hydrogeology Journal, 27: 3063–3080. doi:10.1007/s10040-019-02026-8.

Nastev, M., Savard, M.M., Lapcevic, P., Lefebvre, R., and Martel, R. 2004. Hydraulic properties and scale effects investigation in regional rock aquifers, south-western Quebec, Canada. Hydrogeology Journal, 12: 257–269. doi:10.1007/s10040-004-0340-6.

Okkonen, J., and Kløve, B. 2010. A conceptual and statistical approach for the analysis of climate impact on ground water table fluctuation patterns in cold conditions. Journal of Hydrology, 388: 1–12. doi:10.1016/j.jhydrol.2010.02.015.

Padilla, C., Onda, Y., Iida, T., Takahashi, S., and Uchida, T. 2014. Characterization of the groundwater response to rainfall on a hillslope with fractured bedrock by creep deformation and its implication for the generation of deep-seated landslides on Mt. Wanitsuka, Kyushu Island. Geomorphology, 204: 444–458. doi:10.1016/j.geomorph.2013.08.024.

Phi, T.T. 2016. Some Results of Quantitative Analysis of Fracture Orientation Distribution along the Segment Tien Yen-Mui Chua of Cao Bang-Tien Yen Fault Zone, Quang Ninh Province, Viet Nam. Journal of Geological Resource and Engineering, 2: 81–88. doi:10.17265/2328-2193/2016.02.004.

Priest, S.D. 1993. Discontinuity Analysis for Rock Engineering. Springer Netherlands. doi:10.1007/978-94-011-1498-1.

Pruzan, P. 2016. Research Methodology: The Aims, Practices and Ethics of Science. Springer International Publishing. Available from //www.springer.com/gp/book/9783319271668 [accessed 22 February 2018].

Raposo, J.R., Molinero, J., and Dafonte, J. 2012. Parameterization and quantification of recharge in crystalline fractured bedrocks in Galicia-Costa (NW Spain). Hydrology and Earth System Sciences, 16: 1667–1683. doi:https://doi.org/10.5194/hess-16-1667-2012.

Rinderer, M., van Meerveld, H.J., and Seibert, J. 2014. Topographic controls on shallow groundwater levels in a steep, prealpine catchment: When are the TWI assumptions valid? Water Resources Research, 50: 6067–6080. doi:10.1002/2013WR015009.

Rinderer, M., van Meerveld, I., Stähli, M., and Seibert, J. 2016. Is groundwater response timing in a pre-alpine catchment controlled more by topography or by rainfall? Hydrological Processes, 30: 1036–1051. doi:10.1002/hyp.10634.

Roques, C., Bour, O., Aquilina, L., and Dewandel, B. 2016. High-yielding aquifers in crystalline basement: insights about the role of fault zones, exemplified by Armorican Massif, France. Hydrogeology Journal, 24: 2157–2170. doi:10.1007/s10040-016-1451-6.

Roques, C., Bour, O., Aquilina, L., Dewandel, B., Leray, S., Schroetter, JM., Longuevergne, L., Le Borgne, T., Hochreutener, R., Labasque, T., Lavenant, N., Vergnaud-Ayraud, V., and Mougin, B. 2014. Hydrological behavior of a deep sub-vertical fault in crystalline basement and relationships with surrounding reservoirs. Journal of Hydrology, 509: 42–54. doi:10.1016/j.jhydrol.2013.11.023.

Salve, R., Rempe, D.M., and Dietrich, W.E. 2012. Rain, rock moisture dynamics, and the rapid response of perched groundwater in weathered, fractured argillite underlying a steep hillslope. Water Resources Research, 48. doi:10.1029/2012WR012583.

Scesi, L., and Gattinoni, P. 2010. Water Circulation in Rocks. Springer Netherlands. Available from https://www.springer.com/us/book/9789048124169 [accessed 8 May 2019].

Schultz, R.A. 2019. Geologic Fracture Mechanics. Cambridge University Press, Cambridge. doi:10.1017/9781316996737.

Şen, Z. 2015. Chapter 2 - Basic Porous Medium Concepts. In Practical and Applied Hydrogeology. Elsevier, Oxford. pp. 43–97. doi:10.1016/B978-0-12-800075-5.00002-9.

Shahbazi, A., Saeidi, A., and Chesnaux, R. 2020. A review of existing methods used to evaluate the hydraulic conductivity of a fractured rock mass. Engineering Geology, 265: 105438. doi:10.1016/j.enggeo.2019.105438.

Sharp, J.M. 2014. Fractured Rock Hydrogeology. CRC Press. doi:10.1201/b17016.

Singhal, B.B.S., and Gupta, R.P. 2010a. Hydrogeology of Crystalline Rocks. In Applied Hydrogeology of Fractured Rocks: Second Edition. Edited by B.B.S. Singhal and R.P. Gupta. Springer Netherlands, Dordrecht. pp. 237–255. doi:10.1007/978-90-481-8799-7_13.

Singhal, B.B.S., and Gupta, R.P. 2010b. Fractures and Discontinuities. In Applied Hydrogeology of Fractured Rocks: Second Edition. Edited by B.B.S. Singhal and R.P. Gupta. Springer Netherlands, Dordrecht. pp. 13–33. doi:10.1007/978-90-481-8799-7_2.

Singhal, B.B.S., and Gupta, R.P. 2010c. Introduction and Basic Concepts. In Applied Hydrogeology of Fractured Rocks: Second Edition. Edited by B.B.S. Singhal and R.P. Gupta. Springer Netherlands, Dordrecht. pp. 1–11. doi:10.1007/978-90-481-8799-7_1.

Singhal, B.B.S., and Gupta, R.P. 2010d. Hydraulic Properties of Rocks. In Applied Hydrogeology of Fractured Rocks: Second Edition. Edited by B.B.S. Singhal and R.P. Gupta. Springer Netherlands, Dordrecht. pp. 139–154. doi:10.1007/978-90-481-8799-7_8.

Singhal, B.B.S., and Gupta, R.P. 2010e. Assessment and Management of Groundwater Resources. In Applied Hydrogeology of Fractured Rocks: Second Edition. Edited by B.B.S. Singhal and R.P. Gupta. Springer Netherlands, Dordrecht. pp. 345–375. doi:10.1007/978-90-481-8799-7_20.

Song, S., Sun, F., Zhang, W., Chen, J., Xu, P., Niu, C., Cao, C., and Zhan, J. 2018. Identification of structural domains by considering multiple discontinuity characteristics: a case study of the Songta Dam. Bulletin of Engineering Geology and the Environment, 77: 1589–1598. doi:10.1007/s10064-017-1024-5.

Stoll, M., Huber, F.M., Trumm, M., Enzmann, F., Meinel, D., Wenka, A., Schill, E., and Schäfer, T. 2019. Experimental and numerical investigations on the effect of fracture geometry and fracture aperture distribution on flow and solute transport in natural fractures. Journal of Contaminant Hydrology, 221: 82–97. doi:10.1016/j.jconhyd.2018.11.008.

Tokhmechi, B., Memarian, H., Moshiri, B., Rasouli, V., and Noubari, H.A. 2011. Investigating the validity of conventional joint set clustering methods. Engineering Geology, 118: 75–81. doi:10.1016/j.enggeo.2011.01.002.

Tsang, C.-F., and Stephansson, O. 1996. A conceptual introduction to coupled thermo-hydro-mechanical processes in fractured rocks. In Developments in Geotechnical Engineering. Elsevier. pp. 1–24. doi:10.1016/S0165-1250(96)80020-4.

Turner, F.J., Turner, F.J., and Weiss, L.E. 1963. Structural Analysis of Metamorphic Tectonites. McGraw-Hill.

de Vries, J.J., and Simmers, I. 2002. Groundwater recharge: an overview of processes and challenges. Hydrogeology Journal, 10: 5–17. doi:10.1007/s10040-001-0171-7.

Wenli, Y., Sharifzadeh, M., Yang, Z., Xu, G., and Fang, Z. 2019. Assessment of fracture characteristics controlling fluid flow in discrete fracture networks (DFN). Journal of Petroleum Science and Engineering,. doi:10.1016/j.petrol.2019.04.011.

Witherspoon, P.A., Wang, J.S.Y., Iwai, K., and Gale, J.E. 1980. Validity of Cubic Law for fluid flow in a deformable rock fracture. Water Resources Research, 16: 1016–1024. doi:10.1029/WR016i006p01016.

Wright, E.P. 1992. The hydrogeology of crystalline basement aquifers in Africa. Geological Society, London, Special Publications, 66: 1–27. doi:10.1144/GSL.SP.1992.066.01.01.

Wright, S.N., and Novakowski, K.S. 2019. Groundwater recharge, flow and sTable isotope attenuation in sedimentary and crystalline fractured rocks: Spatiotemporal monitoring from multi-level wells. Journal of Hydrology, 571: 178–192. doi:10.1016/j.jhydrol.2019.01.028.

Zhang, L. 2013. Aspects of rock permeability. Frontiers of Structural and Civil Engineering, 7: 102–116. doi:10.1007/s11709-013-0201-2.

Zhang, W., Zhao, Q., Huang, R., Chen, J., Xue, Y., and Xu, P. 2016. Identification of structural domains considering the size effect of rock mass discontinuities: A case study of an underground excavation in Baihetan Dam, China. Tunnelling and Underground Space Technology, 51: 75–83. doi:10.1016/j.tust.2015.10.026.

Zhou, X., Chen, J., Ruan, Y., Zhang, W., Song, S., and Zhan, J. 2018. Demarcation of Structural Domains in Fractured Rock Masses Using a Three-Parameter Simultaneous

Analysis Method. Advances in Civil Engineering, 2018: e9358098. Hindawi. doi:https://doi.org/10.1155/2018/9358098.

Zhou, X., Chen, J., Zhan, J., Song, S., and Cao, C. 2019. Identification of structural domains considering the combined effect of multiple joint characteristics. Quarterly Journal of Engineering Geology and Hydrogeology, 52: 375–385. Geological Society of London. doi:10.1144/qjegh2018-091.

Zhu, J. 2019. Effective hydraulic conductivity of discrete fracture network with aperture-length correlation. Geosciences Journal,. doi:10.1007/s12303-019-0025-8.

CHAPTER 2

A CLUSTER-BASED MULTIPARAMETRIC SIMILARITY TEST FOR THE COMPARTMENTALIZATION OF CRYSTALLINE ROCKS INTO STRUCTURAL DOMAINS

Attoumane Abi, Julien Walter, Ali Saeidi and Romain Chesnaux

Department of Applied Sciences, Université du Québec à Chicoutimi, 555 Boulevard de l'Université, Chicoutimi, QC G7H 2B1, Canada

Quarterly Journal of Engineering Geology and Hydrogeology (2022)

https://doi.org/10.1144/qjegh2021-136

2.1 Mise en contexte

Ce chapitre présente une méthodologie répondant à l'objectif de recherche qui consiste à subdiviser la roche cristalline en compartiments de réseaux de fractures statistiquement homogènes. Ces réseaux de fractures sont ensuite utilisés au chapitre 4 pour déterminer l'influence du réseau de fractures et des paramètres de fractures sur la fluctuation du niveau des eaux souterraines et le temps de réponse dans les aquifères cristallins.

Habituellement, les études d'échantillonnage de fractures comprennent la collecte de plusieurs échantillons de fractures, qui incluent généralement de nombreuses familles de fractures. Le regroupement de ces échantillons de fractures en domaines structuraux est généralement utile aux géologues, aux hydrogéologues et aux géomécaniciens, car une région de roches fracturées est subdivisée en sous-régions au comportement statistiquement homogène par rapport aux propriétés hydromécaniques. L'une des méthodes couramment utilisées pour regrouper les échantillons de fractures en domaines structuraux considère l'orientation des familles de fractures, et ignore plusieurs caractéristiques des familles de fractures telles que l'espacement, l'ouverture et la persistance, qui sont importants pour la circulation des fluides dans le massif rocheux. Dans ce chapitre, nous proposons une nouvelle méthode de regroupement des échantillons de fracture en domaine structuraux, basée sur l'analyse des familles de fractures, qui prend en compte l'orientation, l'ouverture, la persistance ainsi que l'espacement. Des travaux de terrain ont été menés dans la province géologique de Grenville du Bouclier canadien, dans la région de Lanaudière, au Québec (Canada), où des fractures ont été échantillonnées à partir de 30 affleurements et quatre forages. La méthode proposée est plus adaptée à des applications en hydrogéologie, en mécanique des roches, et surtout dans les études de la circulation des fluides dans le massif rocheux. En outre, une méthode permettant de

compartimentaliser une zone d'étude donnée en domaines structuraux au moyen de diagrammes de Voronoï est présentée.

2.2 Introduction

For the purposes of this paper, it is useful to define some frequently used keywords. A fracture sample is a subset of a fracture population within a given study area. A fracture sample is usually composed of fractures from the same sampling station. Several fracture samples may come from the same fracture population, but adding up the fracture samples doesn't yield the fracture population. The term "fracture samples" is sometimes shortened to "sample" for readability.

It is well established that fracture networks play a significant role in promoting fluid flow in the rock mass (Hamm et al. 2007, Scesi and Gattinoni 2012). Because crystalline rocks have negligible primary permeability (permeability of the rock matrix) in comparison to their secondary permeability (permeability due to fractures), fracture characterization commonly represents one of the most important steps for determining the hydraulic characteristics of the rock mass (Singhal and Gupta 2010f, Roques et al. 2014). Fracture characterization generally requires several fracture samples, which can be from different data sources, such as outcrops, tunnels and boreholes (Singhal and Gupta 2010f). In a given study area with many fracture samples, geologists tend to compartmentalize the study area into structurally homogeneous subareas by comparing and grouping fracture samples. Those homogeneous subareas are called structural domains (Miller 1983, Kulatilake et al. 1990).

The notion of structural domains can be traced back to the early 1900s when the foundations of modern structural petrology was emerging in Austria through the work of Bruno Sander and collaborators (Turner et al. 1963). Back then, a structural domain was

defined as a statistically homogeneous section of the rock mass with regard to a given structural element (e.g., fracture, foliation) at a particular scale (Turner et al. 1963). Sensu stricto, a structurally homogeneous rock body would be identical in all structural elements that compose the rock mass; however, such a phenomenon is rare in nature (Turner et al. 1963). Delineating structural domains is the first step in studying the hydromechanical behavior of a rock mass (Kulatilake et al. 1990). Structural domains are also essential in hydrological studies as the hydraulic properties of the geological formation may vary between different domains (Kulatilake et al. 1990).

From the definition, it can be inferred that several structural elements may be used in the process of defining structural domains (Turner et al. 1963). Generally, depending on the investigated geological processes, researchers may use fold and foliation data (Turner et al. 1963, Vollmer 1990) to determine structural domains, and fracture data is also used (Miller 1983, Martin and Tannant 2004, Song et al. 2015a, Guo et al. 2020), especially in engineering fields. Independently of the considered structural element, the orientation of the structural element is the main characteristic that is utilized to explore homogeneity and delimit the domains (Turner et al. 1963).

Generally, major structural elements such as faults, folds, weathering effects, volcanic vents, and mineralized zones constitute the boundaries of structural domains (Turner et al. 1963, Priest 1993). However, there are conditions where major structural elements are very limited (Li et al. 2014a). In such circumstances, an analysis of structural elements such as fracture samples and foliation data may be sufficient to delineate the domains, as introduced by Miller (1983) and Vollmer (1990), respectively.

In studies involving the hydromechanical behavior of the rock mass, structural domains are derived from fracture data. Early statistical methods were proposed by Miller (1983) and Mahtab and Yegulalp (1984) for the assessment of the homogeneity of fracture samples by comparing pairs of fracture samples with one another. Those methods are

especially useful when stereographic plots of fracture orientations are distributed in ways that do not allow an efficient comparison of fracture samples (Kulatilake et al. 1990, Phi 2016).

Miller's method consists of subdividing the stereonet into several equal-area patches and comparing the number of fracture poles that fall into corresponding patches between the two fracture samples that are being compared. In order to assess the homogeneity of the compared samples, Miller (1983) used the chi-square test to evaluate whether there is a statistically significant difference between the stereonets of the samples. The main limitations of Miller's method are that the subdivision of the stereonets into equal area patches is arbitrary, and different patch design may lead to different results (Miller 1983). Because of the enumerated limitations and that orientation is the only fracture parameter that Miller (1983) used, several researchers have proposed variants of his method. These variants involved other equal area subdivision patches of the stereonet (QuocPhi et al. 2012, Song et al. 2015a), other fracture parameters (Li et al., 2015; Song et al., 2015, 2018; Zhang et al., 2016; Zhou et al., 2018), or other statistical tests (Martin and Tannant 2004, Li et al. 2014b, 2015, Song et al. 2015a, 2018). In the literature, it was noted that Miller's test and its variants were mainly used in engineering studies, which did not exceed a few square kilometers.

Mahtab and Yegulalp (1984) developed a method based on fracture orientations known as a similarity test. This method had a different approach, in which fracture clusters (or fracture sets) were compared between two adjacent fracture samples; however, their method was only justifiable when fracture poles on the stereonet formed clusters and were not randomly distributed. A fracture cluster is a group of fractures that share the same characteristics within a fracture sample and, in general, it is the orientation of the fractures that is considered using stereographic analysis (Tokhmechi et al. 2011). Two adjacent samples are considered to be of the same structural domain when they share at least one

similar fracture cluster (Mahtab and Yegulalp 1984, Kulatilake et al. 1997). From adjacent samples, the assessment of the similarity between clusters is based on the acute angle between clusters' resultant vectors (i.e. the acute angle between mean poles) (Mahtab and Yegulalp 1984).

While Mahtab and Yegulalp (1984) considered only fracture orientation in their method, several researchers (Zhou and Maerz 2002, Tokhmechi et al. 2011, Liu et al. 2020) proved that the use of multiple fracture parameters made clustering more accurate. Using only fracture orientations in the similarity test may lead to inaccuracies in judging the similarity between clusters from two adjacent samples. To illustrate this, in a simulated fracture sample composed of eight clusters, Tokhmechi et al. (2011) identified only three clusters when fracture orientation was alone considered. However, all eight clusters were identified when fracture orientation, infilling material, and infilling percentage parameters were considered and the principal component analysis was adopted.

Although many methods of assessing the homogeneity of fracture samples exist, none of them is based on the similarity of clusters with regard to a set of fracture parameters that control the fluid circulation in the rock mass. This study proposes a new cluster-based method for defining and delimiting structural domains by taking into consideration the fracture orientation, aperture, spacing, and persistence (also called fracture size) in the similarity analysis of clusters from adjacent samples. Fracture orientation influences fluid direction, fracture aperture influences the flow through fractures, and fracture spacing and persistence influence fracture density and interconnectivity, which in return impact the capacity of the rock mass to transmit water (Scesi and Gattinoni 2010b, Wenli et al. 2019b). This method is based on the Mahtab and Yegulalp (1984) similarity test and further explores testing the similarity of clusters with regard to fracture spacing, persistence, and aperture.

To illustrate the developed methodology, a study site with accessible outcrops of crystalline rocks and non-exploited boreholes was selected for fracture samples. Fieldwork

was conducted in Lanaudière, an administrative region located in the southern part of the province of Quebec, Canada. The data was collected under the scope of the project PACE-Lanaudière (Programme d'acquisition de connaissances sur les eaux souterraine), which aims to map the groundwater resources of Lanaudière. The application of the proposed method led to the compartmentalization of the study area into homogeneous structural domains with distinct hydrogeological properties. Such compartmentalization is particularly useful for building large-scale groundwater models in crystalline rocks, which takes into consideration the heterogeneity of different fracture networks in the study area.

2.3 Study area

The study area is bounded by the Sainte-Julienne Fault in the north-west, the river Maskinongé in the north, the Saint-Cuthbert Fault in the south-east, and the river Bayonne in the south-west (Figure 2.1). Two main geological provinces are found in the study area, namely, the Grenville Province and the Saint-Laurent Platform (Béland 1967, Clark and Globensky 1976). The Saint-Laurent Platform is composed of sedimentary rocks and occupies less than 14% of the study area. The remaining 86% is occupied by the Grenville Province, which is located in the southeastern part of the Canadian shield and is mainly composed of Precambrian metamorphic and igneous rocks (Rivers et al. 1989).

The data used in this study is collected in the Grenville Province. In this area specifically, the Saint-Didace granite, charnockitic granulite, and hornblende-biotite-plagioclase gneisses are the most common types of rocks. Saint-Didace granites are characterized by coarse-grained pinkish-gray porphyroid type of granite, with large feldspar grains (Béland 1967). Charnockitic granulites are composed of sodic-calcic plagioclase, orthoclase, perthitic microline, diopsidic clinopyroxene, green hornblende, and biotite (Béland, 1967). The hornblende-biotite-plagioclase gneisses are medium to coarse-grained

rocks with black hornblende and biotite (Béland, 1967). The foliation of the gneisses and granites are gently dipping at an average angle of 20 to 30 degrees, and lithological units are generally parallel to them (Béland 1967). Field investigations and aerial photographic interpretations by Béland (1967) show that the pattern of the foliation is very sinuous with complex variations in the orientation. Béland (1967) said, “The final impression is that of plastically deformed gneisses, somewhat like a sheet of paper which is wrinkled and rumpled and not systematically folded along regularly arranged hinge lines.”

The Sainte-Julienne Fault is located to the east of the Lanaudière region, where it is marked by a clearly visible escarpment oriented around 25°N and crosses the Saint-Julienne region (CERM-UQAC 2019). Field investigations by Clark and Globensky (1976) followed the fault line of the Saint-Cuthbert Fault in the southeast direction of the Lanaudière region, particularly on outcrops observed at the following localities: Chicot, Bayonne, Rowboat, L’Assomption, along the crest of Joliette, and in many quarries. Southwest of Joliette, the Saint-Cuthbert Fault is located in the limestones of Trenton with an estimated longitudinal displacement of 131 m (Clark and Globensky 1976).

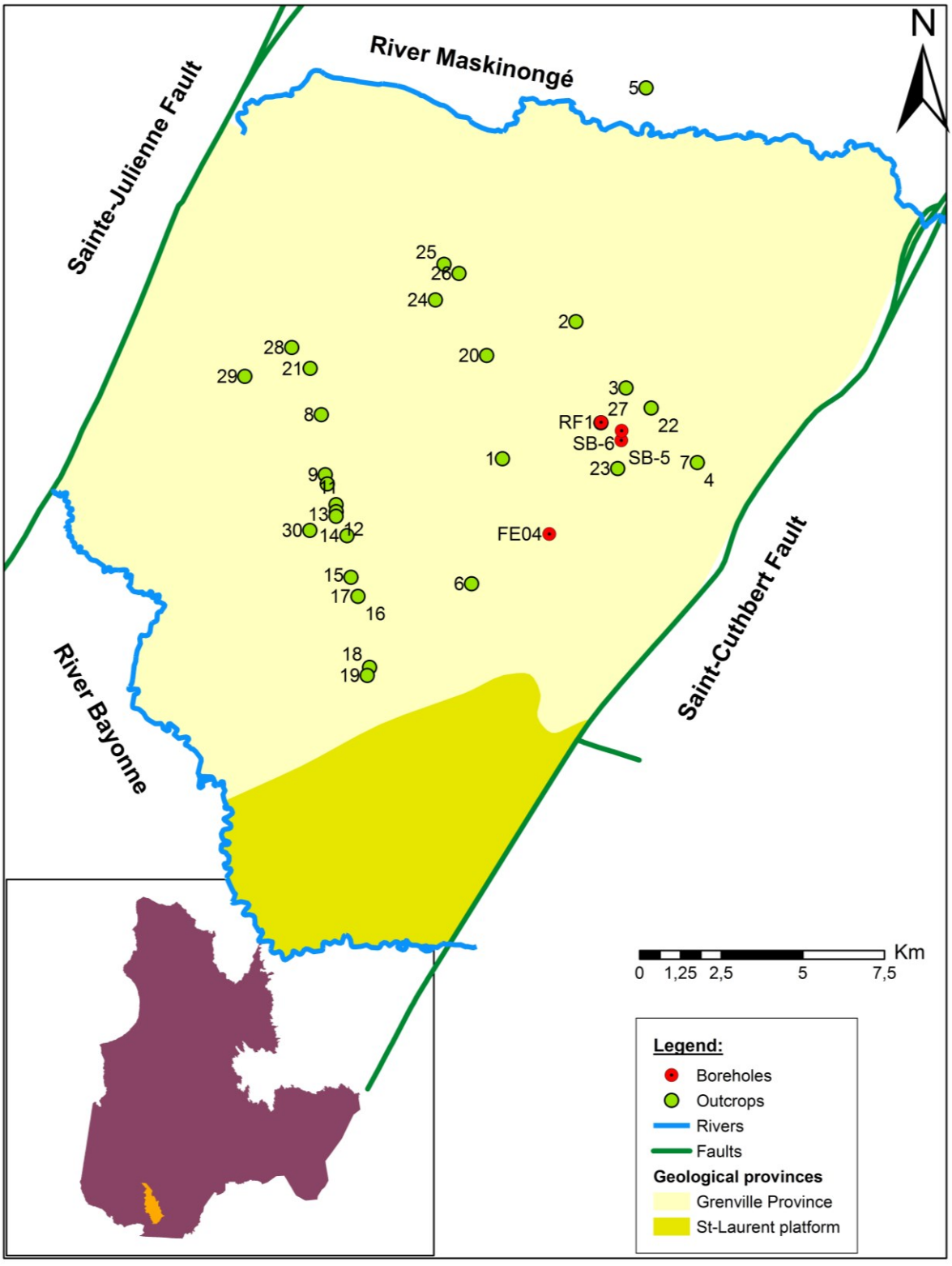


Figure 2.1. Illustration of the geological formations in the study area along with the spatial distribution of the sampling points within the study area, the type of data.

2.4 Data acquisition and processing

2.4.1 Outcrop data set

Field investigations for data collection was conducted during summer 2019 and summer 2020. Outcrops where fractures and other structural elements were not covered by mosses and lichens were investigated. The sampling method used was the scanline technique, which consists of stretching a tape on an outcrop face and systematically sampling all fractures intersecting the tape from point A to B (Priest and Hudson 1981, Priest 2004, Chesnaux et al. 2009, Manda and Mabee 2010). For each intersected fracture in the scanline direction, fracture characteristics such as the dip, dip direction, aperture and persistence were recorded.

A total of 1421 fractures were sampled on 30 outcrops. These sampled fractures were unevenly distributed throughout the study area (Figure 2.1). The number of sampled fractures by outcrop varied between 8 and 349, while scanline lengths varied between 7 m and 235 m. The number of scanlines per outcrop varied according to the outcrop configuration. If the outcrop had more than one reachable outcrop face, the number of scanlines would be at least equal to the number of outcrop faces. The numbering system for outcrops on Figure 2.1 is related to the order of sampling.

2.4.2 Borehole geophysical survey

Boreholes were another source of fracture data, which provided insight on the fracture population and distribution in the subsurface. In this study, four boreholes (Figure 2.1) were investigated using acoustic and optical borehole imaging. The investigated depths were 43, 106, 79, and 28 m for boreholes SB5, SB6, RF1, and FE04, respectively.

SB5 and FE04 had unstable and collapsing borehole walls which only allowed a partial investigation of the total borehole depth. All boreholes were unidirectional and vertical. A total number of 543 fractures were collected from RF1, SB5, FE04, and SB6, which respectively count for 206, 76, 96, and 165 fractures. In terms of fracture characteristics, the borehole imaging methods provided the dip, dip direction, and aperture of every encountered fracture.

2.4.3 Processing of structural data

Fracture sampling is, in itself, a biased activity, but necessary for engineering and geological studies as it provides essential information on the fracture network of the rock mass. Among the different sources of biases, the geologists' own judgment during data collection (Andrews et al. 2019), the methods and approaches used to collect the data (Zeeb et al. 2013), and biases in the measurement of fracture parameters (Priest 1993, Jing and Stephansson 2007b) can be also included. Depending on the sampling methods and fracture parameters, bias corrections have been proposed to attenuate their effect on analysis and interpretation of the data.

Given that the scanline sampling method and borehole imaging were used to collect fracture data for the present study, orientation biases existed and were reduced by applying a correction factor introduced by Terzaghi (1965). The orientation biases are related to the probability of sampling fractures, which is low for fracture with traces nearly parallel to the borehole or scanline direction (Terzaghi 1965, Priest 1993), while the sampling probability is high for fractures with traces almost perpendicular to the borehole or scanline direction (Terzaghi 1965, Priest 1993). Terzaghi (1965) proposed a geometrical correction factor, which considers the angle between the sampling line and the normal of the specific fracture. If we denote δ as the acute angle between the sampling line and the normal of a specific

fracture, α_n/β_n as the trend/plunge of the fracture's normal vector, and α_s/β_s as the trend/plunge of the sampling line, the bias is reduced by assigning a weight w to the fracture (Priest 1993):

$$w = 1/\cos\delta \quad \text{(Eq 2.1)}$$

where $$\cos\delta = |\cos(\alpha_n - \alpha_s)\cos\beta_n\cos\beta_s + \sin\beta_n\sin\beta_s| \quad \text{(Eq 2.2)}$$

Further information on the correction of orientation biases is provided in Priest (1993) and Terzaghi (1965). As with the aforementioned orientation biases, the true spacing of a given fracture cluster (set) was different from the apparent spacing, particularly if fractures were not perfectly perpendicular to the sampling line (Priest 1993). If X_t and X_a denote the true and apparent spacing respectively, the true spacing is given by:

$$X_t = X_a\cos\delta \quad \text{(Eq 2.3)}$$

2.5 Multiparametric similarity test

The different steps of the methodology are presented as a flowchart in Figure 2.2. The first step consisted of defining a cutoff threshold. Mahtab and Yegulalp (1984) defined cutoff threshold as a percentage of the sample size. Clusters that contain more fractures than the cutoff threshold are the main clusters of that sample. In this study, a cutoff threshold of 10% was used as proposed by Mahtab and Yegulalp (1984) and confirmed by the work of Kulatilake et al. (1990). Another important criterion that should be fulfilled by the clusters is linked to their dispersion parameter K (also called concentration factor or Fisher's constant) that should be greater than six (Mahtab and Yegulalp 1984). The parameter K assesses the concentration of the pole vectors around the mean pole vector of the cluster (Marcotte and Henry 2002). The smaller the K values the disperse the cluster (Mahtab et al.

1972, Priest 1993). The threshold of six (K≥6), defined by Mahtab and Yegulalp (1984), ensures the distribution of the pole vectors is in the neighborhood of the mean pole vector of the cluster. Details on the topic may be found in more specialized texts such as Mahtab et al. (1972) and Priest (1993). An estimate of the dispersion parameter K, of a cluster, is provided by Fisher (1953) as follows:

$$K = (N - 1)/(N - R) \qquad \text{(Eq 2.4)}$$

Where N represents the total number of pole vectors in the fracture cluster and R is the resultant vector magnitude of the fracture cluster (resultant vector is also called mean vector). In our case, K values are computed using the program DIPS® (Rocsience Inc.). Hereafter, the term “cluster” designates “main cluster”, as the multiparametric similarity test involved the use of main clusters only.

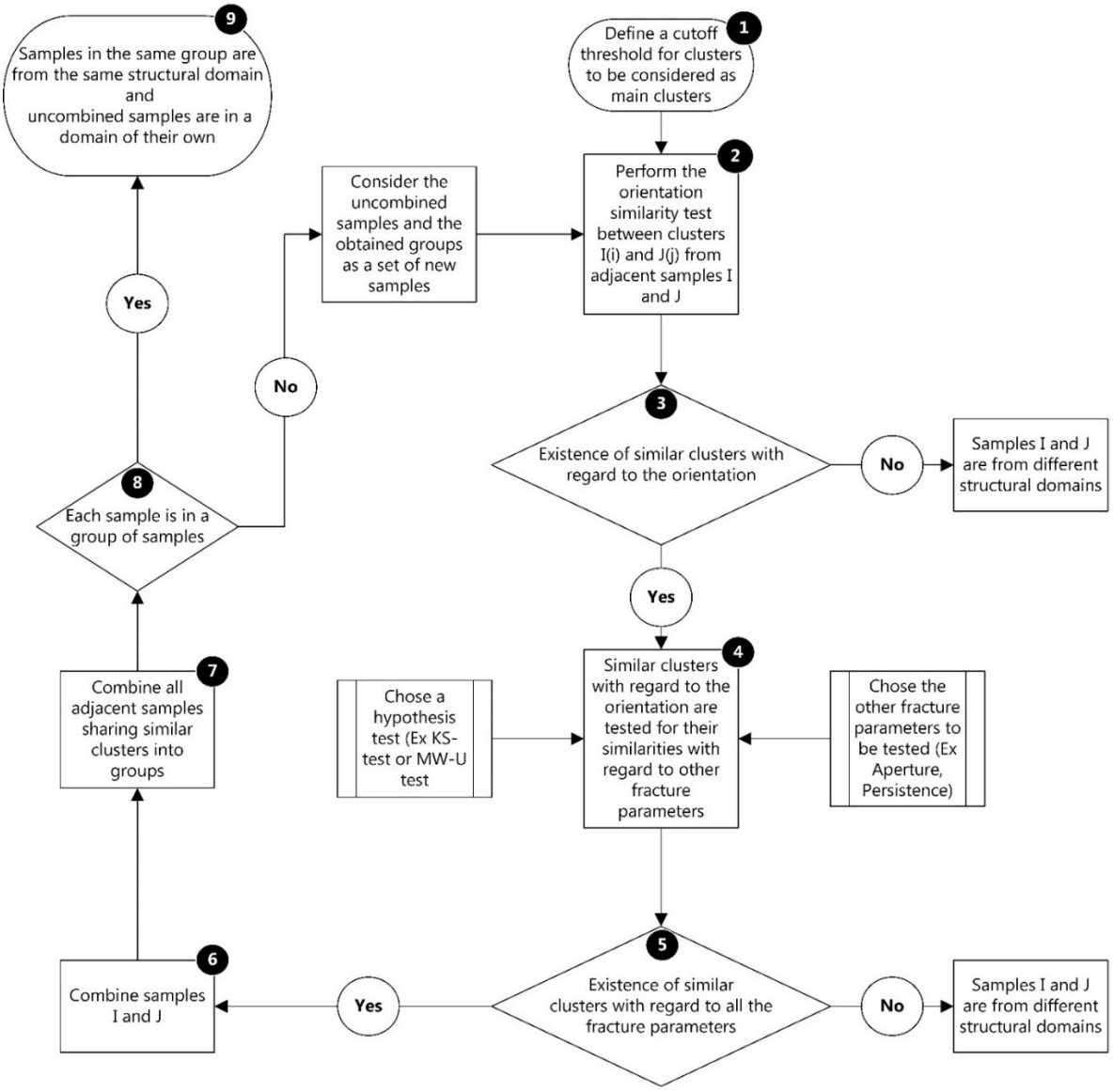


Figure 2.2. Flowchart for the multiparametric similarity test proposed in this study. The numbers in black circles indicate the steps in the flowchart.

The second step consisted of performing the orientation similarity test between all clusters from adjacent samples. We considered I(i) and J(j) as the respective clusters from sample I and J. If $\alpha_{[I(i),J(j)]}$ represents the acute angle between the mean vectors of clusters I(i) and J(j), the sample I and J are said to be similar with regard to their orientation, if Equation 2.5 is fulfilled. An illustration of the concept is provided in Figure 2.3.

$$\left|\cos \alpha_{[I(i),J(j)]}\right| > \cos \Psi_{I(i)} \; or \; \cos \Psi_{J(j)} \qquad \text{(Eq 2.5)}$$

Where $\Psi_{I(i)}$ and $\Psi_{J(j)}$ are respectively the cones of confidence of the means of clusters I(i) and J(j). The cone of confidence provides the cone angle within which the true cluster mean lies, for a chosen degree of certainty (Mahtab and Yegulalp 1984). The degree of certainty is commonly selected between 90% and 99% (Priest 1993). In the present study, the chosen degree of certainty is of 95%. A more comprehensive discussion on the cone of confidence can be found in Priest (1993) and Mahtab et al. (1972).

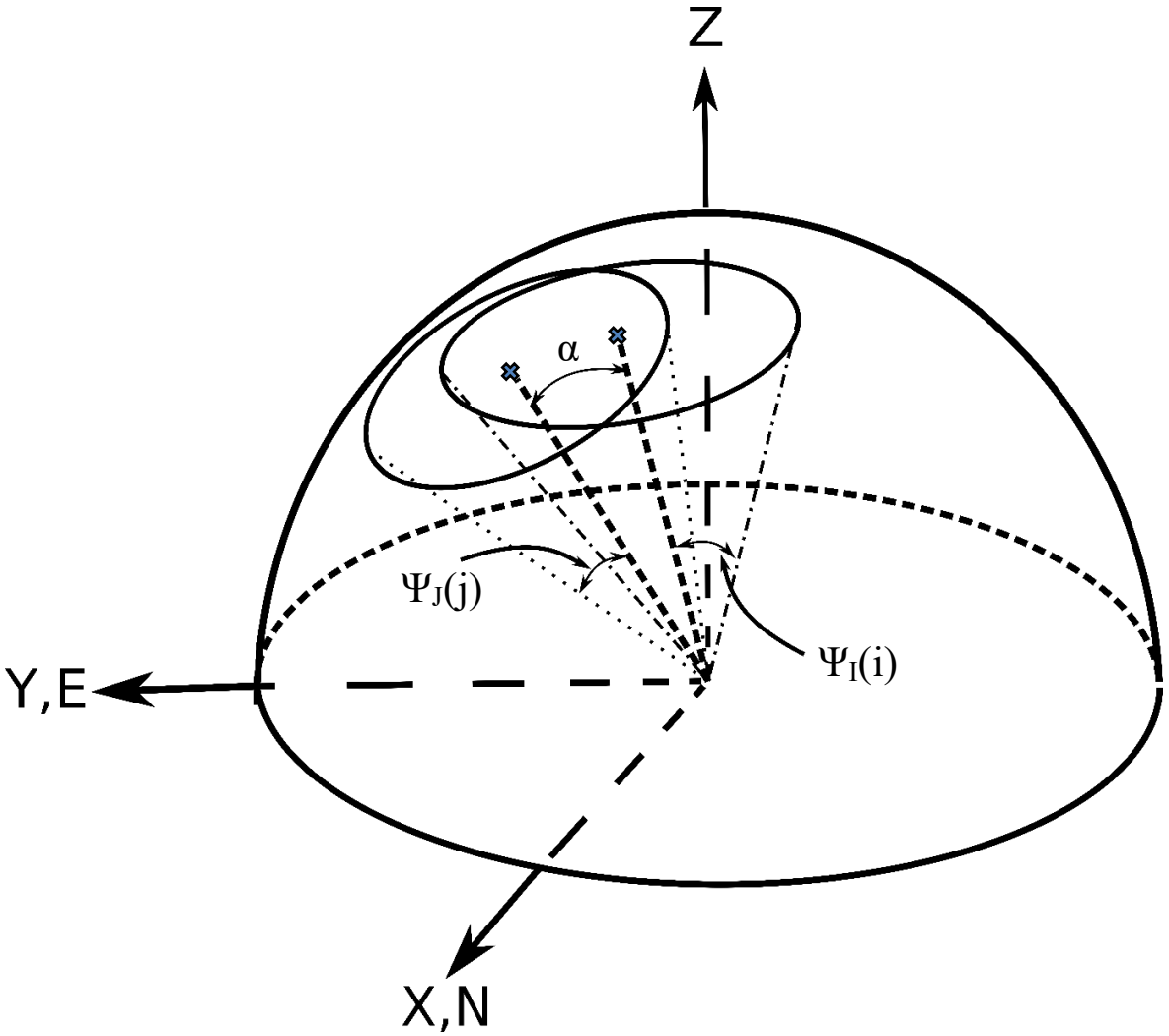


Figure 2.3. Illustration of the orientation similarity test via the comparison of the acute angle α and the means of a pair clusters and their cones of confidences $\Psi_{I(i)}$ and $\Psi_{J(j)}$ (modified after Mahtab and Yegulalp, 1984)

The third step consisted of selecting only adjacent samples I and J that had similar clusters I(i) and J(j), with regard to their orientations, for further analysis in step 4. The fourth step examined the similarities of clusters I(i) and J(j) with respect to other fracture parameters. This examination was performed by using the hypothesis test and statistical analysis on clusters I(i) and J(j). The considered fracture parameters in the hypothesis testing were fracture spacing, aperture, and persistence. Clusters I(i) and J(j) were similar with regard to the enumerated fracture parameters, if they had the same distribution functions. However, fracture parameters such as spacing (Priest and Hudson 1976), aperture (Snow 1970) and persistence (Bonnet et al. 2001b) tended to follow different distribution functions. Hence, in order to avoid an assumption about the distribution functions of clusters I(i) and J(j), nonparametric tests were performed. The most prominent feature of nonparametric tests is that no assumption is required on the underlying distribution functions of the fracture parameters for clusters I(i) and J(j) (Zimmerman 1987, 2011, Tanizaki 1997). The chosen nonparametric tests for the present study were the Mann-Whitney U (MW-U) test and the Kolmogorov-Smirnov (KS) test. The hypothesis tested by the two nonparametric tests was that the clusters I(i) and J(j), from the adjacent samples I and J, were from the same fracture population. Thus, they had the same distribution for each fracture parameter. For readability purposes, details on the MW-U and KS tests are provided in Appendix 1 and Appendix 2.

The fifth step consisted of selecting only adjacent samples I and J that had similar clusters I(i) and J(j), with regard to the chosen fracture parameters, for further analysis in step 6. The sixth step consisted of combining those adjacent samples I and J that had similar clusters. The seventh step consisted of repeating the combination process in step 6 for all the adjacent samples. The combination of samples led to the formation of groups of samples containing similar clusters. The eighth step consisted of examining whether each collected sample was within a group of samples. If yes, we proceeded to step 9, where samples in

the same group were considered to be in the same structural domain. If no, the uncombined samples and the obtained groups were considered as a new set of samples and we repeated the process from step 2. The process was iterated until new combinations were no longer found.

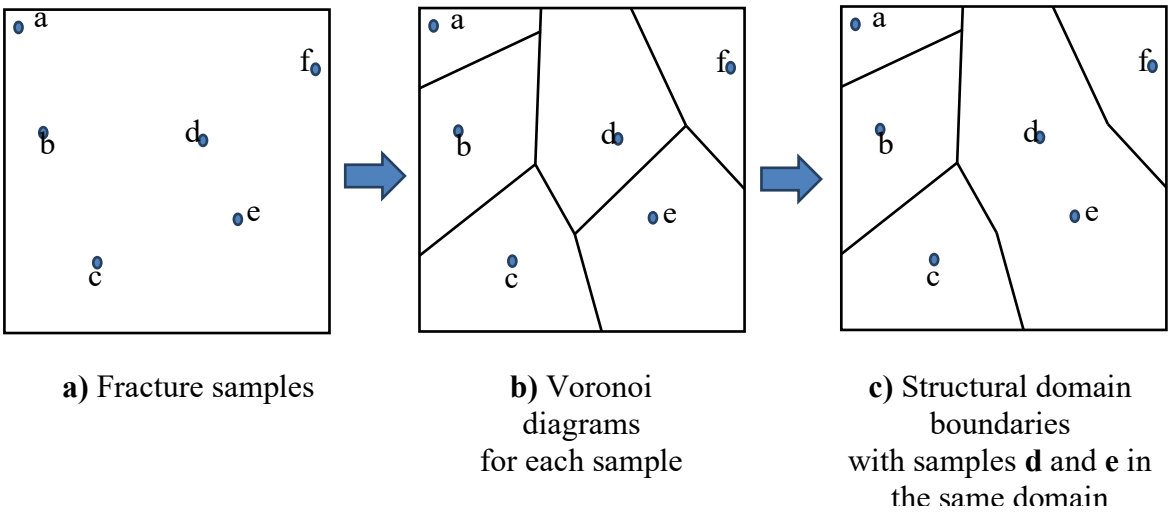
2.6 Structural domain boundaries

The method described in the previous section did not allow the identification of boundaries between the different structural domains. The identification of structural domain boundaries mainly depends on the type of data. When structural domains are investigated along a borehole (Barnes et al. 1998), a tunnel (Zhan et al. 2020) or any other continuous rock exposure, the domain boundaries are relatively well defined. However, when domain boundaries are explored within a study site involving scattered fracture samples (Mahtab and Yegulalp 1984, Song et al. 2015a), the location of domain boundaries become complex. The complexity is due to the lack of information between the samples and the fact that structural domain boundaries do not follow the geological unit boundaries (Mahtab and Yegulalp 1984, Vollmer 1990). It is therefore difficult to have unbiased and\or non-subjective domain boundaries.

In some cases, lineaments, which are linear features observable at the regional scale from satellite images or digital elevation models (DEMs), can be linked to major structural elements like faults (Gleeson and Novakowski 2009). In such circumstances, the use of lineaments in the definition of the domain boundaries would be conceivable. However, in our case, the extraction of lineaments from a one-meter resolution LiDAR DEM brought out the main glacial erosion marks, which were judged unusable for the delimitation of the domain boundaries.

Voronoi diagrams (also called Thiessen polygons) are a non-subjective alternative in the delimitation of structural domain boundaries. Voronoi diagrams are used in several disciplines, including geography (Rhynsburger 1973) and hydrology (Fiedler 2003), to divide an area into n subareas, given n points (Aurenhammer 1991). The subdivision is done according to the nearest-neighbor rule, which affects each subarea to the nearest point (Aurenhammer 1991). Usually, polygons like subareas (see Figure 2.4) are created by computing perpendicular bisectors between all neighboring points (Brassel and Reif 1979). Further information on Voronoi diagrams can be found in Aurenhammer (1991). Voronoi diagrams also have the advantage of being implemented in GIS software such as QGIS (Beyhan et al. 2020) and ArcGIS (Dong 2008, Beyhan et al. 2020).

An application of the Voronoi diagram to the present study considered each fracture sample for the computation of the polygon-like subareas. Each sample was surrounded by a polygon that constituted the boundary between it and its adjacent samples (Figure 2.4 a, b). If two adjacent samples were similar, based on the methodology, the lines of the polygon that separated the two samples were removed (Figure 2.4 c). Conversely, if the samples were not similar, the line between the fracture samples remained as illustrated in Figure 2.4 c. In this example, samples d and e were considered from the same structural domain, while samples a, b, c and f were not similar to one another.



a) Fracture samples

b) Voronoi diagrams for each sample

c) Structural domain boundaries with samples **d** and **e** in the same domain

Figure 2.4. Use of Voronoi diagrams in the delimitation of structural domain boundaries. From left to right, the maps represent, respectively, a) the fracture samples, b) the Voronoi diagrams for each sample, and c) the use of Voronoi diagrams for structural domain boundaries.

2.7 Structural domain determination

2.7.1 Orientation similarity test

To perform the similarity test, potentially similar fracture clusters of adjacent samples needed to be identified. Those potentially similar fracture clusters were clusters whose resultant vectors were visually almost parallel (at a low angle to each other) on the stereonet (Figure 2.5). However, in order to be considered similar with regard to the orientation, they needed to fulfill the angular criteria provided in Equation 2.5. Figure 2.5 considers a subregion containing many fracture samples and plots the resultant vectors of all main fracture clusters on a stereonet; the poles of the resultants of potentially similar fracture clusters tend to form clusters as illustrated. If we denote I(i) and J(j), a pair of potential fracture clusters i and j (from samples I and J, respectively), one can compute the angle

between I(i) and J(j), the cone of confidence of I(i) and J(j), and use Equation 2.5 to perform the orientation similarity test between the pair of fracture clusters.

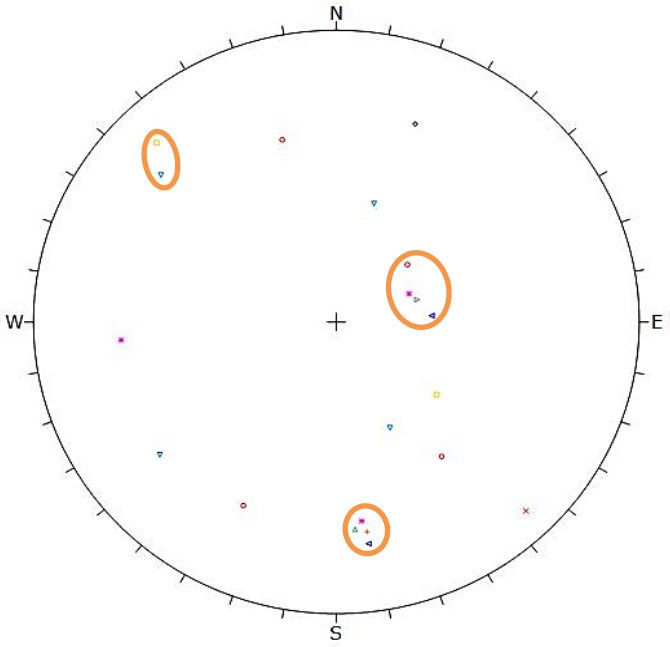


Figure 2.5. Identification of potential fracture clusters of adjacent samples for the similarity analysis on a stereonet. The circled poles of the cluster's resultants contain potentially similar clusters. Each geometrical form represents a sample which can have many clusters.

For readability purposes, partial results of the orientation similarity test are presented in Table 2.1, while the complete Table of the orientation similarity test, between all potential similar clusters, is presented in Appendix 3. The results considered fracture clusters from 30 outcrops and four boreholes. The orientation similarity test was performed between fracture clusters of two adjacent outcrop samples, outcrop-borehole samples, or borehole samples. For each test, the similarity between the pair of potentially similar clusters was either rejected (R) or fail to be rejected (FTR).

Table 2.1. Partial results of the orientation similarity tests between clusters I(i) and J(j) respectively from adjacent samples I and J. R and FTR refer to reject and fail to reject, respectively.

Fracture sets I(i)–J(j)	Angle between I(i) and J(j)	Cone of confidence I(i)	Cone of confidence J(j)	Orientation Similarity test
2(1)–3(1)	22,6	11,7	7	R
27(2)–3(1)	4	15,2	7	FTR
22(1)–3(1)	36,3	6,2	7	R
22(1)–2(1)	22,1	6,1	11,7	R

At this stage, fracture samples that contained similar clusters could be grouped to form structural domains as a result of the orientation similarity test. However, many samples were not grouped given that the orientation similarity of several potentially similar clusters was rejected (see Table 2.1 and Appendix 3). In Table 2.1 and Appendix 3, outcrop samples are identified from numbers 1 to 30, while borehole samples are named SB5, SB6, FE04, and RF1. Comparing cluster i from sample I with cluster j from sample J is denoted I(i)–J(j). Thus, 2(1)–3(1) means comparing cluster 1 from sample 2 with cluster 1 from sample 3. Five groups of similar samples were formed, containing between two and ten fracture samples, whilst nine samples remained ungrouped (Figure 2.6). Similar samples were put in parentheses; the following groups were formed: M2 = (2, 20, 24, 25, and 26), M3 = (3, 27, SB6, SB5), M6 = (6, FE04), M7 = (7, 23) and M8 = (8, 10, 11, 12, 13, 14, 15, 16, 17, and 30). Two of the groups showed orientation similarities between adjacent borehole samples and outcrop samples, which demonstrates the common practice in engineering geology of sampling fractures on outcrops for a better understanding of the subsurface fracture network.

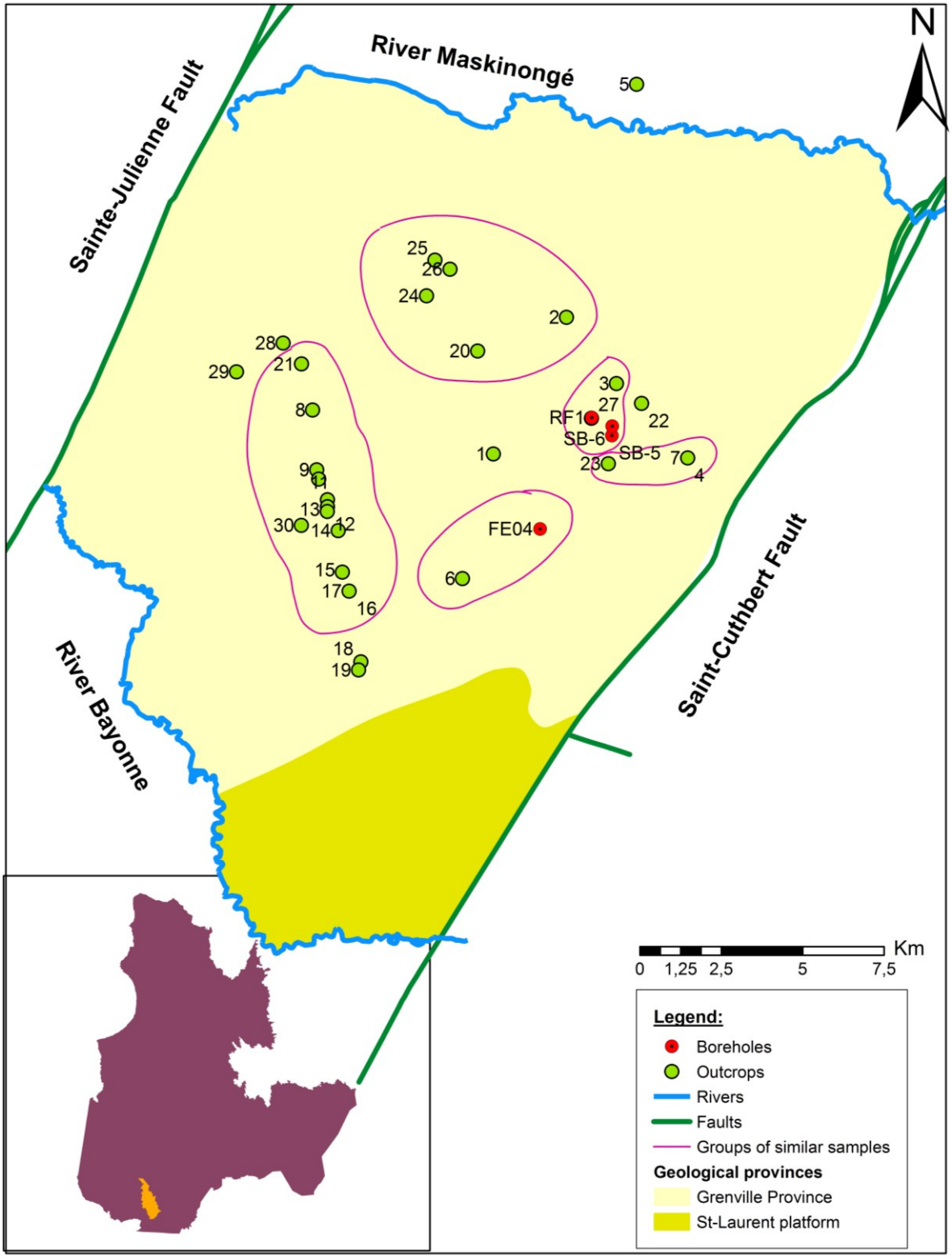


Figure 2.6. Grouping of fracture samples into structural domains according to their orientation similarities (Mahtab and Yegulalp's method). Circled samples are similar, and non-circled samples are not similar to any adjacent sample.

2.7.2 Multiparametric similarity test

This section presents the results of the multiparametric similarity test, which further investigated the similarity of pairs of clusters with regard to the fracture parameters such as persistence, spacing, and aperture. For readability purposes, partial results of the multiparametric similarity test are presented in Table 2.2, while the complete Table of the multiparametric similarity test between all clusters is provided in Appendix 4. In this study, the KS test and MW-U test were the two selected hypothesis tests used in the multiparametric similarity test, and only pairs of clusters which failed to be rejected in the orientation similarity test were considered. The level of significance α chosen for the KS test and the MW-U test was 0.05. Thus, a p-value that was greater than α indicated that the clusters were probably similar for the particular fracture parameter being tested. The overall similarity column, in Table 2.2 and Appendix 4, assesses the similarity between two clusters by taking into consideration the p-values of the two hypothesis tests for all the fracture parameters. Thus, the similarity between the two clusters was rejected, if either the KS test or the MW-U test rejected the similarity of the clusters for one parameter. The similarity between outcrop and borehole clusters was assessed without taking into consideration the persistence, as it could not be obtained from borehole data.

Table 2.2. Partial results of the first iteration multiparametric similarity test using orientation, persistence, spacing, and aperture parameters. KS test and MW-U test refer to the Kolmogorov-Smirnov test and the Mann-Whitney U test. R and FTR stand for rejected and failed to be rejected. P-values in bold are higher than 0.05.

	Persistence-p values		Spacing-p values		Aperture-p values			
Fracture set I(i)–J(j)	**KS test**	**MW-U test**	**KS test**	**MW-U test**	**KS test**	**MW-U test**	**Orientation Similarity test**	**Overall Similarity**
27(2)–3(1)	**0,079**	**0,171**	**0,121**	**0,164**	**0,231**	**0,284**	FTR	FTR
23(1)–7(1)	0,042	0,027	**0,522**	**0,550**	2,8E-4	1,1E-5	FTR	R
20(2)–24(3)	**0,055**	**0,076**	**0,657**	**0,686**	**1,000**	**0,912**	FTR	FTR

It is observed in Table 2.2 and Appendix 4 that the KS test and the MW-U test provide generally analogous results. The similarity between clusters was either rejected for both tests or failed to be rejected. However, some exceptions occurred where the similarity was rejected by one test and failed to be rejected by the other. As the two tests behaved differently, if the similarity was rejected by the KS test while failing to be rejected by the MW-U test, this should indicate that the difference between the two clusters' distribution functions was not of central tendency (i.e. not due to differences in means and medians). Conversely, if the similarity was rejected by the MW-U test while failing to be rejected by the KS test, the difference between the two distribution functions was most likely of central tendency nature. Li et al. (2014b), using a modified Miller's method, found comparable results between the KS test and the Wilcoxon rank sum test (equivalent to the MW-U test) in the assessment of the homogeneity between adjacent tunnel samples. In their results, the KS test rejected the homogeneity between samples that failed to be rejected by the Wilcoxon rank sum test, while the opposite did not occur. In the example illustrated in this paper, the similarity between clusters is accepted if it fails to be rejected by both hypothesis tests (KS-test and MW-U test). Although using two hypothesis tests increase the confidence of the similarity between clusters, as the result of one hypothesis test is verified by the other, the choice of using one or two hypothesis tests is left under the appreciation of the practitioner. It is a common practice in engineering geology to use one hypothesis test (Miller 1983, Song et al. 2015a, Zhou et al. 2018) or two hypothesis tests (Li et al. 2014a, Zhang et al. 2016b) in analysis of similarities between data sets. Arguably, choosing the right hypothesis test is much more important than the number of utilized hypothesis tests. Details on hypothesis tests are provided in more specialized books of statistics such as Johnson and Bhattacharyya (2010) and Alvo and Yu (2018).

In terms of fracture parameters, the rejection rate of the similarity between clusters was different from one parameter to the other. The rejection of the similarity between a pair

of clusters, with regard to a given fracture parameter, indicated that they were most likely not similar. Thus, a pair of clusters could wrongly be considered similar if only one fracture parameter was considered. Similar rejection rates (from hypothesis tests analysis) for different fracture parameters were reported in studies by Zhang et al. (2016) and Zhou et al. (2019).

From the first iteration results (Table 2.2 and Appendix 4) of the multiparametric similarity test, three groups of similar samples G3, G13, and G20 were formed; these groups were constituted by samples (3, 27, SB6), (13, 14, 15, 16, 17), and (20, 24, 26), respectively. A second iteration of the multiparametric similarity test was subsequently performed to integrate the ungrouped samples into the formed groups G3, G13, and G20. The overall similarity column presented in Table 2.3 shows that the overall similarities between adjacent ungrouped samples and the first iteration groups (G3, G13, and G20) were rejected. Under those circumstances, G3, G13, and G20 remained the same as in the first iteration; a spatial representation of the structural domains formed by those groups is illustrated in Figure 2.7.

Table 2.3. Second iteration similarity test using orientation, persistence, spacing, and aperture parameters. KS test and MW-U test refer to the Kolmogorov-Smirnov test and the Mann-Whitney U test. R and FTR stand for rejected and failed to be rejected. P-values in bold are higher than 0.05.

	Persistence-p values		Spacing-p values		Aperture-p values			
Fracture set I(i)–J(j)	**KS test**	**MW-U test**	**KS test**	**MW-U test**	**KS test**	**MW-U test**	**Orientation Similarity test**	**Overall Similarity**
G13(1)–12(1)	0,000	0,000	**0,785**	**0,761**	**0,853**	**0,738**	FTR	R
G13(2)–12(2)	0,021	**0,097**	**0,432**	**0,328**	**0,506**	**0,382**	FTR	R
G13(3)–30(2)	**0,253**	**0,136**	**0,954**	**0,477**	0,025	**0,093**	FTR	R
G20(4)–2(1)	**0,577**	**0,457**	**0,891**	**0,648**	0,000	0,000	FTR	R

The difference between Mahtab and Yegulalp's method (the orientation similarity test) and the multiparametric similarity test can be visually assessed by comparing Figure 2.6 and Figure 2.7. The groups of samples obtained from the multiparametric similarity test were shrunken groups of those obtained from the orientation similarity test, presented in the previous section. Thus, G3, G13, and G20 were the shrunken groups of the groups of samples M3, M8, and M2 respectively. Notably, two groups of samples (M6 and M7) that were similar according to Mahtab and Yegulalp's method were not similar with regard to the multiparametric similarity test. As the multiparametric similarity test takes into consideration other fracture parameters apart from just the fracture orientation, it was more capable of depicting discrepancies in the clusters' similarity. Hence, the multiparametric similarity test rejected the similarity of several adjacent samples that failed to be rejected through Mahtab and Yegulalp's method.

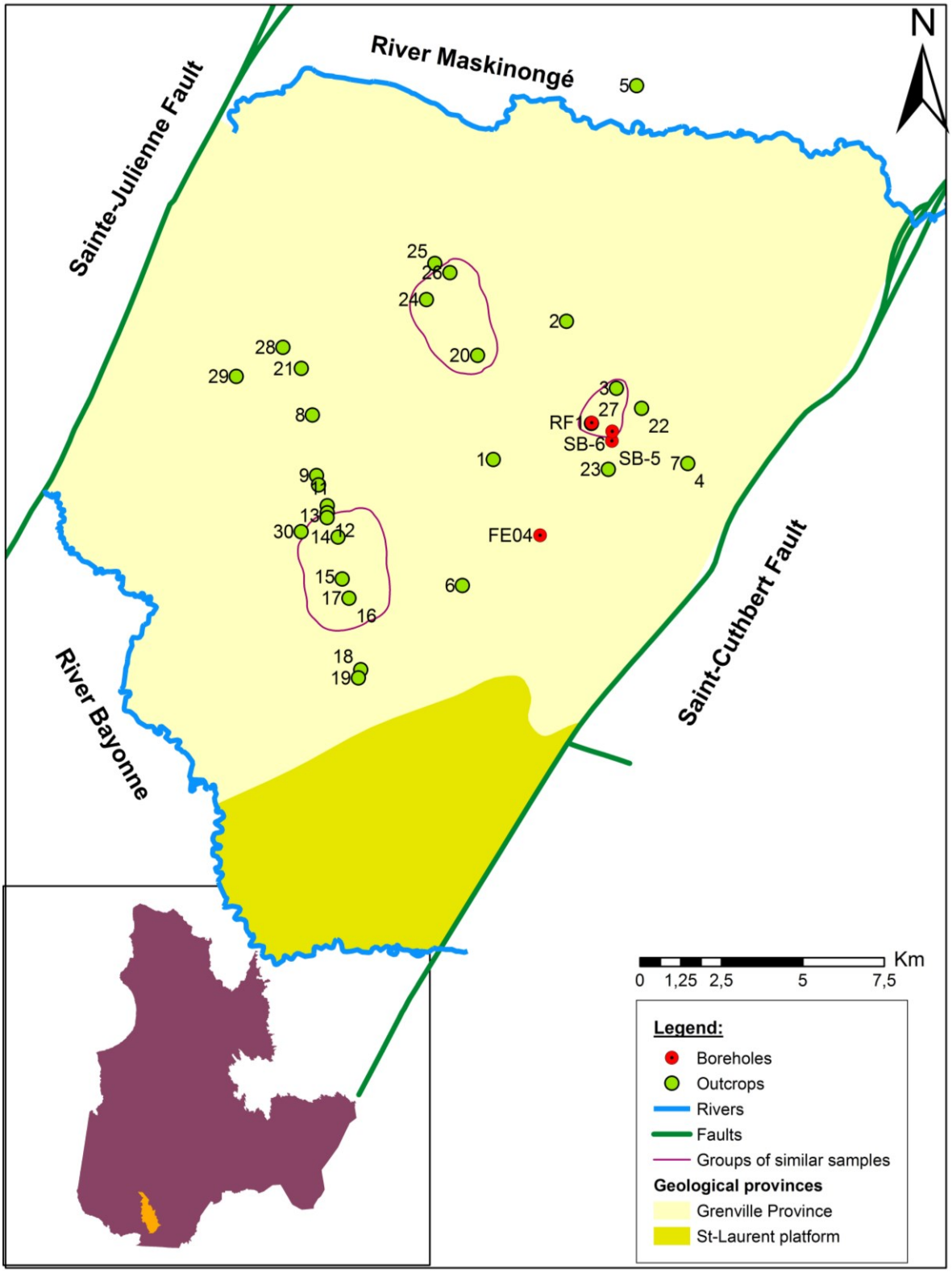


Figure 2.7. Grouping of fracture samples into structural domains according to their multiparametric similarities (fracture orientation, persistence, spacing, and aperture). Similar samples are circled, while non-circled samples are not similar to any adjacent sample.

2.7.3 Fracture parameters combinations and structural domains

In the identification of fracture clusters, the reliability and benefit of adding other fracture parameters (infilling, aperture, persistence, roughness, hardness, and termination) to fracture orientation has been demonstrated in several studies (e.g., Liu et al., 2020; Tokhmechi et al., 2011). The practitioner should be aware of the biases that each fracture parameter is subjected to. These biases have an impact on the rejection rate of the similarity between clusters. Using multiple fracture parameters increases the reliability of the similarity between clusters but also adds up the biases. Zhou et al. (2019) employed a modified Miller's method, which includes fracture orientation, persistence, and aperture and performed an experiment with simulated fracture data and real fracture data. The experiment consisted of identifying similar samples in both data sets. Zhou et al. (2019) did not find pairs of similar samples in the real fracture data, while they found similar samples in the simulated fracture data. In general, the more parameters we use the less likely that pairs of similar samples would be identified in complex natural fracture systems.

Information from Table 2.2 and Table 2.3 revealed that different fracture parameter combinations would lead to different results of similar sample groups. Figure 2.8 illustrates the effect of different combinations of fracture parameters on the grouping of fracture samples into structural domains by performing the multiparametric similarity test with two combinations of fracture parameters. The first combination (illustrated with red circles in Figure 2.8) considered fracture orientation and spacing, while the second combination (illustrated with black circles) considered all the parameters (fracture orientation, persistence, spacing, and aperture). In practice, choosing the fracture parameters combination to be used in the multiparametric similarity test should be dictated by the project objectives. For hydrogeological studies, the hydraulic conductivity in fractured rocks is mainly influenced by the fracture orientation, persistence, spacing, and aperture (Scesi and

Gattinoni 2010b, Wenli et al. 2019b). Thus, it is necessary to consider all these fracture parameters in the multiparametric similarity test for studies of fluid circulation in the rock mass. For geotechnical, mining, and engineering geology studies, several “rock quality” indexes and parameters are used to assess the physical properties of rocks such as the Rock Quality Design (RQD) and the rock mass quality Q (Barton et al. 1974, Schön 2015). The RQD and Q assess the rock mass strength. For a project investigating the rock mass strength, the present method should be modified accordingly, to take into consideration the relevant parameters.

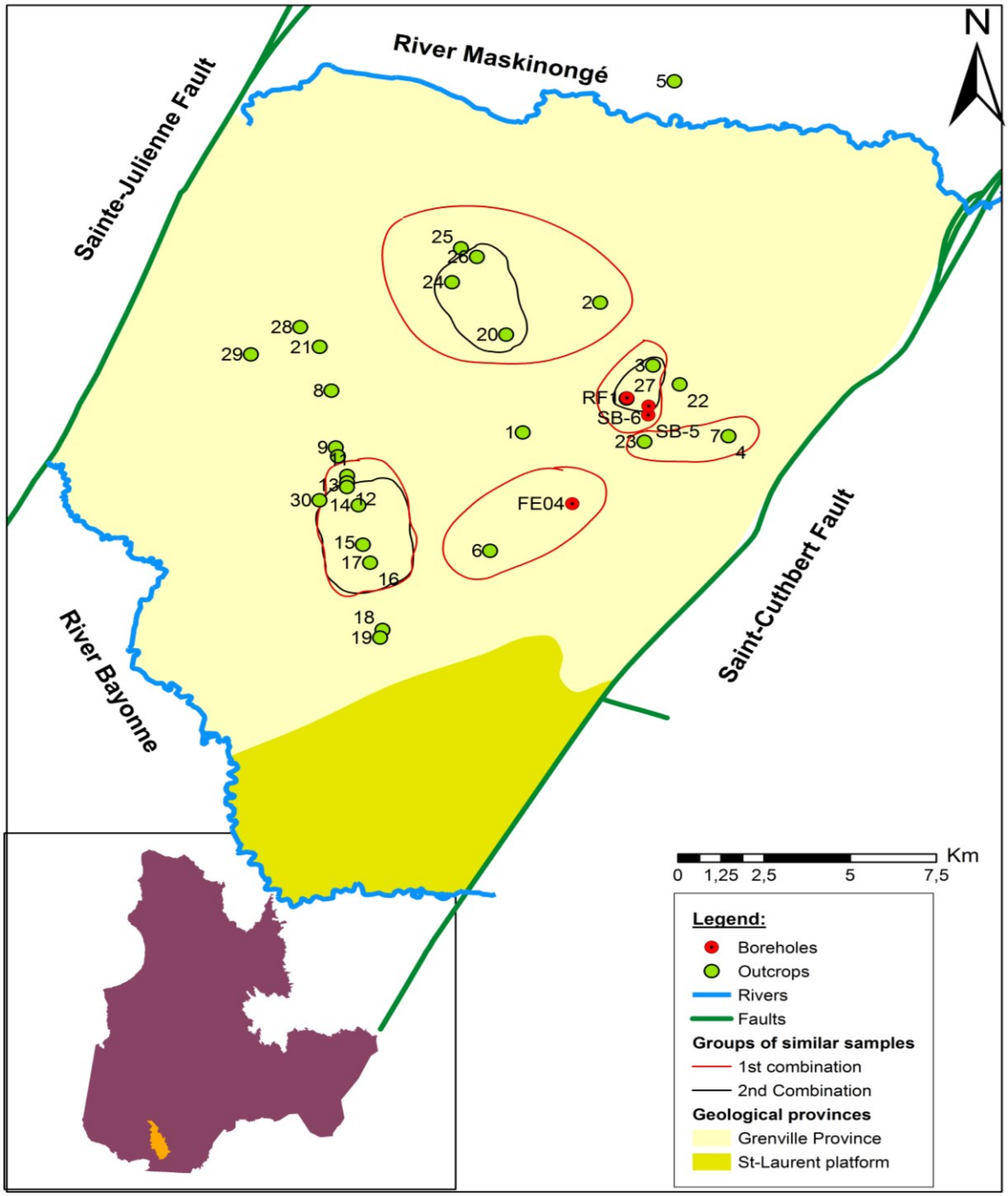


Figure 2.8. Grouping of fracture samples into structural domains according to their multiparametric similarities. Similar samples are circled, while non-circled samples are not similar to any adjacent sample. Red circles use fracture orientation and spacing, whereas black circles use fracture orientation, persistence, spacing, and apertures for the multiparametric similarity test.

In a study area with complex fracture networks, using many fracture parameters in the multiparametric similarity test may lead to a “lack of domaining”. The lack of domaining

is the lack of finding similar samples. Depending on the practitioner, the lack of domaining can be problematic, but others would see it as evidence of complex fracture networks. For instance, in hydrogeological studies, the lack of domaining can help explain why adjacent boreholes in crystalline aquifers can yield very different exploitation discharges and rock mass permeabilities. The issue of borehole yield in crystalline rocks is well known to hydrogeologists and several researchers (Bridge Moore et al. 2002, Banks et al. 2010, Holland and Witthüser 2011, Roques et al. 2016b) have investigated factors that are related to borehole yield in crystalline aquifers; however, the results are far from being generalizable. If one is confronted with a lack of domaining in complex fracture networks and still want to achieve many similar samples, the practitioner can reduce the number of fracture parameters or choose a lower α value in the hypothesis test. It should be noted that the method (Figure 2.2) proposed in this paper is flexible enough to allow the practitioner to choose the hypothesis test (step 4 of the multiparametric similarity test) that is well suited to the research objectives and the data collected.

Voronoi diagrams could be used for the delimitation of structural domain boundaries. Figure 9 illustrates the use of Voronoi diagrams in the delimitation of structural domain boundaries by considering fracture orientation, persistence, spacing, and aperture. Similar maps could be produced for each combination of fracture parameters. The resulting compartmentalization (Figure 2.9) subdivided the study area into 24 compartments of homogeneous fracture networks. Therefore, each compartment could be characterized by the properties of its fracture network such as the number of fracture clusters and the fracture cluster parameters (e.g., fracture orientation, persistence, aperture, and spacing). A summary of each compartment's fracture characteristics is provided in Appendix 5. The compartmentalization was particularly useful in regional studies as it considered the heterogeneity and complexity of the existing fracture networks in the study area. For hydrogeological studies, knowing the fracture clusters and fracture characteristics allows the

assessment of both the hydraulic conductivity and the anisotropy of the hydraulic conductivity, as exemplified by DesRoches et al. (2014). Thus, for each structural domain (compartment), the hydraulic properties of the rock mass could be estimated.

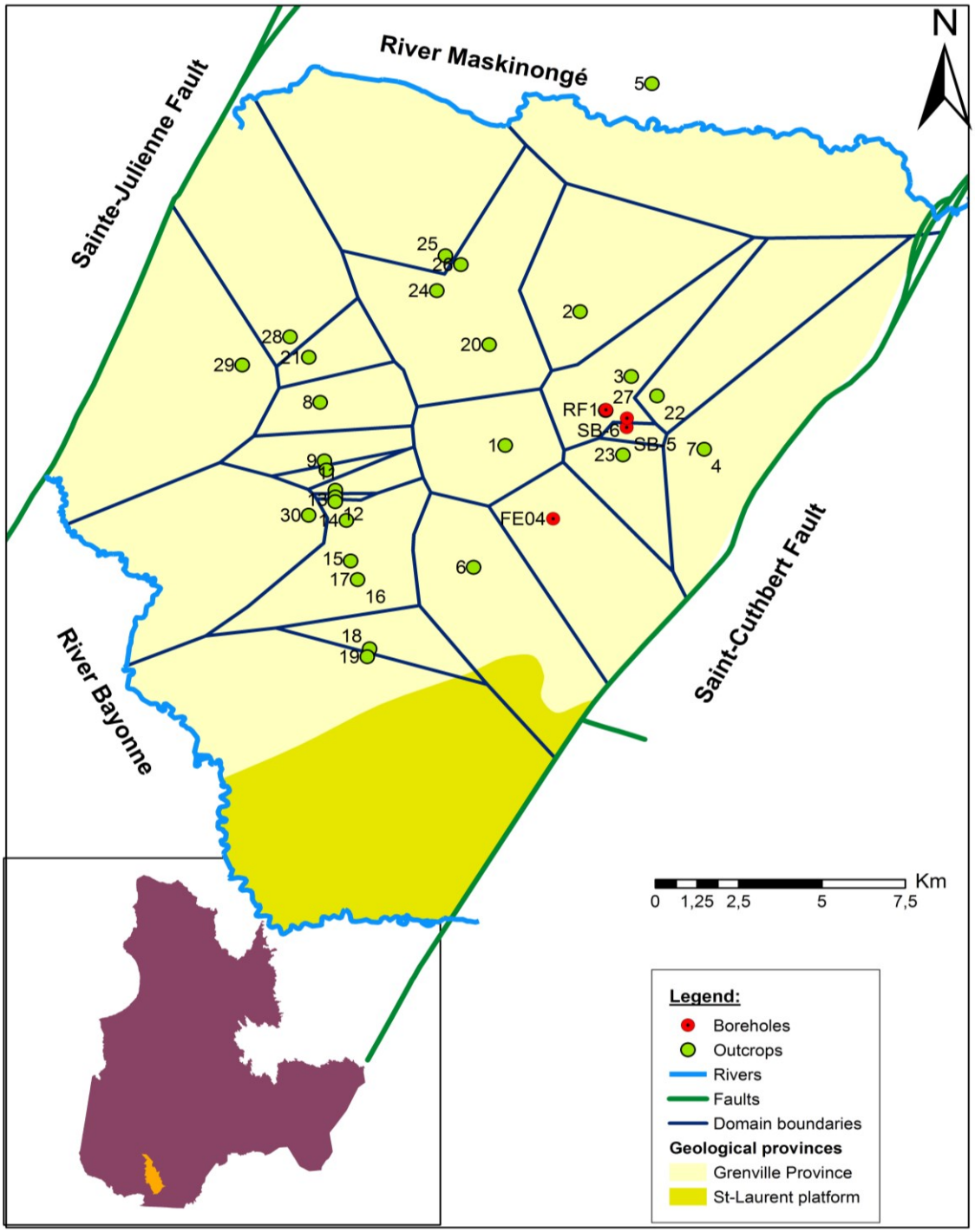


Figure 2.9. Use of Voronoi diagrams to delineate the domain boundaries according to the multiparametric similarity test, using fracture orientation, persistence, spacing, and aperture as fracture parameters.

2.8 Conclusions

Fracture samples comprise many fractures, which are described by several fracture parameters. Among the most important fracture parameters for fluid circulation are fracture orientation, aperture, persistence, and spacing. In this study we developed a cluster-based method to group and delineate fracture samples into structural domains in a rock mass. The method improved the similarity test proposed by Mahtab and Yegulalp (1984) by adding the fracture parameters, such as spacing, persistence, and aperture to the original method, which used only fracture orientation. The multiparametric similarity test is more reliable as it adds a degree of certainty to the similarity of samples by testing their similarity on multiple fracture parameters. Consequently, it provides higher confidence in the delineation of structural domains, which in return implies a better characterization of fracture networks in a study area.

The assessment of the similarity between two adjacent samples is made via clusters of samples. Ultimately, adjacent samples are similar if they share clusters that are similar with regard to their orientation, spacing, persistence, and aperture. Two clusters from adjacent samples are similar with regard to their orientation if their resultant vectors have the same orientation according to a defined criterion. The similarity between the two clusters with regard to their spacing, persistence, and aperture is assessed by means of nonparametric hypothesis testing, such as the KS test and the MW-U test.

The multiparametric similarity test has been applied to a data set comprising 30 outcrop samples and four borehole samples. The results showed that the KS test and the MW-U test perform almost equally in the identification of similar samples. Performing the multiparametric similarity test with different combinations of fracture parameters leads to different groups of similar samples. In general, the more parameters used in the multiparametric similarity test, the fewer similar samples will be identified. This highlights

that, when using only one parameter such as fracture orientation, it is possible to mistakenly consider two adjacent samples similar, while in reality they are dissimilar with regard to spacing or persistence or aperture. However, the practitioner should be conscious of the biases of fracture parameters. It is advisable that corrections are applied to reduce the fracture parameter biases, which is discussed in detail in the studies by Chilès and Marsily (1993), Priest (1993), and Jing and Stephansson (2007).

To remove subjectivity in the delineation of domain boundaries, we also proposed using the Voronoi diagram, which divides the study area into a number of subareas corresponding to the number of samples. Each fracture sample is assigned a subarea. If two adjacent samples are similar, the boundary between the two is removed and retained if the two samples are dissimilar. In general, the more fracture samples we have, the more accurate the domain boundaries.

The multiparametric similarity test has the potential to be used in all studies involving a thorough analysis of fluid circulation in the rock mass and other geomechanical studies (e.g., tunneling, and dam construction). As the method exclusively uses fracture parameters to determine the structural domains, the underlying hypothesis of the method is that the hydromechanical properties of the rock mass are homogenous in the same fracture network, hence in the same structural domain. While this hypothesis is commonly accepted and theoretically sound, further studies are required to investigate the homogeneity of the hydromechanical properties of the rock mass within the structural domains. In this context, the practice of delimiting the structural domains of a study area would be an iterative process where the practitioner finds the right fracture combination that leads to structural domains of homogeneous hydraulic (e.g., hydraulic conductivity) and/or mechanical (e.g., RQD) properties.

2.9 References

Alvo, M., and Yu, P. 2018. A Parametric Approach to Nonparametric Statistics. Springer International Publishing. doi:10.1007/978-3-319-94153-0.

Andrews, B.J., Roberts, J.J., Shipton, Z.K., Bigi, S., Tartarello, M.C., and Johnson, G. 2019. How do we see fractures? Quantifying subjective bias in fracture data collection. Solid Earth, **10**: 487–516. doi:https://doi.org/10.5194/se-10-487-2019.

Aurenhammer, F. 1991. Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Computing Surveys, **23**: 345–405. doi:10.1145/116873.116880.

Banks, D., Gundersen, P., Gustafson, G., Mäkelä, J., and Morland, G. 2010. Regional similarities in the distributions of well yield from crystalline rocks in Fennoscandia. Norges Geologiske Undersokelse Bulletin, **450**: 33–47.

Barnes, R.P., Dee, S.J., Sanderson, D.J., and Bowden, R.A. 1998. Interpretation of structural domains in discontinuity data from Nirex deep boreholes at Sellafield. Proceedings of the Yorkshire Geological Society, **52**: 177–187. Geological Society of London. doi:10.1144/pygs.52.2.177.

Barton, N., Lien, R., and Lunde, J. 1974. Engineering classification of rock masses for the design of tunnel support. Rock mechanics, **6**: 189–236. doi:10.1007/BF01239496.

Béland, R. 1967. Région de Saint-Gabriel-de-Brandon, Comtes de Joliette, Berthier et Maskinongé. Geological reports, Ministère des Richesses Naturelles du Québec, Québec.

Beyhan, B., Güler, C., and Tağa, H. 2020. An algorithm for maximum inscribed circle based on Voronoi diagrams and geometrical properties. Journal of Geographical Systems, **22**: 391–418. doi:10.1007/s10109-020-00325-3.

Bonnet, E., Bour, O., Odling, N.E., Davy, P., Main, I., Cowie, P., and Berkowitz, B. 2001. Scaling of fracture systems in geological media. Reviews of Geophysics, **39**: 347–383. doi:10.1029/1999RG000074.

Bonnini, S., Corain, L., Marozzi, M., and Salmaso, L. 2014. Nonparametric Hypothesis Testing: Rank and Permutation Methods with Applications in R. *In* 1st édition. Wiley.

Brassel, K.E., and Reif, D. 1979. A Procedure to Generate Thiessen Polygons. Geographical Analysis, **11**: 289–303. doi:https://doi.org/10.1111/j.1538-4632.1979.tb00695.x.

Bridge Moore, R., E Schwarz, G., F Clark, S., Walsh, G., and Degnan, J. 2002. Factors Related to Well Yield in the Fractured-Bedrock Aquifer of New Hampshire. US Geological Survey Professional Paper,.

CERM-UQAC. 2019. Rapport D'étape De La Phase I Basée Sur Les Données Existantes Et Planification Révisée De La Phase II. CERM-UQAC, Chicoutimi, QC.

Chesnaux, R., Allen, D.M., and Jenni, S. 2009. Regional fracture network permeability using outcrop scale measurements. Engineering Geology, **108**: 259–271. doi:10.1016/j.enggeo.2009.06.024.

Chilès, J.-P., and Marsily, G. de. 1993. 4 - Stochastic Models of Fracture Systems and Their Use in Flow and Transport Modeling. *In* Flow and Contaminant Transport in Fractured Rock. *Edited by* J. Bear, C.-F. Tsang, and G. de Marsily. Academic Press, Oxford. pp. 169–236. doi:10.1016/B978-0-12-083980-3.50008-5.

Clark, T.H., and Globensky, Y. 1976. Région de Sorel et partie sud-est de Saint-Gabriel-de-Brandon. Geological reports, Ministère des Richesses Naturelles du Québec, Québec.

DesRoches, A., Danielescu, S., and Butler, K. 2014. Structural controls on groundwater flow in a fractured bedrock aquifer underlying an agricultural region of northwestern New Brunswick, Canada. Hydrogeology Journal, **22**: 1067–1086. doi:10.1007/s10040-014-1134-0.

Dong, P. 2008. Generating and updating multiplicatively weighted Voronoi diagrams for point, line and polygon features in GIS. Computers & Geosciences, **34**: 411–421. doi:10.1016/j.cageo.2007.04.005.

Feltovich, N. 2003. Nonparametric Tests of Differences in Medians: Comparison of the Wilcoxon–Mann–Whitney and Robust Rank-Order Tests. Experimental Economics, **6**: 273–297. doi:10.1023/A:1026273319211.

Fiedler, F.R. 2003. Simple, Practical Method for Determining Station Weights Using Thiessen Polygons and Isohyetal Maps. Journal of Hydrologic Engineering, **8**: 219–221. American Society of Civil Engineers. doi:10.1061/(ASCE)1084-0699(2003)8:4(219).

Fisher, R.A. 1953. Dispersion on a sphere. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, **217**: 295–305. Royal Society. doi:10.1098/rspa.1953.0064.

Gleeson, T., and Novakowski, K. 2009. Identifying watershed-scale barriers to groundwater flow: Lineaments in the Canadian Shield. GSA Bulletin, **121**: 333–347. doi:10.1130/B26241.1.

Guo, L., Wu, L., Zhang, J., Liao, M., and Ji, Y. 2020. Identification of homogeneous region boundaries of fractured rock masses in candidate sites for Chinese HLW repository. Bulletin of Engineering Geology and the Environment, **79**: 4221–4243. doi:10.1007/s10064-020-01837-4.

Hamm, S.-Y., Kim, M., Cheong, J.-Y., Kim, J.-Y., Son, M., and Kim, T.-W. 2007. Relationship between hydraulic conductivity and fracture properties estimated from packer tests and borehole data in a fractured granite. Engineering Geology, **92**: 73–87. doi:10.1016/j.enggeo.2007.03.010.

Hart, A. 2001. Mann-Whitney test is not just a test of medians: differences in spread can be important. BMJ, **323**: 391–393. doi:10.1136/bmj.323.7309.391.

Holland, M., and Witthüser, K.T. 2011. Evaluation of geologic and geomorphologic influences on borehole productivity in crystalline bedrock aquifers of Limpopo Province, South Africa. Hydrogeology Journal, **19**: 1065–1083. doi:10.1007/s10040-011-0730-5.

Jing, L., and Stephansson, O. 2007. 5 - The Basics of Fracture System Characterization – Field Mapping and Stochastic Simulations. *In* Developments in Geotechnical Engineering. *Edited by* L. Jing and O. Stephansson. Elsevier. pp. 147–177. doi:10.1016/S0165-1250(07)85005-X.

Johnson, R.A., and Bhattacharyya, G.K. 2010. Statistics - Principles and Methods. *In* 6th edition. John Wiley & Sons, Incorporated.

Kolassa, J.E. 2020. An Introduction to Nonparametric Statistics. *In* 1st edition. Chapman and Hall/CRC, Boca Raton. doi:10.1201/9780429202759.

Krzywinski, M., and Altman, N. 2014. Nonparametric tests. Nature Methods, **11**: 467–468. Nature Publishing Group. doi:10.1038/nmeth.2937.

Kulatilake, P.H.S.W., Fiedler, R., and Panda, B.B. 1997. Box fractal dimension as a measure of statistical homogeneity of jointed rock masses. Engineering Geology, **48**: 217–229. doi:10.1016/S0013-7952(97)00045-8.

Kulatilake, P.H.S.W., Wathugala, D.N., Poulton, M., and Stephansson, O. 1990. Analysis of structural homogeneity of rock masses. Engineering Geology, **29**: 195–211. doi:10.1016/0013-7952(90)90050-B.

Li, Y., Wang, Q., Chen, J., Han, L., and Song, S. 2014a. Identification of structural domain boundaries at the Songta dam site based on nonparametric tests. International Journal of Rock Mechanics and Mining Sciences, **70**: 177–184. doi:10.1016/j.ijrmms.2014.04.018.

Li, Y., Wang, Q., Chen, J., Han, L., Zhang, W., and Ruan, Y. 2014b. Determination of structural domain boundaries in jointed rock masses: An example from the Songta dam site, China. Journal of Structural Geology, **69**: 179–188. doi:10.1016/j.jsg.2014.10.011.

Li, Y., Wang, Q., Chen, J., Song, S., Ruan, Y., and Zhang, Q. 2015. A Multivariate Technique for Evaluating the Statistical Homogeneity of Jointed Rock Masses. Rock Mechanics and Rock Engineering, **48**: 1821–1831. doi:10.1007/s00603-014-0678-6.

Liu, T., Zheng, J., and Deng, J. 2020. A new iteration clustering method for rock discontinuity sets considering discontinuity trace lengths and orientations. Bulletin of Engineering Geology and the Environment,. doi:10.1007/s10064-020-01921-9.

Mahtab, M.A., Bolstad, D.D., Alldredge, J.R., and Shanley, R.J. 1972. Analysis of Fracture Orientations for Input to Structural Models of Discontinuous Rock. U.S. Department of Interior, Bureau of Mines.

Mahtab, M.A., and Yegulalp, T.M. 1984. Similarity Test For Grouping Orientation Data In Rock Mechanics. American Rock Mechanics Association. Available from https://www.onepetro.org/conference-paper/ARMA-84-0495 [accessed 20 October 2020].

Manda, A.K., and Mabee, S.B. 2010. Comparison of three fracture sampling methods for layered rocks. International Journal of Rock Mechanics and Mining Sciences, **47**: 218–226. doi:10.1016/j.ijrmms.2009.12.004.

Marcotte, D., and Henry, E. 2002. Automatic joint set clustering using a mixture of bivariate normal distributions. International Journal of Rock Mechanics and Mining Sciences, **39**: 323–334. doi:10.1016/S1365-1609(02)00033-3.

Martin, M.W., and Tannant, D.D. 2004. A technique for identifying structural domain boundaries at the EKATI Diamond Mine. Engineering Geology, **74**: 247–264. doi:10.1016/j.enggeo.2004.04.001.

Mehta, C.R., and Patel, N.R. 2011. IBM SPSS exact tests. IBM Corporation, Armonk, NY.

Miller, S.M. 1983. A statistical method to evaluate homogeneity of structural populations. Journal of the International Association for Mathematical Geology, **15**: 317–328. doi:10.1007/BF01036073.

Phi, T.T. 2016. Some Results of Quantitative Analysis of Fracture Orientation Distribution along the Segment Tien Yen-Mui Chua of Cao Bang-Tien Yen Fault Zone, Quang Ninh Province, Viet Nam. Journal of Geological Resource and Engineering, **2**: 81–88. doi:10.17265/2328-2193/2016.02.004.

Priest, S.D. 1993. Discontinuity Analysis for Rock Engineering. Springer Netherlands. doi:10.1007/978-94-011-1498-1.

Priest, S.D. 2004. Determination of Discontinuity Size Distributions from Scanline Data. Rock Mechanics and Rock Engineering, **37**: 347–368. doi:10.1007/s00603-004-0035-2.

Priest, S.D., and Hudson, J.A. 1976. Discontinuity spacings in rock. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, **13**: 135–148. doi:10.1016/0148-9062(76)90818-4.

Priest, S.D., and Hudson, J.A. 1981. Estimation of discontinuity spacing and trace length using scanline surveys. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, **18**: 183–197. doi:10.1016/0148-9062(81)90973-6.

QuocPhi, N., SangGi, H., TruongThanh, P., and Phuong, N. 2012. Structural Domain Identification by Fracture Orientation and Fracture Density in Rock Mass | IJG:International Journal of Geoinformatics. International Journal of Geoinformatics, **8**: 35–40.

Rhynsburger, D. 1973. Analytic Delineation of Thiessen Polygons*. Geographical Analysis, **5**: 133–144. doi:https://doi.org/10.1111/j.1538-4632.1973.tb01003.x.

Rivers, T., Martignole, J., Gower, C.F., and Davidson, A. 1989. New tectonic divisions of the Grenville Province, Southeast Canadian Shield. Tectonics, **8**: 63–84. doi:https://doi.org/10.1029/TC008i001p00063.

Roques, C., Bour, O., Aquilina, L., and Dewandel, B. 2016. High-yielding aquifers in crystalline basement: insights about the role of fault zones, exemplified by Armorican Massif, France. Hydrogeology Journal, **24**: 2157–2170. doi:10.1007/s10040-016-1451-6.

Roques, C., Bour, O., Aquilina, L., Dewandel, B., Leray, S., Schroetter, JM., Longuevergne, L., Le Borgne, T., Hochreutener, R., Labasque, T., Lavenant, N., Vergnaud-Ayraud, V., and Mougin, B. 2014. Hydrological behavior of a deep sub-vertical fault in

crystalline basement and relationships with surrounding reservoirs. Journal of Hydrology, **509**: 42–54. doi:10.1016/j.jhydrol.2013.11.023.

Scesi, L., and Gattinoni, P. 2010. Water Circulation in Rocks. Springer Netherlands. Available from https://www.springer.com/us/book/9789048124169 [accessed 8 May 2019].

Scesi, L., and Gattinoni, P. 2012. Methods and models to determine the groundwater flow in rock masses: Review and examples. Environmental Science, Engineering and Technology,: 1–54.

Schön, J.H. 2015. Geomechanical Properties. *In* Developments in Petroleum Science. Elsevier. pp. 269–300. doi:10.1016/B978-0-08-100404-3.00007-X.

Siegel, S. 1956. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill.

Singhal, B.B.S., and Gupta, R.P. 2010. Applied Hydrogeology of Fractured Rocks: Second Edition. *In* 2nd edition. Springer Netherlands. Available from https://www.springer.com/gp/book/9789048187980 [accessed 3 May 2019].

Snow, D.T. 1970. The frequency and apertures of fractures in rock. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, **7**: 23–40. doi:10.1016/0148-9062(70)90025-2.

Song, S., Sun, F., Zhang, W., Chen, J., Xu, P., Niu, C., Cao, C., and Zhan, J. 2018. Identification of structural domains by considering multiple discontinuity characteristics: a case study of the Songta Dam. Bulletin of Engineering Geology and the Environment, **77**: 1589–1598. doi:10.1007/s10064-017-1024-5.

Song, S., Wang, Q., Chen, J., Cao, C., Li, Y., and Zhou, X. 2015. Demarcation of homogeneous structural domains within a rock mass based on joint orientation and trace length. Journal of Structural Geology, **80**: 16–24. doi:10.1016/j.jsg.2015.08.006.

Tanizaki, H. 1997. Power comparison of non-parametric tests: Small-sample properties from Monte Carlo experiments. Journal of Applied Statistics, **24**: 603–632. doi:10.1080/02664769723576.

Terzaghi, R.D. 1965. Sources of Error in Joint Surveys. Géotechnique, **15**: 287–304. doi:10.1680/geot.1965.15.3.287.

Tokhmechi, B., Memarian, H., Moshiri, B., Rasouli, V., and Noubari, H.A. 2011. Investigating the validity of conventional joint set clustering methods. Engineering Geology, **118**: 75–81. doi:10.1016/j.enggeo.2011.01.002.

Turner, F.J., Turner, F.J., and Weiss, L.E. 1963. Structural Analysis of Metamorphic Tectonites. McGraw-Hill.

Vollmer, F.W. 1990. An application of eigenvalue methods to structural domain analysis. GSA Bulletin, **102**: 786–791. GeoScienceWorld. doi:10.1130/0016-7606(1990)102<0786:AAOEMT>2.3.CO;2.

Wenli, Y., Sharifzadeh, M., Yang, Z., Xu, G., and Fang, Z. 2019. Assessment of fracture characteristics controlling fluid flow in discrete fracture networks (DFN). Journal of Petroleum Science and Engineering,. doi:10.1016/j.petrol.2019.04.011.

Wilcox, R.R. 1997. Some practical reasons for reconsidering the Kolmogorov-Smirnov test. British Journal of Mathematical and Statistical Psychology, **50**: 9–20. doi:10.1111/j.2044-8317.1997.tb01098.x.

Zeeb, C., Gomez-Rivas, E., Bons, P.D., Virgo, S., and Blum, P. 2013. Fracture network evaluation program (FraNEP): A software for analyzing 2D fracture trace-line maps. Computers & Geosciences, **60**: 11–22. doi:10.1016/j.cageo.2013.04.027.

Zhan, J., Pang, Y., Chen, J., Cao, C., Song, S., and Zhou, X. 2020. A Progressive Framework for Delineating Homogeneous Domains in Complicated Fractured Rock Masses: A Case Study from the Xulong Dam Site, China. Rock Mechanics and Rock Engineering, **53**: 1623–1646. doi:10.1007/s00603-019-01999-y.

Zhang, W., Zhao, Q., Huang, R., Chen, J., Xue, Y., and Xu, P. 2016. Identification of structural domains considering the size effect of rock mass discontinuities: A case study of an underground excavation in Baihetan Dam, China. Tunnelling and Underground Space Technology, **51**: 75–83. doi:10.1016/j.tust.2015.10.026.

Zhou, W., and Maerz, N.H. 2002. Implementation of multivariate clustering methods for characterizing discontinuities data from scanlines and oriented boreholes. Computers & Geosciences, **28**: 827–839. doi:10.1016/S0098-3004(01)00111-X.

Zhou, X., Chen, J., Ruan, Y., Zhang, W., Song, S., and Zhan, J. 2018. Demarcation of Structural Domains in Fractured Rock Masses Using a Three-Parameter Simultaneous Analysis Method. Advances in Civil Engineering, **2018**: e9358098. Hindawi. doi:https://doi.org/10.1155/2018/9358098.

Zhou, X., Chen, J., Zhan, J., Song, S., and Cao, C. 2019. Identification of structural domains considering the combined effect of multiple joint characteristics. Quarterly Journal of Engineering Geology and Hydrogeology, **52**: 375–385. Geological Society of London. doi:10.1144/qjegh2018-091.

Zhu, J. 2019. Effective hydraulic conductivity of discrete fracture network with aperture-length correlation. Geosciences Journal,. doi:10.1007/s12303-019-0025-8.

Zimmerman, D.W. 1987. Comparative Power of Student T Test and Mann-Whitney U Test for Unequal Sample Sizes and Variances. The Journal of Experimental Education, **55**: 171–174. Routledge. doi:10.1080/00220973.1987.10806451.

Zimmerman, D.W. 2011. A simple and effective decision rule for choosing a significance test to protect against non-normality. British Journal of Mathematical and Statistical Psychology, **64**: 388–409. doi:https://doi.org/10.1348/000711010X524739.

CHAPTER 3

GROUNDWATER LEVEL MONITORING USING EXPLOITED DOMESTIC WELLS: OUTLIER REMOVAL AND IMPUTATION OF MISSING VALUES

Attoumane Abi, Julien Walter, Romain Chesnaux and Ali Saeidi

Department of Applied Sciences, Université du Québec à Chicoutimi, 555 Boulevard de l'Université, Chicoutimi, QC G7H 2B1, Canada

Submitted to hydrogeology journal (2022)

3.1 Mise en contexte

Ce chapitre porte sur l'acquisition et le traitement des données des séries temporelles du niveau des eaux souterraines à partir d'un réseau de surveillance mixte composé de piézomètres et de forages privés exploités en zone rurale. L'inclusion de forages privés dans ces réseaux de surveillance peut fournir un moyen peu coûteux d'obtenir un ensemble plus large de données ; cependant, l'utilisation de ces sites est limitée, car le pompage fréquent de ces forages génère des valeurs aberrantes dans les séries temporelles enregistrées. Dans le cadre de cette étude, un critère basé sur la pente du niveau des eaux souterraines dans la série temporelle est utilisé pour identifier et éliminer les valeurs aberrantes des séries temporelles du niveau des eaux souterraines provenant de forages privés exploités. Néanmoins, l'élimination des valeurs aberrantes crée un problème de valeurs manquantes, ce qui biaise l'analyse ultérieure des séries temporelles. Ainsi, 14 méthodes d'imputation ont été utilisées pour remplacer les valeurs manquantes. L'approche proposée est appliquée aux séries temporelles du niveau des eaux souterraines provenant d'un réseau de surveillance de 20 forages dans la région de Lanaudière, Québec, Canada. Le critère de pente s'avère très efficace pour identifier les valeurs aberrantes dans les forages privés exploités. Parmi les caractéristiques du modèle de valeurs manquantes, la taille et la position des lacunes le long de la série chronologique sont les paramètres les plus importants qui affectent les performances des 14 méthodes d'imputation. Parmi les méthodes d'imputation testées, l'interpolation linéaire et l'interpolation de Stineman, ainsi que le filtrage de Kalman ont été les plus efficaces. La présente étude démontre que les forages privés exploités peuvent être utilisés pour la surveillance des eaux souterraines en éliminant les valeurs aberrantes et en imputant les valeurs manquantes.

3.2 Introduction

Groundwater level monitoring plays an essential role in hydrogeological studies, and the data are usually collected from a monitoring network of boreholes. Zhou et al. (2013) enumerated seven potential uses of groundwater level data over the long-term: (1) the characterization of groundwater systems; (2) the analysis of quantitative groundwater status; (3) the identification of changes in the groundwater system storage, inflow, and outflow; (4) the detection of climate change effects on groundwater resources; (5) the assessment of groundwater exploitation impacts; (6) the calibration of groundwater flow models; and (7) the evaluation of effective groundwater management and protection measures. A borehole monitoring network generally comprises unexploited boreholes whose locations and density are carefully designed to meet specific goals (Ling et al. 2003, Jørgensen and Stockmarr 2009, Thorslund and van Vliet 2020). Although these boreholes are generally drilled and maintained by governmental agencies, the costs involved in building and maintaining a groundwater monitoring network requires careful planning (Andricevic 1990, Zhou et al. 2013, Hosseini and Kerachian 2017).

Generally, private boreholes are numerous in rural areas, which have dispersed populations and where groundwater is the primary household water resource (Fornés et al. 2005b, Raposo et al. 2012b). However, the spatial density of private boreholes is even greater in relatively urbanized areas without water supply systems (Tremblay et al. 2015). In such areas where numerous private boreholes exist, the use of these boreholes in the monitoring network can provide a unique and cost-effective solution for establishing new monitoring networks or extending existing networks. Nonetheless, groundwater levels recorded from exploited boreholes are subject to noise because of the frequent pumping of water (Peterson et al. 2018). Water pumping generates an abrupt drop in the groundwater level within the borehole, a change that is not representative of the natural groundwater level

fluctuations of the aquifer (Peterson et al. 2018). These abrupt drops are undesirable and can be treated as outliers in the groundwater level time series (Li et al. 2016, Peterson et al. 2018). Outliers complicate time-series analysis because they undermine, for example, trend analysis and model calibration (Peterson et al. 2018). Outliers also bias the spatial interpolation of groundwater static water depth (Tremblay et al. 2015).

Outliers, in this context, can be defined as observations that do not match expected values for a groundwater level time series (Peterson et al. 2018, Blázquez-García et al. 2022). In other words, outlier values abruptly deviate from the expected evolution of a time series (Chen et al. 2008, Tremblay et al. 2015). Several factors can produce outliers in groundwater level time series, including a flooded borehole, borehole collapse, rapid pumping, and errors in data management and handling (Peterson et al. 2018); however, the problem of identifying and dealing with outliers in groundwater level time series has received minimal attention to date (Peterson et al. 2018). Nevertheless, in a review article by Blázquez-García et al. (2022), it is recalled that outlier detection in time series has been studied in other disciplines since 1972 by Fox (1972). Outlier detection methods can be classified into two categories, namely univariate and multivariate methods (Blázquez-García et al. 2022). Univariate methods are the most studied and developed in the literature (Blázquez-García et al. 2022), including in groundwater studies (Li et al. 2016, Peterson et al. 2018). Although a review of outlier detection methods is beyond the scope of this article, since outlier detection is not the main objective of this study and outlier detection is only a step of the proposed methodology in this paper, it is worth mentioning the few main works dealing with the subject in hydrogeology. A detailed review of outlier detection methods can be found in Blázquez-García et al. (2022).

The limited research addressing the identification of outliers in groundwater levels includes that of Li et al. (2016). Li et al. (2016) applied a self-learning Pauta-criterion algorithm which detected as outliers, groundwater levels that differed from the mean at three

times the standard deviation of the time series. The authors replaced the outliers with values obtained using a moving average; however, the resulting groundwater level time series still contained outliers, and the performance metrics of the method were not presented in the paper. The other important work is that of Peterson et al. (2018), which presented several tools to identify outliers and errors in groundwater level time series. These tools included identifying rapid change in groundwater levels and improbably high or low groundwater level values. The approach of Peterson et al. (2018) for identifying rapid changes in groundwater level time series relied on a maximum slope threshold of the time series. This maximum slope threshold is also called “constraints on the speed of data changes” in data cleaning, in computer science (Song et al. 2015b, Zhang et al. 2016a). Peterson et al. (2018) have proven the efficacy of using such a slope threshold in detecting outliers due to rapid changes in groundwater level time series.

When identified, the outliers are typically removed from the time series; however, their removal create gaps, i.e., missing values, in groundwater level time series (Li et al. 2016, Peterson et al. 2018). Although missing values in a time series can be ignored—or one can choose to remove the time series entirely from the analysis—there have been attempts to estimate and replace missing values (Gill et al. 2007, Farhangfar et al. 2007, Hron et al. 2010). Data analysis on a time series with missing values generates efficiency losses (Farhangfar et al. 2007), complicates data handling and analysis (Farhangfar et al. 2007), and biases the calculation of the mean value required for many statistical analyses (Schneider 2001). Furthermore, most statistical software and analyses, including those of time series, require a complete data (Farhangfar et al. 2007). Raaijmakers (1999) showed that 10% missing values in a time series result in an efficiency loss in statistical analysis of up to 35%. This observation led Lodder (2013) to suggest using imputation methods to replace missing values in data sets having more than 10% missing values.

The replacement of missing values in the time series is generally referred to as the imputation of missing values (Moritz and Bartz-Beielstein 2017, Dwivedi et al. 2022). The imputation of missing values in time series is common to several disciplines, including climatology (Schneider 2001, Demirhan and Renwick 2018, Afrifa-Yamoah et al. 2020), hydrogeology (Asgharinia & Petroselli, 2020; Dax & Zilberbrand, 2018; Dwivedi et al., 2022), statistics (Farhangfar et al. 2007, Hron et al. 2010), clinical epidemiology (van der Heijden et al. 2006). Most methods used to impute missing values are statistical approaches, soft computing techniques, or forecasting methods (Demirhan and Renwick 2018). As with outlier detection methods, missing value imputation methods come in univariate and multivariate forms, with univariate methods being the most widely used (Moritz et al. 2015). In hydrogeology, statistical methods such as linear interpolation (Sakizadeh et al. 2019, Asgharinia and Petroselli 2020, Noori and Singh 2021, Dwivedi et al. 2022), random forest (Dwivedi et al. 2022), and regression analysis (Asgharinia and Petroselli 2020) are commonly used. Soft computing techniques via artificial neural networks (Gill et al. 2007), support vector machines (Gill et al. 2007), and machine-learning approaches (Evans et al. 2020) have also been applied. It is worth mentioning that a simple imputation method such as linear interpolation was proven to perform as well as random forest (Dwivedi et al. 2022) or Kalman filtering with smoothing (Demirhan and Renwick 2018). Likewise for the outlier detection methods, a full review of imputation methods is beyond the scope of this article. A thorough presentation of the topic can be found in specialized books such as Buuren (2012) and Little and Rubin (2019).

Regardless of the imputation method and discipline, most methods have been tested and validated on annual, monthly, daily, and—rarely— subhour data sets (Dwivedi et al. 2022). Dwivedi et al. (2022) suggested that this rarity in analyzing dense subhour data stems from technological advances only recently being able to offer sensors that can record dense time series over extended periods. To the knowledge of the authors, the only paper in

hydrogeology that has used a subhour data set to impute missing values is Dwivedi et al. (2022), in which the authors used linear interpolation and random-forest methods to impute missing values in time series (30 min resolution) having up to 90% missing values. However, the results of Dwivedi et al. (2022) are far from generalizable, as the missing value pattern in the data sets greatly influences the effectiveness of the imputation method (Demirhan and Renwick 2018, Dwivedi et al. 2022). The term missing value pattern here refers to the gap size, i.e., the number of consecutive missing values, and the gap distribution. For instance, Demirhan and Renwick (2018) demonstrated, using several solar radiation time series from various countries and variable time steps (from minutes to weeks), that the optimal imputation method for a given data set having missing values varied among data sets and time steps.

This study presents a methodology to address the noise generated by frequent pumping in groundwater level time series, which allows optimizing the use of exploited private boreholes in monitoring networks. A step-by-step approach is proposed for determining and removing outliers in groundwater level time series and imputing the missing values in exploited private boreholes data set. This approach uses a slope criterion to identify outliers, and 14 statistical imputation methods are applied to impute missing values. A 20-borehole monitoring network consisting of 14 exploited private boreholes and 6 unexploited boreholes is used to demonstrate the applicability and effectiveness of the proposed methodology on a mixed monitoring network of exploited and unexploited boreholes. The 20-borehole monitoring network, established in the Lanaudière region, Québec, Canada, was instrumented from summer 2019 to summer 2021, with a 15 min recording time step. The application of the proposed methodology led to the reconstruction of groundwater level time series of exploited private boreholes with acceptable errors. The application of the proposed methodology led to the reconstruction of groundwater level time series of exploited private boreholes with acceptable errors. Such a use of exploited private

boreholes is particularly useful for groundwater resources management, as it allows monitoring the evolution of the water table in relatively urbanized areas, which may be prone to overexploitation.

The 20 borehole monitoring network was established as part of the PACES–Lanaudière project (Québec program for the acquisition of knowledge on groundwater), which aims to develop a portrait of the groundwater resources of the Lanaudière region (CERM-PACES 2022). The project goals included 1) determining the state of the resource, its recharge, and its vulnerability; 2) developing partnerships between water stakeholders and managers of the territory to promote the sound management of the resource; and 3) responding to the concerns of the community in regard to the groundwater resource

3.3 Study area

The 465 km^2 study site is found in the southern part of the Lanaudière region of Québec and has been described thoroughly by Abi et al. (2022). It is bordered to the north by the Maskinongé River, to the northwest by the Sainte-Julienne Fault, to the southwest by the Bayonne River, and to the southeast by the Saint-Cuthbert Fault. Two geological provinces can be found in the area (Figure 3.1): the Grenville Province and the St. Lawrence Platform (Béland 1967, Clark and Globensky 1976). The Grenville Province comprises typical gneisses and granitic rocks of the Canadian Shield and covers about 86% of the study area. The St. Lawrence Platform occupies the remaining 14% and is composed of sedimentary rocks. The study area is uplifted with elevations ranging between 16 and 301 m a.s.l. (Figure 3.1). The area is likely a recharge zone of the Laurentide Plateau (CERM-PACES 2022).

The local population depends on groundwater for its water supply, and most households have drilled a borehole to access water. There are 986 enumerated boreholes

in the study area. Two types of aquifers are exploited in the area: shallow aquifers comprising glacial sediment deposits (tills, sand, and clay) and deeper crystalline aquifers (CERM-PACES 2022). Sediment deposit thickness varies highly but tends to be thin, and the bedrock outcrops in many places.

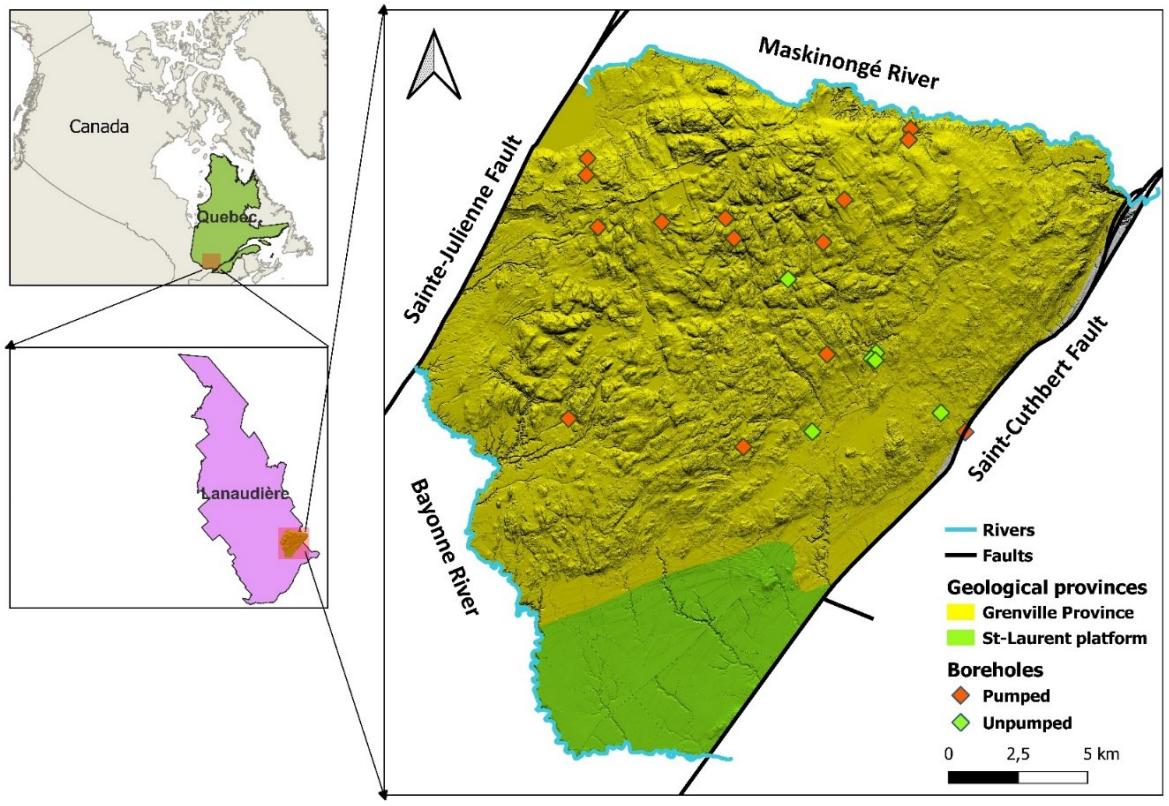


Figure 3.1. Representation of the spatial distribution of the boreholes used in the study area overlain on the geological provinces. The boreholes are divided into pumped and unpumped boreholes. The elevation of the digital elevation model in the background is exaggerated by a factor of five for illustration purposes.

3.4 Groundwater monitoring network

Groundwater levels were recorded from a monitoring network of 20 boreholes drilled into the Canadian Shield (Figure 3.1). The boreholes were instrumented with two dataloggers: a Solinst model 3001 Levelogger® 5 (Solinst Canada Ltd., Georgetown, Ontario), which measures the total pressure head overlying the pressure sensor, including

the water column and the air column, and a Solinst model 3001 Barologger 5, which only measures the air column pressure. Both dataloggers measure the pressure and temperature of their environment. The water column of each borehole was obtained by removing the air pressure from the level logger measurements. For each datalogger, the pressure and temperature sensors had an accuracy of ±0.05% FS (full scale) and ±0.05 ºC, respectively (Solinst Canada Ltd. 2021). The recording time step was set at 15 min, as successfully implemented by Gleeson et al. (2009) for monitoring an extremely rapid recharge in a crystalline rock aquifer.

Groundwater level was monitored between summer 2019 and summer 2021; however, the monitoring periods were not identical for all 20 boreholes because the data loggers were not installed and removed simultaneously. The collected groundwater level data reflected monitoring periods having a 15 min time step. Differences in the total data points recorded per borehole stemmed from differences in the length of the monitoring periods (Appendix 6). The percentage of missing values per borehole varied between 0% and 0.8%. Missing values occurred because of routine field activities, such as hydrogeological testing within the boreholes, downloading data, and reprogramming the datalogger for future data collection. The spatial distribution of the monitoring borehole network was uneven across the study area because of the lack of boreholes in some uninhabited areas. The monitoring network comprised 6 unpumped boreholes in which 3 were drilled specifically for this study, and 14 private boreholes subject to regular pumping for domestic use (Figure 3.1). It's worth mentioning that the water supply system installed in domestic wells is equipped with a pump, which is directly connected to a small tank, to keep the water pressure relatively stable in the house. To fill the small tank, the pump takes just a few seconds (less than a minute). Consequently, the cone of depression generated by regular pumping in this system should be very limited in the vicinity of exploited domestic wells.

3.5 Methodology

The initial step of the methodology presented in Figure 3.2 consisted of dividing the borehole database into two data sets: data set A containing the borehole time series with outliers (in this case, time series from private boreholes with pumping effects) and data set B containing borehole time series without outliers (time series from observation boreholes). In the second step, the outliers are then identified and removed from data set A by applying an outlier identification method. In this study, because outliers stemmed from pumping effects that rapidly reduced the groundwater level in the borehole, it was chosen to apply the slope criterion provided in Equation 1 (Peterson et al. 2018) for outlier identification. Such slope criterion was proven to be efficient in detecting outliers due to rapid changes in groundwater level time series by Peterson et al. (2018).

$$Remove\ h_{i+1}\ from\ the\ time\ series\ if\ \alpha_i > \alpha_{max,}\ where\ \alpha_i = |h_{i+1} - h_i|/|t_{i+1} - t_i| \quad \text{(Eq 3.1)}$$

In Equation 3.1, α_i is the slope at the time step ti, αmax is the slope threshold, h_i and h_{i+1} refer to the groundwater level at the time step t_i and t_{i+1}, respectively. It should be noted that the threshold (α_{max}) generally differs among boreholes and that this is a trial-and-error process. The thresholds depend mainly on the hydrogeological properties of the aquifer, the pumping rate, and the pumping frequency. It should be noted that although the slope criterion was deemed appropriate in this case study for the detection of outliers in time series, it is clear from the methodology presented in Figure 3.2 that the practitioner must choose the method of outlier identification that is suitable for the data set collected.

The third step consisted of verifying, for each borehole, whether the outliers had been successfully removed from the borehole time series. If outliers remained, another threshold value (α_{max}) was set, and Step 2 was repeated until all outliers were identified and removed. The time series with the outliers removed formed data set C; however, removing the outliers produced several missing values in these time series.

Step 4 consisted of reproducing the missing value pattern of each borehole time series in data set C within the borehole time series lacking outliers (data set B), i.e., missing values were created in the data set B time series, according to the missing value patterns of the data set C time series. The process involved simply deleting the data points in the data set B time series that corresponded to the time when an outlier was identified in the time series of data set C. This process permits assessing the performance of the missing value imputation methods using time series for which the missing values were known. The time series having the generated missing values was labeled data set D. It should be noted that there is a serial correlation between the groundwater level time series recorded from the exploited and unexploited boreholes in the monitoring network, which allows the unexploited boreholes to be used to replicate the pattern of missing values from the exploited boreholes and to calculate the performance parameters of the imputation methods. The correlation matrix between the exploited and unexploited boreholes time seris were provided in Appendix 7.

The fifth step consisted of selecting and applying missing value imputation methods to each time series in data set D. It should be noted that at this stage of the methodology, the practitioner will select the imputation methods deemed appropriate. In this study, 14 univariate imputation methods were chosen and applied to each time series having missing values in data set D. These imputation methods, summarized in Table 3.1, were implemented using the R package *imputeTS* (Moritz & Bartz-Beielstein, 2017). For detailed mathematical and computational descriptions of the imputation methods, the reader is referred to Moritz and Bartz-Beielstein (2017), Demirhan & Renwick (2018), and Forsythe et al. (1977). Step 5 led to the creation of data set E, containing borehole time series with imputed values. The number of time series in data set E equaled the number of time series in data set D multiplied by the number of imputation methods.

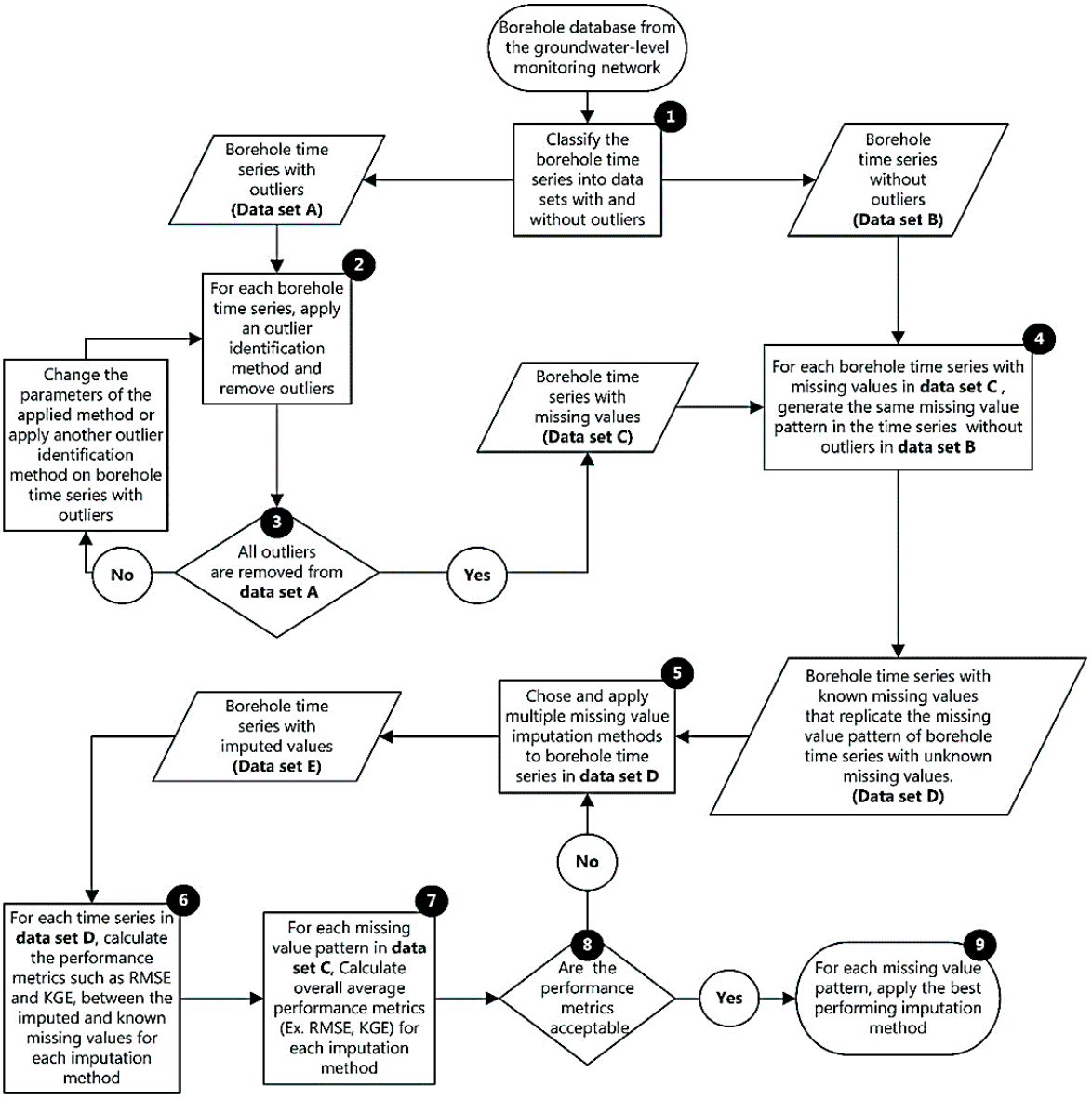


Figure 3.2. Flowchart detailing the identification, removal, and imputation of missing values in a monitoring network of groundwater level in boreholes. The numbers in black circles indicate the specific steps presented in the text.

Step 6 consisted of computing the performance metrics of each borehole time series in data set E for the imputed and the known missing values. The performance metrics used in this study were the root mean square error (RMSE), coefficient of determination (R^2), and Kling–Gupta efficiency (KGE). The mathematical expressions of RMSE (Chicco et al. 2021), R^2 (Chicco et al. 2021), and KGE (Knoben et al. 2019) are respectively provided in Equations 3.2–3.4. The perfect replication of the known missing values by the imputation method for a given borehole time series should result in a RMSE = 0, R^2 = 1, and KGE = 1. Conversely, the worst replication should yield a RMSE = +∞, R^2 = 0, and KGE = −∞.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y'_i)^2}, \qquad (Eq\ 3.2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y'_i - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \mu_y\right)^2}, \qquad (Eq\ 3.3)$$

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{y'}}{\sigma_y} - 1\right)^2 + \left(\frac{\mu_{y'}}{\mu_y} - 1\right)^2}, \qquad (Eq\ 3.4)$$

where n is the number of missing values, y′ is the imputed missing value, y is the known value of the missing value, y_i and y'_i are the ith value of y and y′ respectively, r is the correlation coefficient between y and y′, σ_y and $\sigma_{y'}$ are the standard deviation of y and y′, respectively, and μ_y and $\mu_{y'}$ are the mean value of y and y′, respectively.

The seventh step identifies the optimal imputation method for each missing value pattern. In this step, average performance metrics were calculated for each imputation method and missing value pattern. Step 8 involved assessing whether there was at least one imputation method that satisfactorily imputed the missing values of each missing value pattern. If the applied imputation method did not effectively impute the missing values of certain borehole time series with their corresponding missing value patterns, the process was repeated from Step 5. However, if the imputation methods were successful in imputing

the missing values of the borehole time series, Step 9 was performed, which consisted of applying the optimal imputation method to each time series having missing values.

Table 3.1. Description of the applied imputation methods and algorithms.

Methods	Algorithm	Description
Interpolation	Linear	The missing values are replaced by simple linear interpolation.
	Spline	The time series is subdivided into multiple intervals, called knots, and cubic polynomial functions are fitted to each interval. In each interval, missing values are determined by the corresponding polynomial function.
	Stineman	The method developed by Stineman (1980) is assumed to be more robust than the spline method in the case of abrupt changes in the slope of the time series. Unlike the spline method, the Stineman method uses a rational function for interpolation rather than a polynomial function.
Kalman filters	Auto ARIMA with smoothing	Kalman filtering, also known as linear quadratic estimation, is an algorithm that gives estimates of some unknown variables from observed measurements over time. The algorithm is recursive and used on the state-space representation of an autoregressive integrated moving average (ARIMA) model.
	Structural model with smoothing	Instead of the ARIMA model, the algorithm uses Kalman filtering on structural time-series models, which are linear Gaussian state-space models for univariate time-series.
Weighted moving average (MA)	Simple	For a given window size, missing values are replaced by the mean, and all observations in the window are weighted equally.
	Linear	For the linear option, the observations of the window receive weights that decrease in arithmetic progression (1/2, 1/3, 1/4...) as one moves away from the central value of the window.
	Exponential	For the exponential option, the observations of the window receive weights that decrease exponentially (1/2, $1/2^2$, $1/2^3$...) as one moves away from the central value of the window.
Seasonal decomposition	Interpolation	The time series are decomposed into a trend, seasonal, and residual components. Linear interpolation is then applied to impute missing values on the deseasoned series. The seasonality component is then reintroduced into the series.

	Kalman filter	For the Kalman filter option, the missing values are imputed by applying a Kalman filter with auto ARIMA rather than through linear interpolation.
	Moving average (MA)	The missing values are imputed for the MA option by applying an exponential weighted moving average algorithm.
Seasonal splitting	Interpolation	The time series is split into seasonal portions ; the linear interpolation imputation method is performed separately for each seasonal portion.
	Kalman filter	The missing values are imputed by applying a Kalman filter with auto ARIMA rather than a linear interpolation for the Kalman filter option.
	Moving average (MA)	The missing values are imputed for the MA option by applying an exponential weighted moving average algorithm.

3.6 Outlier identification and removal

As part of the first step of the presented methodology (Figure 3.2), the borehole database was divided into two data sets, one with outliers and one without. The monitoring network of 20 boreholes was split into data set A (with outliers; private boreholes D1, D2, D3, D4, D5, D6, D7, D8, D10, D13, D15, D16, D20, D24) and data set B (without outliers; unexploited boreholes D9, D11, D12, D17, D22, D23).

Outliers can represent a relatively important portion of a time series. The portion of outliers relates closely to the data type, recording conditions, and data handling. In the case of groundwater level time series in exploited boreholes, the data set was strongly influenced by frequent pumping, which generated repetitive and rapid groundwater level drops in the borehole (Figure 3.3). Outliers in borehole time series were identified using the slope criterion (α_{max}) defined by Equation 3.1. This slope criterion worked very well for borehole time series heavily affected by frequent pumping events, as illustrated in Figure 3.3, where 94.2% of the data from borehole D20 was identified as outliers. Peterson et al. (2018) applied such a slope criterion to identify outliers in borehole time series characterized by varying trends, sequential outliers, and periods of marked variability.

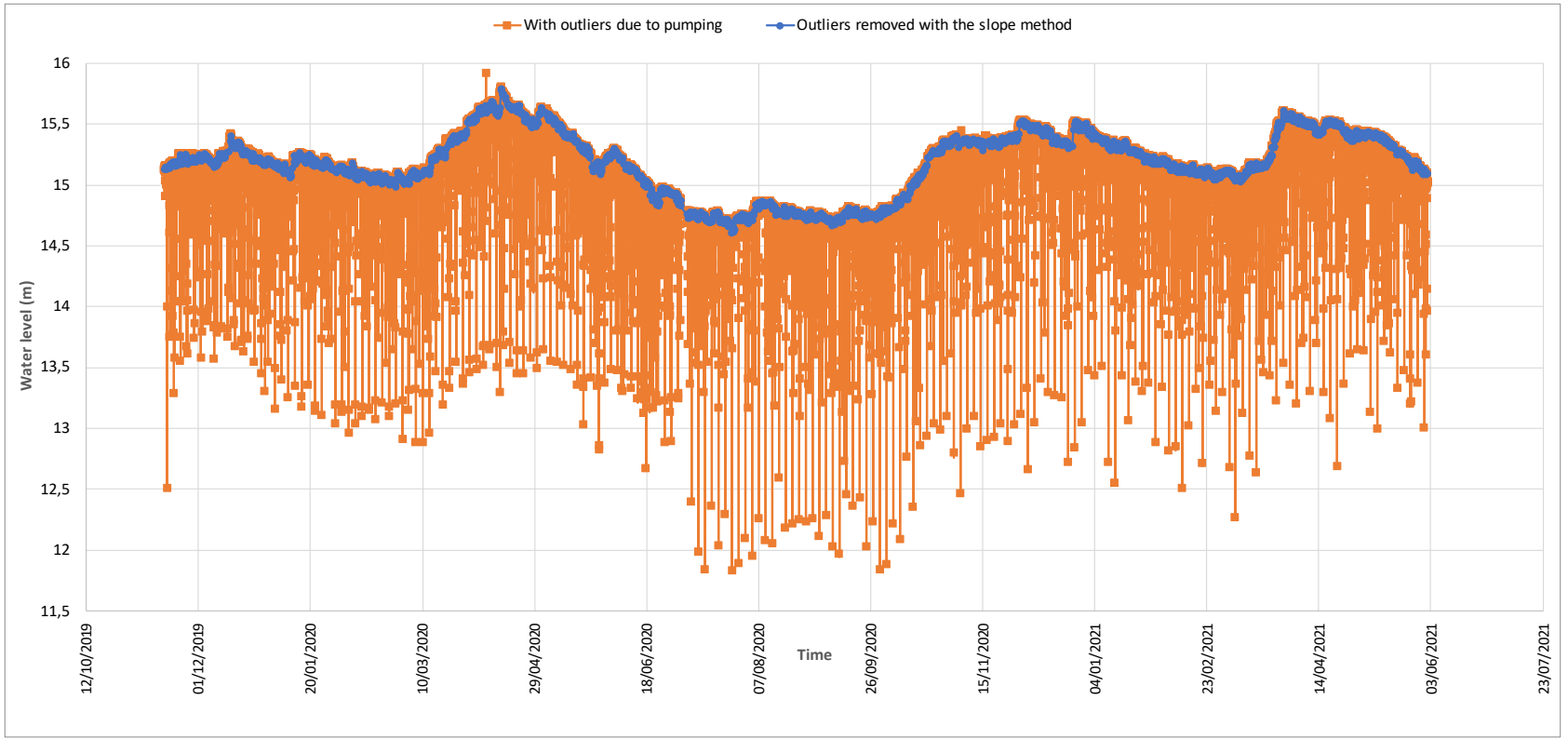


Figure 3.3. Groundwater-level time series from borehole D20 in which 94.2% of the data was identified as outlier. The graph presents the recorded groundwater-level time series of borehole D20 (orange data points) and the groundwater-level time series after outlier removal (blue data points).

The maximum slopes generally differed among boreholes (Table 3.2), and when the maximum slopes were similar, the percentage of identified outliers also tended to vary among boreholes. These differences in maximum slope and the outlier percentage are related to the local hydrogeological properties of the aquifers and the water demand of each household operating the borehole. For example, an intensely exploited private borehole in an aquifer characterized by low storage and transmissivity would produce many outliers in the recorded groundwater level, as the water table would take time to return to its initial state after each pumping. Moreover, if the borehole owner continued to pump water from the borehole while the water table had yet to return to its original state, the number of outliers would be massive. Peterson et al. (2018) used the same maximum slope of 10 $m \cdot day^{-1}$ for the four boreholes examined in their studies, but recognized the need to adapt this threshold to local conditions. However, the borehole data investigated by Peterson et al. (2018) were not strongly affected by rapid changes of groundwater level. In this study, the maximum slope ranged from 0.015 to 1 $m \cdot day^{-1}$, and the percentage of outliers identified in the borehole time series varied between 44.4% and 99.1%. As the effect of frequent pumping to generate outliers increased, the maximum slope αmax needed to identify these outliers in the data set decreased (Table 3.2). An advantage of the high recording frequency of 15 min is that even when 99.1% of the recorded data were identified as outliers and removed, the remaining number of non-outlier data remained greater than the quantity of data in the cases of one recording every two days without outliers.

Table 3.2. Borehole groundwater level time series highlighting the initial number of recorded data (N), the number of identified outliers (n), the percentage of outliers, and the maximum slope criterion αmax; the maximum slope is the slope threshold calculated at time step i, over which the ith groundwater level is considered to be an outlier. The most frequent gap size, the median gap size, the average gap size, and the standard deviation of the missing value gap sizes for each borehole.

	N	Maximum slope α_{max} ($m \cdot day^{-1}$)	Outliers		Missing values			
			n	%	Most frequent gap size	Median Gap size	Average Gap Size	Standard deviation
D1	62456	0.15	31716	50.8	1	2.1	4.0	8.9
D2	62468	0.04	60486	96.8	1	6.1	42.4	86.2
D3	62441	0.1	49915	79.9	1	5.2	11.3	41.6
D4	56410	0.1	34821	61.7	1	1	5.2	97.4
D5	58229	1	39537	67.9	14	32.5	46.8	214.7
D6	54216	0.015	53731	99.1	1	69	139.8	512.5
D7	31751	0.6	14107	44.4	5	7	35.7	911.8
D8	62454	0.4	39188	62.7	1	3.1	33.2	852.4
D10	62542	1	31944	51.1	1	17.9	28.7	75.6
D13	62437	1	36916	59.1	7	16.2	22.8	62.8
D15	54018	0.02	53524	99.1	1	52.3	147.2	611.3
D16	54104	0.1	42578	78.7	1	2.8	7.7	130.5
D20	54109	0.02	50972	94.2	1	10.8	21.3	202.7
D24	54116	0.07	49096	90.7	1	2.9	16.6	186.5

Removing identified outliers from borehole time series created missing values in the time series. The missing values were distributed along the time series with different gap sizes. This distribution of gap sizes characterized the missing value pattern of each borehole. It should be noted that for all of the exploited private boreholes the missing values were exponentially distributed as shown in Appendix 8 (Figures S1 through S14). Such exponential distribution of the missing values is reported to be common in real life data set by Moritz et al. (2015). For each exploited private borehole, summary statistics of missing values, in which common measures of central tendency, such as mode (most frequent deviation size), median and average gap sizes, but also standard deviation as a measure of dispersion were reported in Table 3.2. Although weak correlations were found between the maximum slope threshold and measures of central tendency and dispersion of missing values, it is worth mentioning that a correlation coefficient of 0.5 was obtained between maximum slope and mode. It can be seen from Figure 3.4 and Appendix 9 (Figure S15 through S26) that the percentage of missing values was variable over the monitoring period, even in the same borehole, due to the variation in water demand over time.

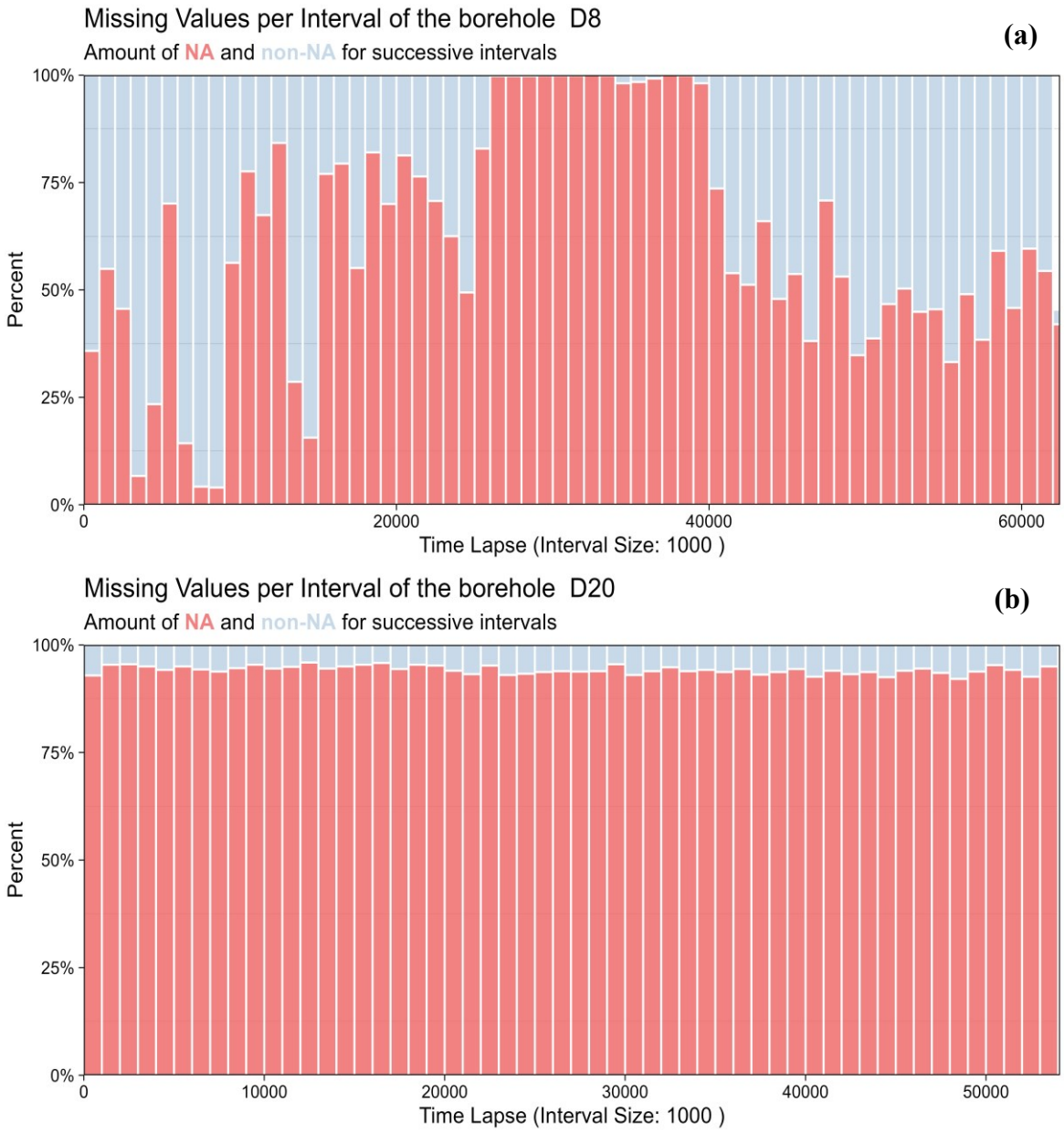


Figure 3.4. Distribution of the percentage of missing values by 1000 time step interval size in borehole time series D8 (a) and D20 (b). NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non missing.

3.7 Imputation of missing values

3.7.1 Performance of the imputation methods

Outliers can be ignored when the missing values are less than 10% of the total time series (Lodder 2013). However, the percentage of missing values in our borehole time series greatly surpassed 10% (Table 3.2); thus, imputation methods were required.

The selected imputation procedures (Section 3.6, Figure 3.2) were applied to data set D, and the performance metrics of each imputation method were calculated. For each of the 14 time series with missing values, the missing value pattern was replicated in the 6 borehole time series without missing values (data set B). Thus, data set D comprised 84 groundwater level time series (14 × 6) having missing data points for which the missing values were known. Because the missing values for all six time series are known (using data set B), the performance metrics could be calculated between the actual and imputed values for all 14 imputation methods. Moreover, as each imputation method was applied to six time series, six calculations for each performance metric were made, and the overall performance of each imputation method for each missing value pattern was averaged from the six borehole time series.

To illustrate, Table 3.3 shows the overall performance metrics of the 14 imputation methods based on the missing value pattern of the D20 borehole time series, i.e., the performance metrics in Table 3.3 represent the average performance metrics for each imputation method calculated from the replication of the missing value pattern from borehole D20 on borehole time series D9, D11, D12, D17, D22, and D23. Note that the overall performance metrics of the 14 imputation methods based on the missing value pattern of borehole time series D1, D2, D3, D4, D5, D6, D7, D8, D10, D13, D15, D16, D24 are provided in Appendix 10 Tables S3–S15. The performance metrics (R^2, RMSE, and KGE) were

calculated between the imputed values and the real observed values of borehole time series D9, D11, D12, D17, D22, and D23 for each imputation method and missing value pattern.

Some researchers (Schaefli and Gupta 2007, Knoben et al. 2019) have highlighted the need to provide benchmarks that indicate when a model performance (imputation method) is accepTable. However, some performance metrics, such as the RMSE, are challenging to interpret (Chicco et al. 2021); for example, a single RMSE value does not indicate how well a model performs. Nevertheless, using performance measures such as the RMSE can be very useful for comparing the performance of imputation methods, where the same missing value patterns are considered. For example, the best-performing imputation methods are those with RMSE values tending toward zero (Table 3.3, Appendix 10 Tables S3–S15). One can link performance metrics such as R^2 and KGE to specific meaning, e.g., for $R^2 = 0$ or KGE ≈ −0.41 (Knoben et al. 2019), the imputation method performs no better than replacing the missing values by the mean value of the time series. As replacing the missing values with the mean value of the time series is an existing imputation method, setting the benchmark at that level ($R^2 = 0$ or KGE ≈ −0.41) can be an attractive option; however, this imputation method does not perform very well (Demirhan and Renwick 2018). In the present study, the R^2 and KGE benchmarks were set at the upper two-thirds of the interval between a perfect imputation of the missing values, i.e., $R^2 = 1$ and KGE = 1, and imputation of the missing values by the mean value of the data set, i.e., $R^2 = 0$ and KGE = −0.41. Therefore, the benchmarks were $R^2 = 0.67$ and KGE = 0.53, over which the imputation method was considered good. Although previous studies did not clearly state their benchmarks, an R^2 greater than 0.4 has been considered a benchmark for a good imputation method (Chicco et al., 2021). In terms of KGE, Knoben et al. (2019) reported that many researchers considered a KGE = 0 as a benchmark, but Prada (2020) used a KGE greater than 0.5.

Table 3.3. Overall average performance metrics for each imputation method on the basis of the missing value pattern of borehole time series D20.

	D20		
	R^2	RMSE	KGE
Linear interpolation	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Spline interpolation	0,987(SD 0,01)	0,035(SD 0,02)	0,993(SD 0,01)
Stineman interpolation	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Kalman filtering with a structural model and smoothing	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Kalman filtering with ARIMA and smoothing	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Simple moving average with *k* = 15	0,995(SD 0,00)	0,022(SD 0,01)	0,997(SD 0,00)
Linear weighted moving average with *k* = 15	0,996(SD 0,00)	0,019(SD 0,01)	0,998(SD 0,00)
Exponential weighted moving average with *k* = 15	0,996(SD 0,00)	0,020(SD 0,01)	0,997(SD 0,00)
Simple moving average with *k* = 5	0,995(SD 0,00)	0,021(SD 0,01)	0,997(SD 0,00)
Linear weighted moving average with *k* = 5	0,996(SD 0,00)	0,019(SD 0,01)	0,998(SD 0,00)
Exponential weighted moving average with *k* = 5	0,996(SD 0,00)	0,020(SD 0,01)	0,997(SD 0,00)
Seasonal decomposition with linear interpolation	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,996(SD 0,00)	0,020(SD 0,01)	0,997(SD 0,00)
Seasonal splitting with linear interpolation	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,996(SD 0,00)	0,020(SD 0,01)	0,997(SD 0,00)

The observed performance metrics values (Table 3,.3. Appendix 10 Tables S3–S15) were generally higher than the benchmarks (R^2 = 0.67 and KGE = 0.53), indicating that most of the imputation methods performed well in imputing the missing values. Comparing the performance metrics of each imputation method across the different missing value patterns of the time series demonstrated that the performance of the imputation method did not necessarily depend on the percentage of missing values in the time series. For instance, the imputation methods performed better in imputing the missing values according to D20 missing value pattern (Table 3.3) having 94% missing values than in D10 and D8 missing value patterns (Appendix 10 Table S10 and S11) having 51% and 62% missing values, respectively. It should also be noted that the opposite situation occurred, whereby for a given

imputation method the imputation of missing values performed better on time series with missing value patterns having lower percentage of missing values than on time series with missing value patterns characterized by a greater number of missing values. The reasons for these apparent discrepancies are related to the missing value patterns in the borehole data, in particular, gap size distribution. The problem of missing value patterns can be easily observed in Table 3.2 and Figure 3.4, where the missing value patterns of the time series recorded in borehole D8 and D20 are compared. The percentage of missing values is highly variable in the D8 time series (Figure 3.4a), with high missing value periods that can reach 99%. On the other hand, the percentage of missing values in the D20 time series (Figure 3.4b) tended to be relatively constant around 94%. In addition, the average and standard deviation of the gap sizes in the D8 missing value pattern were larger than those of the D20 missing value pattern, with a standard deviation difference of more than four times (Table 3.2). Therefore, imputation of missing values according to the missing value model for the D8 time series did not perform as well as that for the D20 series. Similar results from Dwivedi et al. (2022), in which several missing value patterns with different gap sizes and distributions were generated in borehole time series, showed that gap size affected the performance of imputation methods to a greater extent than gap distribution. Dwivedi et al. (2022) found a better performance for an imputation method obtained from a time series with 90% randomly distributed gaps than that obtained from a time series having a gap size representing 10% of the time series. The position of gaps in the time series also influences the imputation method's performance (Demirhan and Renwick 2018, Dwivedi et al. 2022). For subhour data, Demirhan & Renwick (2018) found no difference in the performance of the imputation methods in relation to the position of randomly distributed gaps, whether at the beginning, middle, or end of the time series. In contrast, Dwivedi et al. (2022) found that the position of the gap (large gaps of 10% or more) affected the performance of the imputation methods, i.e., whether the gaps were located on the highs and lows of the time

series or away from them. For borehole time series having missing values generated by frequent pumping, such as those presented in this study, imputation method performance is affected by a combination of all of the abovementioned factors, i.e., gap distribution, gap size, and the position of the gap in the time series with respect to the high and low values. Thus, when the imputation methods performed better on borehole data having a high percentage of missing values than a time series with a lower percentage of missing values, the missing value pattern in the borehole data having a high percentage contained few or no large size gaps, and the position of those large gaps was away from the highs and lows, whereas the missing value pattern in the borehole time series with a lower percentage contained large gaps that could be positioned at the highs or lows of the time series. This is well illustrated in Figure 3.5, where the imputation of missing values in a borehole time series replicating the missing value pattern of borehole D20 (Figure 3.5a), with 94.2% missing data, outperformed the imputation of missing values in the same borehole time series replicating the missing value pattern of borehole D8 (Figure 3.5b), with only 62.7% missing data. The large gaps in the missing value pattern of borehole D8 (Figure 3.5b) affected the performance of the imputation method, particularly in the highs and lows of the time series, which in turn affected the performance measures of the imputation method.

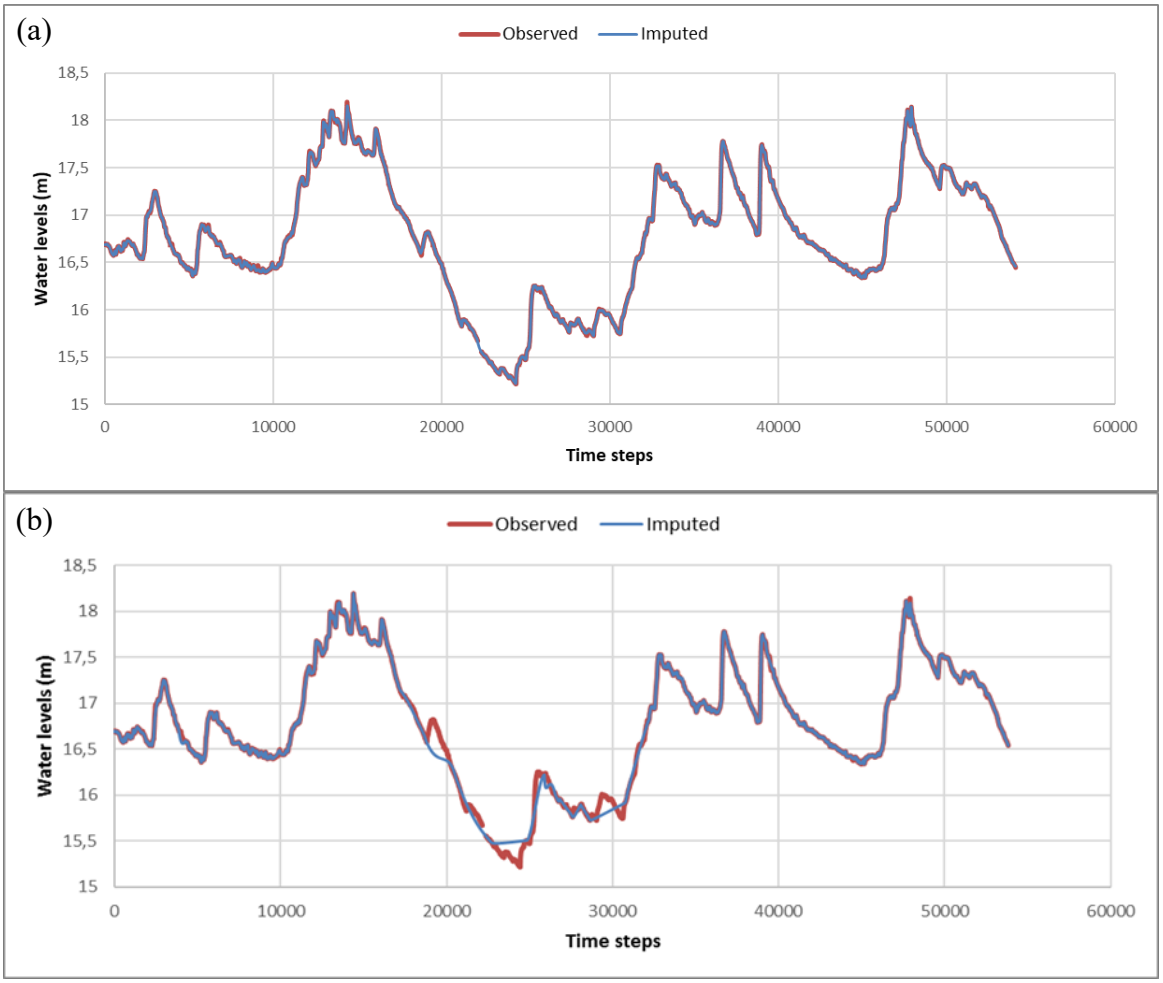


Figure 3.5. An illustration of the effect of gap size on the performance of imputation methods. (a) Imputation of missing values in borehole D11 time series using the missing value pattern of D20, having 94.2% missing data with an R2 = 0.997; (b) imputation of missing values in borehole D11 time series using the missing value pattern of D8, having 62.7% missing data with an R2 = 0.987.

3.7.2 Selecting the optimal imputation method

The selection of the optimal imputation method for a given missing value pattern was guided by the previously defined performance metrics. The ranking of the three best-performing imputation methods for a given missing value pattern was straightforward when all performance metrics provided the same ranking. In the case of conflicting rankings between performance metrics, however, we considered a ranking confirmed by two of the three performance metrics. A ranking summary of the top three performing imputation

methods for each borehole with a missing value pattern is presented in Table 3.4. This ranking is derived from the results of Table 3.3 and Appendix 10 Tables S3-S15, in which the average overall performance parameters for each imputation method are provided.

The five most frequent imputation methods listed in the ranking of the three best-performing imputation methods per borehole missing value pattern (Table 3.4) were linear interpolation, followed by Stineman interpolation, Kalman filtering with a structural model, Kalman filtering with ARIMA, and a linear weighted moving average, each occurring 14, 9, 9, 6, and 4 times, respectively. However, considering only the first class of best-performing imputation methods presented in Table 3.4, linear interpolation and Stineman interpolation each appeared four times, whereas Kalman filtering with ARIMA and linear weighted moving average each appeared three times. Comparable results were obtained by Demirhan & Renwick (2018) for subhour solar irradiance time series having missing values, in which the linear and Stineman interpolations were the optimal imputation methods, along with their seasonal splitting/decomposition versions. Dwivedi et al. (2022) found linear interpolation to be a well-performing imputation method for borehole time series having missing values. It should be noted that the performance of the seasonal versions of the imputation methods (Table 3.4) was equal to that of the non-seasonal versions, as no seasonality was found in the borehole time series. However, for borehole data exhibiting seasonality, the use of the seasonal versions of the imputation methods tends to yield better results (Noori and Singh 2021, Dwivedi et al. 2022). As introduced in the methodology section (Step 9), the optimal imputation for each missing value pattern was always applied.

Table 3.4. Borehole groundwater-level time series having missing value patterns and the three best imputation methods for each time series. Lin-Int refers to linear interpolation, Stin-Int refers to Stineman interpolation, Kal-Str refers to Kalman filtering with a structural model, Kal-ARIMA refers to Kalman filtering with an ARIMA model, and MA-Lin refers to linear weighted moving average. The summary is derived from the results in Table 3.3 and Appendix 10, in which the overall average performance metrics for each imputation method are provided.

Borehole with Missing value pattern	Missing values (%)	Optimal imputation methods		
		1st	2nd	3rd
D1	50.78	Lin-Int	Kal-Str	Stin-Int
D2	96.83	Stin-Int	Lin-Int	Kal-Str
D3	79.94	Lin-Int	Stin-Int	Kal-Str
D4	61.73	Lin-Int	Stin-Int	Kal-ARIMA
D5	67.90	Stin-Int	Lin-Int	Kal-ARIMA
D6	99.11	MA-Lin	Kal-Str	Lin-Int
D7	44.43	Kal-ARIMA	Kal-Str	Lin-Int
D8	62.75	MA-Lin	Lin-Int	Kal-ARIMA
D10	51.08	MA-Lin	Stin-Int	Lin-Int
D13	59.13	Kal-ARIMA	Lin-Int	Kal-Str
D15	99.09	Stin-Int	Lin-Int	Kal-Str
D16	78.70	Lin-Int	Stin-Int	Kal-Str
D20	94.20	Kal-ARIMA	Lin-Int	Kal-Str
D24	90.72	Stin-Int	Lin-Int	MA-Lin

3.8 Conclusions

The use of private boreholes in groundwater level monitoring networks offers a very advantageous and cost-effective approach; however, abrupt pumping-related changes in groundwater levels within these boreholes generate outliers in the recorded time series, causing these boreholes to be ignored for use in groundwater level monitoring. This study proposed a step-by-step approach to identify and remove outliers from borehole groundwater level time series and impute missing values, thereby rendering exploited private boreholes suitable for groundwater level monitoring.

The proposed methodology was applied to a groundwater monitoring network within a 465 km^2 region of the Lanaudière region of Québec (Canada) underlain by crystalline rocks. The network comprised 20 boreholes, 14 exploited private boreholes and 6 unexploited boreholes, with data points recorded at 15 min time steps. For the 14 exploited borehole time series, a slope criterion—threshold of the maximum slope of the groundwater level time series—was applied to identify and remove any outliers from the data. Although the slope criterion effectively identified outliers, removing the outliers created missing values in the time series, representing as much as 99% of the data for the most affected boreholes. The missing values could be described in terms of missing value pattern, which reflects the combination of the gap distribution and gap size, i.e., the number of consecutive missing values. The missing value patterns differed among the exploited private boreholes because of locally varying hydrogeological properties of the aquifer (e.g., storage, transmissivity) and household water usage.

Fourteen imputation methods were applied to each time series having missing values. The performance of the imputation methods was evaluated by replicating each missing value pattern in the time series of the six unexploited private boreholes for which the generated missing values were known. The missing values were satisfactorily imputed, thereby allowing the reconstruction of the groundwater level time series of the exploited private boreholes. Gap size was the factor having the greatest influence on the performance of the imputation methods; the larger the gap size, the worse the performance of the imputation method. Overall, among the 14 imputation methods considered, linear interpolation, Stineman interpolation, and Kalman filtering with a structural model performed best for imputing missing values.

Using a combination of exploited and unexploited boreholes in a monitoring network optimizes implementation costs and allows for an increase in the number of boreholes in the monitoring networks. This study shows that to implement such a monitoring network, the

practitioner is advised to check the missing value gap size distribution after a monitoring period (e.g. of one year), and replace boreholes that are highly affected by large gaps sizes in the monitoring network. The use of data loggers able to record the groundwater levels at high frequency (every 15 min or less) is advisable, since a lot of data will be detected as outliers. The proposed methodology can be adapted by the practitioner to different types of groundwater level time series recorded in different conditions by selecting alternative outlier identification and imputation methods. Also, the proposed methodology uses groundwater level time series of unexploited boreholes to evaluate the performance of different imputation methods for each missing value pattern. Therefore, appropriate modifications of the methodology are essential to be applied to a groundwater monitoring network consisting solely of exploited private boreholes. Future advances in outlier identification methods and imputation methods can be expected to help expand the use of private boreholes in monitoring networks, as these advances will make the proposed methodology, or its subsequent modifications, more effective.

3.9 References

Abi, A., Walter, J., Saeidi, A., and Chesnaux, R. 2022. A cluster-based multiparametric similarity test for the compartmentalization of crystalline rocks into structural domains. Quarterly Journal of Engineering Geology and Hydrogeology, **55**: qjegh2021-136. doi:10.1144/qjegh2021-136.

Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M., and Fisher, A.J. 2020. Missing data imputation of high-resolution temporal climate time series data. Meteorological Applications, **27**: e1873. doi:10.1002/met.1873.

Andricevic, R. 1990. Cost-effective network design for groundwater flow monitoring. Stochastic Hydrology and Hydraulics, **4**: 27–41. doi:10.1007/BF01547730.

Asgharinia, S., and Petroselli, A. 2020. A comparison of statistical methods for evaluating missing data of monitoring wells in the Kazeroun Plain, Fars Province, Iran. Groundwater for Sustainable Development, **10**: 100294. doi:10.1016/j.gsd.2019.100294.

Béland, R. 1967. Région de Saint-Gabriel-de-Brandon, Comtes de Joliette, Berthier et Maskinongé. Geological reports, Ministère des Richesses Naturelles du Québec, Québec.

Blázquez-García, A., Conde, A., Mori, U., and Lozano, J.A. 2022. A Review on Outlier/Anomaly Detection in Time Series Data. ACM Computing Surveys, **54**: 1–33. doi:10.1145/3444690.

Buuren, S. van. 2012. Flexible Imputation of Missing Data. Chapman and Hall/CRC, New York. doi:10.1201/b11826.

CERM-PACES. 2022. Résultats du programme d'acquisition de connaissances sur les eaux souterraines du territoire municipalisé de Lanaudière. Centre d'études sur les ressources minérales, Université du Québec à Chicoutimi, Quebec, Canada.

Chen, D., Lu, C.-T., Kou, Y., and Chen, F. 2008. On Detecting Spatial Outliers. GeoInformatica, **12**: 455–475. doi:10.1007/s10707-007-0038-8.

Chicco, D., Warrens, M.J., and Jurman, G. 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, **7**: e623. PeerJ Inc. doi:10.7717/peerj-cs.623.

Clark, T.H., and Globensky, Y. 1976. Région de Sorel et partie sud-est de Saint-Gabriel-de-Brandon. Geological reports, Ministère des Richesses Naturelles du Québec, Québec.

Dax, A., and Zilberbrand, M. 2018. Imputing missing groundwater observations. Hydrology Research, **49**: 831–845. doi:10.2166/nh.2017.220.

Demirhan, H., and Renwick, Z. 2018. Missing value imputation for short to mid-term horizontal solar irradiance data. Applied Energy, **225**: 998–1012. doi:10.1016/j.apenergy.2018.05.054.

Dwivedi, D., Mital, U., Faybishenko, B., Dafflon, B., Varadharajan, C., Agarwal, D., Williams, K.H., Steefel, C.I., and Hubbard, S.S. 2022. Imputation of contiguous gaps and extremes of subhourly groundwater time series using random forests. Journal of Machine Learning for Modeling and Computing, **3**. Begel House Inc. doi:10.1615/JMachLearnModelComput.2021038774.

Evans, S., Williams, G.P., Jones, N.L., Ames, D.P., and Nelson, E.J. 2020. Exploiting Earth Observation Data to Impute Groundwater Level Measurements with an Extreme Learning Machine. Remote Sensing, **12**: 2044. Multidisciplinary Digital Publishing Institute. doi:10.3390/rs12122044.

Farhangfar, A., Kurgan, L.A., and Pedrycz, W. 2007. A Novel Framework for Imputation of Missing Values in Databases. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, **37**: 692–709. doi:10.1109/TSMCA.2007.902631.

Fornés, J.M., la Hera, Á., and Llamas, M.R. 2005. The silent revolution in groundwater intensive use and its influence in Spain. Water Policy, **7**: 253–268. doi:10.2166/wp.2005.0016.

Forsythe, G.E., Malcolm, M.A., and Moler, C.B. 1977. Computer Methods for Mathematical Computations. Prentice Hall Professional Technical Reference.

Fox, A.J. 1972. Outliers in Time Series. Journal of the Royal Statistical Society. Series B (Methodological), **34**: 350–363. [Royal Statistical Society, Wiley].

Gill, M.K., Asefa, T., Kaheil, Y., and McKee, M. 2007. Effect of missing data on performance of learning algorithms for hydrologic predictions: Implications to an imputation technique. Water Resources Research, **43**. doi:10.1029/2006WR005298.

Gleeson, T., Novakowski, K., and Kurt Kyser, T. 2009. Extremely rapid and localized recharge to a fractured rock aquifer. Journal of Hydrology, **376**: 496–509. doi:10.1016/j.jhydrol.2009.07.056.

van der Heijden, G.J.M.G., T. Donders, A.R., Stijnen, T., and Moons, K.G.M. 2006. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. Journal of Clinical Epidemiology, **59**: 1102–1109. doi:10.1016/j.jclinepi.2006.01.015.

Hosseini, M., and Kerachian, R. 2017. A data fusion-based methodology for optimal redesign of groundwater monitoring networks. Journal of Hydrology, **552**: 267–282. doi:10.1016/j.jhydrol.2017.06.046.

Hron, K., Templ, M., and Filzmoser, P. 2010. Imputation of missing values for compositional data using classical and robust methods. Computational Statistics & Data Analysis, **54**: 3095–3107. doi:10.1016/j.csda.2009.11.023.

Jørgensen, L.F., and Stockmarr, J. 2009. Groundwater monitoring in Denmark: characteristics, perspectives and comparison with other countries. Hydrogeology Journal, **17**: 827–842. doi:10.1007/s10040-008-0398-7.

Knoben, W.J.M., Freer, J.E., and Woods, R.A. 2019. Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrology and Earth System Sciences, **23**: 4323–4331. Copernicus GmbH. doi:10.5194/hess-23-4323-2019.

Li, L., Wen, Z., and Wang, Z. 2016. Outlier Detection and Correction During the Process of Groundwater Lever Monitoring Base on Pauta Criterion with Self-learning and Smooth Processing. *In* Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems. *Edited by* L. Zhang, X. Song, and Y. Wu. Springer, Singapore. pp. 497–503. doi:10.1007/978-981-10-2663-8_51.

Ling, M., S. Rifai, H., J. Newell, C., J. Aziz, J., and R. Gonzales, J. 2003. Groundwater monitoring plans at small-scale sites—an innovative spatial and temporal methodology. Journal of Environmental Monitoring, **5**: 126–134. Royal Society of Chemistry. doi:10.1039/B207682A.

Little, R.J.A., and Rubin, D.B. 2019. Statistical Analysis with Missing Data. John Wiley & Sons.

Lodder, P. 2013. To impute or not impute: That's the question. Advising on research methods: Selected topics,: 1–7. Johannes van Kessel Publishing.

Moritz, S., and Bartz-Beielstein, T. 2017. imputeTS: Time Series Missing Value Imputation in R. The R Journal, **9**: 207. doi:10.32614/RJ-2017-009.

Moritz, S., Sardá, A., Bartz-Beielstein, T., Zaefferer, M., and Stork, J. 2015. Comparison of different Methods for Univariate Time Series Imputation in R. arXiv:1510.03924 [cs, stat],. Available from http://arxiv.org/abs/1510.03924 [accessed 19 April 2022].

Noori, A.R., and Singh, S.K. 2021. Spatial and temporal trend analysis of groundwater levels and regional groundwater drought assessment of Kabul, Afghanistan. Environmental Earth Sciences, **80**: 698. doi:10.1007/s12665-021-10005-0.

Peterson, T.J., Western, A.W., and Cheng, X. 2018. The good, the bad and the outliers: automated detection of errors and outliers from groundwater hydrographs. Hydrogeology Journal, **26**: 371–380. doi:10.1007/s10040-017-1660-7.

Prada, M.C.R. 2020. Operational Flood Forecasting, Warning and Response for Multi-Scale Flood Risks in Developing Cities. CRC Press. doi:10.1201/9780138745011.

Raaijmakers, Q.A.W. 1999. Effectiveness of Different Missing Data Treatments in Surveys with Likert-Type Data: Introducing the Relative Mean Substitution Approach. Educational and Psychological Measurement, **59**: 725–748. doi:10.1177/0013164499595001.

Raposo, J.R., Molinero, J., and Dafonte, J. 2012. Parameterization and quantification of recharge in crystalline fractured bedrocks in Galicia-Costa (NW Spain). Hydrology and Earth System Sciences, **16**: 1667–1683. doi:https://doi.org/10.5194/hess-16-1667-2012.

Sakizadeh, M., Mohamed, M.M.A., and Klammler, H. 2019. Trend Analysis and Spatial Prediction of Groundwater Levels Using Time Series Forecasting and a Novel Spatio-Temporal Method. Water Resources Management, **33**: 1425–1437. doi:10.1007/s11269-019-02208-9.

Schaefli, B., and Gupta, H.V. 2007. Do Nash values have value? Hydrological Processes, **21**: 2075–2080. doi:10.1002/hyp.6825.

Schneider, T. 2001. Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values. Journal of Climate, **14**: 853–871. American Meteorological Society. doi:10.1175/1520-0442(2001)014<0853:AOICDE>2.0.CO;2.

Solinst Canada Ltd. 2021, October 19. Dataloggers & Telemetry Systems. Solinst Canada Ltd. Available from https://www.solinst.com/products/data/brochure-dataloggers-telemetry.pdf [accessed 22 March 2022].

Song, S., Zhang, A., Wang, J., and Yu, P.S. 2015. SCREEN: Stream Data Cleaning under Speed Constraints. *In* Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. ACM, Melbourne Victoria Australia. pp. 827–841. doi:10.1145/2723372.2723730.

Stineman, R.W. 1980. A consistently well-behaved method of interpolation. Creative Computing, **6**: 54–57.

Thorslund, J., and van Vliet, M.T.H. 2020. A global dataset of surface water and groundwater salinity measurements from 1980–2019. Scientific Data, **7**: 231. doi:10.1038/s41597-020-0562-z.

Tremblay, Y., Lemieux, J.-M., Fortier, R., Molson, J., Therrien, R., Therrien, P., Comeau, G., and Talbot Poulin, M.-C. 2015. Semi-automated filtering of data outliers to improve spatial analysis of piezometric data. Hydrogeology Journal, **23**: 851–868. doi:10.1007/s10040-015-1257-y.

Zhang, A., Song, S., and Wang, J. 2016. Sequential Data Cleaning: A Statistical Approach. *In* Proceedings of the 2016 International Conference on Management of Data. ACM, San Francisco California USA. pp. 909–924. doi:10.1145/2882903.2915233.

Zhou, Y., Dong, D., Liu, J., and Li, W. 2013. Upgrading a regional groundwater level monitoring network for Beijing Plain, China. Geoscience Frontiers, **4**: 127–138. doi:10.1016/j.gsf.2012.03.008.

CHAPTER 4

INFERRING FACTORS CONTROLLING THE RESPONSE TIME AND VARIABILITY OF GROUNDWATER LEVELS IN REGIONAL CRYSTALLINE AQUIFERS: A STATISTICAL DATA DRIVEN ANALYSIS

Attoumane Abi, Julien Walter, Romain Chesnaux and Ali Saeidi

Department of Applied Sciences, Université du Québec à Chicoutimi, 555 Boulevard de l'Université, Chicoutimi, QC G7H 2B1, Canada

Submitted to hydrogeology journal (2022)

4.1 Mise en contexte

Ce chapitre traite des facteurs contrôlant le temps de réponse et la variabilité du niveau piézométrique dans les aquifères de roche cristalline. À l'échelle locale et régionale, le niveau des eaux souterraines est contrôlé par plusieurs facteurs. Les facteurs les plus couramment étudiés sont les contrôles climatiques, géologiques et géomorphologiques sur la variabilité du niveau des eaux souterraines et le temps de réponse, et dans de nombreux cas, un seul facteur de contrôle est pris en compte dans l'analyse. Cependant, de nombreux autres facteurs peuvent affecter la variabilité du niveau des eaux souterraines et le temps de réponse, tels que les propriétés des dépôts de sédiments et les caractéristiques du réseau de fractures dans les aquifères cristallins. Dans cette étude, une approche plus globale est utilisée pour considérer les paramètres climatiques, géomorphologiques et du réseau de fractures comme des facteurs de contrôle potentiels. Au total, 18 paramètres ont été analysés pour déterminer les interrelations, sachant que chaque facteur de contrôle est décrit par plusieurs paramètres. L'étude a analysé un jeu de données de deux ans sur les niveaux d'eau souterraine dans 20 forages réalisés dans la roche cristalline du Bouclier canadien dans la région de Lanaudière, au Québec, Canada Les facteurs liés à la géomorphologie et au réseau de fractures sont liés à la variabilité du niveau des eaux souterraines et à son temps de réponse. Parmi les divers paramètres analysés dans chaque facteur de contrôle, l'épaisseur des sédiments et la pente locale du facteur géomorphologique, ainsi que la persistance moyenne et la conductivité hydraulique équivalente du facteur réseau de fractures, sont les plus étroitement liés à la variabilité du niveau des eaux souterraines et à son temps de réponse. Toutefois, des études complémentaires sont nécessaires pour élucider les processus physiques à l'origine de certaines interrelations entre les paramètres du réseau de fractures et les paramètres de variabilité du niveau des eaux souterraines

4.2 introduction

One of the most commonly used methods for determining recharge is the well-known water table fluctuation method, given its simplicity and availability of groundwater level data (Crosbie et al. 2005, Jie et al. 2011, Delottier et al. 2018). Groundwater levels variability analysis can be related to recharge, since its increase is generally due to recharge (Dewandel et al. 2010, Mizan et al. 2019). However, analysis of groundwater level variability provides much more information than recharge because groundwater level variability inform on the storage variation, which is very valuable for sustainable groundwater management (de Vries and Simmers 2002, Dickinson et al. 2004, Kotchoni et al. 2019). In this regard, several researchers have studied the factors controlling groundwater level variability and response time of different aquifers under a variety of hydrological and geomorphological conditions (Padilla et al. 2014, Rinderer et al. 2016, Lorenzo-Lacruz et al. 2017, Kotchoni et al. 2019, Boumaiza et al. 2021, Dong et al. 2022). From those studies, it has been shown that factors influencing the groundwater level and its response time are sediment thickness and sediment hydraulic properties (Gleeson et al. 2009, Miles and Novakowski 2016, Wright and Novakowski 2019), precipitation patterns (Jocson et al. 2002, Dickinson et al. 2004, Lorenzo-Lacruz et al. 2017), flood events (Saveca et al. 2022, Dong et al. 2022), preferential flow in the subsurface (Padilla et al. 2014, Nicolas et al. 2019), and geomorphological characteristics such as the slope and other topographic indexes (Forster and Smith 1988, Detty and K. J. McGuire 2010, Rinderer et al. 2016). Given the different types of aquifers, the varieties of hydrological conditions and geomorphological settings, factors influencing the groundwater level and its response are far from being well understood, especially for crystalline aquifers (Salve et al. 2012, Gabrielli et al. 2012, Cai and Ofterdinger 2016).

Previous works on crystalline aquifers have focused on studying a single factor influencing groundwater level and its response time (Salve et al. 2012, Padilla et al. 2014, Cai and Ofterdinger 2016) and a few others have combined the analysis of two factors such as topographic and rainfall controls on the variability of groundwater levels and its response time (Forster and Smith 1988, Rinderer et al. 2014, 2017). Furthermore, the majority of studies in crystalline aquifers ignore the effect of the fracture characteristics, while the fracture network may play an important role in the form of preferential flow that can trigger a rapid response of the groundwater level to snowmelt (Gleeson et al. 2009). In fact, Gleeson et al. (2009) have demonstrated that rapid response time of the groundwater level in crystalline aquifers is triggered by highly conductive fractures. However, Wright and Novakowski (2019) reported that the rapid response time is marginal in comparison to the relatively slow response time of the groundwater level that is attributed to regional flow. Wright and Novakowski (2019) have further emphasized on the need for more regional studies on the understanding of groundwater level behavior in crystalline aquifers.

The purpose of this study is to investigate the factors controlling the response time and variability of groundwater levels in regional crystalline aquifers, while considering multiple factors such as the geomorphology, sediment deposits, fracture network characteristics, and rainfall and snowmelt patterns. It employed time series analysis for the examination of the response time of the aquifer to rainfall and snowmelt, and the use of spearman correlation analysis for the analysis of the interrelationship between the different parameters describing the geomorphology, the sediment deposits, fracture network characteristics, the response time and variability of groundwater levels. Eighteen parameters were analyzed for finding factors controlling the response time and variability of groundwater levels in regional crystalline aquifers. The study benefitted from a data set of a monitoring network composed of 20 boreholes drilled in the crystalline rock and instrumented between summer 2019 to summer 2021. The study was conducted on a 465

km^2 uplifted area, with relatively thin sediments deposits, underlain by crystalline rocks of the Canadian Shield, located in the Lanaudière region, Quebec, Canada.

4.3 Study area

This study was conducted under the umbrella of the PACES-Lanaudière project (Programme d'acquisition de connaissances sur les eaux souterraines), which aimed to provide a quantitative and qualitative assessment of groundwater resources across the Lanaudière region; establish the state of this resource, the nature of its recharge, and the vulnerability of the regional aquifers; develop partnerships among stakeholders and managers of water resources within the territory to promote the sound management of the resource; and respond to the concerns of the community with respect to the groundwater resource (CERM-PACES 2022). The present study took place in the Lanaudière region see Abi et al. (2022) for a detailed portrait of the region). The study area is an uplifted area delimited to the north by the Maskinongé River, to the southwest by the Bayonne River, to the northwest by the Sainte-Julienne Fault, and to the southwest by the Saint-Cuthbert Fault. The hilly landscapes of the study area vary from 16 m to 301 m asl, with slopes ranging from 0° to 55°. The area has been subjected to glacial erosion that has shaped the landforms and left glacial striations oriented south and southwest on the outcropping rocks (Béland 1967, CERM-PACES 2022).

Most surficial sediments are of glacial origin deposited during the Late Wisconsinan deglaciation, which occurred over 10,000 years ago (Parent and Occhietti 1999). Sediment thickness is variable across the study area (0–100 m thick), with most sediment consisting of tills varying in their composition of gravel, silt, sand, and boulders (Fig. 4.1). Tills comprise about 41% of the surficial sediments, followed by glaciomarine sediments (about 40%) and other Quaternary sediments (19%). The sediment layer is commonly used for sand mining and serves as a granular aquifer for water consumption (CERM-PACES 2022). Some

municipalities, such as Saint-Barthélemy, have several boreholes to exploit this granular aquifer for its water supply.

Another important geological formation exploited as an aquifer is the underlying crystalline rock. In the study region, this rock includes the Grenville Province and the St. Lawrence platform. The Grenville Province occupies about 86% of the study area (Fig. 4.1) and forms part of the southeastern region of the Canadian Shield. All field investigations occurred in the Grenville Province portion of the study area. The Canadian Shield is mainly composed of Precambrian metamorphic and igneous rocks (Rivers et al. 1989). In the study area specifically, St. Didace granite, charnockitic granulite, and hornblende-biotite-plagioclase gneiss are the most common rock types (Béland 1967).

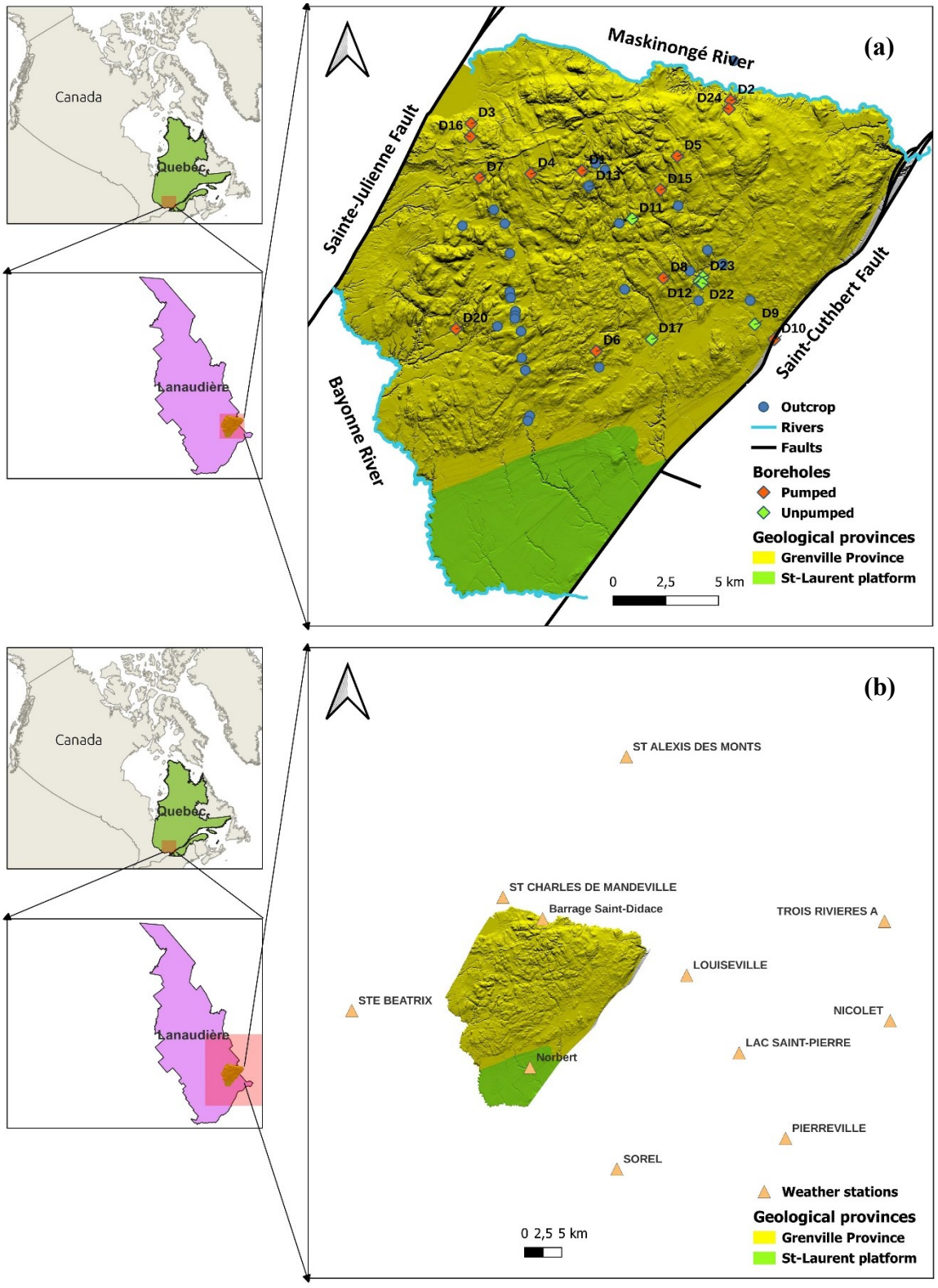


Figure 4.1. (a) Illustration of the study area with the borehole monitoring network and (b) the location of weather stations used for the meteorological data.

4.4 Data acquisition and processing

4.4.1 Groundwater monitoring network

A monitoring network of 20 boreholes was set up to record regional groundwater levels from summer 2019 to summer 2021. The boreholes were instrumented with Solinst Model 3001 Levelogger® 5 and Barologger 5 pressure transducers, which had a pressure sensor accuracy of ±0.05% FS (of full scale) and a temperature sensor accuracy of ±0.05 °C (Solinst Canada Ltd. 2021). Groundwater levels were recorded every 15 minutes. To obtain the true groundwater level in the boreholes, the data recorded by the level loggers were compensated for air pressure using the barologger data. The barologgers recorded air pressure only, whereas the level loggers recorded both water level and air pressure.

The monitoring network comprised 14 exploited private boreholes that were subjected to frequent pumping (pumped) and 6 normal piezometers from unexploited (unpumped) boreholes (Fig. 4.1). Although groundwater levels recorded from the piezometers reflect the natural fluctuation of groundwater levels, levels recorded from private boreholes were subjected to outliers (or noise) because of frequent pumping. A methodology proposed by Abi et al. (Submitted) was applied to solve the issue of outliers in private boreholes and allow to effectively use the private boreholes in the monitoring network. Abi et al. (Submitted) proposed to identify and remove the outliers through a slope criterion (Peterson et al. 2018), and the missing values (the removed outliers) are then filled via imputation methods through the R package *imputeTS* (Moritz and Bartz-Beielstein 2017).

4.4.2 Meteorological data

In order to make the meteorological data more representative of local conditions around the boreholes, a total of 10 weather stations were used to interpolate the daily values of rainfall, snowfall, and temperature It is the ordinary kriging spatial interpolation method that is used through the function *autoKrige* of the R package *automap* (Hiemstra et al. 2009). The other meteorological data of wind speed, relative humidity, and solar radiation were obtained from the Trois-Rivières weather station, located 43 km from the study area center. A map locating the used weather stations is provided in figure 1b.

Although the meteorological stations record snowfall, it is snowmelt that contributes to the groundwater recharge through infiltration and causes groundwater levels in the aquifer to rise (French and Binley 2004, Zhou et al. 2021). Unless the conditions favor snowmelt, snowfall will tend to accumulate on the ground over winter. Snowmelt conditions are complex and relate to the energy balance of the snowpack, although temperature is the main triggering factor (Walter et al. 2005, Zhou et al. 2021). The relative importance of temperature to the snowmelt process has led to the development of many snowmelt models that rely on a linear relationship between snowmelt and temperature (Zhou et al. 2021). However, the most reliable snowmelt models are those physically based on the energy balance (Walter et al. 2005, Zhou et al. 2021). In this study, the daily snowmelt was computed using the method of Walter et al. (2005), which is implemented in the R package *EcoHydRology* using the function *SnowMelt*. This calculation is based on an energy balance equation, but estimate the majority of the components of the equation with common meteorological and location parameters such as temperature, precipitation, and latitude. To produce more reliable estimates of snowmelt, one can also provide wind speed, slope, and aspect. The R code can be retrieved from the Cran GitHub repository of EcoHydRology, and

the mathematical and physical aspects of the snowmelt function are detailed in Walter et al. (2005).

4.4.3 Sediment and geological data

The 20 boreholes of the monitoring network were drilled into the crystalline rock of the Canadian Shield through sedimentary deposits of varying thickness. Several types of sediments lie upon the bedrock (Table 4.1). Sedimentary and geological data used in this study are mainly from the data base of the PACES-Lanaudière project, which is a collaboration between the Université du Québec à Chicoutimi and the Québec Ministère de l'Environnement et de la Lutte contre les changements climatiques (CERM-PACES 2022). The PACES-Lanaudière database comprises data from various sources, namely SIGEOM (Système d'information géominière du Québec) from the Québec Ministère de l'Énergie et des Ressources naturelles; SIH (Système d'information hydrogéologique) from the Québec Ministère de l'Environnement et de la Lutte contre les changements climatiques; as well as from new field investigations in the study area and several consultant reports collected from local governmental institutions, e.g., regional municipalities. More specifically, the SIGEOM database provided the bedrock and Quaternary geological maps as well as numerous geological reports. The SIH database provided the logs of all boreholes drilled in the province of Québec for hydrogeological purposes, particularly in the study area. In the study area, the PACES-Lanaudière collected a total number of 1901 boreholes with lithological data and a map of the Quaternary geology (Fig. 4.2).

Table 4.1. Sediment thickness and lithologic data of the monitoring network boreholes. Note that the complete Table of the 20 boreholes is provided in Appendix 11.

Boreholes	Sediment thickness (m)	Slope (°)	Lithology			
			Layer 1	Layer 2	Layer 3	Layer 4
1	1.2	9.3	0-1.2 m Sand	1.2-68.6 m Bedrock		
2	26.6	4.4	0-1.5 m Sand	1.5-16.7 m Silty sand	16.7-26.2 m Gravel	26.2-57.9 m Bedrock
3	7.2	3.4	0-7.2 m Sand	7.2-80.8 m Bedrock		
4	21.3	3.1	0-1.5 m Sand	1.5-21.3 m Clay	21.3-105.1 m Bedrock	
5	8.8	3.9	0-8.8 m Clay with gravel	8.8-92.9 m Bedrock		

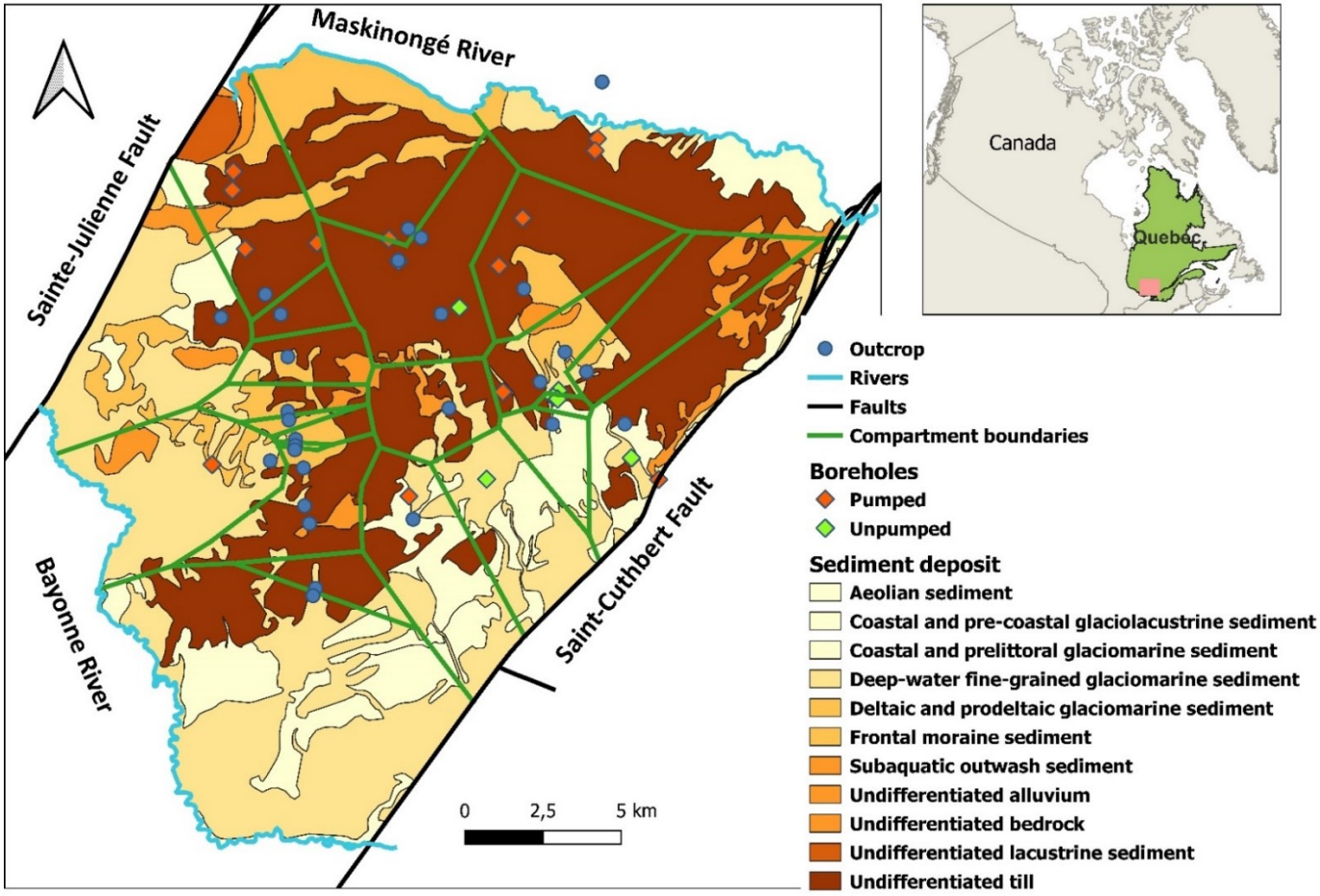


Figure 4.2. Illustration of the sediment types along with the fracture sample locations and the compartmentalization of the homogeneous areas according to the multiparametric similarity test, using fracture orientation, persistence, spacing, and aperture as fracture parameters.

4.4.4 Fracture sampling and network

The fracture network of the crystalline bedrock underlying the study area was thoroughly characterized by Abi et al. (2022). The scanline method was utilized to sample fractures on the outcrops, whereas acoustic and borehole imaging were used for boreholes. In total, 1421 and 543 fractures were respectively sampled from 30 outcrops and 4 boreholes across the study area. Using these outcrop and borehole fracture data, Abi et al. (2022) developed a methodology that compartmentalized the study area into regions of homogeneous fracture networks. The methodology is based on evaluating the similarity between fracture sets (or clusters) of adjacent samples through their fracture parameters. The fracture parameters used in the multiparametric similarity test of Abi et al. (2022) were fracture orientation, aperture, spacing, and persistence. If two samples are statistically similar, they are considered to share the same fracture network. If no similarity is identified, the two samples are considered to have different fracture networks. Figure 4.2 presents the results of the compartmentalization of the study area, and details of the methodology are detailed in Abi et al. (2022). Thus, each compartment can be described by the number of fracture sets and their parameters, i.e., mean orientation, aperture, spacing, and persistence. Using the fracture set parameters, the hydraulic conductivity can be estimated for each compartment, as detailed in the next section.

4.5 Methodology

4.5.1 Cross-correlation analysis of time series

The analysis of relationship between time series are usually performed using the cross-correlation analysis, which assesses whether there is a linear relationship between a pair of time series (Lee and Lee 2000; Cai and Ofterdinger 2016; Lorenzo-Lacruz et al. 2017; Fronzi et al. 2020; Dong et al. 2022). Cross-correlation analysis also identifies any time lag between the time series pair; this lag informs on the response time of the output time series in relation to the input time series (Lee and Lee 2000; Cai and Ofterdinger 2016). The delay between the two time series is generally considered the system's average response time for the studied input and output (Chen et al. 2004; Cai and Ofterdinger 2016). Generally, it is the peak correlation with its corresponding time lag that is reported in a cross-correlation analysis and negative time lags are ignored (Chen et al. 2004; Lee et al. 2006; Rathay et al. 2017). However, researchers such as Cai and Ofterdinger (2016) take negative time lags into account and have interpreted cases where the maximum correlation corresponds to a negative time lag as an indication of non-association between the input and output time series. In what follows, only positive time lags are considered, as suggested by Rathay et al. (2017), neglecting time lags where the groundwater response precedes the input time series.

The time series being analyzed for cross-correlation must fulfill specific requirements to avoid spurious correlations (Olden and Neff 2001; Yue et al. 2002; Cryer and Chan 2008; Probst et al. 2012; Dean and Dunsmuir 2016; Li and Guo 2019). The time series (input and output) being analyzed for cross-correlation must be stationary, and at least one of the time series must be independent white noise (Cryer and Chan 2008; Box et al. 2015; Dean and Dunsmuir 2016). In time series analysis, the second-order stationarity is generally

considered and is defined by a constant mean and variance at all times (Cryer and Chan 2008; Box et al. 2015; Shumway and Stoffer 2017). Independent white noise is defined as a sequence of independent, identically distributed (i.i.d) random variables (Cryer and Chan 2008), i.e., independent uncorrelated random variables (Shumway and Stoffer 2017).

An aquifer can be considered as a system in which rainfall and snowmelt are the inputs and groundwater level the output. Cross-correlation thus serves to identify the level of dependency between the input and output signals and the system's average response time. The cross-correlation function is expressed mathematically as (Cai and Ofterdinger 2016)

$$C_{xy}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(x_t - \bar{x})(y_{t+k} - \bar{y})\,, k \geq 0 \qquad (Eq\ 4.1)$$

$$\gamma_{xy}(k) = \frac{C_{xy}(k)}{\sigma_x \sigma_y}\,, \qquad (Eq\ 4.2)$$

where $C_{xy}(k)$ represents the cross-correlation, n is the number of measurements in the time series, k is the time lag, x_t and y_t represent the input and output time series values at time t, $\bar{x}$ and $\bar{y}$ are respectively the mean values of the time series x_t and y_t, $\gamma_{xy}(k)$ is the cross-correlation function (CCF), and σ_x and σ_y are the standard deviation of the input and output time series, respectively. confidence interval of 95% was selected, and the significance limits of the correlation are given by $\pm 1.96\sqrt{n}$ (Cai and Ofterdinger 2016).

For the sake of convenience, the methods and analysis of the applicability conditions of the cross-correlation method are provided in Appendix 12 texts S1 and S2, and the reader is referred to these documents for more detailed information. In hydrological studies, time series such as groundwater level (Lee and Lee 2000, Cai and Ofterdinger 2016), annual precipitation (Chen et al. 2004), and atmospheric pressure (Lee and Lee 2000) are often autocorrelated and nonstationary (Chen et al. 2004). When a time series is autocorrelated, the measurement at time step t+1 depends on that at time step t (Dean and Dunsmuir 2016).

Therefore, in hydrological studies, time series most generally do not fulfill the prerequisites for cross-correlation analysis.

Numerous statistical tests have been developed for testing, more objectively, time series stationarity (Hyndman and Athanasopoulos 2021) and whether a time series is independent white noise (Dean and Dunsmuir 2016, Li et al. 2021, Hyndman and Athanasopoulos 2021). Among the most commonly used ones, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (Kwiatkowski et al. 1992, Li et al. 2021, Hyndman and Athanasopoulos 2021) is utilized in this study to test for stationarity and the Ljung–Box test statistic to test for independent white noise. In the case of nonstationary time series, it is the commonly used differencing method (Box et al. 2015, Li et al. 2021, Hyndman and Athanasopoulos 2021) that was utilized for transforming nonstationary time series into stationary ones. Similarly, in the case where none of the input and output time series is independent white noise, a prewhitening method is commonly used (Probst et al. 2012, Shumway and Stoffer 2017, Li et al. 2021). Details of the abovementioned methods are provided in Appendix12 texts S1 and S2.

If one defines vertical inflow as the potential water from the precipitation available for infiltration, the vertical inflow will be the time series comprising the sum of rainfall and snowmelt. The use of the vertical inflow is important in temperate areas, as it considers both snowmelt and rainfall (Boumaiza et al. 2021). However, it is also useful to study the individual influences of rainfall and snowmelt on groundwater levels. Therefore, the input time series considered in this study are the rainfall, snowmelt, and vertical inflow. In this study, it is also demonstrated that differencing is required to make the groundwater level time series stationary and that prewhitening was required for cross-correlation that involved snowmelt and vertical inflow (Appensix 12 texts S1 and S2).

4.5.2 Spearman correlation

The Spearman correlation is a nonparametric method, which does not require that the analyzed variables be normally distributed (Kolassa 2020, Gibbons and Chakraborti 2020). In addition, the Spearman correlation can identify a nonlinear association between two variables, whereas the commonly used Pearson correlation measures only the linear association between the two variables (Kolassa 2020). In this study, the various parameters analyzed for their degree of association are not always normally distributed but rather show a mostly lognormal distribution (Appendix 13 Figure S28 and S29). A Spearman correlation value of 1 indicates perfect concordance between the variables, a value of -1 indicates perfect discordance between the variables, and a value of 0 indicates that the variables are most likely independent of each other (Kolassa 2020, Gibbons and Chakraborti 2020). Pairwise Spearman correlations were used in previous analyses of the relationship between geomorphological indices and groundwater levels.

4.5.3 Hydrogeological signatures, geomorphology, and fracture parameters

To assess the effect of geomorphology, the sediment deposit, and the fracture network characteristics on groundwater levels, common geomorphological indices, sediment deposit metricsindices, fracture parameters, and hydrogeological signatures were computed. Similar to hydrologic signatures definition linked to streamflow (McMillan 2021), hydrogeological signatures are here defined as quantitative metrics indices that describe the statistical or dynamic properties of groundwater levels. Therefore, the used hydrogeological signatures are the amplitude, median water table depth, recharge slope and discharge slope. Amplitude is the difference between the minimum and maximum groundwater level

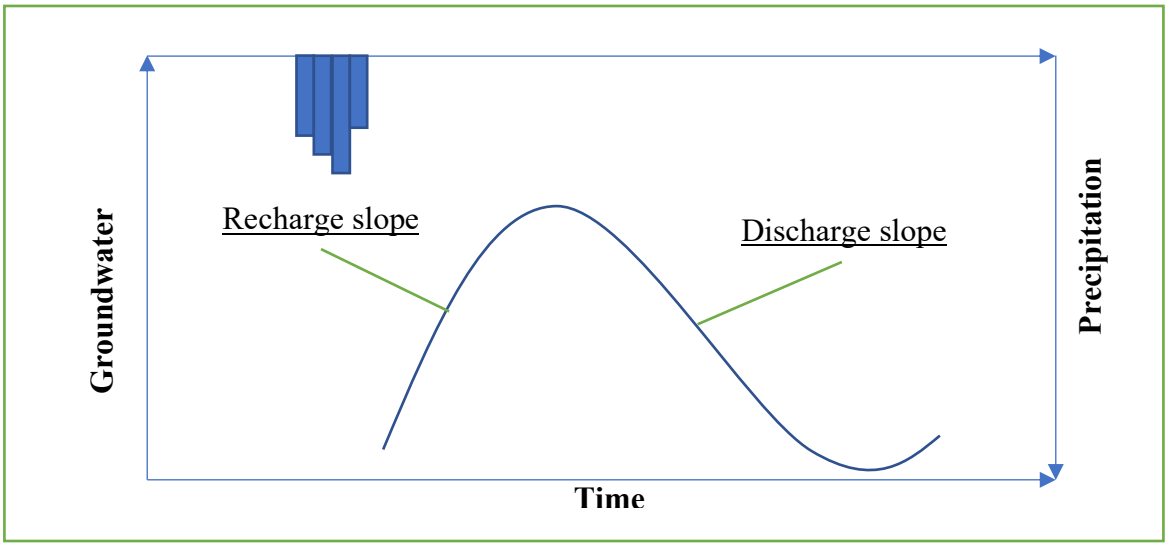


Figure 4.3. Illustration of the recharge and discharge slope of the groundwater level hydrograph as a function of time (recharge event) in response to precipitations

recorded in a borehole. During a recharge event, the recharge slope is the average slope of the groundwater level when it rises in the aquifer, and discharge slope is the average slope when groundwater level falls (see Figure 3). Recharge and discharge slopes are calculated using the formula $S = |\Delta h|/|\Delta t|$, where S is the slope and $|\Delta h|$ is the variation of groundwater level over time $|\Delta t|$. Similar discharge slopes were already used by Rinderer et al. (2016) and Bachmair and Weiler (2012). The absolute number ensures that the calculated slope is

always positive; therefore, a steep recharge slope relates to a rapid increase in groundwater level and a steep discharge slope reflects a rapid decrease in groundwater levels.

By geomorphology, it is the landform that is considered here. Landform has been defined as any physical element of the Earth's surface having a characteristic and recognizable shape (MacMillan and Shary 2009). The geomorphological indices comprised the local and watershed slope, the channel initiation threshold (CIT) index, and the topographic wetness index (TWI). All geomorphological indices were derived from a 1 m resolution LiDAR-based digital elevation model. The system for automated geoscientific analyses (SAGA) v. 8.3.0, which is a free and open-source software (FOSS), was the geographic information system (GIS) tool used to calculate the indices. Local slope was calculated at the pixel level, whereas watershed slope was the average of the slopes computed in the upstream watershed area of each pixel. The CIT index is a stream-power index derivative, given by $CIT = SCA \cdot (\tan \beta)^2$, where β is the local slope and SCA is the specific catchment area, defined as the upstream watershed area per unit contour width (Moore et al. 1991; Montgomery and Foufoula-Georgiou 1993). A stream-power index derivative, such as CIT, can be used to locate places where surface runoff may concentrate (Moore et al. 1991). The TWI is one of the most used topographical indices in hydrological studies (Mattivi et al. 2019; Riihimäki et al. 2021) and is given by $TWI = \ln (SCA/\beta)$. The TWI indicates the tendency of a location to accumulate water (Mattivi et al. 2019), making it particularly useful for identifying the spatial distribution of areas prone to saturation (Moore et al. 1991). The sediment deposit index is sediment thickness.

The fracture network parameters of the crystalline rock were calculated for each homogeneous zone, as it is assumed the parameters were constant within the respective homogeneous zones. For each homogeneous zone, the average aperture, persistence, and spacing were calculated by averaging the values of the geometric mean of the clusters for each parameter, given that spacing, persistence, and aperture were lognormally distributed.

Calculating the equivalent hydraulic conductivity of fractured rock from the aperture and spacing is much more complex and tends to overestimate the hydraulic conductivity even when the aperture is corrected for asperities generated by the roughness of the fracture walls (Scesi and Gattinoni 2010; Sun et al. 2011). The effective aperture, which participates in the circulation of water in the fractured rock, is much smaller than the measured mechanical aperture; the latter must be corrected for the asperities of the fracture walls (Scesi and Gattinoni 2010; Zhang 2013; Zhu 2019). Scesi and Gattinoni (2009) and Zhang (2013) provide a detailed presentation of the calculation of hydraulic conductivity, and a thorough summary of the calculation of the effective aperture can be found in (Scesi and Gattinoni 2010; Shahbazi et al. 2020). The method applied in this study is the one presented in Scesi and Gattinoni (2009), which consists of calculating the hydraulic conductivity of each fracture set by compensating for the asperities using the roughness of the fracture walls, and then calculating the equivalent hydraulic conductivity of a homogeneous zone by taking the summation of the hydraulic conductivity of all fracture sets within the homogeneous zone. Details related to this calculation are provided in the Appendix 16 Table S19. Using this method, the calculated hydraulic conductivity (Appendix 16 Table S19) served for comparative purposes of the various homogeneous zones and was not representative of the actual hydraulic conductivity of the rock mass, which must be determined by an appropriate in situ hydraulic test.

4.6 Results and discussion

4.6.1 Time series Analysis and Aquifer Response Time

Rainfall, snowmelt and vertical inflow are considered as inputs to the system and groundwater level as the output. Complete cross-correlation results are provided in Table 4.2, which shows for each combination of input and output time series, the maximum

correlation as well as the corresponding time lag and the correlation level beyond which the maximum correlation is considered significant. All the time lags are reported in days. Cross-correlation showed a generally significant correlation between the input and output time series, although the level of significance varied greatly. The input having the highest correlation with groundwater level varied among boreholes. Of the 20 groundwater level time series, cross-correlation values were most often higher when rainfall was considered as an input, followed by vertical inflow, and then snowmelt, with 14, 5, and 1 cases, respectively. Furthermore, rainfall had the strongest relationship with groundwater level at borehole D13 ($r = 0.34$, time lag = 0); snowmelt showed the strongest relationship with groundwater level at borehole D16 ($r = 0.20$, time lag = 1), and vertical inflow had the strongest correlation with groundwater level at borehole D13 ($r = 0.24$, time lag = 16).

Some cross-correlations were not significant between climate time series and groundwater level, and for the same borehole, the cross-correlation could be significant for one or two climate time series and not significant for the remaining climate time series. Cases where cross-correlations were not significant can be seen in Table 4.2, for boreholes 5, 10, 22 and 23. When considering rainfall or vertical inflow, only one case of non-significant correlation was observed for each, at boreholes D5 and D23, respectively. For snowmelt, cases of non-significant correlations were observed in boreholes D10, D22 and D23.

Similar cross-correlations between rainfall and groundwater level have been reported in the literature, where the maximum correlation values differ among sites and also among boreholes within the same site (Lee and Lee 2000; Chen et al. 2004; Lee et al. 2006; Cai and Ofterdinger 2016; Rathay et al. 2018). Reported cross-correlation values between rainfall and groundwater level match those observed in this study. In fractured hard rock aquifers, for instance, Cai and Ofterdinger (2016) reported cross-correlation values between 0.05 and 0.54, Rathay et al. (2018) reported values between 0.11 and 0.28, whereas in this study, cross-correlation values ranged from 0.04 to 0.34.

Cross-correlation analysis between snowmelt or vertical inflow as inputs and groundwater level are less common in the literature than rainfall–groundwater level cross-correlations; nonetheless, Flerchinger et al. (1992) reported cross-correlation values between snowmelt and groundwater level of 0.1 to 0.54, and Okkonen and Kløve (2010) reported values between –0.5 to 0.45. In this study, cross-correlation values between snowmelt and groundwater level ranged from 0.01 to 0.20. Cross-correlation analysis by Boumaiza et al. (2021) showed a significant correlation between vertical inflow and groundwater level, with a correlations greater than 0.5, but the maximum correlation values are not provided. Comparatively, the maximum correlation values obtained in this study are lower, ranging from 0.06 to 0.26.

As noted above, although the cross-correlation is statistically significant, the cross-correlation values are low between the output groundwater level and the inputs of rainfall, snowmelt, and vertical inflow (generally less than 0.5). These low values occur because the cross-correlation analysis considers all values of the input and output time series to calculate the correlation values, whereas some minor rainfall, snowmelt, or vertical inflow events may not generate a response in groundwater level. This is the *damping effect*, expressed as a loss of signal between the input and output time series, and is primarily caused by the hydrogeological properties of the sediment deposit (Corona et al. 2018; Liesch and Wunsch 2019). The damping effect is particularly important in thick sediment deposits with low hydraulic conductivity (Liesch and Wunsch 2019).

Table 4.2 also highlights the aquifer response time as a function of the climate variables for each borehole. Overall, the response time of groundwater levels was shorter when rainfall was considered the input to the cross-correlation, followed by vertical inflow and then snowmelt. In fact, for each input time series, the response time that appears most (the mode) is 0 day for rainfall (16 times out of 20), 1 day for snowmelt (9 times out of 20), and 0 day for vertical inflow (8 times out of 20). The average aquifer response times with

respect to the input time series were 1.7 (standard deviation [SD], 4.4), 11.7 (SD, 11.1), and 7.1 (SD, 8.8) days for rainfall, snowmelt, and vertical inflow, respectively. This difference is likely related to the fact that the soil is partially frozen in winter and early spring, which limits the rate of infiltration into the aquifer, creating larger lags when snowmelt and vertical inflow are considered. (French and Binley 2004; Wright and Novakowski 2022). Studies in Ireland (Cai and Ofterdinger 2016) , Canada (Rathay et al. 2017), South Korea (Lee and Lee 2000), Benin (Kotchoni et al. 2019) and England (Lee et al. 2006) have shown relatively similar aquifer response times, ranging from 0 day to several days, depending on sediment deposit, geological complexity, and hydrological processes. Rathay et al. (2017) reported response times of groundwater levels to rainfall that are proportional to sediment thickness in relatively thin sediment deposits (less than 3.7 m). However, such proportionality is not observed in relatively complex geologic settings (Cai and Ofterdinger 2016; Boumaiza et al. 2021). In this study, the association between response time and geomorphology is discussed in the next section.

Table 4.2. Maximum cross-correlation and the lag for when the maximum cross-correlation is reached for each monitoring borehole. The input time series are the rainfall, snowmelt, and vertical inflow; the output is groundwater level

	Rainfall			Snowmelt			Vertical inflow		
Boreholes	**Corr**	**Bound**	**Lag (days)**	**Corr**	**Bound**	**Lag (days)**	**Corr**	**Bound**	**Lag (days)**
D1	0,30	0,08	0	0,17	0,07	16	0,22	0,08	0
D2	0,13	0,08	0	0,08	0,08	37	0,10	0,08	12
D3	0,25	0,08	0	0,16	0,08	15	0,16	0,08	0
D4	0,28	0,08	0	0,17	0,08	1	0,20	0,08	0
D5	0,04	0,08	1	0,09	0,08	16	0,12	0,08	16
D6	0,09	0,08	0	0,09	0,08	15	0,10	0,08	15
D7	0,23	0,11	0	0,17	0,11	1	0,15	0,11	12
D8	0,11	0,08	0	0,10	0,08	1	0,11	0,08	1
D9	0,13	0,08	0	0,09	0,08	18	0,09	0,07	0
D10	0,09	0,08	0	0,05	0,08	28	0,08	0,08	16
D11	0,25	0,08	0	0,13	0,08	1	0,21	0,08	0
D12	0,14	0,08	0	0,15	0,08	1	0,16	0,08	1
D13	0,34	0,08	0	0,16	0,08	16	0,24	0,08	0
D15	0,11	0,08	12	0,13	0,09	31	0,14	0,08	31
D16	0,33	0,09	0	0,20	0,09	1	0,26	0,09	0
D17	0,18	0,12	0	0,20	0,11	18	0,13	0,12	16
D20	0,23	0,09	0	0,12	0,08	1	0,15	0,08	0
D22	0,11	0,08	4	0,01	0,08	15	0,08	0,08	4
D23	0,08	0,08	17	0,01	0,09	1	0,06	0,08	17
D24	0,16	0,08	0	0,14	0,08	1	0,14	0,09	1

4.6.2 Effect of Geomorphology and the Sediment deposit on Groundwater Level

In the previous section, groundwater level time series were studied in response to rainfall, snowmelt and vertical inflow time series. However, groundwater level is also influenced by factors such as geomorphology (Bachmair and Weiler 2012; Rinderer et al. 2016) and sediment deposits (Gleeson et al. 2009; Corona et al. 2018; Liesch and Wunsch 2019). In addition to the geomorphological and sediment deposit indices, the time lags obtained from the previous analyses (Section 4.6.1) were also included as indices indicating the response time of groundwater level to rainfall, snowmelt, and vertical inflow.

The results of the degree of pairwise association between the different indices through Spearman correlation are presented in Table 4.3 and Appendix 13 Figure S27, while the data of the different parameters used for the calculation of the Spearman correlation are provided in Appendix 15 Table S18. Correlation values that are statistically significant are highlighted in Table 4.3 and Appendix 13 Figure S27. In the following, we only focus on pairwise parameters whose association through Spearman correlation (Table 4.3) are meaningful for a better understanding of the effect of geomorphology and sediment deposit on groundwater level.

Spearman correlation identified that local slope was significantly associated with multiple indices. Local slope was positively correlated with the recharge and discharge slopes and negatively correlated with sediment thickness and the aquifer response time to rainfall. Sediment thickness was positively correlated with the median water table depth and negatively correlated with the local and watershed slopes. The watershed slope is also positively correlated with CIT and recharge slope. Additionally, CIT was positively correlated with amplitude and negatively correlated with TWI. While TWI was not correlated to any additional index, amplitude was positively correlated with recharge slope. Furthermore, recharge slope was negatively correlated with median water table depth and the aquifer

response time to vertical inflow, whereas discharge slope was not correlated with any other index except local slope.

Having a local slope that is positively correlated with the recharge and discharge slopes indicate that, for a given point, the speed at which the water flux travels into and out of the aquifer system is influenced by the local slope. Similar results were reported by other researchers such as Rinderer et al. (2016), who found the local slope to be correlated with the discharge slope in pre-alpine catchment. However, the degree of association between local slope and recharge slope was greater than that between local slope and discharge slope. Such degree of association can be explained by the different hydrogeological processes occurring during recharge and discharge. During recharge, a combination of vertical flow through the sediment deposit and lateral flow from upstream generates the increase of the water table that yield the recharge slope. During discharge, however, the decrease of the water table is mainly due to lateral flow. The increase of the water table (recharge slope) was an order of magnitude greater than its decrease (discharge slope) (Appendix 15 Table S18).

As also reported by Rinderer et al. (2016), in this study, aquifer response time to rainfall was negatively correlated with local slope Moreover, the local and watershed slopes were negatively correlated with sediment thickness. This negative correlation indicates that the aquifer response time to rainfall is slower in gently sloping areas where sediment deposits are thicker than on areas with relatively steep slopes. The present observations match those of Gleeson et al. (2009) and Corona et al. (2018) in demonstrating that thicker sediment deposits favor a longer time for rainfall, snowmelt and vertical inflow to infiltrate the aquifer. It should be noted that the response time is not only controlled by the sediment thickness, since sediment hydraulic conductivity (Gleeson et al. 2009) and the complexity of the aquifer system play an important role (Padilla et al. 2014; Corona et al. 2018). This can be easily illustrated by comparing the results from boreholes D4 and D5. Table S18

(Appendix 15) shows that the aquifer response time to rainfall in borehole D4 was relatively fast compared to that in borehole D5, despite the fact that the sediment thickness in borehole D4 was greater than in borehole D5 and the local slope in borehole D5 was greater than in borehole D4. This discrepancy may be explained in part by the complexity of the aquifer layers in the two boreholes, as shown in Table S16 of the Appendix 11. It can be seen in Table S16 that boreholes D4 and D5 have 3 and 2 geologic layers, respectively, and both contain a clayey formation.

Regarding the results presented in Table 4.3, only the aquifer response time to rainfall was found to be associated with geomorphological indices. In fact, the geomorphological indices were not associated with the aquifer response time to snowmelt and vertical inflows. A possible explanation for this non-correlation should be related to the freezing and thawing of soils during winters and springs, which makes recharge from snowmelt and vertical inflow dependent on the energy balance (primarily temperature) and aquifer characteristics, rather than geomorphologic indices (French and Binley 2004; Wright and Novakowski 2022).

When considering the median water table depth, it was found to be negatively correlated with recharge slope and sediment thickness and moderately correlated with the local slope. Therefore, local slope and sediment characteristics can help predict the median water table depth. Comparable findings have also been reported for the relationship between median water table depth and the local slope (Bachmair and Weiler 2012; Rinderer et al. 2014) as well as between median water table depth and sediment thickness (Rinderer et al. 2014). However, while Rinderer et al. (2014) found that geomorphological indices such as slope and sediment thickness were good predictors of median water table depth in pre-Alpine catchments in switzeland, Bachmair and Weiler (2012) found that geomorphological indices were not sufficient to predict water table depth in hillslope catchments in Ireland. Bachmair and Weiler (2012) argued that the lack of predictability using geomorphological

parameters occurs because they didn't include other relatively important parameters, such as preferential flow in the subsurface and sediment thickness.

The significant association between the TWI and hydrogeological signatures has been well demonstrated (Detty and McGuire 2010; Rinderer et al. 2014, 2016). However, in this study the TWI was not significantly associated with any hydrogeological signature (Table 4.3). In contrast, the CIT was positively associated with amplitude. CIT and TWI are both functions of slope and specific catchment area. However, CIT reflects the potential of a site to concentrate water as runoff, while TWI reflects the potential for water stagnation at a site. In this regard, it can be inferred that the boreholes examined in this study were located in areas subject to surface runoff rather than water stagnation. The positive association between CIT and amplitude, which in turn was positively associated with recharge slope, suggests that in this study, not only surface water but also groundwater concentrates to generate water table rise in the boreholes, since local slope and watershed slope are positively associated with recharge slope.

Table 4.3 . Spearman rho correlation matrix of the hydrogeological signatures and the geomorphological and sediment deposit indices. The rho correlation values in bold are significant at a p-value of 0.05; CIT, channel initiation threshold index; TWI, topographic wetness index

	Sediment thickness	Local Slope	Watershed Slope	CIT	TWI	Amplitude	Recharge slope	Discharge slope	Median water table depth	Vertical inflow time lag	Rainfall time lag	Snowmelt time lag
Sediment thickness	1.00	**-0.49**	**-0.69**	-0.46	-0.02	-0.12	-0.42	0.24	**0.65**	0.10	0.16	-0.01
Local Slope	**-0.49**	1.00	**0.57**	0.41	-0.25	0.37	**0.62**	**0.48**	-0.41	-0.33	**-0.50**	0.12
Watershed Slope	**-0.69**	**0.57**	1.00	**0.66**	-0.15	0.44	**0.47**	-0.03	-0.18	-0.18	-0.28	0.04
CIT	-0.46	0.41	**0.66**	1.00	**-0.57**	**0.52**	0.17	-0.03	-0.02	0.14	-0.11	0.04
TWI	-0.02	-0.25	-0.15	**-0.57**	1.00	-0.36	-0.05	0.05	-0.19	0.04	0.09	-0.07
Amplitude	-0.12	0.37	0.44	**0.52**	-0.36	1.00	**0.57**	0.26	0.02	-0.05	-0.28	0.11
Recharge slope	-0.42	**0.62**	**0.47**	0.17	-0.05	**0.57**	1.00	0.36	**-0.46**	**-0.45**	-0.41	-0.01
Discharge slope	0.24	**0.48**	-0.03	-0.03	0.05	0.26	0.36	1.00	-0.11	-0.01	-0.22	0.01
Median water table depth	**0.65**	-0.41	-0.18	-0.02	-0.19	0.02	**-0.46**	-0.11	1.00	0.15	0.15	0.09
Vertical inflow time lag	0.10	-0.33	-0.18	0.14	0.04	-0.05	**-0.45**	-0.01	0.15	1.00	**0.60**	0.36
Rainfall time lag	0.16	**-0.50**	-0.28	-0.11	0.09	-0.28	-0.41	-0.22	0.15	**0.60**	1.00	0.14
Snowmelt time lag	-0.01	0.12	0.04	0.04	-0.07	0.11	-0.01	0.01	0.09	0.36	0.14	1.00

4.6.3 Effect of Crystalline Rock and Fracture Characteristics on Groundwater Level

The Spearman correlation was also used in this section for analyzing the relationship between groundwater level variability and fractures network characteristics of the crystalline rocks. The results of the pairwise association between the different indices were presented in Table 4.4 and Appendix 14 Figure S28, while the data used for these Spearman correlation analyses were provided in the Appendix 16 Table S19. In what follows, only relevant correlations will be reported and discussed. For instance, significant correlations between fracture parameters are not relevant to the purpose of this study because each fracture parameter cannot be a causal factor for another fracture parameter. Specifically, fracture orientation, persistence, spacing, number of fracture sets, and aperture are inherently independent properties used in fracture characterization, and these characteristics are dependent on the rock fracturing process (Priest 1993; Fossen 2016). Therefore, significant correlations such as those between the number of fracture sets and average aperture or between the number of fracture sets and average spacing are ignored. Significant correlations between hydrogeological signatures are also not discussed, since they were discussed in section 4.6.2.

The Spearman correlation results showed that the number of fracture sets index was negatively associated with equivalent hydraulic conductivity and aquifer response time to snowmelt. Average aperture was positively correlated with equivalent hydraulic conductivity and median water table depth. Average spacing was negatively correlated with equivalent hydraulic conductivity and aquifer response time to snowmelt. Average persistence was negatively associated with equivalent hydraulic conductivity, discharge slope and aquifer response time to snowmelt. Additionally, equivalent hydraulic conductivity was also positively correlated with median water table depth, aquifer response time to vertical inflow

and negatively correlated with recharge slope. Aquifer response time to vertical flow was also positively correlated with rock thickness.

To facilitate the interpretation of the results presented in Table 4.4 and Appendix 14 Figure S28, it was deemed necessary to present the following equations describing the hydraulic conductivity in fractured rocks, toward the direction N, having n fracture sets (Shahbazi et al. 2020):

$$K_N = \sum_{i=1}^{n} K_i \cos^2 \alpha_i \qquad (Eq\ 4.3)$$

where K_i is the hydraulic conductivity of the fracture set *i*, and α_i is the angle between the fracture set *i* and the direction N.

$$K_i = \frac{g(a_{ave})^3}{12\nu C_{ave} S_{ave}} \qquad (Eq\ 4.4)$$

where g is the gravitational acceleration, v is the dynamic viscosity of water, a_{ave} is the average aperture of the fracture set, C_{ave} is the average correction factor of the fracture set, and S_{ave} is the average spacing of the fractures within the fracture set. Equation 4.3 reveals that the equivalent hydraulic conductivity K_N is given by the sum of the hydraulic conductivity contributions of each fracture set *i*. However, the contribution of each fracture set *i* depends not only on the hydraulic conductivity (K_i) of the fracture set, but also on the angle α_i between the fracture set and the direction N. In this respect, depending on the combination of fracture set hydraulic conductivities (K_i) and angles α_i, the resulting equivalent hydraulic conductivity K_N may be higher for a small number of fracture sets than for a larger number of fracture sets. From equation 4.4, it can be deduced that for a fracture set, the greater the average aperture, the higher the hydraulic conductivity, and conversely, the greater the average spacing, the lower the hydraulic conductivity.

The Spearman correlation results showed that the number of fracture sets index was negatively associated with equivalent hydraulic conductivity and aquifer response time to snowmelt. Average aperture was positively correlated with equivalent hydraulic conductivity and median water table depth. Average spacing was negatively correlated with equivalent hydraulic conductivity and aquifer response time to snowmelt. Average persistence was negatively associated with equivalent hydraulic conductivity, discharge slope and aquifer response time to snowmelt. Additionally, equivalent hydraulic conductivity was also positively correlated with median water table depth, aquifer response time to vertical inflow and negatively correlated with recharge slope. Aquifer response time to vertical flow was also positively correlated with rock thickness.

As previously indicated, equivalent hydraulic conductivity is positively correlated with average aperture and negatively correlated with the number of fracture sets and average spacing. In this respect, Equations 4.3 and 4.4 perfectly express the fact that average aperture and spacing should be positively and negatively correlated with equivalent hydraulic conductivity respectively, since average aperture is in the numerator and average spacing in the denominator of the fracture set hydraulic conductivity formula (Equation 3). Furthermore, given Equation 4.3 and the structural domains, it can be deduced that in this study, structural domains with a small number of fracture sets tend to produce higher equivalent hydraulic conductivities than structural domains with a larger number of fracture sets. This could be explained by a combination of factors in favor of calculated equivalent hydraulic conductivities from structural domains with a small number of fracture sets over structural domains with a larger number of fracture sets, such as favorable orientations (in this case near vertical), larger apertures and smaller spacing between fractures.

The rate at which the water table rises in the aquifer (recharge slope) is negatively correlated with equivalent hydraulic conductivity, whereas it should be positively correlated, since the water table should rise faster when hydraulic conductivity is high. This may suggest

that the rate at which the water table rises is more closely related to another factor, such as the local slope, which is linked to the thickness of the sediment deposit, as discussed in the previous section 4.6.2.

Considering the different climate imputs and the fracture characteristics indices, it is the response of the aquifer to snowmelt that is negatively related to number of fracture sets, average spacing and average persistence, whereas the equivalent hydraulic conductivity is positively correlated with the response of the aquifer to vertical inflow. However, as seen in equation 4.3 and equation 4.4, the average persistence is expected to be positively correlated with snowmelt time lag and the equivalent hydraulic conductivity negatively correlated with vertical inflow time lag. The correlation between the number of fracture sets and response time of the aquifer to snowmelt is positive, as expected (equation 4.3).

It should be noted that the physical processes behind some of the significant correlations shown in Table 4.4 are not obvious and require further investigation. To cite a few other such correlations, the negative correlation between mean persistence and discharge slope, the positive correlation between average aperture and median water table depth.

Table 4.4. Spearman rho correlation matrix of the hydrogeological signatures, crystalline rock, and fracture characteristics. The rho correlation values in bold are significant at a p-value of 0.05

	Number of fracture sets	Average Aperture	Average Spacing	Average persistence	Equivalent hydraulic conductivity	Rock thickness	Amplitude	Recharge slope	Discharge slope	Median water table depth	Vertical inflow time lag	Rainfall time lag	Snowmelt time lag
Number of fracture sets	1.00	**-0.67**	**0.51**	0.29	**-0.59**	-0.30	-0.01	0.12	0.00	-0.24	-0.31	-0.33	**-0.49**
Average Aperture	**-0.67**	1.00	-0.42	-0.46	**0.93**	0.27	-0.04	-0.42	-0.32	**0.56**	0.35	0.27	0.20
Average Spacing	**0.51**	-0.42	1.00	**0.73**	**-0.51**	-0.38	-0.35	-0.01	-0.12	-0.32	-0.34	-0.21	**-0.58**
Average persistence	0.29	-0.46	**0.73**	1.00	**-0.64**	0.05	-0.21	0.01	**-0.55**	-0.18	-0.41	-0.12	**-0.55**
Equivalent hydraulic conductivity	**-0.59**	**0.93**	**-0.51**	**-0.64**	1.00	0.22	-0.16	**-0.57**	-0.21	**0.56**	**0.49**	0.37	0.19
Rock thickness	-0.30	0.27	-0.38	0.05	0.22	1.00	0.24	-0.04	-0.16	0.07	**0.51**	0.39	0.20
Amplitude	-0.01	-0.04	-0.35	-0.21	-0.16	0.24	1.00	**0.56**	0.20	0.02	-0.05	-0.28	0.11
Recharge slope	0.12	-0.42	-0.01	0.01	**-0.57**	-0.04	**0.56**	1.00	0.27	**-0.46**	**-0.47**	-0.44	-0.01
Discharge slope	0.00	-0.32	-0.12	**-0.55**	-0.21	-0.16	0.20	0.27	1.00	-0.08	0.07	-0.08	0.10
Median water table depth	-0.24	**0.56**	-0.32	-0.18	**0.56**	0.07	0.02	**-0.46**	-0.08	1.00	0.15	0.15	0.09
Vertical inflow time lag	-0.31	0.35	-0.34	-0.41	**0.49**	**0.51**	-0.05	**-0.47**	0.07	0.15	1.00	**0.60**	0.36
Rainfall time lag	-0.33	0.27	-0.21	-0.12	0.37	0.39	-0.28	-0.44	-0.08	0.15	**0.60**	1.00	0.14
Snowmelt time lag	**-0.49**	0.20	**-0.58**	**-0.55**	0.19	0.20	0.11	-0.01	0.10	0.09	0.36	0.14	1.00

4.7 Conclusion

Within the aquifer system, recharge and discharge are expressed respectively as an increase or decrease of water table height, and this level varies over time in response to various input factors. These input factors include geomorphological, sediment deposit, meteorological, and crystalline rock fracture characteristics. Although the effects of geomorphological and meteorological input factors on groundwater levels have been studied previously, this study also adds the role of the sediment deposit and crystalline rock fracture characteristics as potential factors controlling aquifer response time and groundwater levels in hillslope crystalline aquifers.

For the analysis of the response time of the aquifer system, the cross-correlation analysis was performed between rainfall and groundwater levels, snowmelt and groundwater levels, vertical inflow and groundwater levels. In general, the analysis showed slightly faster response time of the aquifer system to rainfall relative to its response time to snowmelt. The most common response time was 0 days for rainfall and 1 day for snowmelt and vertical inflow; these lags suggest that recharge is almost instantaneous in the case of rainfall and slightly delayed in the case of snowmelt. The cross-correlation between rainfall, snowmelt, and vertical inflow in relation to groundwater level was generally statistically significant

To evaluate the effect of geomorphology, the sediment deposits, and crystalline rock fracture characteristics on the variability of groundwater levels and response time, several parameters characterizing each factor were considered. These parameters were the local and watershed slope, the channel initiation threshold (CIT) index, the topographic wetness index (TWI), and sediment thickness for geomorphological factors; number of fracture sets, average aperture, average spacing, average persistence, equivalent hydraulic conductivity and rock thickness for the fractured crystalline rock. The parameters used to characterize

groundwater level time series were amplitude, recharge slope, discharge slope, Median water table depth, and the variable response times of the aquifer to rainfall, snowmelt, and vertical inflow.

Spearman correlation identified that the variability of groundwater levels and response time were mainly linked to geomorphological parameters and fracture parameters. More specifically, local slope, sediment thickness were the most important parameters controlling groundwater levels in the studied crystalline aquifers of the Canadian Shield. Local slope and sediment thickness are associated with recharge slope, discharge slope, and the variable response times of the aquifer. Moreover, sediment thickness and local slope are intercorrelated, in that sediment deposits are thicker in areas having a gentle local slope. Fracture characteristics of the crystalline rock do not significantly affect the response time or variability of groundwater levels. The link between the fracture parameters and the indices related to the variability of groundwater levels and response time was also significant; however, the physical processes behind the correlations were not always abvious. This suggest that further investigations are needed to clarify the relation between the fracture parameters and hydrogeological indices.

It should be noted that the present study investigated the direct links between input and output variables through various parameters in the form of a black box model, neglecting the effects of evapotranspiration and capillary rise. Nonetheless, present approach is useful for identifying the important input variables needed to be included in a statistical regression model for predicting groundwater levels. Such statistical regression models have proven successful for predicting groundwater levels when considering rainfall, snowmelt, and evapotranspiration as input variables (Chen et al. 2002, Okkonen and Kløve 2010). This study that it would also be interesting to include geomorphological parameters in such regression models, especially given the advantage of including these variables in terms of their simple calculations that do not require marked amounts of computing power and time.

4.8 References

Abi, A., Walter, J., Chesnaux, R., and Saeidi, A. (Submitted). Regional groundwater level monitoring with exploited private boreholes: outlier removal and imputation of missing values. Hydrogeology Journal,.

Abi, A., Walter, J., Saeidi, A., and Chesnaux, R. 2022. A cluster-based multiparametric similarity test for the compartmentalization of crystalline rocks into structural domains. Quarterly Journal of Engineering Geology and Hydrogeology, 55: qjegh2021-136. doi:10.1144/qjegh2021-136.

Bachmair, S., and Weiler, M. 2012. Hillslope characteristics as controls of subsurface flow variability. Hydrology and Earth System Sciences, 16: 3699–3715. Copernicus GmbH. doi:10.5194/hess-16-3699-2012.

Béland, R. 1967. Région de Saint-Gabriel-de-Brandon, Comtes de Joliette, Berthier et Maskinongé. Geological reports, Ministère des Richesses Naturelles du Québec, Québec.

Boumaiza, L., Chesnaux, R., Walter, J., and Meghnefi, F. 2021. Assessing response times of an alluvial aquifer experiencing seasonally variable meteorological inputs. Groundwater for Sustainable Development, 14: 100647. doi:10.1016/j.gsd.2021.100647.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M. 2015. Time Series Analysis: Forecasting and Control. John Wiley & Sons.

Bresciani, E., Goderniaux, P., and Batelaan, O. 2016. Hydrogeological controls of water table-land surface interactions. Geophysical Research Letters, 43: 9653–9661. doi:10.1002/2016GL070618.

Cai, Z., and Ofterdinger, U. 2016. Analysis of groundwater-level response to rainfall and estimation of annual recharge in fractured hard rock aquifers, NW Ireland. Journal of Hydrology, 535: 71–84. doi:10.1016/j.jhydrol.2016.01.066.

CERM-PACES. 2022. Résultats du programme d'acquisition de connaissances sur les eaux souterraines du territoire municipalisé de Lanaudière. Centre d'études sur les ressources minérales, Université du Québec à Chicoutimi, Quebec, Canada.

Chen, Z., Grasby, S.E., and Osadetz, K.G. 2002. Predicting average annual groundwater levels from climatic variables: an empirical model. Journal of Hydrology, 260: 102–117. doi:10.1016/S0022-1694(01)00606-0.

Chen, Z., Grasby, S.E., and Osadetz, K.G. 2004. Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada. Journal of Hydrology, 290: 43–62. doi:10.1016/j.jhydrol.2003.11.029.

Corona, C.R., Gurdak, J.J., Dickinson, J.E., Ferré, T.P.A., and Maurer, E.P. 2018. Climate variability and vadose zone controls on damping of transient recharge. Journal of Hydrology, 561: 1094–1104. doi:10.1016/j.jhydrol.2017.08.028.

Crosbie, R.S., Binning, P., and Kalma, J.D. 2005. A time series approach to inferring groundwater recharge using the water table fluctuation method. Water Resources Research, 41. doi:10.1029/2004WR003077.

Cryer, J.D., and Chan, K.-S. 2008. Time Series Analysis. Springer New York, New York, NY. doi:10.1007/978-0-387-75959-3.

Dean, R.T., and Dunsmuir, W.T.M. 2016. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. Behavior Research Methods, 48: 783–802. doi:10.3758/s13428-015-0611-2.

Delottier, H., Pryet, A., Lemieux, J.M., and Dupuy, A. 2018. Estimating groundwater recharge uncertainty from joint application of an aquifer test and the water-table fluctuation method. Hydrogeology Journal, 26: 2495–2505. doi:10.1007/s10040-018-1790-6.

Detty, J.M. and K. J. McGuire. 2010. Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment. Hydrological Processes, 24: 2222–2236. doi:10.1002/hyp.7656.

Dewandel, B., Perrin, J., Ahmed, S., Aulong, S., Hrkal, Z., Lachassagne, P., Samad, M., and Massuel, S. 2010. Development of a tool for managing groundwater resources in semi-arid hard rock regions: application to a rural watershed in South India. Hydrological Processes, 24: 2784–2797. doi:10.1002/hyp.7696.

Dickinson, J.E., Hanson, R.T., Ferré, T.P.A., and Leake, S.A. 2004. Inferring time-varying recharge from inverse analysis of long-term water levels. Water Resources Research, 40. doi:10.1029/2003WR002650.

Dong, L., Guo, Y., Tang, W., Xu, W., and Fan, Z. 2022. Statistical Evaluation of the Influences of Precipitation and River Level Fluctuations on Groundwater in Yoshino River Basin, Japan. Water, 14: 625. Multidisciplinary Digital Publishing Institute. doi:10.3390/w14040625.

Flerchinger, G.N., Cooley, K.R., and Ralston, D.R. 1992. Groundwater response to snowmelt in a mountainous watershed. Journal of Hydrology, 133: 293–311. doi:10.1016/0022-1694(92)90260-3.

Forster, C., and Smith, L. 1988. Groundwater flow systems in mountainous terrain: 2. Controlling factors. Water Resources Research, 24: 1011–1023. doi:10.1029/WR024i007p01011.

Fossen, H. 2016. Structural Geology. In 2nd edition. Cambridge University Press.

French, H., and Binley, A. 2004. Snowmelt infiltration: monitoring temporal and spatial variability using time-lapse electrical resistivity. Journal of Hydrology, 297: 174–186. doi:10.1016/j.jhydrol.2004.04.005.

Fronzi, D., Di Curzio, D., Rusi, S., Valigi, D., and Tazioli, A. 2020. Comparison between Periodic Tracer Tests and Time-Series Analysis to Assess Mid- and Long-Term Recharge Model Changes Due to Multiple Strong Seismic Events in Carbonate Aquifers. Water, 12: 3073. Multidisciplinary Digital Publishing Institute. doi:10.3390/w12113073.

Gabrielli, C.P., McDonnell, J.J., and Jarvis, W.T. 2012. The role of bedrock groundwater in rainfall–runoff response at hillslope and catchment scales. Journal of Hydrology, 450–451: 117–133. doi:10.1016/j.jhydrol.2012.05.023.

Gibbons, J.D., and Chakraborti, S. 2020. Nonparametric Statistical Inference. In 6th edition. Chapman and Hall/CRC, New York. doi:10.1201/9781315110479.

Gleeson, T., Novakowski, K., and Kurt Kyser, T. 2009. Extremely rapid and localized recharge to a fractured rock aquifer. Journal of Hydrology, 376: 496–509. doi:10.1016/j.jhydrol.2009.07.056.

Haitjema, H.M., and Mitchell-Bruker, S. 2005. Are Water Tables a Subdued Replica of the Topography? Groundwater, 43: 781–786. doi:10.1111/j.1745-6584.2005.00090.x.

Hiemstra, P.H., Pebesma, E.J., Twenhöfel, C.J.W., and Heuvelink, G.B.M. 2009. Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. Computers & Geosciences, 35: 1711–1721. doi:10.1016/j.cageo.2008.10.011.

Houben, T., Pujades, E., Kalbacher, T., Dietrich, P., and Attinger, S. 2022. From Dynamic Groundwater Level Measurements to Regional Aquifer Parameters— Assessing the Power of Spectral Analysis. Water Resources Research, 58: e2021WR031289. doi:10.1029/2021WR031289.

Hyndman, R.J., and Athanasopoulos, G. 2021. Forecasting: Principles and Practice. In 3rd ed. édition. Otexts.

Jie, Z., van Heyden, J., Bendel, D., and Barthel, R. 2011. Combination of soil-water balance models and water-table fluctuation methods for evaluation and improvement of groundwater recharge calculations. Hydrogeology Journal, 19: 1487–1502. doi:10.1007/s10040-011-0772-8.

Jocson, J.M.U., Jenson, J.W., and Contractor, D.N. 2002. Recharge and aquifer response: Northern Guam Lens Aquifer, Guam, Mariana Islands. Journal of Hydrology, 260: 231–254. doi:10.1016/S0022-1694(01)00617-5.

Kolassa, J.E. 2020. An Introduction to Nonparametric Statistics. In 1st edition. Chapman and Hall/CRC, Boca Raton. doi:10.1201/9780429202759.

Kotchoni, D.O.V., Vouillamoz, J.-M., Lawson, F.M.A., Adjomayi, P., Boukari, M., and Taylor, R.G. 2019. Relationships between rainfall and groundwater recharge in seasonally humid Benin: a comparative analysis of long-term hydrographs in sedimentary and crystalline aquifers. Hydrogeology Journal, 27: 447–457. doi:10.1007/s10040-018-1806-2.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics, 54: 159–178. doi:10.1016/0304-4076(92)90104-Y.

Lee, J.-Y., and Lee, K.-K. 2000. Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system. Journal of Hydrology, 229: 190–201. doi:10.1016/S0022-1694(00)00158-X.

Lee, L.J.E., Lawrence, D.S.L., and Price, M. 2006. Analysis of water-level response to rainfall and implications for recharge pathways in the Chalk aquifer, SE England. Journal of Hydrology, 330: 604–620. doi:10.1016/j.jhydrol.2006.04.025.

Li, G., Chen, J., Peng, D., and Gu, X. 2021. Short communication: The lag response of daily milk yield to heat stress in dairy cows. Journal of Dairy Science, 104: 981–988. doi:10.3168/jds.2020-18183.

Li, Y., and Guo, Z. 2019. Comprehensive Division of Rock Mass Structure in Granite Area–Taking Tianhu Rock Mass as An Example. IOP Conference Series: Earth and Environmental Science, 304: 052053. IOP Publishing. doi:10.1088/1755-1315/304/5/052053.

Liesch, T., and Wunsch, A. 2019. Aquifer responses to long-term climatic periodicities. Journal of Hydrology, 572: 226–242. doi:10.1016/j.jhydrol.2019.02.060.

Lorenzo-Lacruz, J., Garcia, C., and Morán-Tejeda, E. 2017. Groundwater level responses to precipitation variability in Mediterranean insular aquifers. Journal of Hydrology, 552: 516–531. doi:10.1016/j.jhydrol.2017.07.011.

MacMillan, R.A., and Shary, P.A. 2009. Chapter 9 Landforms and Landform Elements in Geomorphometry. In Developments in Soil Science. Elsevier. pp. 227–254. doi:10.1016/S0166-2481(08)00009-3.

Mattivi, P., Franci, F., Lambertini, A., and Bitelli, G. 2019. TWI computation: a comparison of different open source GISs. Open Geospatial Data, Software and Standards, 4: 6. doi:10.1186/s40965-019-0066-y.

McMillan, H.K. 2021. A review of hydrologic signatures and their applications. WIREs Water, 8: e1499. doi:10.1002/wat2.1499.

Miles, O.W., and Novakowski, K.S. 2016. Large water-table response to rainfall in a shallow bedrock aquifer having minimal overburden cover. Journal of Hydrology, 541: 1316–1328. doi:10.1016/j.jhydrol.2016.08.034.

Mizan, S.A., Ahmed, S., and Selles, A. 2019. Spatial estimation of groundwater storage from a 2D specific yield in the crystalline aquifer of the Maheshwaram watershed. Journal of Earth System Science, 128: 185. doi:10.1007/s12040-019-1218-2.

Montgomery, D.R., and Foufoula-Georgiou, E. 1993. Channel network source representation using digital elevation models. Water Resources Research, 29: 3925–3934. doi:10.1029/93WR02463.

Moore, I.D., Grayson, R.B., and Ladson, A.R. 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. Hydrological Processes, 5: 3–30. doi:10.1002/hyp.3360050103.

Moritz, S., and Bartz-Beielstein, T. 2017. imputeTS: Time Series Missing Value Imputation in R. The R Journal, 9: 207. doi:10.32614/RJ-2017-009.

Nicolas, M., Bour, O., Selles, A., Dewandel, B., Bailly-Comte, V., Chandra, S., Ahmed, S., and Maréchal, J.-C. 2019. Managed Aquifer Recharge in fractured crystalline rock aquifers: Impact of horizontal preferential flow on recharge dynamics. Journal of Hydrology, 573: 717–732. doi:10.1016/j.jhydrol.2019.04.003.

Okkonen, J., and Kløve, B. 2010. A conceptual and statistical approach for the analysis of climate impact on ground water table fluctuation patterns in cold conditions. Journal of Hydrology, 388: 1–12. doi:10.1016/j.jhydrol.2010.02.015.

Olden, J.D., and Neff, B.D. 2001. Cross-correlation bias in lag analysis of aquatic time series. Marine Biology, 138: 1063–1070. doi:10.1007/s002270000517.

Padilla, C., Onda, Y., Iida, T., Takahashi, S., and Uchida, T. 2014. Characterization of the groundwater response to rainfall on a hillslope with fractured bedrock by creep deformation and its implication for the generation of deep-seated landslides on Mt. Wanitsuka, Kyushu Island. Geomorphology, 204: 444–458. doi:10.1016/j.geomorph.2013.08.024.

Parent, M., and Occhietti, S. 1999. Late Wisconsinan deglaciation and glacial lake development in the Appalachians of southeastern Québec. Géographie physique et Quaternaire, 53: 117–135. Les Presses de l'Université de Montréal. doi:10.7202/004859ar.

Peterson, T.J., Western, A.W., and Cheng, X. 2018. The good, the bad and the outliers: automated detection of errors and outliers from groundwater hydrographs. Hydrogeology Journal, 26: 371–380. doi:10.1007/s10040-017-1660-7.

Priest, S.D. 1993. Discontinuity Analysis for Rock Engineering. Springer Netherlands. doi:10.1007/978-94-011-1498-1.

Probst, W.N., Stelzenmüller, V., and Fock, H.O. 2012. Using cross-correlations to assess the relationship between time-lagged pressure and state indicators: an exemplary analysis of North Sea fish population indicators. ICES Journal of Marine Science, 69: 670–681. doi:10.1093/icesjms/fss015.

Rathay, S.Y., Allen, D.M., and Kirste, D. 2017. Response of a fractured bedrock aquifer to recharge from heavy rainfall events. Journal of Hydrology,. doi:10.1016/j.jhydrol.2017.07.042.

Riihimäki, H., Kemppinen, J., Kopecký, M., and Luoto, M. 2021. Topographic Wetness Index as a Proxy for Soil Moisture: The Importance of Flow-Routing Algorithm and Grid Resolution. Water Resources Research, 57: e2021WR029871. doi:10.1029/2021WR029871.

Rinderer, M., McGlynn, B.L., and van Meerveld, H.J. 2017. Groundwater similarity across a watershed derived from time-warped and flow-corrected time series. Water Resources Research, 53: 3921–3940. doi:10.1002/2016WR019856.

Rinderer, M., van Meerveld, H.J., and Seibert, J. 2014. Topographic controls on shallow groundwater levels in a steep, prealpine catchment: When are the TWI assumptions valid? Water Resources Research, 50: 6067–6080. doi:10.1002/2013WR015009.

Rinderer, M., van Meerveld, I., Stähli, M., and Seibert, J. 2016. Is groundwater response timing in a pre-alpine catchment controlled more by topography or by rainfall? Hydrological Processes, 30: 1036–1051. doi:10.1002/hyp.10634.

Rivers, T., Martignole, J., Gower, C.F., and Davidson, A. 1989. New tectonic divisions of the Grenville Province, Southeast Canadian Shield. Tectonics, 8: 63–84. doi:https://doi.org/10.1029/TC008i001p00063.

Rouleau, A., and Gale, J.E. 1985. Statistical characterization of the fracture system in the Stripa granite, Sweden. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 22: 353–367. doi:10.1016/0148-9062(85)90001-4.

Salve, R., Rempe, D.M., and Dietrich, W.E. 2012. Rain, rock moisture dynamics, and the rapid response of perched groundwater in weathered, fractured argillite underlying a steep hillslope. Water Resources Research, 48. doi:10.1029/2012WR012583.

Saveca, P.S.L., Abi, A., Stigter, T.Y., Lukas, E., and Fourie, F. 2022. Assessing groundwater dynamics and hydrological processes in the sand river deposits of the Limpopo River, Mozambique. Frontiers in Water, 3. doi:10.3389/frwa.2021.731642.

Scesi, L., and Gattinoni, P. 2010. Water Circulation in Rocks. Springer Netherlands. Available from https://www.springer.com/us/book/9789048124169 [accessed 8 May 2019].

Scesi, L., and Gattinoni, P. 2012. Methods and models to determine the groundwater flow in rock masses: Review and examples. Environmental Science, Engineering and Technology,: 1–54.

Shahbazi, A., Saeidi, A., and Chesnaux, R. 2020. A review of existing methods used to evaluate the hydraulic conductivity of a fractured rock mass. Engineering Geology, 265: 105438. doi:10.1016/j.enggeo.2019.105438.

Shumway, R.H., and Stoffer, D.S. 2017. Time Series Analysis and Its Applications: With R Examples. Springer.

Solinst Canada Ltd. 2021, October 19. Dataloggers & Telemetry Systems. Solinst Canada Ltd. Available from https://www.solinst.com/products/data/brochure-dataloggers-telemetry.pdf [accessed 22 March 2022].

Sun, J.P., Zhao, Z.Y., Li, J., and Fu, S. 2011. Influences of Fracture Aperture and Roughness on Hydraulic Conductivity in Fractured Rock Mass. Guangzhou, (China). pp. 372–374. doi:10.1063/1.3651921.

de Vries, J.J., and Simmers, I. 2002. Groundwater recharge: an overview of processes and challenges. Hydrogeology Journal, 10: 5–17. doi:10.1007/s10040-001-0171-7.

Walter, M.T., Brooks, E.S., McCool, D.K., King, L.G., Molnau, M., and Boll, J. 2005. Process-based snowmelt modeling: does it require more input data than temperature-index modeling? Journal of Hydrology, 300: 65–75. doi:10.1016/j.jhydrol.2004.05.002.

Wright, S.N., and Novakowski, K.S. 2019. Groundwater recharge, flow and stable isotope attenuation in sedimentary and crystalline fractured rocks: Spatiotemporal monitoring from multi-level wells. Journal of Hydrology, 571: 178–192. doi:10.1016/j.jhydrol.2019.01.028.

Wright, S.N., and Novakowski, K.S. 2022. Numerical analysis of midwinter infiltration along the soil-rock interface: A pathway for enhanced bedrock recharge. Advances in Water Resources, 166: 104261. doi:10.1016/j.advwatres.2022.104261.

Yue, S., Pilon, P., Phinney, B., and Cavadias, G. 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrological Processes, 16: 1807–1829. doi:10.1002/hyp.1095.

Zhang, L. 2013. Aspects of rock permeability. Frontiers of Structural and Civil Engineering, 7: 102–116. doi:10.1007/s11709-013-0201-2.

Zhou, G., Cui, M., Wan, J., and Zhang, S. 2021. A Review on Snowmelt Models: Progress and Prospect. Sustainability, 13: 11485. Multidisciplinary Digital Publishing Institute. doi:10.3390/su132011485.

Zhu, J. 2019. Effective hydraulic conductivity of discrete fracture network with aperture-length correlation. Geosciences Journal,. doi:10.1007/s12303-019-0025-8.

CHAPTER 5

DISCUSSIONS AND CONCLUSIONS

5.1 Structural domains and crystalline aquifers

It is well demonstrated in chapters 1 and 2, the advantages of using fracture based structural domains in the characterization of fractured rock for hydromechanical studies. These advantages lie on the capacity that fracture based structural domain concept has in considering the heterogeneity of the fracture networks in fractured rock. Since the introduction of homogeneity tests (also called similarity tests) between fracture samples by Miller (1983) and Mahtab and Yegulalp (1984), the concept of fracture based structural domain has been widely adopted in rock mechanics and geotechnics, especially in projects of tunnel engineering, dam construction and mining. Different structural domain methods are reported in chapter 2. Despite such wide adoption of structural domain methods in engineering, fracture based structural domains were mainly utilized in small-scale studies that goes from hundreds of square meters to a few square kilometers.

Since we have not encountered a single study in hydrogeology using fracture-based structural domain methods and most studies in the literature are limited to the presentation of structural domain methods without going further into the construction of conceptual models, it is deemed useful to discuss in this section the basis for such an exercise, especially for regional hydrogeological studies.

5.2 Conceptual model and controlling factors

In this thesis, not only a multiparameter similarity test between fracture samples was proposed, but also an objective method for delineating structural domains was developed.

The practitioner should also remember the major structural elements (see chapter 2) such as faults in the conceptual model, since they delineate the structural domains. Indeed, major structural elements can act as barriers or conduits for groundwater flow (Gleeson and Novakowski 2009, Bense et al. 2013). In the case where major structural elements do not exist in a given region, the use of Voronoi diagrams, as exemplified in Chapter 2, would be even more relevant, since it allows the region to be compartmentalized into zones of distinct fracture networks. It is worth mentioning that the use of Voronoi diagrams is independent of the similarity test utilized between the fracture samples. Therefore, one can combine the similarity test proposed by Miller (1983) or its derivative (Martin and Tannant 2004, Zhang et al. 2016b, Song et al. 2018, Zhou et al. 2018) to the Voronoi diagrams to delineate the structural domain.

In the case of existing major structural elements in the study region, the compartments delineated by the Voronoi diagrams would be split into sub compartments by the major structural elements. A simple conceptual model illustrating the mutual use of Voronoi diagrams and major structural elements is provided in Figure 5.1. Figure 5.1 illustrates a watershed underlain by a fractured rock with relatively thin sediment deposits and a major structural element (e.g., fault). It is supposed that an application of the methodology presented in chapter 2 led to the definition of the Voronoi diagrams as illustrated in Figure 2. The Voronoi diagrams have compartmentalized the watershed into three compartments *a*, *b* and *c*. However, since the major structural element passes through compartments *b* and *c*, each has been split in two sub compartments, namely (*b1,b2*) and (*c1,c2*), respectively, the sub compartments of *b* and *c*. In such circumstance, the fracture network in each sub compartment can be characterized according to the fracture samples in it. In the case of total absence of fracture samples in one of the sub compartments, the practitioner would report the fracture network characteristics of the other sub compartment in which the fracture samples were collected. The practitioner should be aware that a certain

homogeneity of the fracture network within the compartment is assumed. However, one could make the opposite assumption that there is no homogeneity. This debate would be inconclusive since the fracture network can never be described at 100% in a given study area. As highlighted in chapter 2, the more fracture samples the better the conceptual model to be built.

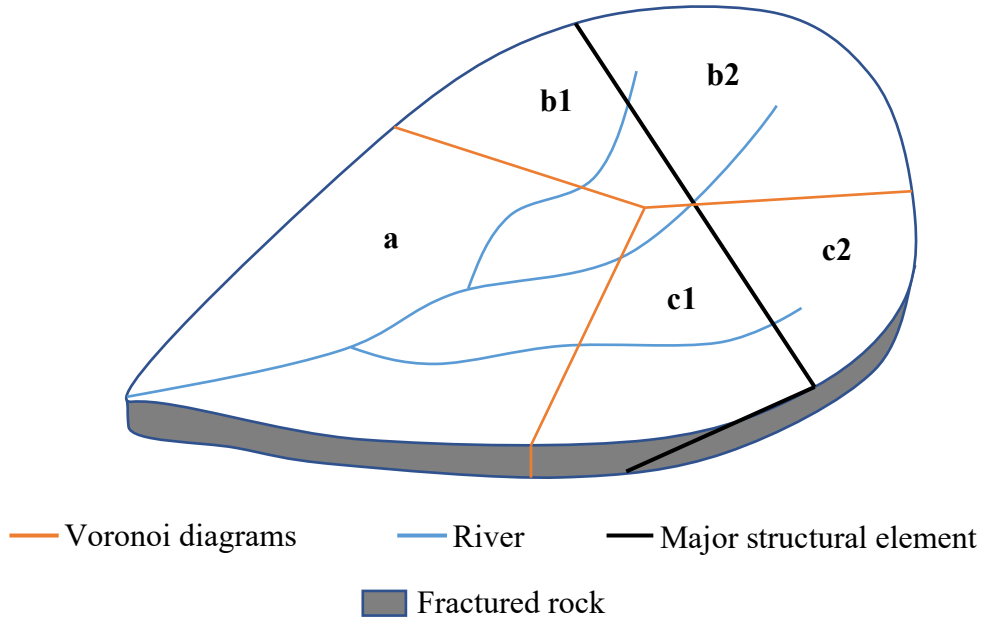


Figure 5.1. Illustration of a watershed with three compartments of fracture based structural domains, delineated with Voronoi diagrams and a major structural element. The letters designate the compartments.

The results obtained in chapter 4 showed that the fracture characteristics were not significantly linked to factors controlling the response time and groundwater level variability in regional crystalline aquifers. However, the landform (topography) and sediment deposits have a significant impact on the response time and groundwater level variability in regional crystalline aquifers. In this regards, the use of equivalent porous medium (EPM) instead of discrete fracture network (DFN) for groundwater modeling seems to be well adapted in regional crystalline aquifers. These findings agree with the common practice of assimilating crystalline aquifers into equivalent porous media (DesRoches et al. 2014, Miles and

Novakowski 2016). However, according to this thesis (chapter 2), it is suggested to apply the EPM with particular attention to structural domains and major structural elements. To do so, each structural domain (or sub compartment), represented in EPM, will be assigned its own specific hydraulic properties, including the anisotropy. By doing so, the heterogeneity and complexity of fracture networks in crystalline aquifers will be considered without compromising on the computational limitation of DFN models. In fact, the needed computational power to simulate very complex DFN models may be a limitation for large watersheds (Cook 2003). Given the significant impact of the landform (topography) and sediment deposits on the response time and groundwater level variability in regional crystalline aquifers, those two factors need to be well characterized and taken into accound in the model.

5.3 Exploited private boreholes in monitoring network

It is demonstrated in this thesis that exploited private boreholes can be used for groundwater level monitoring. However, their use is highly impacted by the water usage of the household and the aquifer characteristics. Although it is expected that most water usage will take place at day time, some water usage can still be found at night time. The speed at which the water table recovers depends on the aquifer characteristics. In this thesis, the pumping capacity of the pumps installed in the boreholes were unknown. Similarly, the hydraulic properties of the aquifers in which the boreholes are drilled are unknown to the users.

Despite these unknown pieces of information, the removal of the noises (outliers) generated by the frequent water pumping could be efficiently removed by a slope criterion (chapter 3), and the missing values satisfactorily imputed. From the results obtained in chapter 3, it seems that the use of linear interpolation is an adequate choice for the

imputation of missing values considering its simplicity of implementation and its performance compared to other methods. Linear interpolation has also given good results in the imputation of missing values of solar irradiation data (Demirhan and Renwick 2018). Therefore, a practitioner who implemented a monitoring network with no observation borehole could opt for the linear interpolation as far as large gap sizes of missing values are limited in the time series after the removal of outliers.

5.4 Outcomes and perspectives for future research

The contributions of this thesis have several outcomes in the form of the conceptualization and modeling of crystalline aquifers with complex fracture networks. Contributions are also in the field of data acquisition and data treatment of groundwater levels, as well as in land use planning and development. More precisely, as highlighted in sections 5.1 and 5.2, the use of fracture based structural domain concept allows building more realistic conceptual models and ultimately more accurate hydrogeological models. In groundwater monitoring, this thesis has proposed a new methodology allowing the use of the many exploited private boreholes that are unused for groundwater monitoring. This implementation avoids the drilling of boreholes required for groundwater monitoring, and also avoids the inconvenience of filling or sealing these boreholes at the end of research projects in order to prevent any risk of subsequent contamination. Another advantage of using exploited private boreholes in the monitoring network is the fact that groundwater level is directly monitored from the boreholes that generate the progressive depletion of the aquifer, which is useful to the decision maker for subsequent water management strategies.

This thesis also showed the relationship between crystalline aquifers groundwater level variability and response time to geomorphological factors and fracture network factors. However not all the interrelationship between the different parameters could be linked to

sound physical hydrogeological processes, especially between the fracture network parameters and the parameters related to the groundwater level variability. In fact, one limitation of the correlation method is that correlation doesn't mean causation. As such, two given parameters can be correlated but not obviously hydrogeological linked. Consequently, further studies are needed to better understand some of the interactions between fracture network parameters and hydrogeological indices in crystalline aquifers.

This thesis also opens up to several other perspectives for future research. To cite a few:

1. The use of structural domain concept, as described in sections 5.1 and 5.2, in groundwater flow model of crystalline aquifers with complex fracture networks. A theoretical complex fracture network could be built at the watershed scale with its corresponding DFN groundwater model of the crystalline aquifer. Since the structural domains are known, a comparison of the DFN and EPM model could be performed.
2. Several methods were created for the analysis of the similarity between fracture samples, allowing to identify if the samples belong to the same structural domain. However, small studies have addressed the delineation of those structural domains and the Voronoi diagrams proposed in this thesis is a tool among other mathematical tools.
3. It is shown on this thesis that exploited private boreholes can be used for groundwater monitoring and the results of imputing the missing values were very variable among the boreholes. These variations are related to the household water usage and aquifer properties. However, those two pieces of information were unknown to us. Thus, further studies combining modeling and field investigations would be required to elaborate guidelines for the implementation of a monitoring network with exploited private boreholes. For

instance, are the missing values linked to the number of inhabitants in the household, the number of rooms, the presence of a garden or swimming pool, or number of toilets. Or maybe it is mainly linked to the aquifer characteristics.

4. In this thesis, statistical approaches were used in the analysis of factors controlling the response time and groundwater level variability in regional crystalline aquifers. The use of a theoretical groundwater model to perform such an analysis is part of an ongoing research studying the response time and variability of groundwater levels within two structural domains in the same aquifer and overlain by a variety of sediment deposits. This allows investigating similarities and differences in the response of aquifer groundwater levels in two adjacent structural domains subjected to similar climatic conditions and sediment deposits.

5.5 References

Bense, V.F., Gleeson, T., Loveless, S.E., Bour, O., and Scibek, J. 2013. Fault zone hydrogeology. Earth-Science Reviews, 127: 171–192. doi:10.1016/j.earscirev.2013.09.008.

Cook, P.G. 2003. A Guide to Regional Groundwater Flow in Fractured Rock Aquifers. Seaview Press, Adelaide, South Australia. Available from https://publications.csiro.au/rpr/pub?list=BRO&pid=procite:054c4b83-a418-4317-a07d-509e3b62f9fa [accessed 12 October 2019].

Demirhan, H., and Renwick, Z. 2018. Missing value imputation for short to mid-term horizontal solar irradiance data. Applied Energy, 225: 998–1012. doi:10.1016/j.apenergy.2018.05.054.

DesRoches, A., Danielescu, S., and Butler, K. 2014. Structural controls on groundwater flow in a fractured bedrock aquifer underlying an agricultural region of northwestern New Brunswick, Canada. Hydrogeology Journal, 22: 1067–1086. doi:10.1007/s10040-014-1134-0.

Gleeson, T., and Novakowski, K. 2009. Identifying watershed-scale barriers to groundwater flow: Lineaments in the Canadian Shield. GSA Bulletin, 121: 333–347. doi:10.1130/B26241.1.

Mahtab, M.A., and Yegulalp, T.M. 1984. Similarity Test For Grouping Orientation Data In Rock Mechanics. American Rock Mechanics Association. Available from https://www.onepetro.org/conference-paper/ARMA-84-0495 [accessed 20 October 2020].

Martin, M.W., and Tannant, D.D. 2004. A technique for identifying structural domain boundaries at the EKATI Diamond Mine. Engineering Geology, 74: 247–264. doi:10.1016/j.enggeo.2004.04.001.

Miles, O.W., and Novakowski, K.S. 2016. Large water-table response to rainfall in a shallow bedrock aquifer having minimal overburden cover. Journal of Hydrology, 541: 1316–1328. doi:10.1016/j.jhydrol.2016.08.034.

Miller, S.M. 1983. A statistical method to evaluate homogeneity of structural populations. Journal of the International Association for Mathematical Geology, 15: 317–328. doi:10.1007/BF01036073.

Song, S., Sun, F., Zhang, W., Chen, J., Xu, P., Niu, C., Cao, C., and Zhan, J. 2018. Identification of structural domains by considering multiple discontinuity characteristics: a case study of the Songta Dam. Bulletin of Engineering Geology and the Environment, 77: 1589–1598. doi:10.1007/s10064-017-1024-5.

Zhang, W., Zhao, Q., Huang, R., Chen, J., Xue, Y., and Xu, P. 2016. Identification of structural domains considering the size effect of rock mass discontinuities: A case study of an underground excavation in Baihetan Dam, China. Tunnelling and Underground Space Technology, 51: 75–83. doi:10.1016/j.tust.2015.10.026.

Zhou, X., Chen, J., Ruan, Y., Zhang, W., Song, S., and Zhan, J. 2018. Demarcation of Structural Domains in Fractured Rock Masses Using a Three-Parameter Simultaneous Analysis Method. Advances in Civil Engineering, 2018: e9358098. Hindawi. doi:https://doi.org/10.1155/2018/9358098.

APPENDICES

Appendix 1: Mann-Whitney U test

The Mann-Whitney U test (MW-U test) is generally considered to be the nonparametric equivalent of the parametric t-test, as both tests are used for identifying discrepancies in central tendency between two samples (Hart 2001, Feltovich 2003, Krzywinski and Altman 2014). However, in terms of statistical power, Zimmerman (2011) found the MW-U test to be considerably more powerful than the t-test (a power difference of 0.1 to 0.3) when the samples' distributions were skewed, such as lognormal and exponential. Conversely, a power difference of 0.01 to 0.04 in favor of the t-test for symmetric distributions like the uniform and normal was reported (Zimmerman 2011). The statistical power is the probability to correctly reject the null hypothesis (H0) when it is false (Feltovich 2003, Kolassa 2020). In this study, the null hypothesis is that the two clusters from the adjacent samples are from the same fracture population. Thus, they have the same distribution for each fracture parameter.

Let us reconsider clusters I(i) and J(j), from samples I and J, containing respectively n and m number of fractures; let X1, …, Xn and Y1, …, Ym be the two independent fracture clusters comprising I(i) and J(j), respectively. The first step in the MW-U test consists of combining and arranging, from least to greatest, the two clusters (I(i) and J(j)) into a cluster of size N = n+m. Let us denote R (Ya), the rank of Ya in the combined cluster and S (Ya) be the rank of Ya among the elements Y1, …, Ym. Then the MW-U statistic can be defined as (Alvo and Yu 2018):

$$U(X,Y) = \sum_{a=1}^{m} R(Y_a) - \sum_{a=1}^{m} S(Y_a) = W - \frac{m(m+1)}{2} \qquad (Eq\ A.1)$$

where $W = \sum_{a=1}^{m} R(Y_a)$ represents the Wilcoxon statistics. It should be noted that for large fracture clusters (n and m > 20 (Siegel 1956)), the distribution of U (X, Y) quickly approaches the normal distribution (Siegel 1956, Alvo and Yu 2018). With a mean and variance (var) of:

$$mean = \mu_U = \frac{nm}{2} \qquad (Eq\ A.2)$$

$$Var = \frac{nm(N+1)}{12} \qquad (Eq\ A.3)$$

Therefore, to determine an estimate of the level of significance of an observed value U (p-value):

$$z = \frac{U - \mu_U}{\sqrt{var}} = \frac{U - \frac{nm}{2}}{\sqrt{\frac{nm(N+1)}{12}}} \qquad (Eq\ A.4)$$

The two-tailed p-value is then given by:

$$p - value = 2P\left(z \leq \frac{U - \frac{nm}{2}}{\sqrt{\frac{nm(N+1)}{12}}}\right) \qquad (Eq\ A.5)$$

As the MW-U test is based on ranks, corrections for eventual ties in the ranking process are required when they occur (Siegel 1956, Alvo and Yu 2018). In this case, changes occur in the variance equation (Eq A.2), while the mean does not change (Eq A.3). Thus, the equation for the estimate of the p-value (Eq A.5) will change accordingly. The new variance which corrects for ties is given by Alvo and Yu (2018) as:

$$Var_{ties} = \frac{nm(N+1)}{12} - \frac{nm\left(\sum_{i=1}^{k} d_i^3 - d_i\right)}{12N(N-1)} \qquad (Eq\ A.6)$$

where k represents the number of tied groups if we consider that observations with the same rank form a group of observation, while di is the number of observations tied in the ith tied group (i =1, 2, …, k).

By using the normal approximation, an estimate of the p-value may lead to inaccurate results (Mehta and Patel 2011). A more accurate method of obtaining the p-value is the exact test, which provides the exact p-value (Mehta and Patel 2011, Alvo and Yu 2018). The

exact test uses a resampling method, involving the permutation of the observed data (within a cluster), in all possible ways, so that a histogram of test statistic values (U value of each possible outcome) that estimates the null distribution can be created (Alvo and Yu 2018). The interested reader is directed to a more specialized book such as the one by Alvo and Yu, (2018). One downside of the exact test is that it is computationally intensive as, in the permutation process, all possible outcomes need to be computed (Mehta and Patel 2011, Alvo and Yu 2018). Therefore, in case of large data, the Monte Carlo method provides a good alternative as it does not compute all possible outcomes (Mehta and Patel 2011). In this research, all the reported p-values are obtained via the exact test or the Monte Carlo method. The used level of significance α, below which the null hypothesis is rejected, is 0.05.

Appendix 2 : Kolmogorov-Simonov test

Unlike the MW-U test, the two-tailed Kolmogorov-Smirnov test (KS test) is a more general test which is responsive to any kind of differences between the distribution functions of the two independent samples (Siegel 1956, Wilcox 1997, Bonnini et al. 2014). The differences are evaluated on the basis of the maximum vertical difference between the two empirical distribution functions (Siegel 1956). Thus, if the distribution functions have a large distance between each other at any point, in comparison to a critical distance of the KS distribution for a given level of significant α, it indicates that the samples may come from different populations (Siegel 1956). In terms of statistical power, the KS test is reported to have a minor advantage over the MW-U test for small samples, while the MW-U test has a minor advantage over the KS test for large samples (Siegel 1956).

Let X1, …, Xn and Y1, …, Ym be the two independent fracture clusters composing I(i) and J(j), respectively. Let us consider $F_n(x)$ and $F_m(x)$ the respective cumulative empirical distribution function of X1, …, Xn and Y1, …, Ym, such as

$$\begin{cases} F_n(x) = \dfrac{number\ of\ samples\ X's \leq x}{n} \\ F_m(x) = \dfrac{number\ of\ samples\ Y's \leq x}{m} \end{cases} \qquad (Eq\ A.7)$$

The KS test concentrates on the largest observed departure between $F_n(x)$ and $F_m(x)$ and is formulated as (Siegel 1956, Wilcox 1997):

$$D = maximum|F_n(x) - F_m(x)| \qquad (Eq\ A.8)$$

where D is the Kolmogorov-Smirnov distance function. A good estimate of the p-value is provided by the exact test, which is described in the previous section. In this research, all the reported Kolmogorov-Smirnov two-tailed test p-values are estimated by means of the exact test. Given that the chosen level of significance α is 0.05 and the null

hypothesis is that the two samples belong to the same population, the null hypothesis is rejected for all p-values that are below 0.05.

Appendix 3 : Results of the orientation similarity tests between clusters

The results of the orientation similarity tests between clusters I(i) and J(j), respectively, from adjacent samples I and J are shown in the Table below. R and FTR stand for rejected and failed to be rejected, respectively.

Fracture sets I(i)–J(j)	Angle between I (i) and J(j)	Cone of confidence I(i)	Cone of confidence J(j)	Orientation Similarity test
2(1)–3 (1)	22,6	11,7	7	R
27(2)–3 (1)	4	15,2	7	FTR
22(1)–3 (1)	36,3	6,2	7	R
22(1)–2 (1)	22,1	6,1	11,7	R
23(2)–4 (1)	27,3	7,2	7	R
23(2)–27 (2)	28,2	7,2	15,2	R
23(1)–7 (1)	13,15	7	13,3	FTR
23(2)–1 (1)	20,2	7,2	5,7	R
20(2)–24 (3)	4	7,2	7,8	FTR
25(1)–26 (2)	6	5,5	6,8	FTR
24(4)–26(1)	4,47	6,02	6	FTR
20(2)–25(1)	11,7	7,2	5,5	R
24(1)–2(1)	6,4	5,8	7	FTR
18(1)–19(1)	9,2	5,8	5,8	R
15(1)–19(1)	5	14,5	5,8	FTR
15(1)–18(1)	7,9	14,5	5,8	FTR
16(1)–17(1)	5,4	7,6	7,5	FTR
16(3)–6(1)	9	8,3	7,7	R
14(1)–16(5)	9,4	9,5	6,7	FTR
14(1)–17(2)	7	9,5	9,7	FTR
15(2)–16(1)	9	13,6	7,6	FTR
16(3)–18(1)	12	8,3	5,8	R
14(1)–13(1)	7,61	9,5	7	FTR
17(2)–13(1)	8,9	9,7	7	FTR
16(1)–13(2)	3	7,6	8,6	FTR
13(1)–1(1)	13,4	7	5,7	R
13(1)–12 (1)	5,1	7	5,6	FTR
14(1)–11 (1)	5,4	9,5	8,7	FTR
14(1)–12 (1)	4,4	9,5	5,6	FTR

12(1)–11 (1)	6	5,6	8,7	FTR
13(1)–30 (3)	16,1	7	3,8	R
12(4)–30 (2)	8	11,4	4,7	FTR
10(2)–30 (1)	4,2	7,3	3,8	FTR
11(2)–30 (5)	9,9	8,8	5	R
21(3)–28 (1)	10	7,8	6,3	R
8(1)–21 (3)	9,4	6,8	7,8	R
10(2)–8 (1)	3	7,3	6,8	FTR
21(3)–29 (2)	14,2	7,8	4,7	R
SB6(2)-3 (1)	3,1	7,7	6,9	FTR
SB6(2)–27 (2)	2,21	7,7	15,1	FTR
SB6(3)–RF1(1)	10	6,7	4,7	R
SB6(3)–SB5(1)	3,2	6,7	5,3	FTR
6(1)–FE04(1)	7	7,7	6,5	FTR

Appendix 4 : Results of the first iteration multiparametric similarity test

The results of the first iteration multiparametric similarity test, using fracture orientation, persistence, spacing and aperture parameters is shown in the Table below. KS test and MW-U test refer to the Kolmogorov-Smirnov test and the Mann-Whitney U test. R and FTR stand for rejected and failed to be rejected, respectively. P-values in bold are higher than 0.05.

	Persistence-p values		Spacing-p values		Aperture-p values			
Fracture set I(i)–J(j)	**KS test**	**MW-U test**	**KS test**	**MW-U test**	**KS test**	**MW-U test**	**Orientation Similarity test**	**Overall Similarity**
27(2)–3 (1)	**0,079**	**0,171**	**0,121**	**0,164**	**0,231**	**0,284**	FTR	FTR
23(1)–7(1)	0,042	0,027	**0,522**	**0,550**	2,8 E-4	1,1E-5	FTR	R
20(2)–24(3)	**0,055**	**0,076**	**0,657**	**0,686**	**1,000**	**0,912**	FTR	FTR
25(1)–26(2)	0,018	0,020	**0,222**	**0,112**	**0,549**	**0,618**	FTR	R
24(4)–26(1)	**0,402**	**0,747**	**0,886**	**0,810**	**0,741**	**0,983**	FTR	FTR
24(1)–2(1)	**0,241**	**0,284**	**0,695**	**0,610**	0,001	0,002	FTR	R
15(1)–19(1)	5,6 E-5	6,8 E-5	1,3 E-5	1,3 E-4	0,006	0,001	FTR	R
15(1)–18(1)	2,4 E-5	7,8 E-5	4,9 E-4	0,001	0,003	0,001	FTR	R
16(1)–17(1)	**0,422**	**0,213**	**0,538**	**0,510**	**0,813**	**0,250**	FTR	FTR
14(1)–16(5)	**0,689**	**0,833**	**0,121**	**0,109**	**0,240**	**0,065**	FTR	FTR
14(1)–17(2)	**0,249**	**0,228**	**0,909**	**0,755**	**1,000**	**0,852**	FTR	FTR
15(2)–16(1)	**0,856**	**0,482**	**0,273**	**0,256**	**0,100**	**0,100**	FTR	FTR
14(1)–13(1)	**0,052**	**0,210**	**0,726**	**0,967**	**1,000**	**0,860**	FTR	FTR
17(2)–13(1)	**0,092**	**0,765**	**0,930**	**0,879**	**1,000**	**0,966**	FTR	FTR
16(1)–13(2)	**0,137**	**0,628**	**0,990**	**0,882**	0,045	**0,119**	FTR	R

13(1)–12(1)	9,3 E-7	4,8 E-5	**0,445**	**0,595**	**0,822**	**0,584**	FTR	R
14(1)–11(1)	**0,096**	**0,084**	**0,959**	**0,929**	**0,699**	**0,515**	FTR	FTR
14(1)–12(1)	3,6 E-5	3,4 E-4	**0,908**	**0,886**	**1,000**	**0,821**	FTR	R
12(1)–11 (1)	2,3 E-7	8,9 E-6	**0,301**	**0,331**	**0,398**	**0,187**	FTR	R
12(4)–30 (2)	5,0 E-3	4,7 E-2	**0,075**	0,035	0,043	0,046	FTR	R
10(2)–30 (1)	**0,234**	**0,917**	0,013	**0,056**	0,044	**0,161**	FTR	R
10(2)–8 (1)	0,000	0,000	0,005	0,009	**0,083**	**0,280**	FTR	R
SB6(2)–3 (1)			**0,617**	**0,267**	0,032	**0,149**	FTR	R
SB6(2)–27 (2)			**0,573**	**0,665**	**1,000**	**1,000**	FTR	FTR
SB6(3)–SB5(1)			**0,314**	**0,207**	0,022	0,022	FTR	R
6(1)–FE04(1)			**0,227**	**0,197**	0,010	0,049	FTR	R

Appendix 5 : Fracture clusters and fracture characteristics of structural domains

The fracture clusters and fracture characteristics of structural domains obtained by the multiparametric similarity test using fracture orientation, persistence, spacing, and aperture parameter. The characteristics of structural domains correspond to the compartmentalization illustrated in Figure 2.9.

	Compartments	Clusters	Dip (°)	Dip direction (°)	Persistence (m)	Aperture (mm)	Spacing (m)
Groups of samples	**G3**	1	35±18	267±18	2─6	0,1─0,1	0,3─1,4
		2	75±18	352±18	1─5	0,1─3	0,1─4,5
	G13	1	87±21	207±21	0,5─4	0,1─2	0,1─9
		2	86±19	246±19	0,5─4	0,1─4	0,1─8
		3	88±20	134±20	0,5─4	0,1─3	0,2─5
		4	17±25	331±25	0,8─4	0,1─3	1,4─3,4
	G20	1	89±21	191±21	0,6─6	0,1─0,1	0,1─5
		2	74±18	43±18	0,5─4	0,1─2	0,1─3
		3	31±22	253±22	0,8─6,5	0,1─0,1	0,1─1
		4	84±10	164±10	0,5─6,2	0,1─0,1	0,1-2
Uncombined samples	**1**	1	72±10	201±10	1,5─5		0,2 ± 2,2
	2	1	89±17	306±17	1─3	1─3	0,1─2
	4	1	77±13	174±13	0,4─1,5	0,1─2	0,1─1,2
	5	1	70±23	171±23	0,6─4	1─3	0,2─3
		2	8±27	311±27	1─4	0,1─2	0,3─1
	6	1	73±19	133±19	1─5	0,1─1	0,1─5
		2	46±14	198±14	1─2	0,1─0,1	2,3─4,7
		3	43±21	334±21	1,5─3	0,1─2	0,4─0,8

	4	74±12	51±12	1,5─2	0,1─1	0,3─0,3
7	1	84±19	278±19	0,5─1,5	0,5─2	0,3─1,3
8	1	89±27	177±27	0,15─2	0,1─2	0─2,2
	2	80±24	250±24	0,3─1,8	0,1─2	0,4─3
	3	46±18	175±18	0,5─2	1─1	0,3─5
	4	51±11	6±11	0,15─1,5	0,1─1	0,7─6
9	1	46±6	50±6	1,2─4	0,1─2	1,6─7
	2	3±29	245±29	1,5─5	0,1─2	0,2─1,7
10	1	5±17	146±17	1,7─8	1─2	2,7─8,7
	2	90±22	180±22	0,7─4	0,1─2	0,1─7,4
	3	87±4	260±4	1,8─3	2─2	3,4─3,4
11	1	88±27	22±27	0,5─2,5	0,1─1	0,1─6
	2	79±20	266±20	0,6─2	0,1─1	0,1─3,4
	3	16±17	328±17	2,5─4	1─2	2─3
12	1	90±20	207±20	0,5─7,5	0,1─2	0,1─8
	2	83±18	242±18	0,3─4	0,1─3	0,3─4,2
	3	5±20	291±20	0,5─6	0,1─3	0,5─9,8
	4	84±23	328±23	1─4	0,1─2	0,7─7
18	1	83±15	148±15	0,5─1,8	0,1─1	0,1─1,6
	2	81±20	359±20	0,5─2	0,1─2	0,15─1,4
	3	87±20	45±20	0,2─2	0,1─2	0,1─5
19	1	31±20	182±20	0,2─2	0,1─1	0,03─1,9
21	1	89±21	61±21	0,8─3	0,1─2	0,13─4,5
	2	78±15	32±15	0,5─3	0,1─3	0,1─7,4
	3	82±20	3±20	0,7─3	0,1─2	0,1─5,13
22	1	83±20	316±20	0,7─8	0,1─0,1	0,01─1,8
23	1	82±23	101±23	0,5─20	0,1─2	0,03─1,4
	2	87±23	18±23	1─20	0,1─0,1	0,03─2,4
25	1	86±20	36±20	1─17	0,1─60	0,05─2,7

28	1	87±21	18±21	0,34─6	0,1─30	0,14─3,15
	2	80±18	239±18	0,7─4	0,1─1,5	0,85─14,9
	3	17±25	211±25	1,2─18,2	0,1─2	0,6─8
29	1	88±20	289±20	0,25─5	0,1─10	0,02─5,3
	2	73±23	13±23	0,25─3	0,1─40	0,02─3,5
30	1	87±23	1±23	0,6─8,6	0,1─10	0,08─16,9
	2	89±27	130±27	0,45─12	0,1─10	0,01─15,3
	3	90±27	47±27	0,3─4,5	0,1─10	0,02─11,2
	4	24±27	244±27	1,75─10,8	0,1─200	0,4─35,6
	5	87±22	270±22	0,5─4,5	0,1─10	0,04─19
SB5	1	31±19	254±19		0─249	0,03─5,9
RF1	1	34±23	230±23		0─68	0,02─11,27
	2	70±22	26±22		0─0	0,04─7,38
	3	66±25	164±25		0─73	0,01─4,53
	4	60±23	323±23		0─214	0,02─7,45
FE04	1	13±28	253±28		0─15	0,01─3,03
	2	65±18	314±18		0─12	0,01─3,05

Appendix 6 : Borehole monitoring periods

Table S1. Descriptive statistics of the groundwater level time-series data recorded from the 20 boreholes; SD, standard deviation; N, number of recorded data points; n, number of missing values.

Boreholes	*N*	Groundwater level (m)				Missing values	
		Min	Max	Mean	SD	*n*	%
D1	62456	13.60	16.74	15.37	0.45	142	0.2
D2	62468	1.82	5.91	5.15	0.38	244	0.4
D3	62441	0.00	6.97	5.30	0.79	116	0.2
D4	56410	0.00	6.18	4.91	0.55	162	0.3
D5	58229	0.00	19.62	15.70	2.29	461	0.8
D6	54216	7.39	15.90	15.20	0.53	227	0.4
D7	31751	0.01	6.79	5.34	0.78	0	0.0
D8	62454	4.19	16.79	16.11	0.63	243	0.4
D9	62543	12.66	15.23	13.85	0.53	40	0.1
D10	62542	14.64	39.37	27.99	8.30	40	0.1
D11	62478	14.68	18.20	16.57	0.83	226	0.4
D12	64949	15.33	17.06	15.87	0.36	412	0.6
D13	62437	0.00	17.23	14.75	1.49	518	0.8
D15	54018	0.00	5.49	4.66	0.62	231	0.4
D16	54104	1.66	6.15	5.08	0.39	146	0.3
D17	28005	4.85	5.68	5.20	0.22	0	0.0
D20	54109	11.83	15.91	15.14	0.33	241	0.4
D22	54141	0.00	17.68	16.59	0.83	372	0.7
D23	54105	15.11	16.36	15.58	0.26	291	0.5
D24	54116	8.76	17.00	14.94	0.70	241	0.4

Appendix 7 : Correlation matrix between the time series of exploited and unexploited boreholes

The correlations were calculated after the removal of the outliers in the exploited private borehole time series.

Table S2. Correlation matrix between the time series of exploited and unexploited boreholes

		Unexploited boreholes					
		D9	**D11**	**D12**	**D17**	**D22**	**D23**
Exploited private boreholes	**D1**	0,712	0,927	0,921	0,880	0,908	0,911
	D2	0,812	0,884	0,928	0,871	0,920	0,926
	D3	0,768	0,912	0,936	0,871	0,927	0,913
	D4	0,740	0,915	0,932	0,850	0,926	0,914
	D5	0,540	0,556	0,765	0,811	0,768	0,717
	D6	0,780	0,956	0,791	0,780	0,787	0,796
	D7	0,740	0,893	0,924	0,812	0,915	0,896
	D8	0,678	0,900	0,709	0,734	0,704	0,713
	D10	0,810	0,823	0,759	0,909	0,754	0,776
	D13	0,632	0,873	0,890	0,773	0,884	0,854
	D15	0,488	0,870	0,736	0,585	0,734	0,706
	D16	0,657	0,925	0,903	0,751	0,900	0,883
	D20	0,846	0,954	0,903	0,880	0,898	0,907
	D24	0,785	0,918	0,889	0,876	0,884	0,892

Appendix 8 : Borehole gap size distributions

One should read the following Figures as follows : For instance, of the 40 most common gap sizes in the borehole time series D20, the most common gap size was that of a single missing value, followed by a gap of two missing values, then three, and so forth (Figure S13). However, the gap size that generated the most missing values in the data set was that of 23 consecutive missing values, followed by a gap size of 19 consecutive missing values.

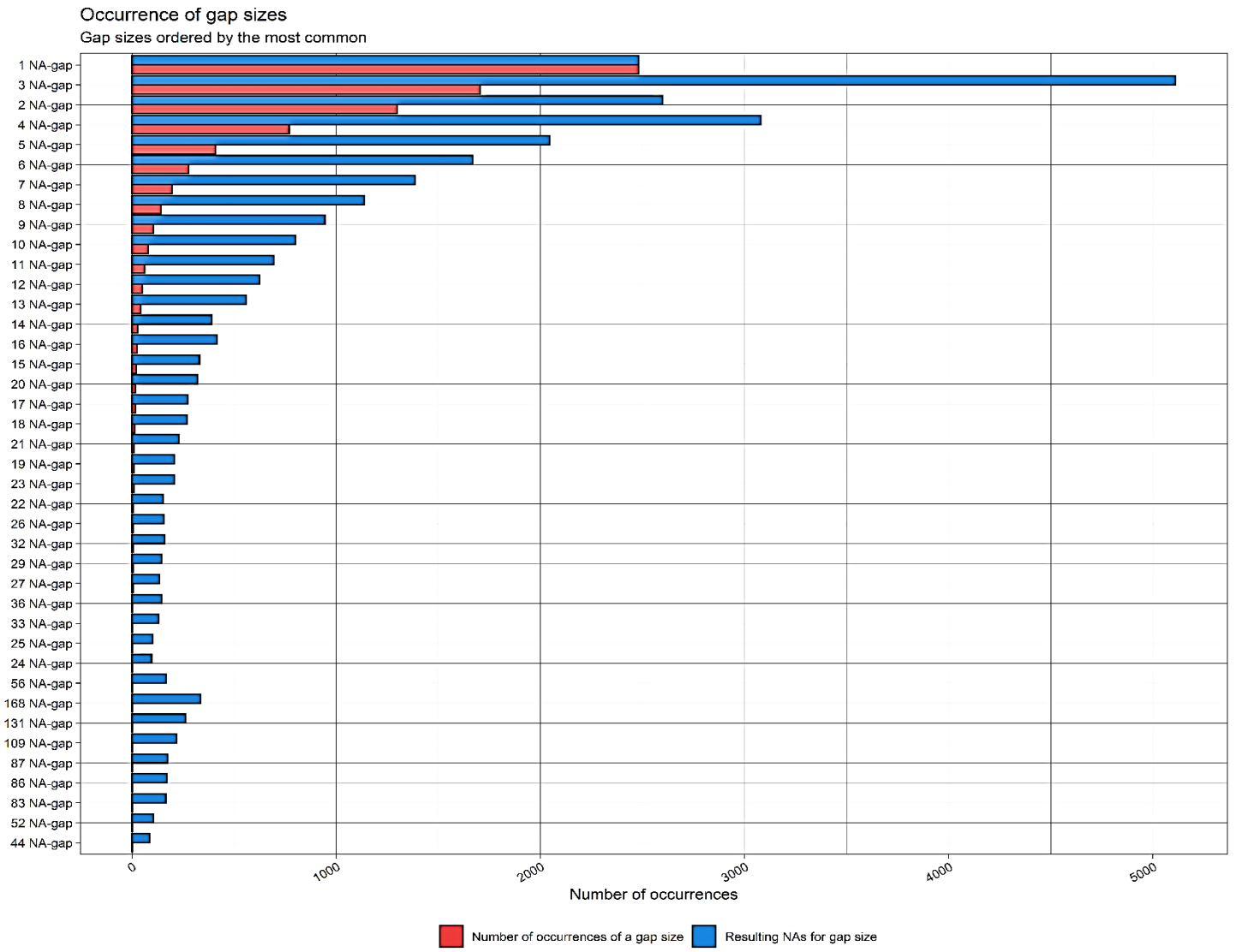


Figure S1. Distribution of the 40 most common gap sizes in the borehole data set D1.

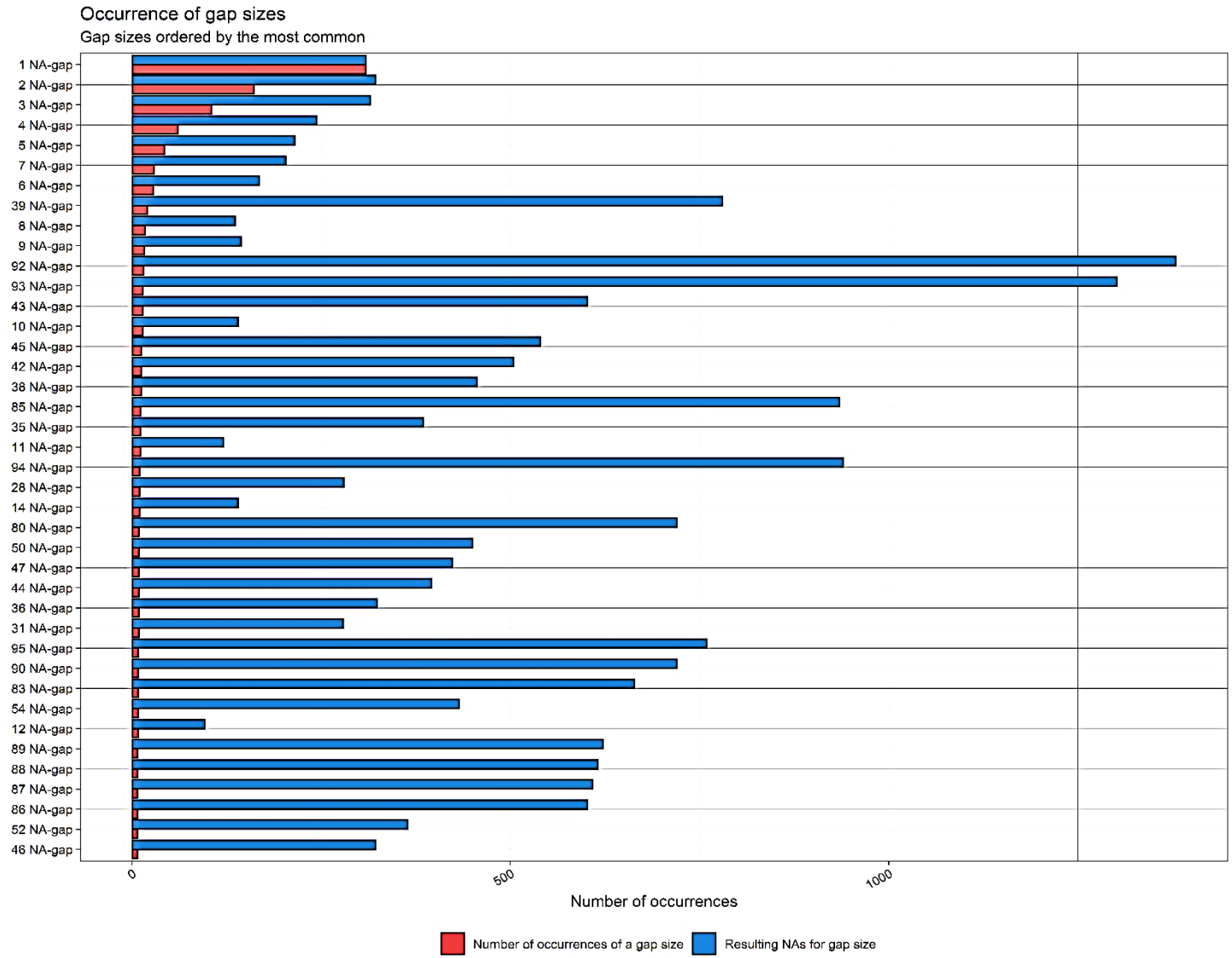


Figure S2. Distribution of the 40 most common gap sizes in the borehole data set D2.

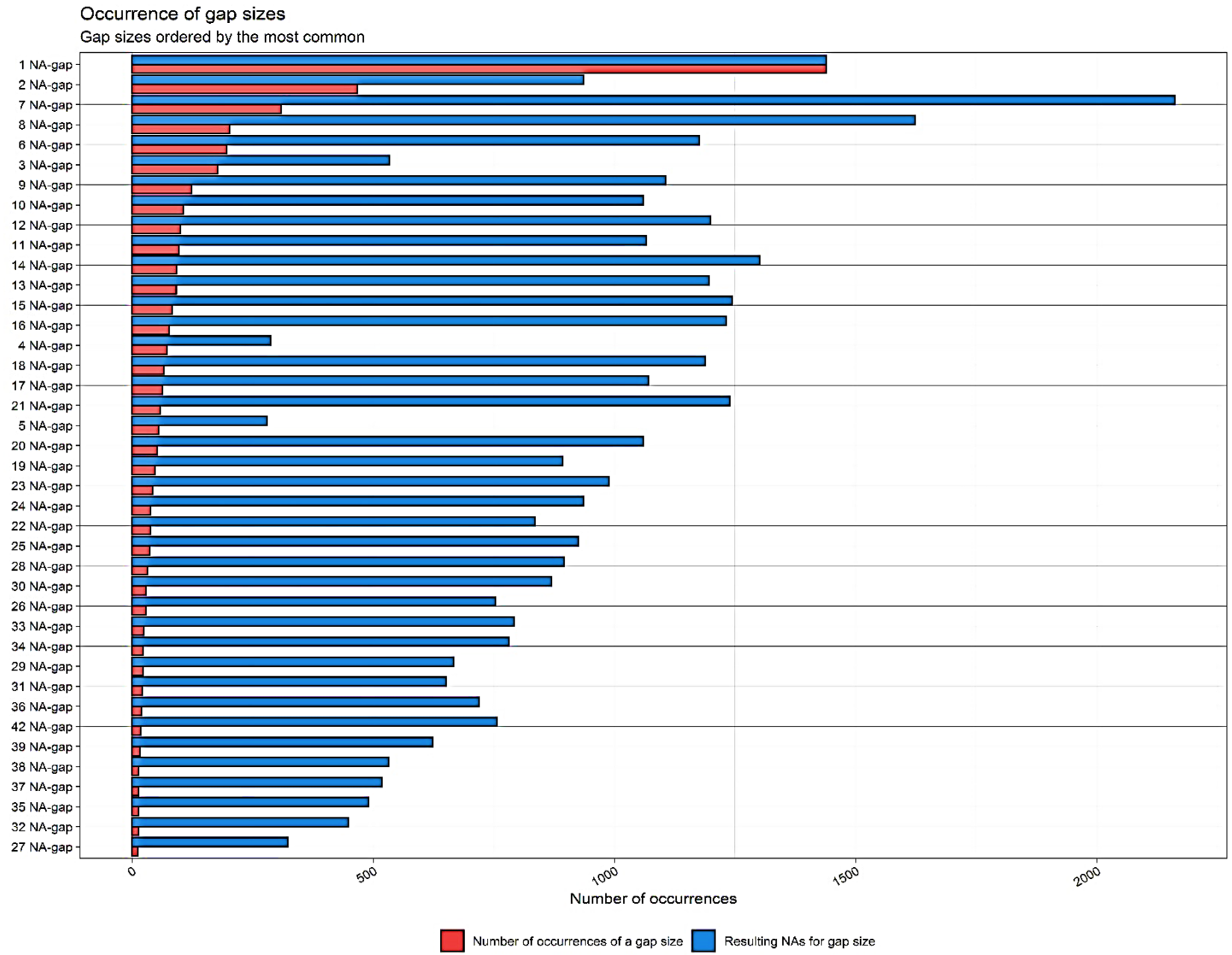


Figure S3. Distribution of the 40 most common gap sizes in the borehole data set D3.

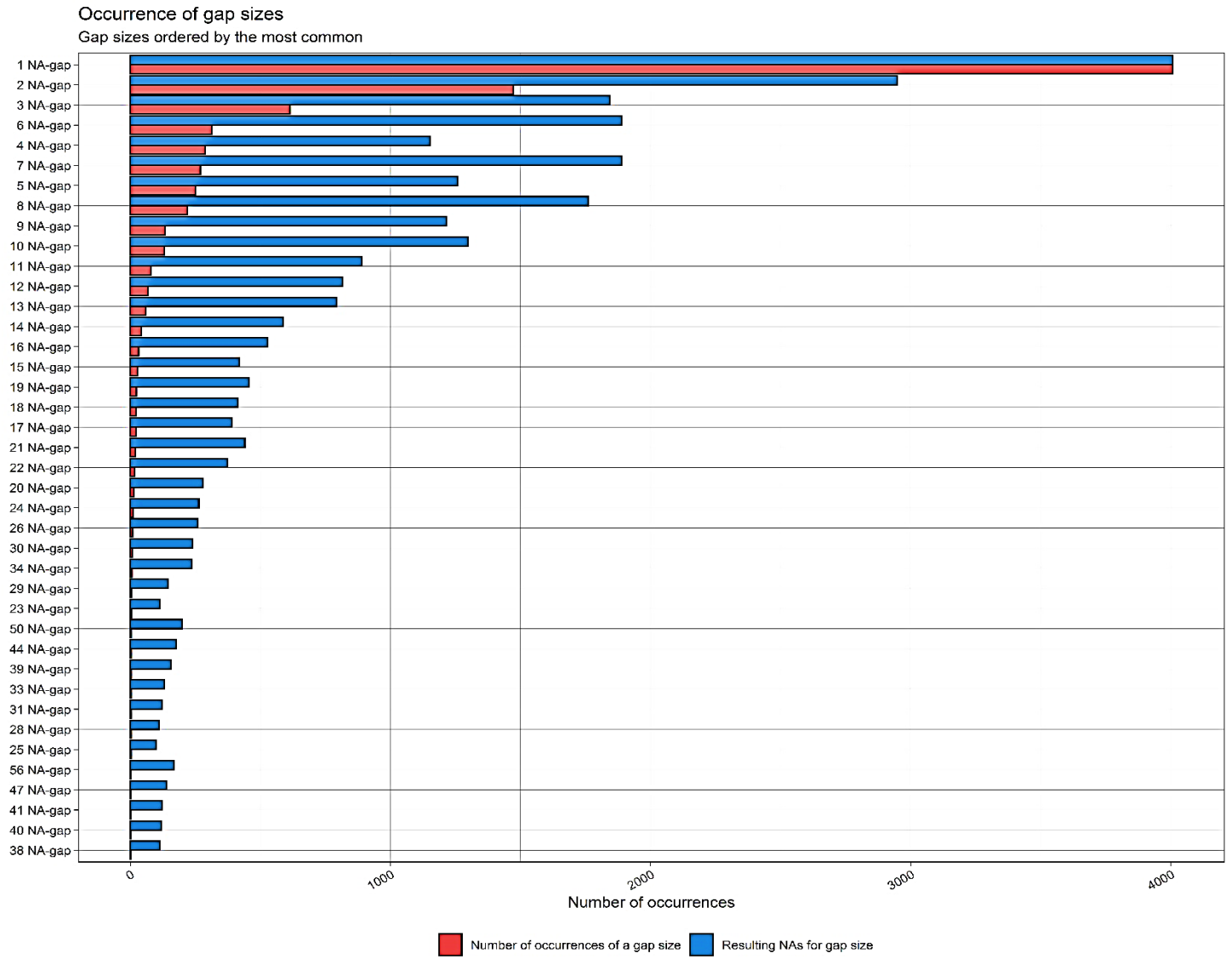


Figure S4. Distribution of the 40 most common gap sizes in the borehole data set D4.

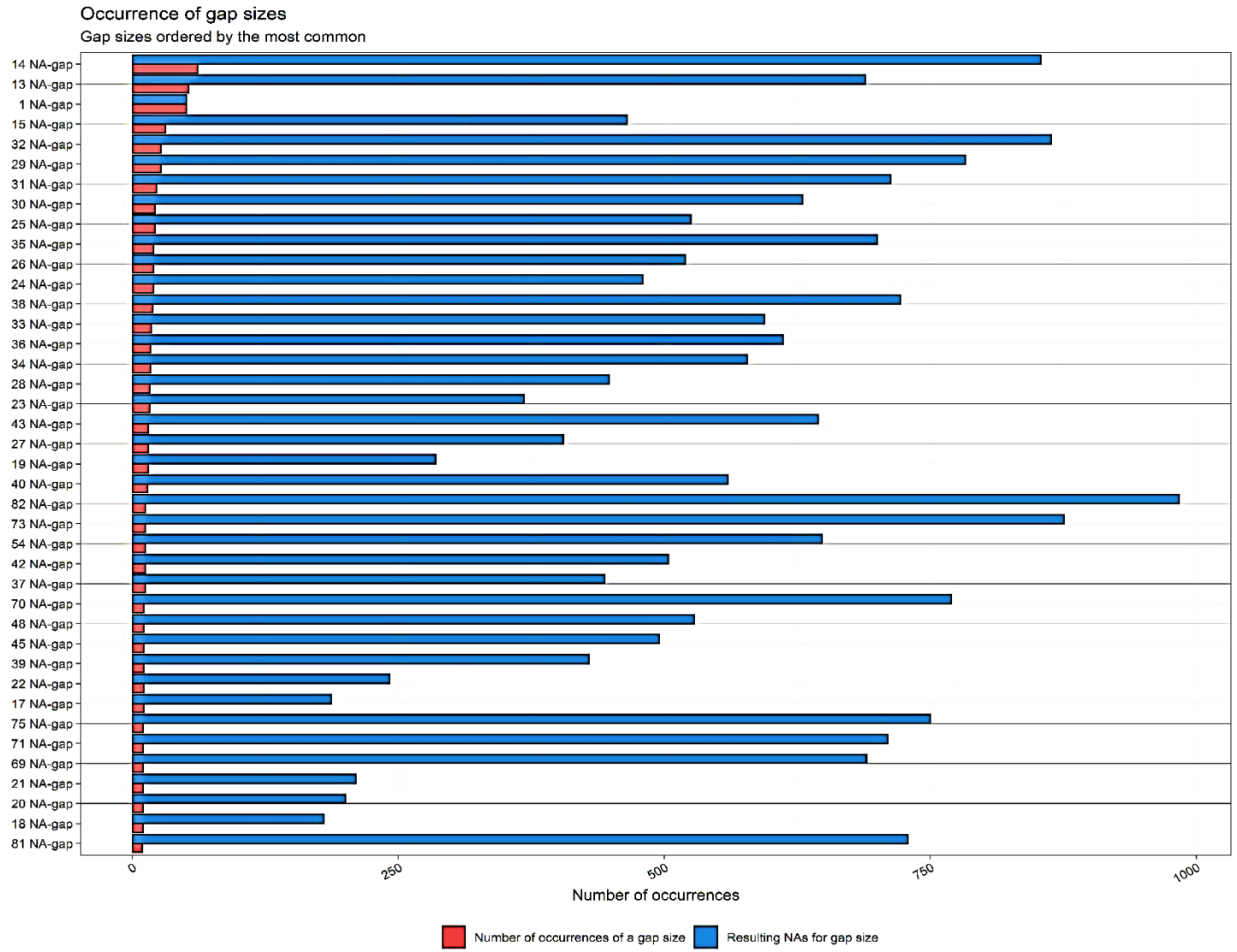


Figure S5. Distribution of the 40 most common gap sizes in the borehole data set D5.

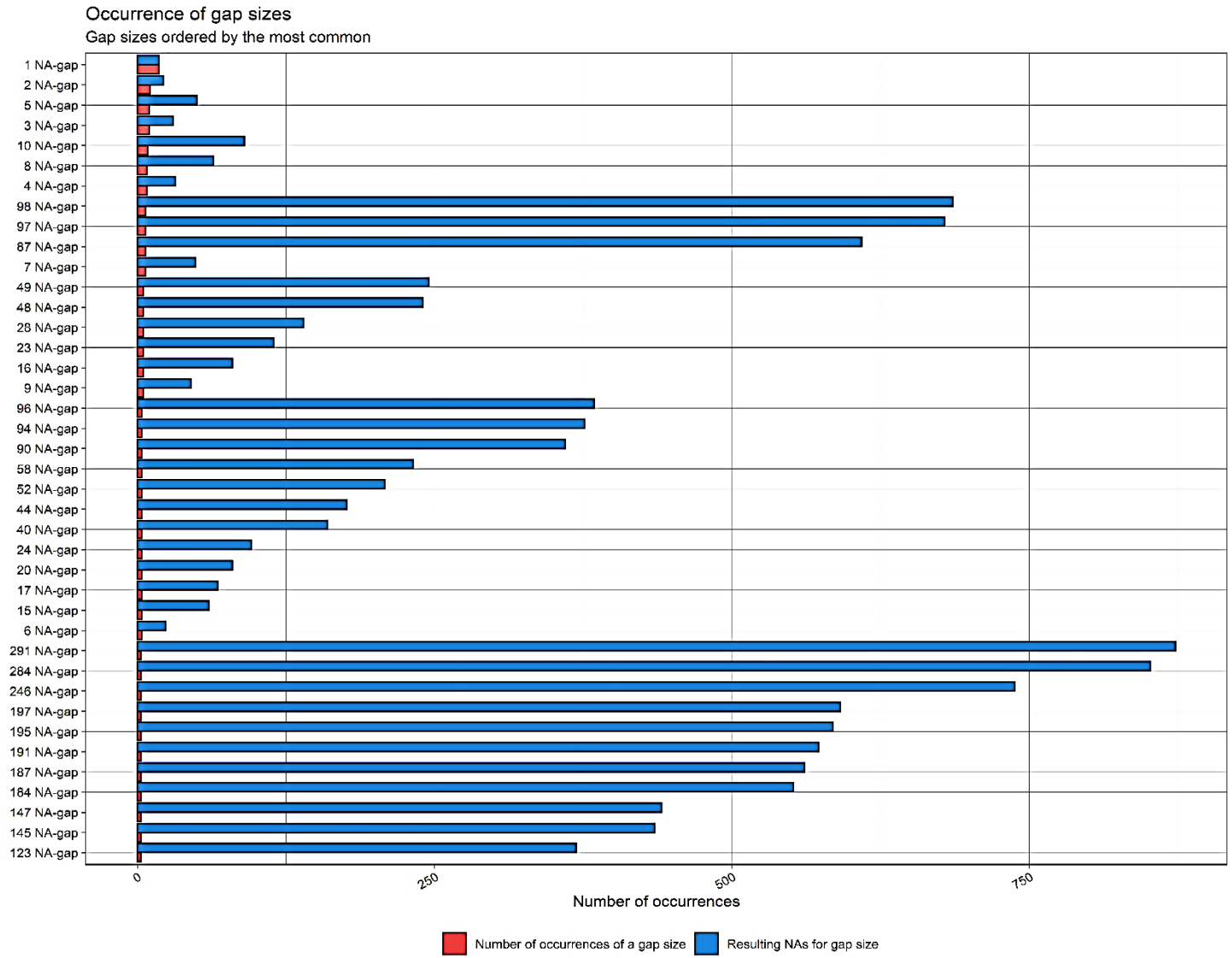


Figure S6. Distribution of the 40 most common gap sizes in the borehole data set D6.

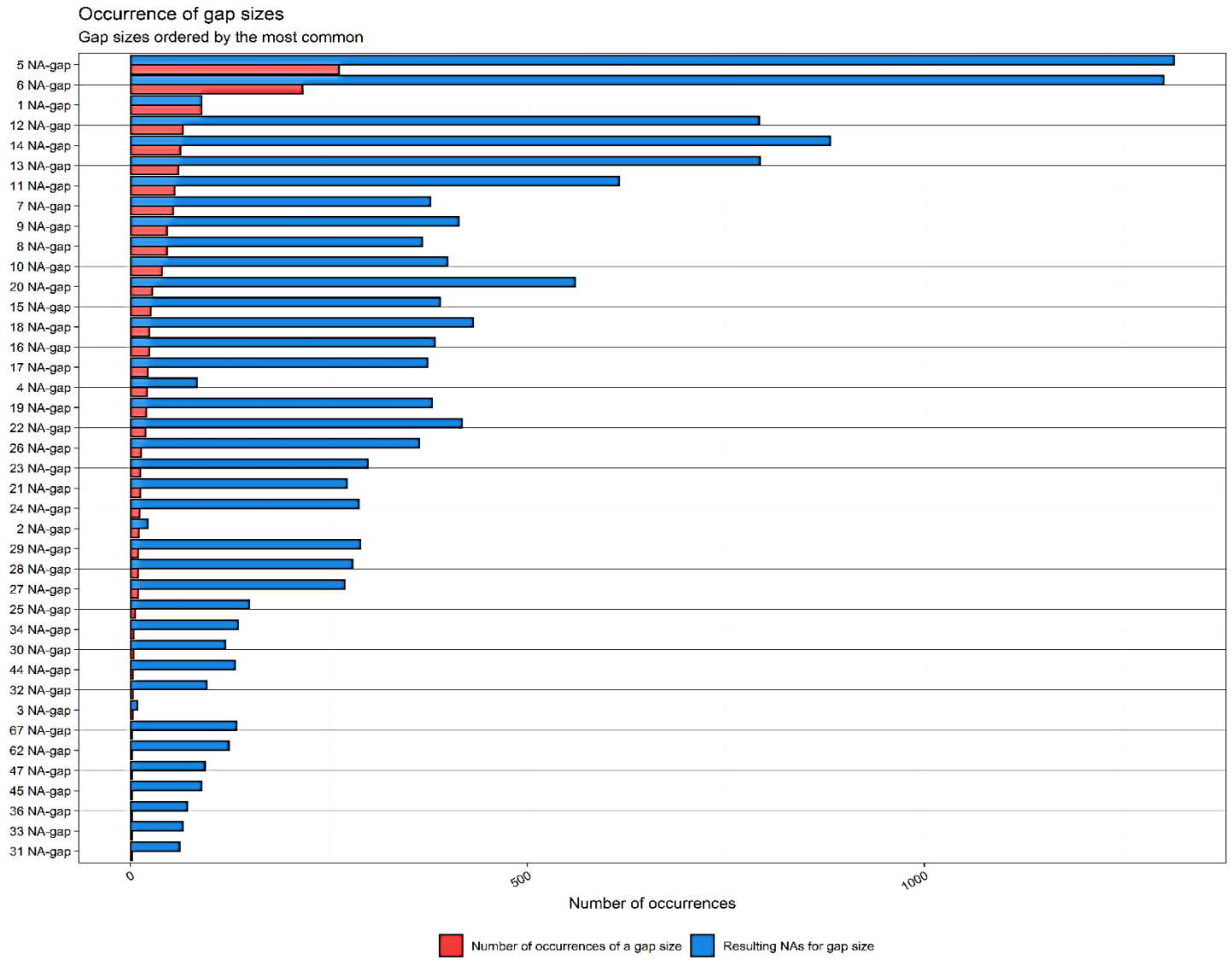


Figure S7. Distribution of the 40 most common gap sizes in the borehole data set D7.

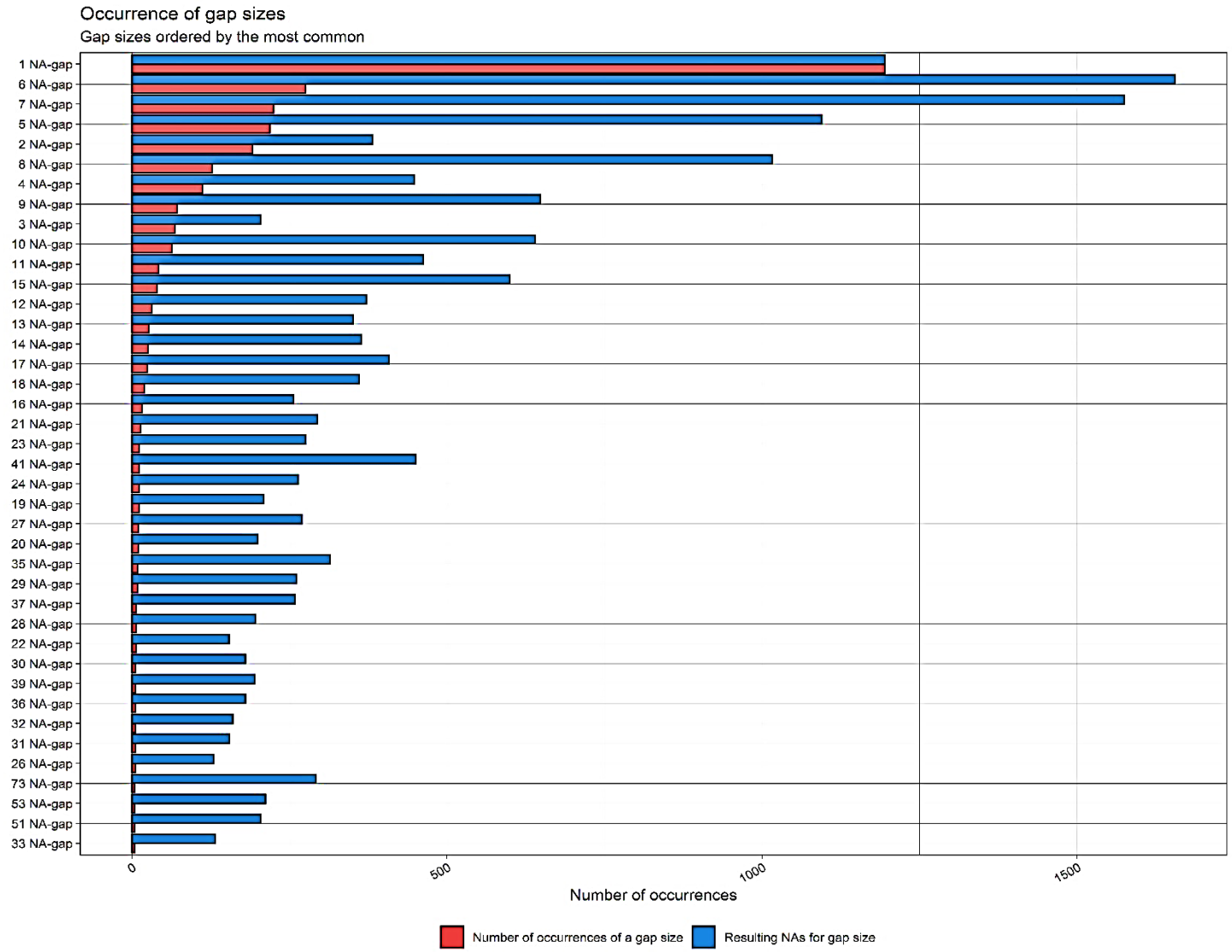


Figure S8. Distribution of the 40 most common gap sizes in the borehole data set D8.

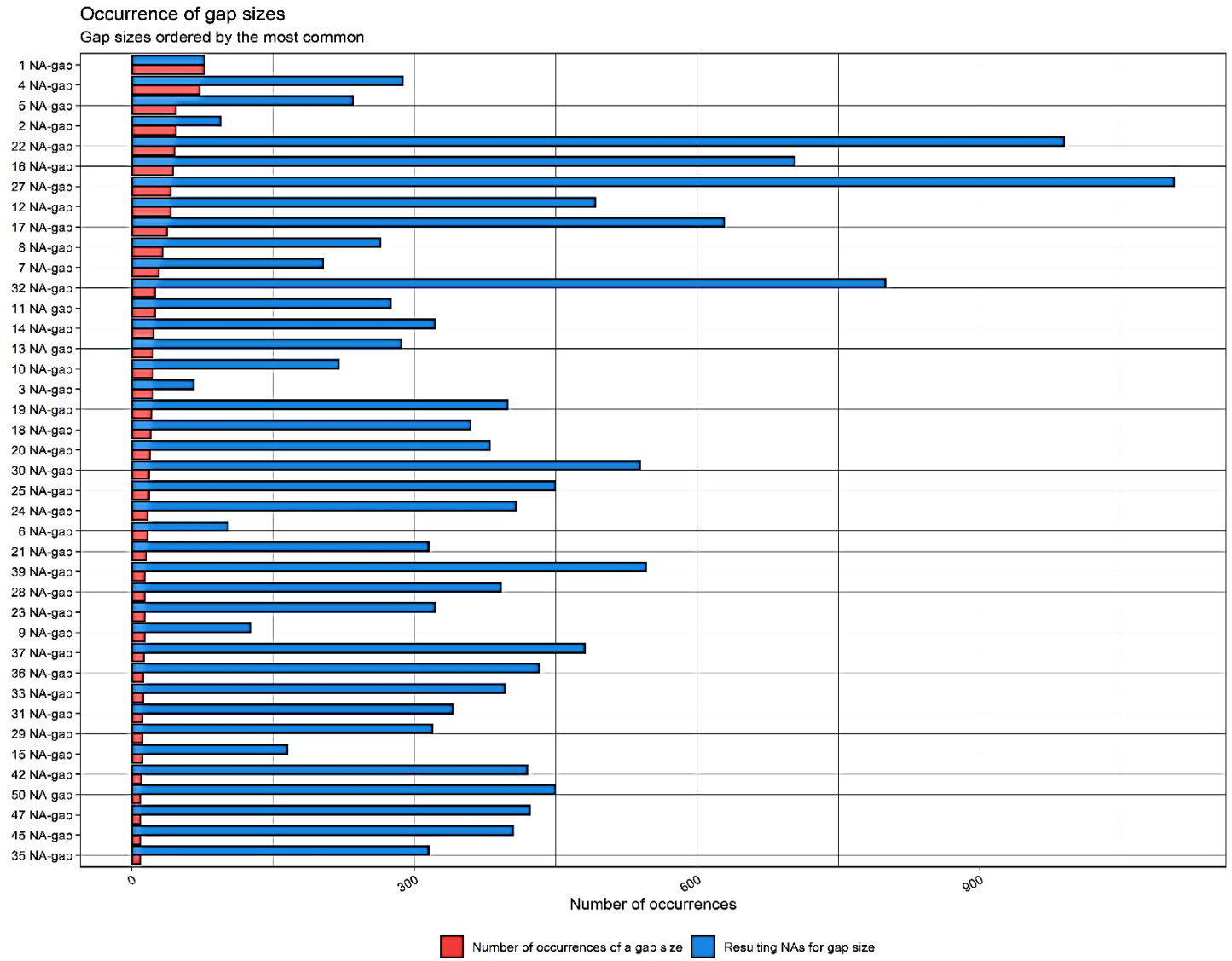


Figure S9. Distribution of the 40 most common gap sizes in the borehole data set D10.

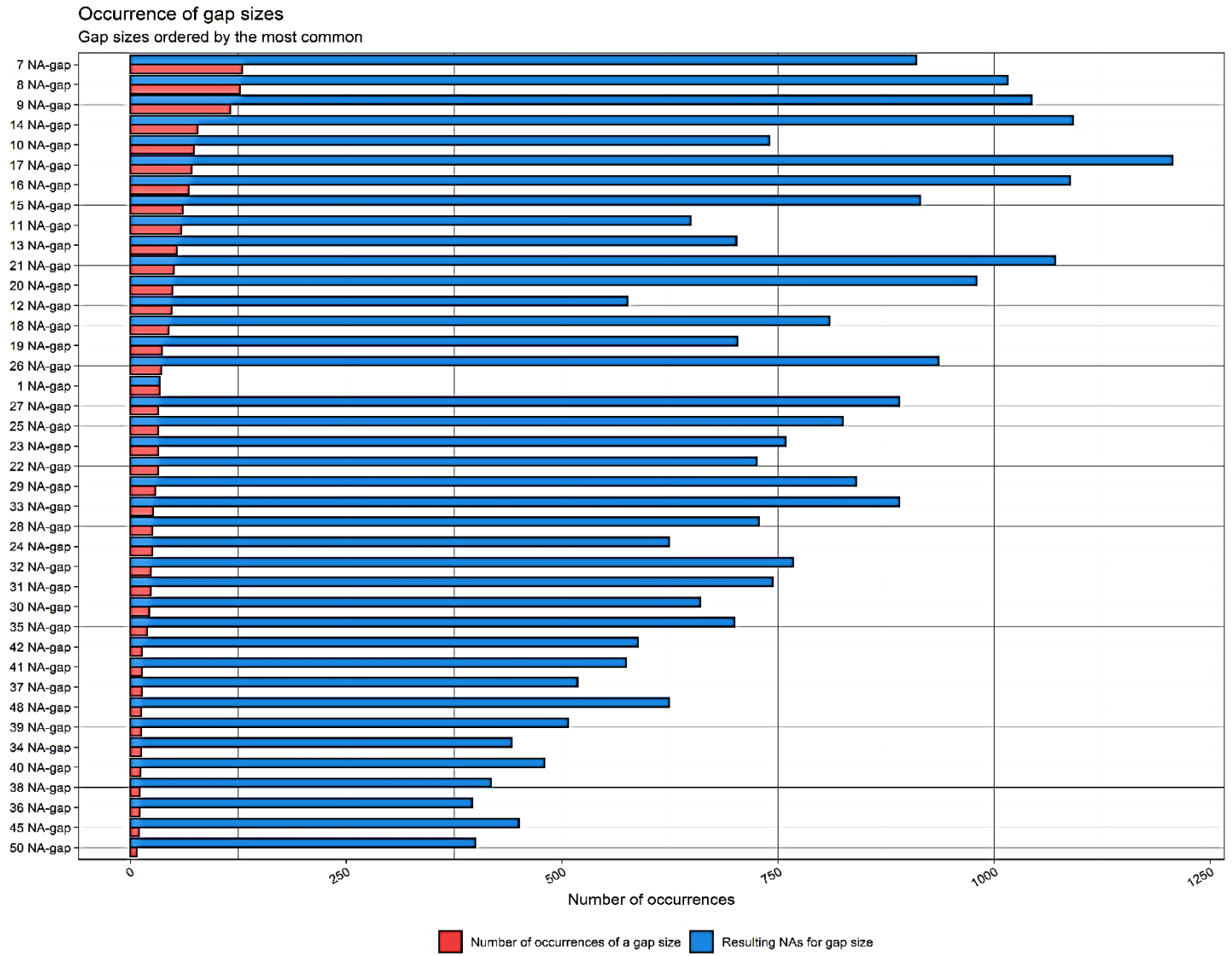


Figure S10. Distribution of the 40 most common gap sizes in the borehole data set D13.

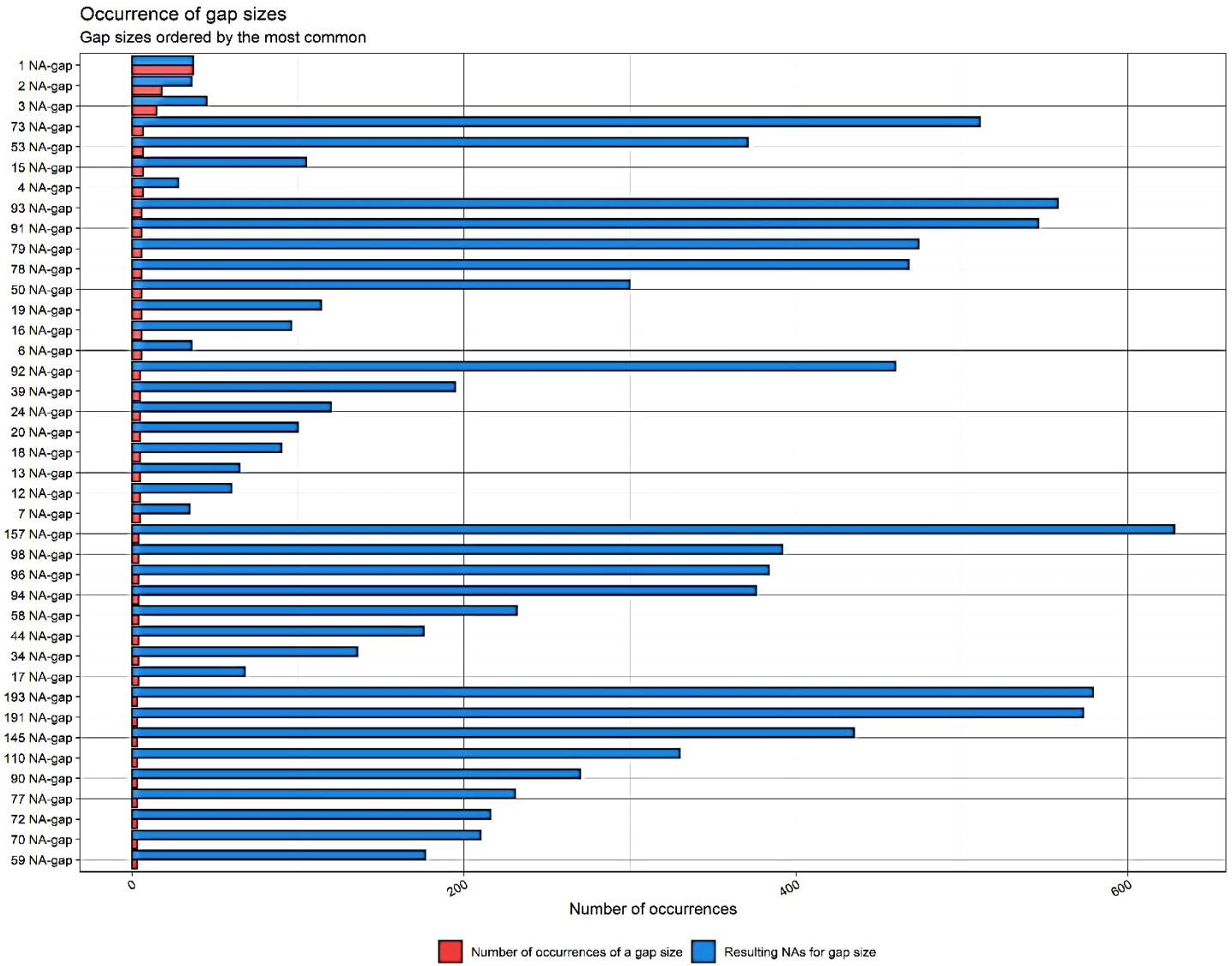


Figure S11. Distribution of the 40 most common gap sizes in the borehole data set D15.

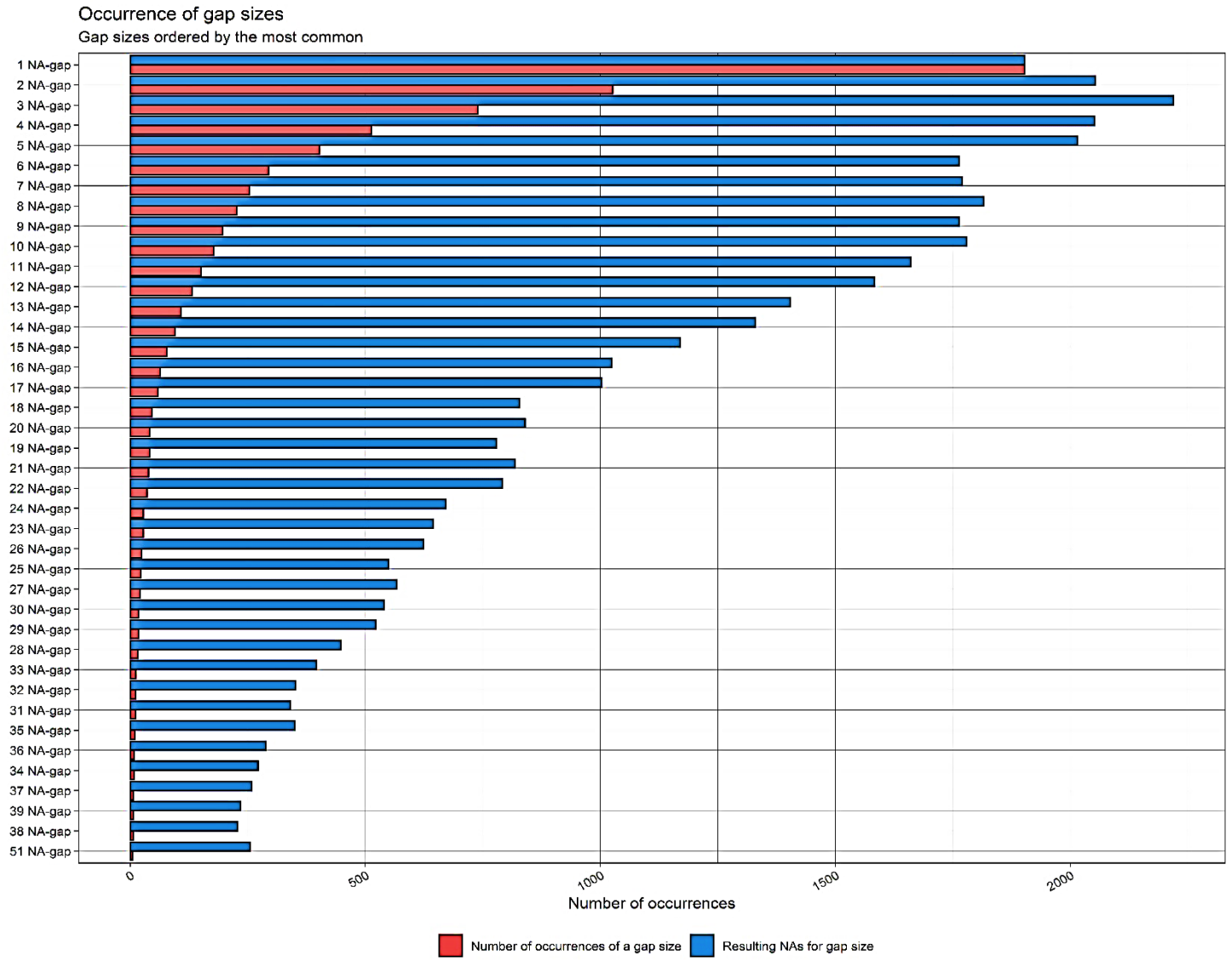


Figure S12. Distribution of the 40 most common gap sizes in the borehole data set D16.

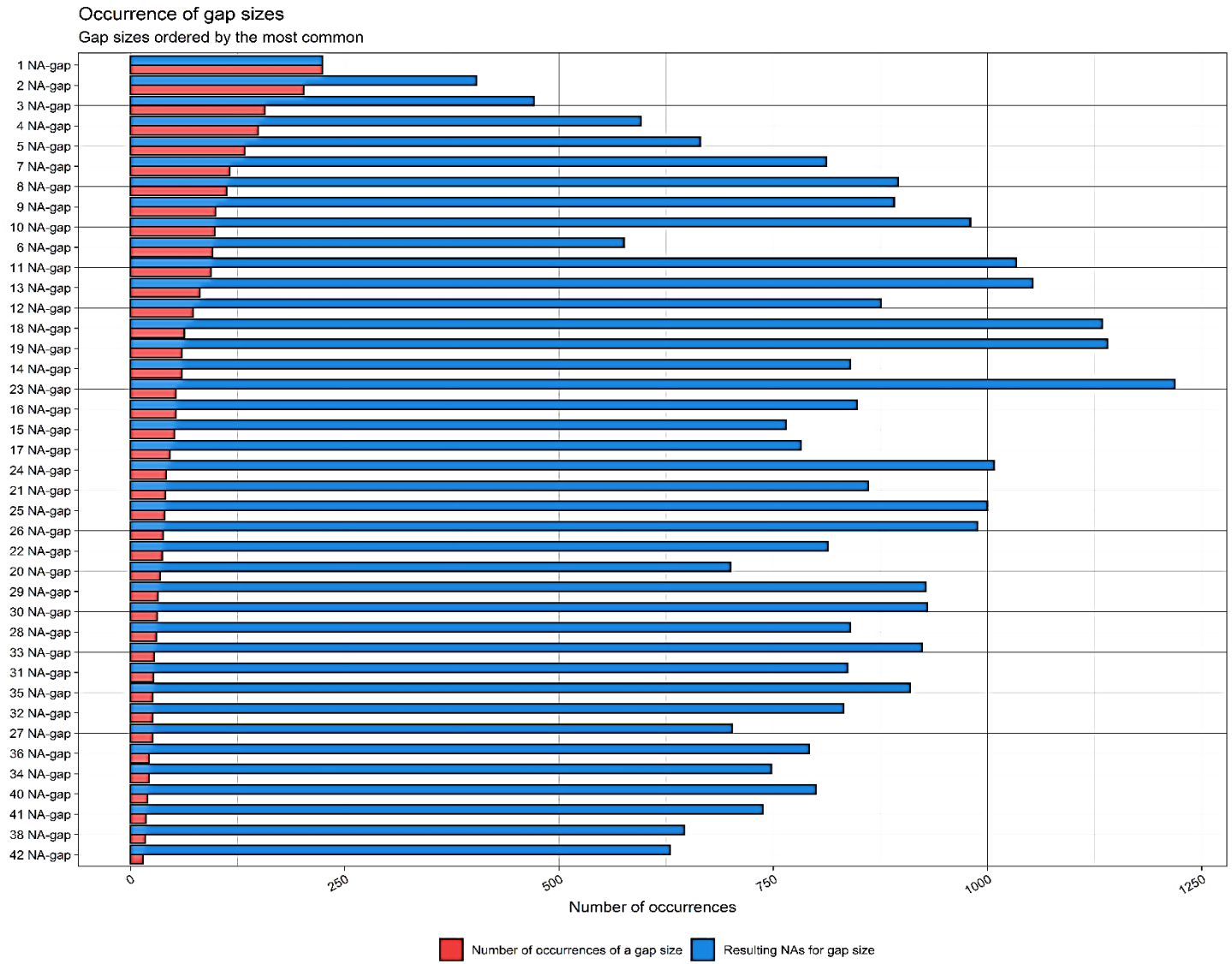


Figure S13. Distribution of the 40 most common gap sizes in the borehole data set D20.

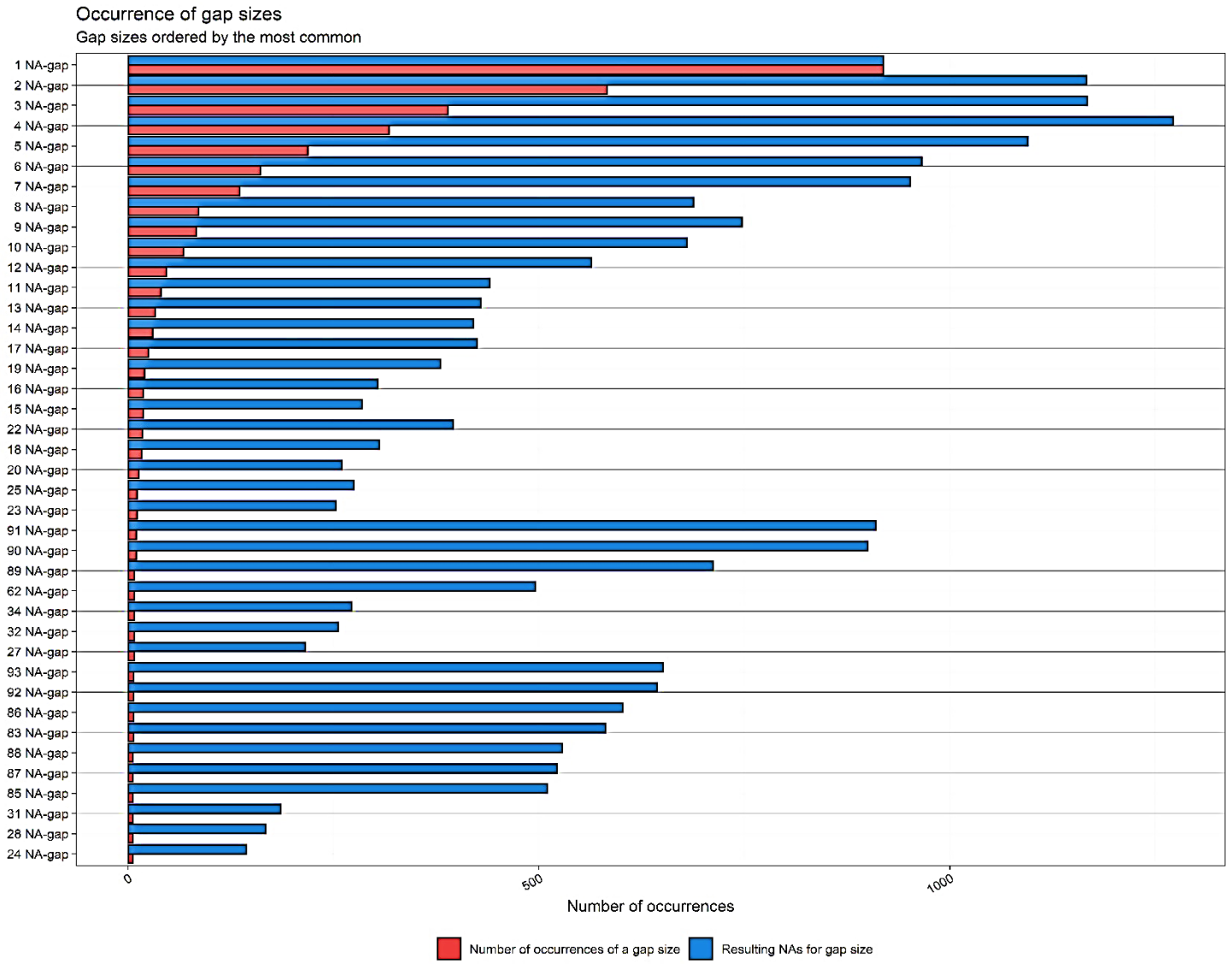


Figure S14. Distribution of the 40 most common gap sizes in the borehole data set D24.

Appendix 9 : Distribution of the percentage of missing values

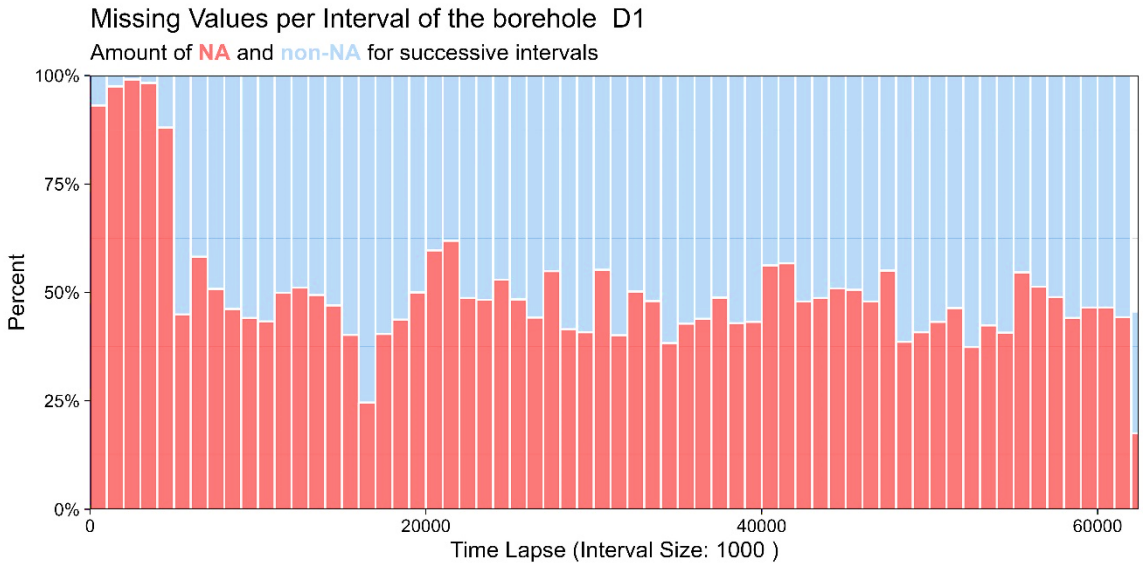


Figure S15. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D1. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

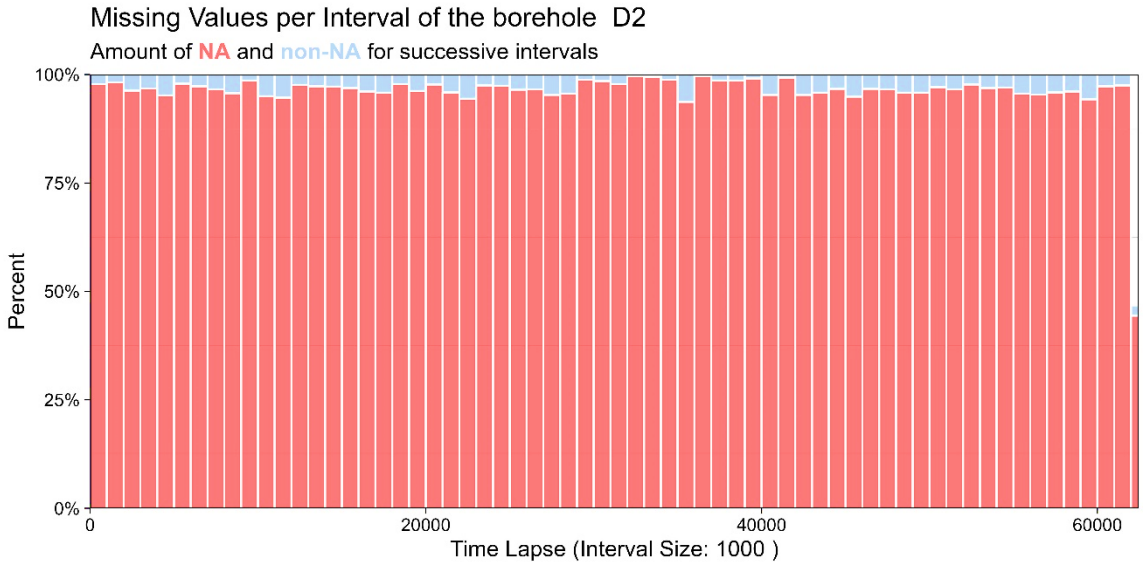


Figure S16. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D2. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

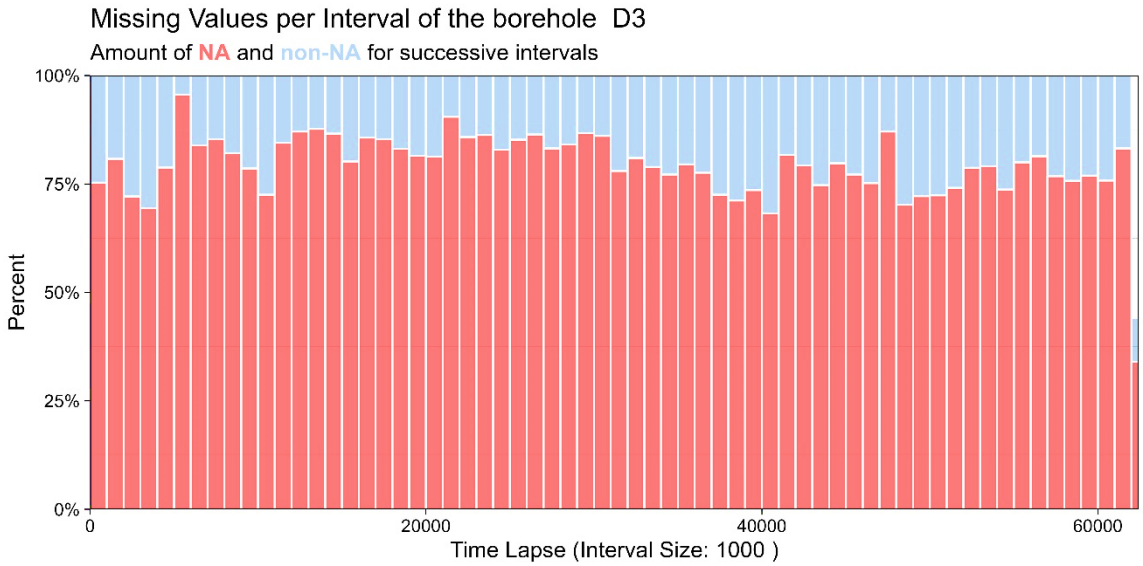


Figure S17. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D3. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

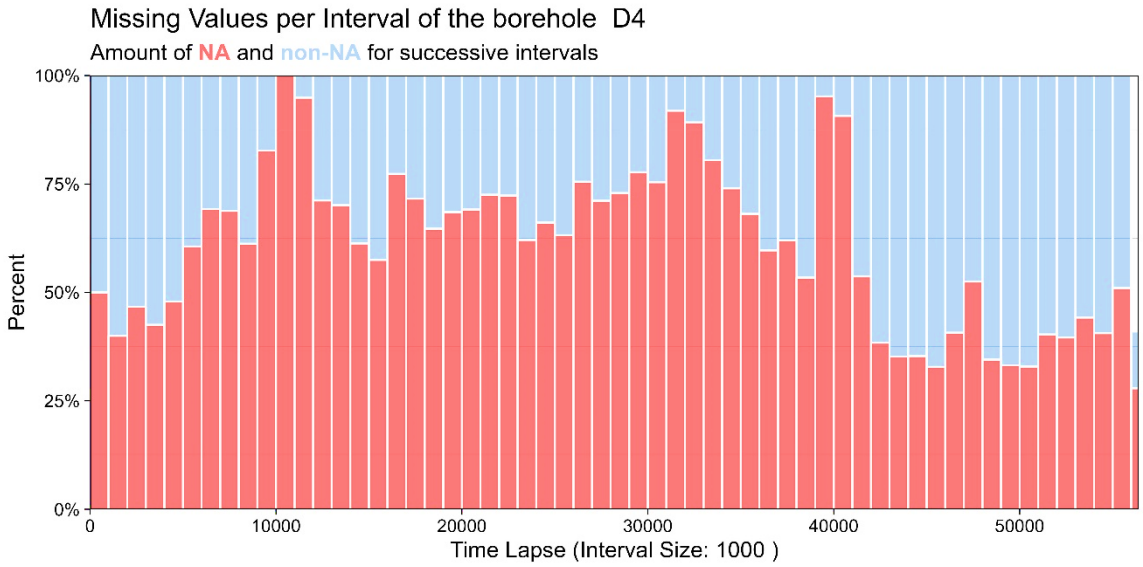


Figure S18. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D4. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

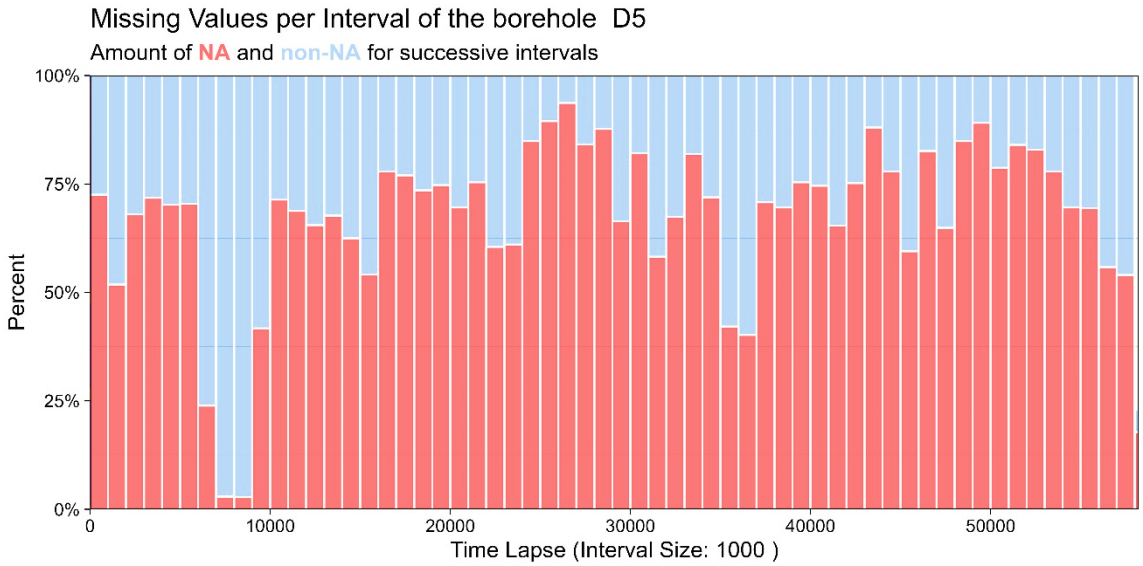


Figure S19. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D5. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

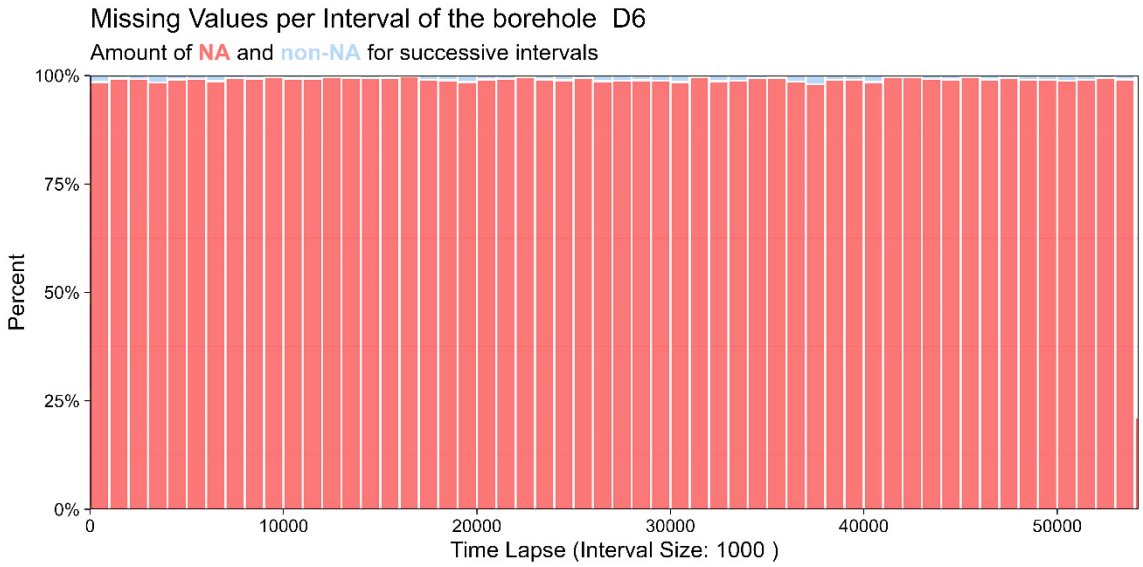


Figure S20. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D6. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

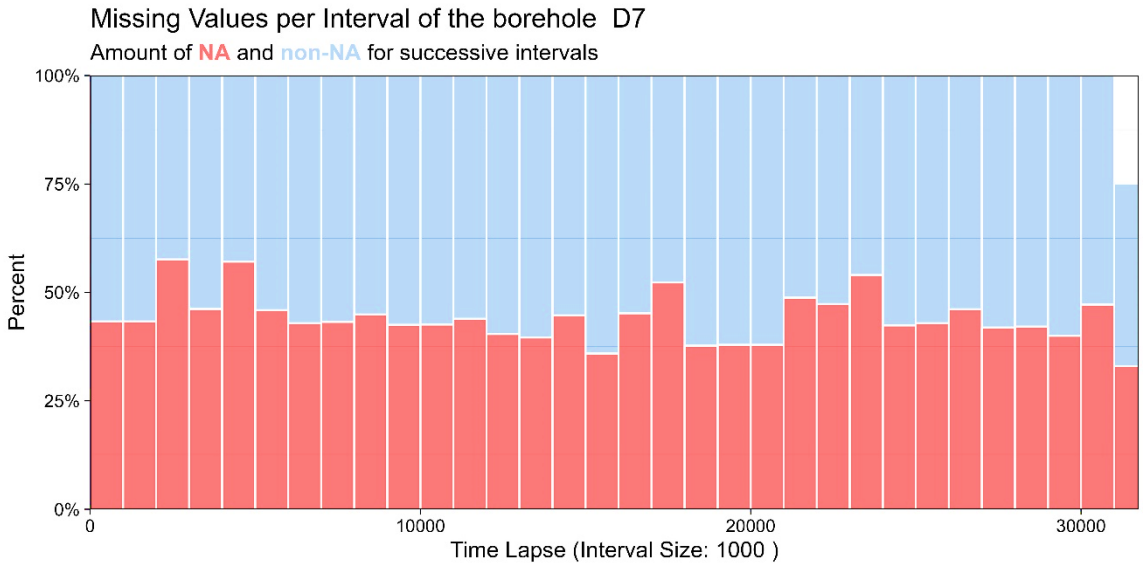


Figure S21. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D7. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

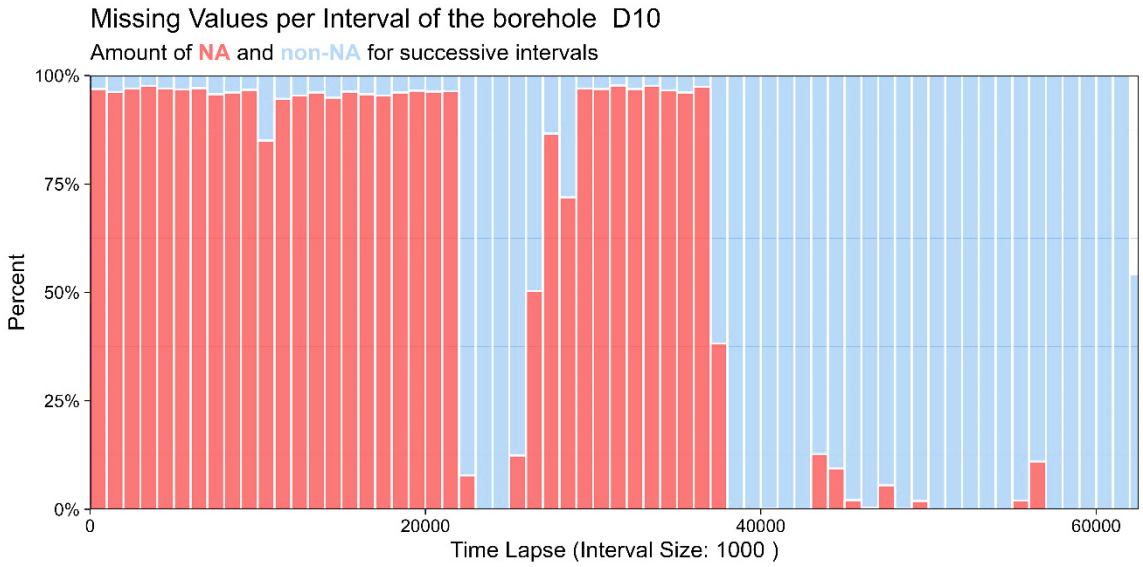


Figure S22. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D10. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

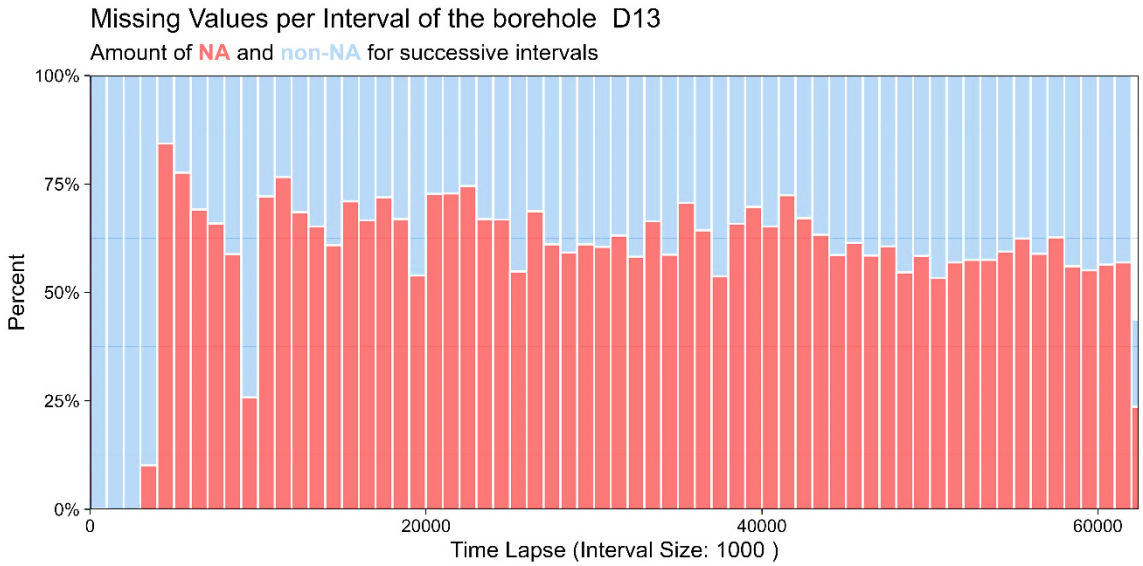


Figure S23. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D13. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

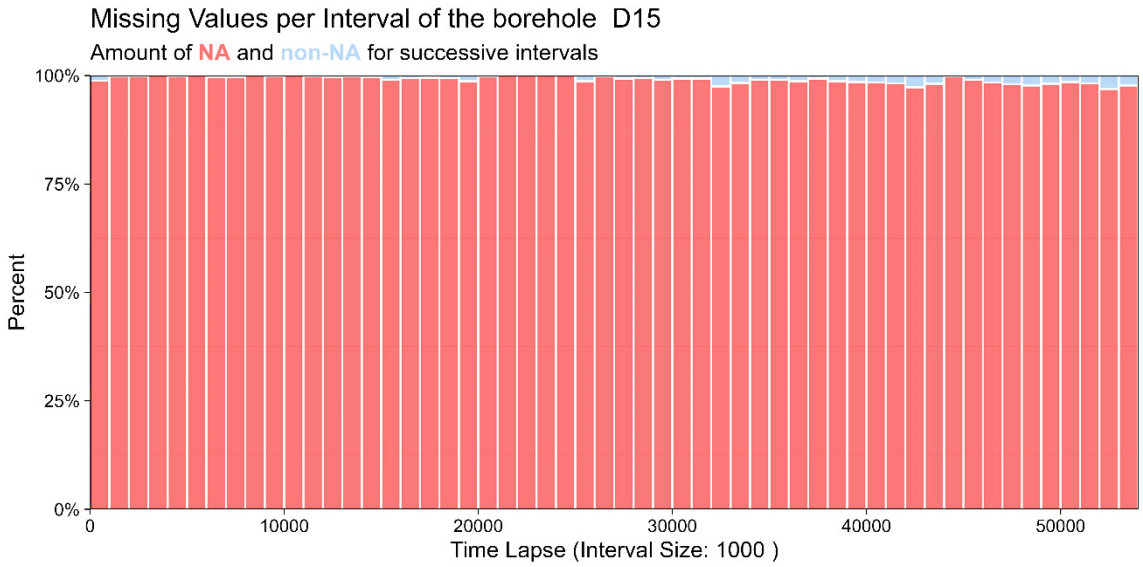


Figure S24. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D14. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

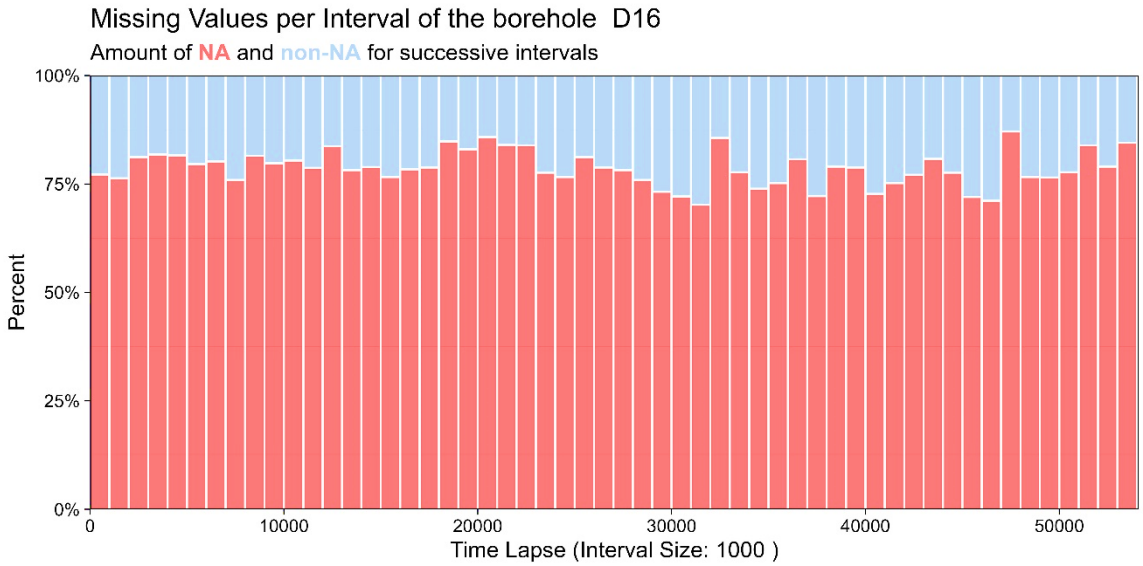


Figure S25. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D16. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

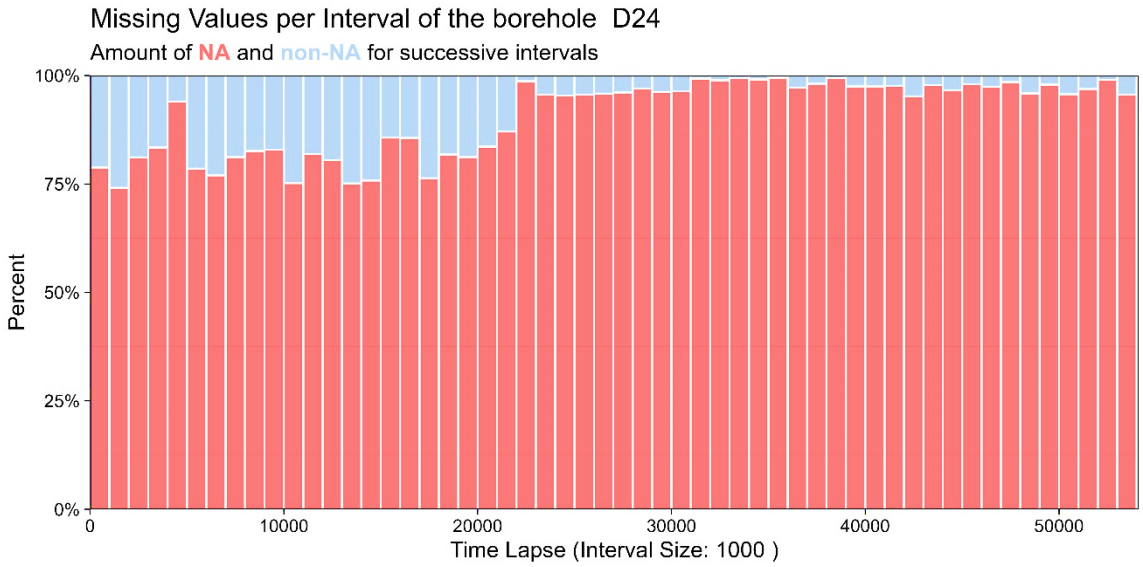


Figure S26. Distribution of the percentage of missing values by 1000-time step interval size in borehole time series D24. NA (non-available) refers to missing values. The percentage of missing values in each interval size is represented in red bars for the missing values and in blue bars for the non-missing.

Appendix 10 : Average performance metrics of the imputation methods tested per borehole

Table S3. Overall average performance metrics for the applied imputation methods on the basis of missing value pattern of borehole data set D1.

	D1		
	R^2	RMSE	KGE
Linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Spline interpolation	0,999(SD 0,00)	0,012(SD 0,00)	0,999(SD 0,00)
Stineman interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Kalman filtering with a structural model and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Simple moving average with *k* = 15	0,998(SD 0,00)	0,015(SD 0,01)	0,999(SD 0,00)
Linear weighted moving average with *k* = 15	0,999(SD 0,00)	0,013(SD 0,00)	0,999(SD 0,00)
Exponential weighted moving average with *k* = 15	0,999(SD 0,00)	0,012(SD 0,00)	0,999(SD 0,00)
Simple moving average with *k* = 5	0,999(SD 0,00)	0,013(SD 0,00)	0,999(SD 0,00)
Linear weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,00)	0,999(SD 0,00)
Exponential weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,00)	0,999(SD 0,00)
Seasonal decomposition with linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,00)	0,999(SD 0,00)
Seasonal splitting with linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,00)	0,999(SD 0,00)

Table S4. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D2.

	D2		
	R^2	RMSE	KGE
Linear interpolation	0,972(SD 0,03)	0,053(SD 0,03)	0,980(SD 0,01)
Spline interpolation	0,881(SD 0,11)	0,120(SD 0,08)	0,912(SD 0,07)
Stineman interpolation	0,970(SD 0,03)	0,054(SD 0,03)	0,981(SD 0,01)
Kalman filtering with a structural model and smoothing	0,972(SD 0,03)	0,053(SD 0,03)	0,980(SD 0,01)
Kalman filtering with ARIMA and smoothing	0,787(SD 0,36)	0,895(SD 1,85)	0,122(SD 1,87)
Simple moving average with *k* = 15	0,967(SD 0,03)	0,061(SD 0,03)	0,978(SD 0,01)
Linear weighted moving average with *k* = 15	0,968(SD 0,03)	0,060(SD 0,03)	0,979(SD 0,01)
Exponential weighted moving average with *k* = 15	0,966(SD 0,03)	0,061(SD 0,03)	0,979(SD 0,01)
Simple moving average with *k* = 5	0,967(SD 0,03)	0,061(SD 0,03)	0,978(SD 0,01)
Linear weighted moving average with *k* = 5	0,968(SD 0,03)	0,060(SD 0,03)	0,979(SD 0,01)
Exponential weighted moving average with *k* = 5	0,966(SD 0,03)	0,061(SD 0,03)	0,979(SD 0,01)
Seasonal decomposition with linear interpolation	0,972(SD 0,03)	0,053(SD 0,03)	0,980(SD 0,01)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,787(SD 0,36)	0,895(SD 1,85)	0,122(SD 1,87)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,966(SD 0,03)	0,061(SD 0,03)	0,979(SD 0,01)
Seasonal splitting with linear interpolation	0,972(SD 0,03)	0,053(SD 0,03)	0,980(SD 0,01)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,787(SD 0,36)	0,895(SD 1,85)	0,122(SD 1,87)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,966(SD 0,03)	0,061(SD 0,03)	0,979(SD 0,01)

Table S5. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D3.

	D3		
	R^2	RMSE	KGE
Linear interpolation	0,996(SD 0,00)	0,021(SD 0,02)	0,995(SD 0,01)
Spline interpolation	0,993(SD 0,01)	0,027(SD 0,02)	0,994(SD 0,01)
Stineman interpolation	0,996(SD 0,00)	0,021(SD 0,02)	0,995(SD 0,01)
Kalman filtering with a structural model and smoothing	0,996(SD 0,01)	0,022(SD 0,02)	0,995(SD 0,01)
Kalman filtering with ARIMA and smoothing	0,972(SD 0,06)	0,048(SD 0,08)	0,978(SD 0,04)
Simple moving average with *k* = 15	0,993(SD 0,01)	0,029(SD 0,03)	0,995(SD 0,00)
Linear weighted moving average with *k* = 15	0,994(SD 0,01)	0,028(SD 0,02)	0,995(SD 0,00)
Exponential weighted moving average with *k* = 15	0,994(SD 0,01)	0,027(SD 0,02)	0,996(SD 0,00)
Simple moving average with *k* = 5	0,994(SD 0,01)	0,028(SD 0,02)	0,995(SD 0,00)
Linear weighted moving average with *k* = 5	0,994(SD 0,01)	0,027(SD 0,02)	0,996(SD 0,00)
Exponential weighted moving average with *k* = 5	0,994(SD 0,01)	0,027(SD 0,02)	0,996(SD 0,00)
Seasonal decomposition with linear interpolation	0,996(SD 0,00)	0,021(SD 0,02)	0,995(SD 0,01)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,972(SD 0,06)	0,048(SD 0,08)	0,978(SD 0,04)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,994(SD 0,01)	0,027(SD 0,02)	0,996(SD 0,00)
Seasonal splitting with linear interpolation	0,996(SD 0,00)	0,021(SD 0,02)	0,995(SD 0,01)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,972(SD 0,06)	0,048(SD 0,08)	0,978(SD 0,04)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,994(SD 0,01)	0,027(SD 0,02)	0,996(SD 0,00)

Table S6. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D4.

	D4		
	R^2	RMSE	KGE
Linear interpolation	0,995(SD 0,00)	0,026(SD 0,01)	0,995(SD 0,00)
Spline interpolation	0,955(SD 0,03)	0,072(SD 0,04)	0,961(SD 0,03)
Stineman interpolation	0,993(SD 0,00)	0,031(SD 0,01)	0,993(SD 0,00)
Kalman filtering with a structural model and smoothing	0,974(SD 0,02)	0,052(SD 0,04)	0,971(SD 0,03)
Kalman filtering with ARIMA and smoothing	0,995(SD 0,00)	0,026(SD 0,01)	0,995(SD 0,00)
Simple moving average with *k* = 15	0,993(SD 0,00)	0,032(SD 0,01)	0,994(SD 0,00)
Linear weighted moving average with *k* = 15	0,993(SD 0,00)	0,031(SD 0,01)	0,994(SD 0,00)
Exponential weighted moving average with *k* = 15	0,993(SD 0,00)	0,031(SD 0,02)	0,995(SD 0,00)
Simple moving average with *k* = 5	0,993(SD 0,00)	0,032(SD 0,01)	0,994(SD 0,00)
Linear weighted moving average with *k* = 5	0,993(SD 0,00)	0,031(SD 0,01)	0,994(SD 0,00)
Exponential weighted moving average with *k* = 5	0,993(SD 0,00)	0,031(SD 0,02)	0,995(SD 0,00)
Seasonal decomposition with linear interpolation	0,995(SD 0,00)	0,026(SD 0,01)	0,995(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,995(SD 0,00)	0,026(SD 0,01)	0,995(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,993(SD 0,00)	0,031(SD 0,02)	0,995(SD 0,00)
Seasonal splitting with linear interpolation	0,995(SD 0,00)	0,026(SD 0,01)	0,995(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,995(SD 0,00)	0,026(SD 0,01)	0,995(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,993(SD 0,00)	0,031(SD 0,02)	0,995(SD 0,00)

Table S7. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D5.

	D5		
	R^2	RMSE	KGE
Linear interpolation	0,994(SD 0,01)	0,023(SD 0,02)	0,996(SD 0,00)
Spline interpolation	0,987(SD 0,01)	0,035(SD 0,01)	0,990(SD 0,01)
Stineman interpolation	0,995(SD 0,01)	0,022(SD 0,02)	0,997(SD 0,00)
Kalman filtering with a structural model and smoothing	0,994(SD 0,01)	0,023(SD 0,02)	0,995(SD 0,00)
Kalman filtering with ARIMA and smoothing	0,995(SD 0,01)	0,022(SD 0,02)	0,995(SD 0,00)
Simple moving average with *k* = 15	0,992(SD 0,01)	0,030(SD 0,02)	0,995(SD 0,00)
Linear weighted moving average with *k* = 15	0,992(SD 0,01)	0,029(SD 0,02)	0,995(SD 0,00)
Exponential weighted moving average with *k* = 15	0,992(SD 0,01)	0,029(SD 0,02)	0,995(SD 0,00)
Simple moving average with *k* = 5	0,992(SD 0,01)	0,029(SD 0,02)	0,995(SD 0,00)
Linear weighted moving average with *k* = 5	0,992(SD 0,01)	0,028(SD 0,02)	0,995(SD 0,00)
Exponential weighted moving average with *k* = 5	0,992(SD 0,01)	0,028(SD 0,02)	0,996(SD 0,00)
Seasonal decomposition with linear interpolation	0,994(SD 0,01)	0,023(SD 0,02)	0,996(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,995(SD 0,01)	0,022(SD 0,02)	0,995(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,992(SD 0,01)	0,028(SD 0,02)	0,996(SD 0,00)
Seasonal splitting with linear interpolation	0,994(SD 0,01)	0,023(SD 0,02)	0,996(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,995(SD 0,01)	0,022(SD 0,02)	0,995(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,992(SD 0,01)	0,028(SD 0,02)	0,996(SD 0,00)

Table S8. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D6.

	D6		
	R^2	RMSE	KGE
Linear interpolation	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Spline interpolation	0,925(SD 0,08)	0,088(SD 0,06)	0,948(SD 0,06)
Stineman interpolation	0,978(SD 0,02)	0,048(SD 0,02)	0,981(SD 0,01)
Kalman filtering with a structural model and smoothing	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Kalman filtering with ARIMA and smoothing	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Simple moving average with *k* = 15	0,973(SD 0,02)	0,055(SD 0,03)	0,979(SD 0,01)
Linear weighted moving average with *k* = 15	0,975(SD 0,02)	0,052(SD 0,02)	0,981(SD 0,01)
Exponential weighted moving average with *k* = 15	0,972(SD 0,02)	0,055(SD 0,03)	0,981(SD 0,01)
Simple moving average with *k* = 5	0,973(SD 0,02)	0,055(SD 0,03)	0,979(SD 0,01)
Linear weighted moving average with *k* = 5	0,975(SD 0,02)	0,052(SD 0,02)	0,981(SD 0,01)
Exponential weighted moving average with *k* = 5	0,972(SD 0,02)	0,055(SD 0,03)	0,981(SD 0,01)
Seasonal decomposition with linear interpolation	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,972(SD 0,02)	0,055(SD 0,03)	0,981(SD 0,01)
Seasonal splitting with linear interpolation	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,979(SD 0,02)	0,046(SD 0,02)	0,981(SD 0,01)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,972(SD 0,02)	0,055(SD 0,03)	0,981(SD 0,01)

Table S9. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D7.

	D7		
	R^2	RMSE	KGE
Linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,998(SD 0,00)
Spline interpolation	0,995(SD 0,00)	0,016(SD 0,01)	0,997(SD 0,00)
Stineman interpolation	0,999(SD 0,00)	0,008(SD 0,00)	0,998(SD 0,00)
Kalman filtering with a structural model and smoothing	0,999(SD 0,00)	0,008(SD 0,00)	0,999(SD 0,00)
Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,008(SD 0,00)	0,999(SD 0,00)
Simple moving average with *k* = 15	0,996(SD 0,00)	0,016(SD 0,01)	0,996(SD 0,00)
Linear weighted moving average with *k* = 15	0,997(SD 0,00)	0,015(SD 0,01)	0,997(SD 0,00)
Exponential weighted moving average with *k* = 15	0,998(SD 0,00)	0,013(SD 0,01)	0,998(SD 0,00)
Simple moving average with *k* = 5	0,997(SD 0,00)	0,013(SD 0,01)	0,998(SD 0,00)
Linear weighted moving average with *k* = 5	0,998(SD 0,00)	0,013(SD 0,01)	0,998(SD 0,00)
Exponential weighted moving average with *k* = 5	0,998(SD 0,00)	0,013(SD 0,01)	0,998(SD 0,00)
Seasonal decomposition with linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,998(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,008(SD 0,00)	0,999(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,998(SD 0,00)	0,013(SD 0,01)	0,998(SD 0,00)
Seasonal splitting with linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,998(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,008(SD 0,00)	0,999(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,998(SD 0,00)	0,013(SD 0,01)	0,998(SD 0,00)

Table S10. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D8.

	D8		
	R^2	RMSE	KGE
Linear interpolation	0,987(SD 0,01)	0,045(SD 0,03)	0,980(SD 0,01)
Spline interpolation	0,878(SD 0,08)	0,136(SD 0,05)	0,902(SD 0,06)
Stineman interpolation	0,984(SD 0,01)	0,049(SD 0,02)	0,985(SD 0,01)
Kalman filtering with a structural model and smoothing	0,929(SD 0,08)	0,090(SD 0,07)	0,945(SD 0,05)
Kalman filtering with ARIMA and smoothing	0,987(SD 0,01)	0,045(SD 0,02)	0,981(SD 0,01)
Simple moving average with *k* = 15	0,978(SD 0,01)	0,061(SD 0,03)	0,980(SD 0,01)
Linear weighted moving average with *k* = 15	0,981(SD 0,01)	0,056(SD 0,03)	0,982(SD 0,01)
Exponential weighted moving average with *k* = 15	0,488(SD 0,31)	0,778(SD 0,87)	-0,651(SD 2,47)
Simple moving average with *k* = 5	0,979(SD 0,01)	0,060(SD 0,03)	0,980(SD 0,01)
Linear weighted moving average with *k* = 5	0,981(SD 0,01)	0,056(SD 0,03)	0,982(SD 0,01)
Exponential weighted moving average with *k* = 5	0,488(SD 0,31)	0,778(SD 0,87)	-0,652(SD 2,47)
Seasonal decomposition with linear interpolation	0,987(SD 0,01)	0,045(SD 0,03)	0,980(SD 0,01)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,987(SD 0,01)	0,045(SD 0,02)	0,981(SD 0,01)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,488(SD 0,31)	0,778(SD 0,87)	-0,652(SD 2,47)
Seasonal splitting with linear interpolation	0,987(SD 0,01)	0,045(SD 0,03)	0,980(SD 0,01)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,987(SD 0,01)	0,045(SD 0,02)	0,981(SD 0,01)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,488(SD 0,31)	0,778(SD 0,87)	-0,652(SD 2,47)

Table S11. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D10.

	D10		
	R^2	RMSE	KGE
Linear interpolation	0,993(SD 0,01)	0,020(SD 0,01)	0,990(SD 0,01)
Spline interpolation	0,982(SD 0,01)	0,033(SD 0,01)	0,977(SD 0,03)
Stineman interpolation	0,992(SD 0,01)	0,020(SD 0,01)	0,992(SD 0,00)
Kalman filtering with a structural model and smoothing	0,989(SD 0,01)	0,024(SD 0,01)	0,984(SD 0,02)
Kalman filtering with ARIMA and smoothing	0,910(SD 0,18)	0,118(SD 0,23)	0,882(SD 0,24)
Simple moving average with *k* = 15	0,990(SD 0,01)	0,025(SD 0,01)	0,989(SD 0,01)
Linear weighted moving average with *k* = 15	0,991(SD 0,01)	0,024(SD 0,01)	0,989(SD 0,01)
Exponential weighted moving average with *k* = 15	0,990(SD 0,01)	0,025(SD 0,01)	0,991(SD 0,01)
Simple moving average with *k* = 5	0,990(SD 0,01)	0,025(SD 0,01)	0,990(SD 0,01)
Linear weighted moving average with *k* = 5	0,991(SD 0,01)	0,024(SD 0,01)	0,990(SD 0,01)
Exponential weighted moving average with *k* = 5	0,990(SD 0,01)	0,025(SD 0,01)	0,991(SD 0,01)
Seasonal decomposition with linear interpolation	0,993(SD 0,01)	0,020(SD 0,01)	0,990(SD 0,01)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,910(SD 0,18)	0,118(SD 0,23)	0,882(SD 0,24)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,990(SD 0,01)	0,025(SD 0,01)	0,991(SD 0,01)
Seasonal splitting with linear interpolation	0,993(SD 0,01)	0,020(SD 0,01)	0,990(SD 0,01)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,910(SD 0,18)	0,118(SD 0,23)	0,882(SD 0,24)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,990(SD 0,01)	0,025(SD 0,01)	0,991(SD 0,01)

Table S12. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D13.

	D13		
	R^2	RMSE	KGE
Linear interpolation	0,998(SD 0,00)	0,013(SD 0,01)	0,997(SD 0,00)
Spline interpolation	0,995(SD 0,00)	0,023(SD 0,01)	0,995(SD 0,00)
Stineman interpolation	0,999(SD 0,00)	0,012(SD 0,01)	0,997(SD 0,00)
Kalman filtering with a structural model and smoothing	0,999(SD 0,00)	0,011(SD 0,00)	0,997(SD 0,00)
Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,011(SD 0,01)	0,997(SD 0,00)
Simple moving average with *k* = 15	0,996(SD 0,00)	0,020(SD 0,01)	0,997(SD 0,00)
Linear weighted moving average with *k* = 15	0,997(SD 0,00)	0,019(SD 0,01)	0,997(SD 0,00)
Exponential weighted moving average with *k* = 15	0,997(SD 0,00)	0,017(SD 0,01)	0,997(SD 0,00)
Simple moving average with *k* = 5	0,997(SD 0,00)	0,018(SD 0,01)	0,997(SD 0,00)
Linear weighted moving average with *k* = 5	0,997(SD 0,00)	0,017(SD 0,01)	0,997(SD 0,00)
Exponential weighted moving average with *k* = 5	0,997(SD 0,00)	0,017(SD 0,01)	0,997(SD 0,00)
Seasonal decomposition with linear interpolation	0,998(SD 0,00)	0,013(SD 0,01)	0,997(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,011(SD 0,01)	0,997(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,997(SD 0,00)	0,017(SD 0,01)	0,997(SD 0,00)
Seasonal splitting with linear interpolation	0,998(SD 0,00)	0,013(SD 0,01)	0,997(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,011(SD 0,01)	0,997(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,997(SD 0,00)	0,017(SD 0,01)	0,997(SD 0,00)

Table S13. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D15.

	D15		
	R^2	RMSE	KGE
Linear interpolation	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Spline interpolation	0,608(SD 0,28)	0,382(SD 0,27)	0,429(SD 0,43)
Stineman interpolation	0,935(SD 0,05)	0,088(SD 0,04)	0,943(SD 0,04)
Kalman filtering with a structural model and smoothing	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Kalman filtering with ARIMA and smoothing	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Simple moving average with *k* = 15	0,942(SD 0,04)	0,092(SD 0,06)	0,949(SD 0,04)
Linear weighted moving average with *k* = 15	0,944(SD 0,04)	0,090(SD 0,06)	0,950(SD 0,04)
Exponential weighted moving average with *k* = 15	0,265(SD 0,34)	2,911(SD 1,30)	-6,217(SD 4,01)
Simple moving average with *k* = 5	0,942(SD 0,04)	0,092(SD 0,06)	0,949(SD 0,04)
Linear weighted moving average with *k* = 5	0,944(SD 0,04)	0,090(SD 0,06)	0,950(SD 0,04)
Exponential weighted moving average with *k* = 5	0,265(SD 0,34)	2,911(SD 1,30)	-6,217(SD 4,01)
Seasonal decomposition with linear interpolation	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,265(SD 0,34)	2,911(SD 1,30)	-6,217(SD 4,01)
Seasonal splitting with linear interpolation	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,952(SD 0,04)	0,081(SD 0,05)	0,952(SD 0,04)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,265(SD 0,34)	2,911(SD 1,30)	-6,217(SD 4,01)

Table S14. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D16.

	D16		
	R^2	RMSE	KGE
Linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Spline interpolation	0,997(SD 0,00)	0,016(SD 0,01)	0,998(SD 0,00)
Stineman interpolation	0,999(SD 0,00)	0,008(SD 0,00)	0,999(SD 0,00)
Kalman filtering with a structural model and smoothing	0,999(SD 0,00)	0,008(SD 0,00)	0,999(SD 0,00)
Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Simple moving average with *k* = 15	0,998(SD 0,00)	0,015(SD 0,01)	0,999(SD 0,00)
Linear weighted moving average with *k* = 15	0,998(SD 0,00)	0,013(SD 0,01)	0,999(SD 0,00)
Exponential weighted moving average with *k* = 15	0,999(SD 0,00)	0,012(SD 0,01)	0,999(SD 0,00)
Simple moving average with *k* = 5	0,998(SD 0,00)	0,013(SD 0,01)	0,999(SD 0,00)
Linear weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,01)	0,999(SD 0,00)
Exponential weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,01)	0,999(SD 0,00)
Seasonal decomposition with linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal decomposition with exponential weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,01)	0,999(SD 0,00)
Seasonal splitting with linear interpolation	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,999(SD 0,00)	0,009(SD 0,00)	0,999(SD 0,00)
Seasonal splitting with exponential weighted moving average with *k* = 5	0,999(SD 0,00)	0,012(SD 0,01)	0,999(SD 0,00)

Table S15. Overall average performance metrics for imputation methods for missing values using the missing value pattern of the groundwater-level time series of borehole D24.

	D24		
	R^2	RMSE	KGE
Linear interpolation	0,986(SD 0,01)	0,037(SD 0,02)	0,992(SD 0,01)
Spline interpolation	0,904(SD 0,11)	0,102(SD 0,08)	0,921(SD 0,10)
Stineman interpolation	0,985(SD 0,01)	0,040(SD 0,02)	0,991(SD 0,01)
Kalman filtering with a structural model and smoothing	0,986(SD 0,01)	0,037(SD 0,02)	0,992(SD 0,01)
Kalman filtering with ARIMA and smoothing	0,960(SD 0,05)	0,059(SD 0,05)	0,970(SD 0,05)
Simple moving average with k = 15	0,982(SD 0,01)	0,045(SD 0,02)	0,989(SD 0,01)
Linear weighted moving average with k = 15	0,983(SD 0,01)	0,043(SD 0,02)	0,990(SD 0,01)
Exponential weighted moving average with k = 15	0,983(SD 0,01)	0,044(SD 0,02)	0,990(SD 0,01)
Simple moving average with k = 5	0,982(SD 0,01)	0,044(SD 0,02)	0,989(SD 0,01)
Linear weighted moving average with k = 5	0,983(SD 0,01)	0,043(SD 0,02)	0,990(SD 0,01)
Exponential weighted moving average with k = 5	0,983(SD 0,01)	0,044(SD 0,02)	0,990(SD 0,01)
Seasonal decomposition with linear interpolation	0,986(SD 0,01)	0,037(SD 0,02)	0,992(SD 0,01)
Seasonal decomposition with Kalman filtering with ARIMA and smoothing	0,960(SD 0,05)	0,059(SD 0,05)	0,970(SD 0,05)
Seasonal decomposition with exponential weighted moving average with k = 5	0,983(SD 0,01)	0,044(SD 0,02)	0,990(SD 0,01)
Seasonal splitting with linear interpolation	0,986(SD 0,01)	0,037(SD 0,02)	0,992(SD 0,01)
Seasonal splitting with Kalman filtering with ARIMA and smoothing	0,960(SD 0,05)	0,059(SD 0,05)	0,970(SD 0,05)
Seasonal splitting with exponential weighted moving average with k = 5	0,983(SD 0,01)	0,044(SD 0,02)	0,990(SD 0,01)

Appendix 11 : Borehole monitoring network lithologic data

Table S16. Borehole monitoring network lithologic data.

Boreholes	Sediment thickness (m)	Slope (°)	Lithology				
			Seq1	Seq2	Seq3	Seq4	Seq5
1	1.2	9.3	0-1.2 m Sand	1.2-68.6 m Bedrock			
2	26.2	4.4	0-1.5 m Sand	1.5-16.7 m Silty sand	16.7-26.2 m Gravel	26.2-57.9 m Bedrock	
3	7.2	3.4	0-7.2 m Sand	7.2-80.8 m Bedrock			
4	21.3	3.1	0-1.5 m Sand	1.5-21.3 m Clay	21.3-105.1 m Bedrock		
5	8.8	3.9	0-8.8 m Clay with gravel	8.8-92.9 m Bedrock			
6	4	4.0	0-4 m Sand	4-47.3 m Bedrock			
7	5.8	10.9	0-5.8 m Sand	5.8-93 m Bedrock			
8		4.1	Not available				
9	13.26	4.2	0-2.13 m Organic matter	2.13-4.27 m Clay	4.27-13.26m Sand with rocks	13.26-72.54 m Bedrock	
10	7.62	4.2	0-1.52 m Organic matter	1.52-3.66 m Clay	3.66-7.62m Sand	7.62-94.49 m Limestone	94.49-109.12 m Bedrock

11	1.83	13.8	0-1.83 m Sand	1.83-60.35 m Bedrock			
12 (SB4)	7.01	2.4	0-3.5 m Medium Sand	3.5-7.01 m Sand with gravel	7.01-106.69 m Bedrock		
13	12	7.9	0-12 m Clay with gravel	12-60.96 m Bedrock			
15	2.4	2.6	0-2.4 m Sand	2.4-111.2 m Bedrock			
16	0.75	11.1	0-0.75 m Clay with rocks	0-38.1 m Bedrock			
17-FE04	9.8	3.5	0-6 m Silty sand with gravel	6-9.8 m Sand with gravel	9.8-91.44 m Bedrock		
20	25	0.9	0-9.1 m Sand	9.1-21.3 m Clay	21.3-25 m Limestone	25-36.6 m Bedrock	
22 (SB5)	12.8	2.8	0-0.3 m Sand with gravel and rocks	0.3-4.57 m Medium sand	4.57-8.84 m Silt	8.84-12.8 m Sand with gravel	12.8-91.45 m Bedrock
23 (SB6)	29.26	0.5	0-4.57 m Medium sand with gravel	4.57-9.14 m Silt	9.14-13.72 m Silty sand	13.72-29.26 m Sand with gravel	29.26-106.69 m Bedrock
24	26.2	3.5	0-1.5 m Sand	1.5-16.7 m Clay	16.7-26.2 m Gravel	26.2-57.9 m Bedrock	

Appendix 12 : Cross-correlation and Autocorrelation Analysis

Text S1 : Applicability of the Cross-correlation Analysis

a) Stationarity and independent random variable

This section presents the conditions for applying cross-correlation analysis to study the relationships between precipitation and snowmelt time series with groundwater level time series. In the following analysis, the precipitation and snowmelt time series are considered as inputs to the system, while the groundwater levels are considered as outputs. The analyses of relationship between time series are usually performed using the cross-correlation analysis, which assesses whether there is a linear relationship between two time series (Lee and Lee 2000, Cai and Ofterdinger 2016, Lorenzo-Lacruz et al. 2017, Fronzi et al. 2020, Dong et al. 2022). Another advantage of the cross-correlation analysis is that it allows us to assess whether there is a time lag between the two time series, which informs us about the response time of the output time series relative to the input time series of rainfall and snowmelt (Lee and Lee 2000, Cai and Ofterdinger 2016). However, the time series being analyzed for cross-correlation need to fulfill certain requirements to avoid spurious correlation results (Olden and Neff 2001, Yue et al. 2002, Cryer and Chan 2008, Probst et al. 2012, Dean and Dunsmuir 2016, Li and Guo 2019). In fact, the time series (input and output) being analyzed for cross-correlation need to be stationary and at least one of them must be independent white noise (Cryer and Chan 2008, Box et al. 2015, Dean and Dunsmuir 2016). Stationarity means that the process described by the time series is in statistical equilibrium, in the sense that the probability laws governing the evolution of the process remain unchanged over time (Cryer and Chan 2008, Shumway and Stoffer 2017). In time series analysis, second order stationarity is generally considered and is defined by a constant mean and variance at all times (Cryer and Chan 2008, Box et al. 2015, Shumway and Stoffer 2017). The independent white noise is defined as a sequence of independent,

identically distributed (i.i.d) random variables (Cryer and Chan 2008), i.e., independent uncorrelated random variables (Shumway and Stoffer 2017).

In hydrological studies, time series such as groundwater levels (Lee and Lee 2000, Cai and Ofterdinger 2016), annual precipitation (Chen et al. 2004) and atmospheric pressure (Lee and Lee 2000) are often autocorrelated and also nonstationary (Chen et al. 2004). However, it should be noted that the time step of the time series is also important; for instance the annual rainfall has been reported to be autocorrelated (Chen et al. 2004) but daily rainfall was reported to be uncorrelated (Cai and Ofterdinger 2016). When a time series is autocorrelated, it means that it is correlated to a lagged version of itself (Lee and Lee 2000, Cai and Ofterdinger 2016). In other words, the measurement at time step t+1 depends on the measurement at time step t (Dean and Dunsmuir 2016). Therefore, in hydrological studies, time series most likely don't fulfill the prerequisite for cross-correlation analysis. Such situations tend to inflate the cross-correlation results between the two time series, which has been described as "nonsense-correlation" by some authors (Yule 1926, Dean and Dunsmuir 2016).

Numerous statistical tests have been developed for testing, more objectively, the stationarity of time series (Hyndman and Athanasopoulos 2021). Among the most commonly used ones, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (Kwiatkowski et al. 1992, Li et al. 2021a, Hyndman and Athanasopoulos 2021) which is used in this study. The null hypothesis of the KPSS test is that the time series is stationary (Kwiatkowski et al. 1992). The chosen level of significance α over which the null hypothesis is accepted is 0.05. Therefore, for calculated level of significance, p-values, that are greater than the threshold level of significance α, the time series are stationary. The KPSS test was implemented using the R package tseries (Trapletti et al. 2022).

In the case of nonstationary of a time series, differencing is the most commonly used method to transform the nonstationary time series into stationary (Box et al. 2015, Li et al.

2021a, Hyndman and Athanasopoulos 2021). It is the first order differencing that was used in this study and it consisted of transforming the nonstationary time series x_t into a time series $x'_t = x_t - x_{t-1}$ (Probst et al. 2012, Shumway and Stoffer 2017, Hyndman and Athanasopoulos 2021). The differencing was performed using inbuilt function of the R code. It is found that a first order differencing was sufficient to make the groundwater level time series stationary in our case. Another interest in the first order differencing of the groundwater level time series is that it can be related to a specific meaning, namely the daily increase or decrease rate of the water table in the aquifer.

A stationary time series is not necessarily independent white noise (Dean and Dunsmuir 2016, Li et al. 2021a, Hyndman and Athanasopoulos 2021). Testing whether the time series is independent white noise is carried out using a portmanteau test, implemented in the R package stats, through the function box.test. The portmanteau test implemented is the Ljung-Box test statistic that assumes a null hypothesis of independent white noise time series (Ljung and Box 1978, Hyndman and Athanasopoulos 2021). The chosen α value over which the null hypothesis is accepted is 0.05. Thus p-values that are lower than 0.05 indicate that the time series are not independent white noise. Since cross-correlation analysis requires at least one of the two time series to be independent white noise, the commonly used solution is the prewhitening of the input time series in the case of two non-white noise time series (Probst et al. 2012, Shumway and Stoffer 2017, Li et al. 2021a).

Prewhitening is a process that consists of transforming the input time series into independent white noise by decorrelating the input time series in order to remove the biases induced by the autocorrelation (Chen et al. 2004). In practice, it generally consists of fitting an autoregressive moving average model to the input time series, then removing the fitted autoregressive moving average model from the input time series and output time series (Chen et al. 2004, Cryer and Chan 2008, Dean and Dunsmuir 2016). Because prewhitening is a linear relationship, any existing linear relationships between the original input and output

time series will be preserved after prewhitening (Cryer and Chan 2008). In this study, the fitted models were autoregressive integrated moving average models (ARIMA model), whose orders were chosen according to the corrected Akaike's Information Criterion (AICc). For a given time series the best fitting ARIMA model is the one having lowest AICc value (Hyndman and Athanasopoulos 2021). Fitting the ARIMA model to the input time series was performed using the function *auto.arima* in the R package *Forecast* and the prewhitening was performed using the function prewhiten in the R package *TSA*. In depth details on fitting ARIMA models to time series are provided in Hyndman and Athanasopoulos (2021) and details on prewhitening are provided in Cryer and Chan (2008) and Box et al. (2015).

b) Autocorrelation of time series

The autocorrelation of a time series is assessed via the autocorrelation analysis provided in equation A9 and equation A10. The autocorrelation analysis shows whether there is a linear relationship between successive measurements within a given time series by calculating the correlation between lagged versions of the time series (Cai and Ofterdinger 2016).

$$C(k) = \frac{1}{n}\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x}), \qquad k \geq 0 \qquad \text{(Eq A. 9)}$$

$$\gamma(k) = \frac{C(k)}{C(0)} \qquad (Eq\ A.10)$$

where C(k) is called the correlogram, n is the number of measurements in the time series, k is the time lag, x_t is the value of the variables studied at time t, $\bar{x}$ is the mean value of the time series x_t, γ(k) is the autocorrelation function (ACF). When a time series is autocorrelated, the autocorrelation function shows a gradual and steady decrease in correlation for a long-time lag (Cai and Ofterdinger 2016). Conversely, when the time series is uncorrelated, the autocorrelation function displays a rapid decrease in correlation to values that are statically insignificant within a short time lag or the autocorrelation function

displays correlation values that are statistically insignificant from the beginning (Chen et al. 2004, Cai and Ofterdinger 2016). A confidence interval of 95% was chosen for this study. In this case, the correlation significance limits are given by $\pm 1{,}96\sqrt{n}$ (Cai and Ofterdinger 2016, Dean and Dunsmuir 2016).

Text S2: Results of the Applicability of the Cross-correlation Analysis: Stationarity and independent random variable

Using the cross-correlation analysis requires the verification of the stationarity and independence of the time series as explained in the previous section. If one defines vertical inflow as the potential water available for infiltration, the vertical inflow will be the time series composed of the sum of rainfall and snowmelt. It is the KPSS hypothesis test presented in the methodology section that is used for analyzing the stationarity of the time series. The chosen level of significance α over which the null hypothesis is accepted is 0.05. Therefore, p-values that are greater than the threshold level of significance α, the time series are stationary. The results of the time series stationarity analysis are provided in Table S17, where it can be seen that the rainfall, snowmelt and vertical inflow time series are not stationary, similar to the majority of the groundwater level time series. In the case of nonstationary of a time series, the first order differencing was used as presented in the methodology. It is shown in Table S17 that this first order differencing was sufficient to make the groundwater level time series stationary because the p-values obtained from the KPSS test are greater than 0.05.

Table S17. Results of the KPSS test and Box test to check for the stationarity and independence of time series, respectively. The p-values highlighted in green are those for which the null hypothesis is accepted.

		White noise test	Stationarity test
Interpolated variable	**Location**	**Ljung-Box test**	**KPSS test**
Rainfall	D1	1,93E-05	0.1
	D2	1,46E-05	0.1
	D3	0,00E+00	0.1
	D4	7,52E-07	0.1
	D5	1,94E-05	0.1
	D6	1,14E-02	0.1
	D7	2,53E-05	0.1
	D8	6,87E-05	0.1
	D9	8,62E-05	0.1
	D10	8,39E-05	0.1
	D11	1,95E-06	0.1
	D12	6,08E-05	0.1
	D13	2,41E-05	0.1
	D15	9,73E-06	0.1
	D16	3,33E-15	0.1
	D17	1,97E-04	0.1
	D20	9,44E-05	0.1
	D22	6,06E-05	0.1
	D23	5,91E-05	0.1
	D24	1,40E-05	0.1
Snowmelt	D1	0,00E+00	0.1
	D2	0,00E+00	0.1
	D3	0,00E+00	0.1
	D4	0,00E+00	0.1
	D5	0,00E+00	0.1
	D6	0,00E+00	0.1
	D7	0,00E+00	0.1
	D8	0,00E+00	0.1
	D9	0,00E+00	0.1
	D10	0,00E+00	0.1
	D11	0,00E+00	0.1
	D12	0,00E+00	0.1
	D13	0,00E+00	0.1

	D15	0,00E+00	0.1
	D16	0,00E+00	0.1
	D17	0,00E+00	0.1
	D20	0,00E+00	0.1
	D22	0,00E+00	0.1
	D23	0,00E+00	0.1
	D24	0,00E+00	0.1
Vertical Inflow	D1	0,00E+00	0.1
	D2	0,00E+00	0.1
	D3	0,00E+00	0.1
	D4	0,00E+00	0.1
	D5	0,00E+00	0.1
	D6	0,00E+00	0.1
	D7	0,00E+00	0.1
	D8	0,00E+00	0.1
	D9	0,00E+00	0.1
	D10	0,00E+00	0.1
	D11	0,00E+00	0.1
	D12	0,00E+00	0.1
	D13	0,00E+00	0.1
	D15	0,00E+00	0.1
	D16	0,00E+00	0.1
	D17	0,00E+00	0.1
	D20	0,00E+00	0.1
	D22	0,00E+00	0.1
	D23	0,00E+00	0.1
	D24	0,00E+00	0.1
Prewhiten Rainfall	D1	8,88E-01	0.1
	D2	8,43E-01	0.1
	D3	8,99E-01	0.1
	D4	7,85E-01	0.1
	D5	8,67E-01	0.1
	D6	9,51E-01	0.1
	D7	7,99E-01	0.1
	D8	8,60E-01	0.1
	D9	8,81E-01	0.1
	D10	8,81E-01	0.1
	D11	8,00E-01	0.1
	D12	8,60E-01	0.1

	D13	8,61E-01	0.1
	D15	8,34E-01	0.1
	D16	7,73E-01	0.1
	D17	8,88E-01	0.1
	D20	8,03E-01	0.1
	D22	8,60E-01	0.1
	D23	8,60E-01	0.1
	D24	8,43E-01	0.1
Prewhiten Snowmelt	D1	9,96E-01	0.1
	D2	9,99E-01	0.1
	D3	1,00E+00	0.1
	D4	1,00E+00	0.1
	D5	1,00E+00	0.1
	D6	9,96E-01	0.1
	D7	9,99E-01	0.1
	D8	1,00E+00	0.1
	D9	9,86E-01	0.1
	D10	9,89E-01	0.1
	D11	1,00E+00	0.1
	D12	9,59E-01	0.1
	D13	9,96E-01	0.1
	D15	9,98E-01	0.1
	D16	8,74E-01	0.1
	D17	1,00E+00	0.1
	D20	1,00E+00	0.1
	D22	9,81E-01	0.1
	D23	9,66E-01	0.1
	D24	9,99E-01	0.1
Prewhiten Vertical Inflow	D1	8,12E-01	0.1
	D2	8,32E-01	0.1
	D3	9,10E-01	0.1
	D4	7,73E-01	0.1
	D5	8,18E-01	0.1
	D6	6,75E-01	0.1
	D7	4,87E-01	0.1
	D8	8,26E-01	0.1
	D9	8,01E-01	0.1
	D10	8,30E-01	0.1
	D11	9,78E-01	0.1
	D12	8,31E-01	0.1

	D13	7,51E-01	0.1
	D15	7,96E-01	0.1
	D16	9,20E-01	0.1
	D17	6,95E-01	0.1
	D20	6,76E-01	0.1
	D22	8,51E-01	0.1
	D23	8,37E-01	0.1
	D24	8,61E-01	0.1

* For p-values that are greater than 0.1, the *kpss.test* function of the R package *tseries* only indicates that the real p-value is greater than 0.1.

As presented in the methodology, it is the Ljung-Box test statistic that is used for testing whether the time series is independent white noise. The chosen α value over which the null hypothesis is accepted is 0.05. Thus p-values that are lower than 0.05 indicate that the time series are not independent white noise. The result of the Ljung-Box test presented in Table S17 shows that not a single input time series is independent white noise among the original time series of rainfall, snowmelt, groundwater level. Since cross-correlation analysis requires at least one of the two time series to be independent white noise, we used the prewhitening process to transform the input time series into independent white noise as explained in the methodology. The results presented in Table S17 show that prewhitened rainfall, snowmelt and vertical inflow are independent white noise time series according to the Ljung-Box test, whereas their non-prewhitened versions are not. As a result, prewhitening is required when rainfall, snowmelt or vertical inflow time series is involved in a cross-correlation analysis.

.

Appendix 13 : Spearman rho correlation matrix of geomorphology and unsaturated zone on groundwater level

It can be seen in Figure S27 that the pairwise scatterplots of the different parameters do not necessarily have a linear relationship, which further justifies the use of a Spearman correlation method rather than the Pearson correlation. The Spearman correlation results presented in Figure S27 were produced using the R packages Hmisc and PerformanceAnalytics are utilized.

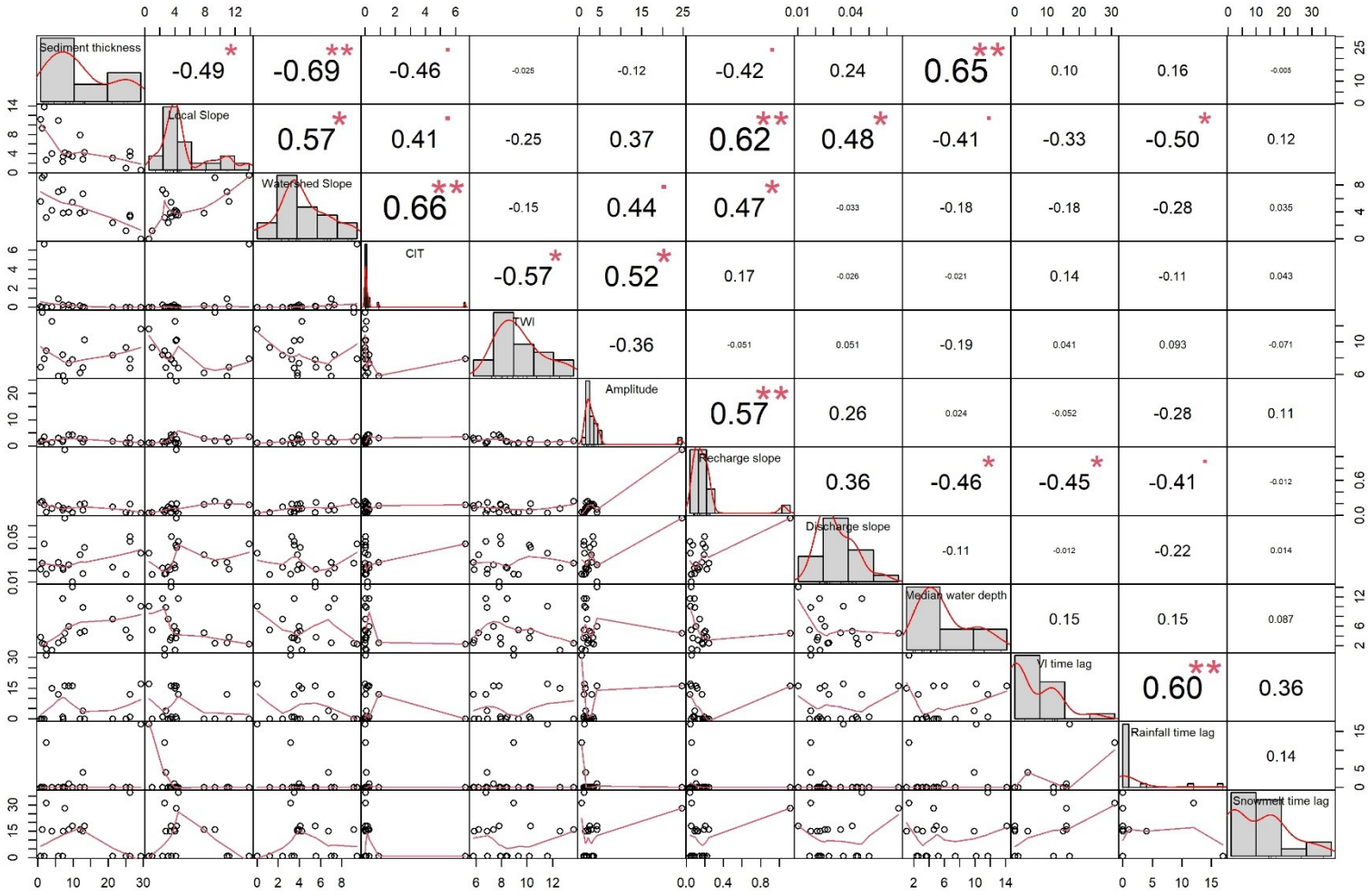


Figure S27. Visualization of the Spearman rho correlation matrix along with pairwise scatterplots and the histograms of each variable. VI time lag refers to vertical inflow time lag, CIT refers to Channel Initiation Threshold and TWI refers to Topographic Wetness Index. **. The correlation is significant at the 0.01 level; *. The correlation is significant at the 0.05 level.

Appendix 14 : Spearman rho correlation matrix of the crystalline rock and fracture characteristics indices on groundwater level

It can be seen in Figure S28 that the pairwise scatterplots of the different parameters do not necessarily have a linear relationship, which further justifies the use of a Spearman correlation method rather than the Pearson correlation. The Spearman correlation results presented in Figure S28 were produced using the R packages Hmisc and PerformanceAnalytics are utilized.

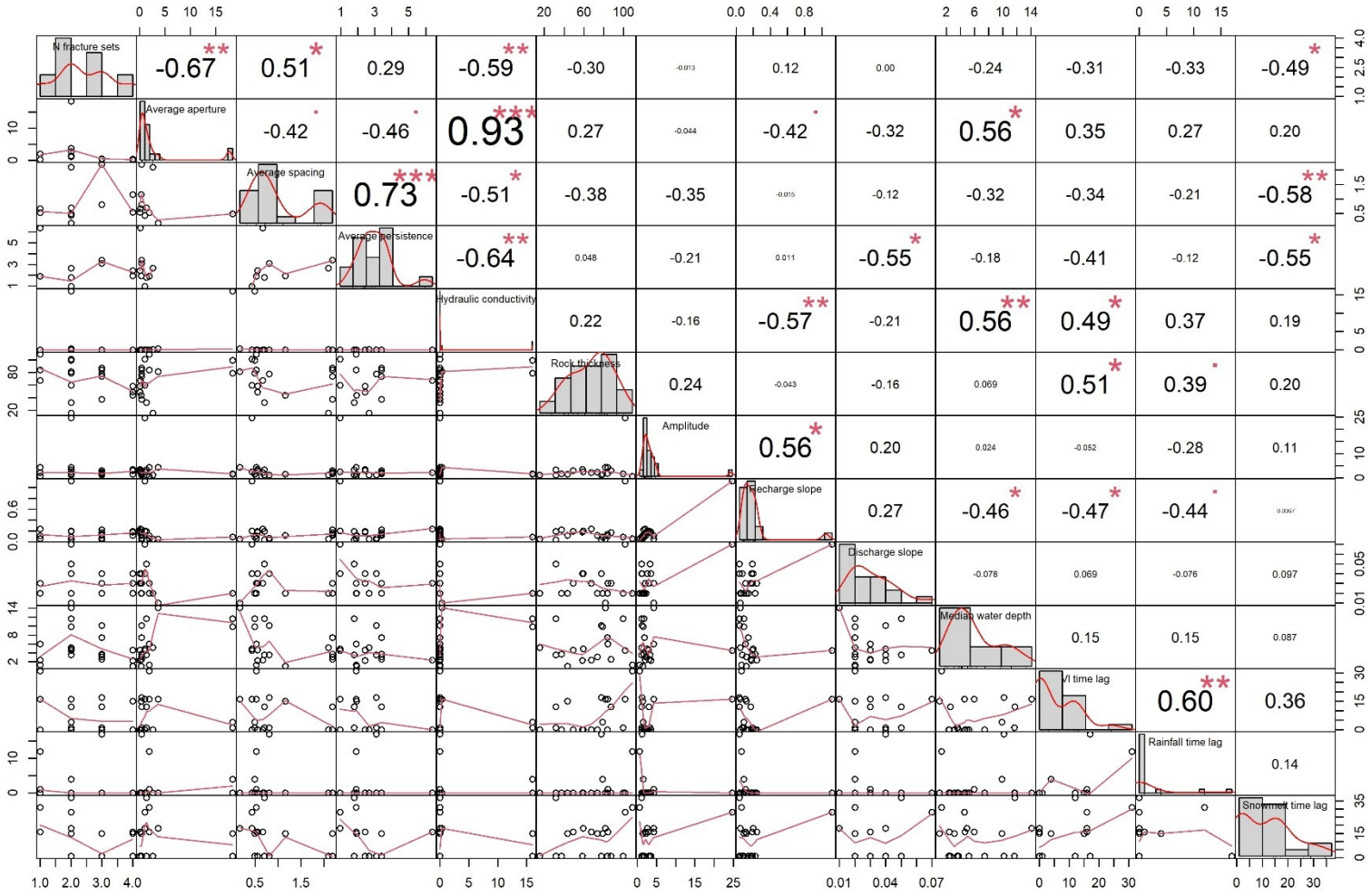


Figure S28. Visualization of the Spearman rho correlation matrix along with pairwise scatterplots and the histograms of each variable. VI time lag refers to vertical inflow time lag. ***, the correlation is significant at the 0.001 level;**, the correlation is significant at the 0.01 level; * the

correlation is significant at the 0.05 level. The abscissa axis (x-axis) and the ordinate axis (y-axis) provide the range of variation of variables represented on the scatterplots. The red lines are fitted trend lines for the scatterplots and fitted distribution for the histograms.

Appendix 15 : Data for the analysis of the effect of Geomorphology and Unsaturated Zone on Groundwater Level

Table S18: Geomorphology, Unsaturated Zone and Groundwater Level parameters for each borehole.

Datologger	Sediment thickness (m)	Local Slope (Deg)	Watershed Slope (Deg)	CIT	TWI	Amplitude (m)	Recharge slope (m/day)	Discharge slope (m/day)	Median water table depth (m)	Vertical inflow time lag (Day)	Rainfall time lag (Day)	Snowmelt time lag (Day)
D1	1.2	9.3	9.1	6.91E-04	13.7	2.1	0.2	0.03	2.4	0	0	16
D2	26.2	4.4	3.5	2.81E-02	6.8	1.1	0.0	0.05	11.7	12	0	37
D3	7.2	3.4	3.8	1.67E-02	5.8	2.3	0.1	0.02	3.1	0	0	15
D4	21.3	3.1	2.4	7.02E-03	8.4	1.9	0.1	0.02	7.4	0	0	1
D5	8.8	3.9	4.0	2.51E-01	7.4	4.2	0.1	0.03	6.0	16	1	16
D6	4	4.0	4.2	6.58E-02	12.6	1.5	0.1	0.02	1.0	15	0	15
D7	5.8	10.9	7.0	8.93E-01	5.8	3.0	0.2	0.03	2.6	12	0	1
D8		4.1	3.5	5.08E-03	10.2	1.1	0.1	0.04	3.5	1	0	1
D9	13.26	4.2	4.1	1.32E-02	10.3	2.6	0.2	0.04	4.9	0	0	18
D10	7.62	4.2				24.7	1.1	0.07	4.6	16	0	28
D11	1.83	13.8	9.4	6.66E+00	7.9	3.5	0.2	0.04	2.3	0	0	1
D12	7.01	2.4	7.3	2.54E-01	8.4	1.7	0.1	0.02	11.7	1	0	1
D13	12	7.9	3.8	1.71E-01	6.2	2.9	0.2	0.03	4.6	0	0	16
D15	2.4	2.6	3.2	2.52E-02	8.9	0.6	0.1	0.02	1.3	31	12	31
D16	0.75	11.1	5.6	7.71E-02	6.9	1.7	0.2	0.02	3.7	0	0	1
D17	9.8	3.5	5.5	1.50E-01	7.8	4.2	0.1	0.01	14.1	16	0	18
D20	25	0.9	1.2	4.89E-04	9.3	1.2	0.1	0.02	5.2	0	0	1
D22	12.8	2.8	6.7	7.24E-02	6.8	1.5	0.1	0.02	9.9	4	4	15
D23	29.26	0.5	0.0	0.00E+00	11.6	1.3	0.0	0.04	10.1	17	17	1
D24	26.2	3.5	3.3	1.89E-02	7.9	3.3	0.2	0.05	3.6	1	0	1

Appendix 16 : Data for the analysis of the effect of Crystalline Rock and Fracture Characteristics on Groundwater Level

Table S19: Geomorphology, Unsaturated Zone and Groundwater Level parameters for each borehole.

Boreholes	Number of fracture sets	Average Aperture	Average Spacing	Average persistence	Equivalent hydraulic conductivity	Rock thickness	Amplitude	Recharge slope	Discharge slope	Median water table depth (m)	Vertical inflow time lag	Rainfall time lag	Snowmelt time lag
D1	1	0.2	0.7	6.4	6.71E-06	67.4	2.1	0.24	0.03	2.43	0	0	16
D2	2	1.4	0.7	1.8	5.44E-03	31.7	1.1	0.04	0.05	11.68	12	0	37
D3	3	0.4	2.2	3.4	2.59E-04	73.6	2.3	0.12	0.02	3.14	0	0	15
D4	3	0.4	2.2	3.4	2.59E-04	83.8	1.9	0.14	0.02	7.42	0	0	1
D5	1	1.9	0.5	1.9	1.04E-02	84.1	4.2	0.11	0.03	5.96	16	1	16
D6	4	0.3	1.2	1.9	2.86E-04	43.3	1.5	0.12	0.02	1.01	15	0	15
D7	3	0.4	2.2	3.4	2.59E-04	87.2	3.0	0.17	0.03	2.58	12	0	1
D8	3	0.4	0.8	3.1	6.11E-04		1.1	0.07	0.04	3.54	1	0	1
D9	2	1.1	0.4	1.0	3.35E-03	59.3	2.6	0.20	0.04	4.94	0	0	18
D10	2	1.1	0.4	1.0	3.35E-03	101.5	24.7	1.12	0.07	4.56	16	0	28
D11	4	0.1	0.6	2.4	8.30E-06	58.5	3.5	0.18	0.04	2.32	0	0	1

D12	2	18.3	0.5		1.60E+01	99.7	1.7	0.09	0.02	11.65	1	0	1
D13	4	0.1	0.6	2.4	8.30E-06	49.0	2.9	0.19	0.03	4.63	0	0	16
D15	1	1.9	0.5	1.9	1.04E-02	108.8	0.6	0.06	0.02	1.34	31	12	31
D16	3	0.4	2.2	3.4	2.59E-04	37.4	1.7	0.22	0.02	3.69	0	0	1
D17	2	3.7	0.2		3.31E-01	81.6	4.2	0.05	0.01	14.11	16	0	18
D20	2	2.7	2.1	2.7	4.12E-02	15.3	1.2	0.05	0.02	5.20	0	0	1
D22	2	18.3	0.5		1.60E+01	78.7	1.6	0.09	0.02	9.86	4	4	15
D23	3	0.4	0.8	3.1	6.11E-04	77.4	1.3	0.04	0.04	10.14	17	17	1
D24	2	1.4	0.7	1.8	5.44E-03	31.7	3.3	0.19	0.05	3.60	1	0	1

The equivalent hydraulic conductivity is here calculated using the following formula:

$$K_N = \sum_{i=1}^{n} K_i \cos^2 \alpha_i \quad (3)$$

where K_i is the hydraulic conductivity of the fracture set i, and α_i is the angle between the fracture set i and the direction N.

$$K_i = \frac{g(a_{ave})^3}{12\nu C_{ave} s_{ave}} \quad (4)$$

where g is the gravitational acceleration, v is the dynamic viscosity of water, a_{ave} is the average aperture of the fracture set, C_{ave} is the average correction factor of the fracture set, and s_{ave} is the average spacing of the fractures within the fracture set.

Here, it is the vertical equivalent hydraulic conductivity that is calculated and provided in Table S19.

References

Alvo, M., and Yu, P. 2018. A Parametric Approach to Nonparametric Statistics. Springer International Publishing. doi:10.1007/978-3-319-94153-0.

Bonnini, S., Corain, L., Marozzi, M., and Salmaso, L. 2014. Nonparametric Hypothesis Testing: Rank and Permutation Methods with Applications in R. In 1st édition. Wiley.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M. 2015. Time Series Analysis: Forecasting and Control. John Wiley & Sons.

Cai, Z., and Ofterdinger, U. 2016. Analysis of groundwater-level response to rainfall and estimation of annual recharge in fractured hard rock aquifers, NW Ireland. Journal of Hydrology, 535: 71–84. doi:10.1016/j.jhydrol.2016.01.066.

Chen, Z., Grasby, S.E., and Osadetz, K.G. 2004. Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada. Journal of Hydrology, 290: 43–62. doi:10.1016/j.jhydrol.2003.11.029.

Cryer, J.D., and Chan, K.-S. 2008. Time Series Analysis. Springer New York, New York, NY. doi:10.1007/978-0-387-75959-3.

Dean, R.T., and Dunsmuir, W.T.M. 2016. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. Behavior Research Methods, 48: 783–802. doi:10.3758/s13428-015-0611-2.

Dong, L., Guo, Y., Tang, W., Xu, W., and Fan, Z. 2022. Statistical Evaluation of the Influences of Precipitation and River Level Fluctuations on Groundwater in Yoshino River Basin, Japan. Water, 14: 625. Multidisciplinary Digital Publishing Institute. doi:10.3390/w14040625.

Feltovich, N. 2003. Nonparametric Tests of Differences in Medians: Comparison of the Wilcoxon–Mann–Whitney and Robust Rank-Order Tests. Experimental Economics, 6: 273–297. doi:10.1023/A:1026273319211.

Fronzi, D., Di Curzio, D., Rusi, S., Valigi, D., and Tazioli, A. 2020. Comparison between Periodic Tracer Tests and Time-Series Analysis to Assess Mid- and Long-Term Recharge Model Changes Due to Multiple Strong Seismic Events in Carbonate Aquifers. Water, 12: 3073. Multidisciplinary Digital Publishing Institute. doi:10.3390/w12113073.

Hart, A. 2001. Mann-Whitney test is not just a test of medians: differences in spread can be important. BMJ, 323: 391–393. doi:10.1136/bmj.323.7309.391.

Hyndman, R.J., and Athanasopoulos, G. 2021. Forecasting: Principles and Practice. In 3rd ed. édition. Otexts.

Kolassa, J.E. 2020. An Introduction to Nonparametric Statistics. In 1st edition. Chapman and Hall/CRC, Boca Raton. doi:10.1201/9780429202759.

Krzywinski, M., and Altman, N. 2014. Nonparametric tests. Nature Methods, 11: 467–468. Nature Publishing Group. doi:10.1038/nmeth.2937.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics, 54: 159–178. doi:10.1016/0304-4076(92)90104-Y.

Lee, J.-Y., and Lee, K.-K. 2000. Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system. Journal of Hydrology, 229: 190–201. doi:10.1016/S0022-1694(00)00158-X.

Li, G., Chen, J., Peng, D., and Gu, X. 2021. Short communication: The lag response of daily milk yield to heat stress in dairy cows. Journal of Dairy Science, 104: 981–988. doi:10.3168/jds.2020-18183.

Li, Y., and Guo, Z. 2019. Comprehensive Division of Rock Mass Structure in Granite Area–Taking Tianhu Rock Mass as An Example. IOP Conference Series: Earth and Environmental Science, 304: 052053. IOP Publishing. doi:10.1088/1755-1315/304/5/052053.

Ljung, G.M., and Box, G.E.P. 1978. On a measure of lack of fit in time series models. Biometrika, 65: 297–303. doi:10.1093/biomet/65.2.297.

Lorenzo-Lacruz, J., Garcia, C., and Morán-Tejeda, E. 2017. Groundwater level responses to precipitation variability in Mediterranean insular aquifers. Journal of Hydrology, 552: 516–531. doi:10.1016/j.jhydrol.2017.07.011.

Mehta, C.R., and Patel, N.R. 2011. IBM SPSS exact tests. IBM Corporation, Armonk, NY.

Olden, J.D., and Neff, B.D. 2001. Cross-correlation bias in lag analysis of aquatic time series. Marine Biology, 138: 1063–1070. doi:10.1007/s002270000517.

Probst, W.N., Stelzenmüller, V., and Fock, H.O. 2012. Using cross-correlations to assess the relationship between time-lagged pressure and state indicators: an exemplary analysis of North Sea fish population indicators. ICES Journal of Marine Science, 69: 670–681. doi:10.1093/icesjms/fss015.

Shumway, R.H., and Stoffer, D.S. 2017. Time Series Analysis and Its Applications: With R Examples. Springer.

Siegel, S. 1956. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill.

Trapletti, A., Hornik, K., and code), B.L. (BDS test. 2022, May 1. tseries: Time Series Analysis and Computational Finance. Available from https://CRAN.R-project.org/package=tseries [accessed 20 July 2022].

Wilcox, R.R. 1997. Some practical reasons for reconsidering the Kolmogorov-Smirnov test. British Journal of Mathematical and Statistical Psychology, 50: 9–20. doi:10.1111/j.2044-8317.1997.tb01098.x.

Yue, S., Pilon, P., Phinney, B., and Cavadias, G. 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrological Processes, 16: 1807–1829. doi:10.1002/hyp.1095.

Yule, G.U. 1926. Why do we Sometimes get Nonsense-Correlations between Time-Series?--A Study in Sampling and the Nature of Time-Series. Journal of the Royal Statistical Society, 89: 1. doi:10.2307/2341482.

Zimmerman, D.W. 2011. A simple and effective decision rule for choosing a significance test to protect against non-normality. British Journal of Mathematical and Statistical Psychology, 64: 388–409. doi:https://doi.org/10.1348/000711010X524739.