PhD dissertation submitted to the GRNE Doctoral school
« Géosciences Ressources Naturelles et Environnement »

Revisiting the concentration-discharge (C-Q) relationships with high-frequency measurements

Relations concentration-débit (C-Q) et mesures haute-fréquence

Análisis de la relación caudal-concentración (C-Q) utilizando medidas de alta-frecuencia

José Manuel Tunqui Neira

Defended on 3 December 2019 before the jury composed by:

Florentina	MOATAR	Referee
Markus	HRACHOWITZ	Referee
Ludovic	OUDIN	Examiner
Chantal	GASCUEL	Examiner
Jérôme	GAILLARDET	Examiner
Jean-Marie	MOUCHEL	Supervisor
Vazken	ANDRÉASSIAN	Supervisor
Gaëlle	TALLEC	Co-supervisor

PhD funded by the Peruvian Government

Remerciements

"Merci" est un mot si simple et pourtant si profond qu'il peut pénétrer les profondeurs de notre être. Bien que plusieurs personnes m'aient dit que j'aime trop dire ce mot à l'oral, à l'écrit c'est un peu plus difficile et surtout ici dans mon manuscrit, le produit final d'un grand effort scientifique et mental qui m'a captivé durant ces 3 années de thèse et qui n'aurait pu être terminé sans le grand soutien (à la fois scientifique, philosophique et moral) de nombreuses personnes à qui je resterai éternellement reconnaissant. J'espère que je ne vais oublier personne (car heureusement pour moi ils étaient nombreux !).

Je veux d'abord commencer par les personnes qui m'ont permis de mener à bien ce projet: mon Directeur de thèse, Jean-Marie Mouchel, une grande personne qui, avec beaucoup de patience et une grande pédagogie, m'a fait comprendre des concepts qui, dans bien des cas, étaient nouveaux pour moi et qui m'ont trop aidé dans le bon assemblage des différents sujets traités dans la thèse. Je voudrais également remercier mon co-Directeur de thèse, Vazken Andréassian. Merci beaucoup Vazken, tout d'abord pour m'avoir aidé à obtenir mon sujet de thèse, pour avoir lutté sans relâche avec la bureaucratie afin que je puisse être accueilli dans l'équipe Hydro (la convention d'accueil arrivera-t-elle un jour?), pour m'avoir appris qu'une idée, aussi simple soit-elle, est valable, et qu'en faisant un bon assemblage de celles-ci on peut obtenir de grands résultats avec un grand impact scientifique. J'aimerais aussi dire un grand merci à mon encadrante Gaëlle Tallec. Gaëlle je veux te remercier pour tout le travail que tu as fourni pour que cette thèse soit un succès. Tu étais toujours là pour m'aider, me conseiller et m'encourager dans les moments les plus difficiles, pour rire et plaisanter dans les moments les plus calmes. Plus qu'une encadrante, tu étais/est/seras une grande amie. Je ne pourrai jamais te remercier assez pour tout ce que tu as fait pour moi.

Je tiens également à remercier les membres du jury d'avoir accepté de participer à l'évaluation du manuscrit et à la discussion qui a suivi le jour de la soutenance : à Florentina Moatar et Markus Hrachowitz en tant que rapporteurs et Chantal Gascuel, Jérôme Gaillardet en tant qu'examinateurs et Ludovic Oudin en tant que président du jury.

Par ailleurs, je remercie aussi les personnes qui ont suivi mes travaux *via* mon comité de thèse : Jean-Luc Probst, Julien Bouchez, Julien Tounebize et Michel Meybeck m'ont apporté une aide précieuse et leurs expertises ont guidé ces travaux dans la bonne direction. J'aimerais également remercier François Bourgin, Patrick Ansart et Cédric Chaumont pour leurs commentaires et leur apport constructif qui m'ont beaucoup aidé dans la partie final de mes travaux.

Ce travail de recherche n'aurait pas pu être réalisé sans le financement du gouvernement péruvien à travers la subvention Cienciactiva de CONCYTEC, ainsi que la base de données haute fréquence utilisée, construite grâce au programme EQUIPEX -CRITEX.

Maintenant, je voudrais remercier les membres de la grande équipe Hydro d'Irstea Antony. Je dois avouer que de toutes mes expériences professionnelles (qui ont été très variées et internationales), c'est le meilleur endroit où j'ai pu travailler. Je voudrais commencer par remercier les plus expérimentés: Charles Perrin et Maria Helena Ramos, grandes références en matière d'hydrologie et de la prévison d'ensemble. Je tiens également à remercier le "coupable " de ma première arrivée dans l'équipe en tant que stagiaire, Guillaume Thirel, une personne magnifique tant sur le plan personnel que professionnel. A Olivier Delaigue pour son aide précieuse dans la réalisation de mes codes sous R ainsi que les bons conseils pour lancer un frisbee correctement! (cela m'a pris 3 ans mais je peux dire que j'ai réussi héhé). A Pierre Nicolle pour les grandes parties de volley-ball et de frisbee ainsi que les magnifiques soirées chez Mamane. Mille mercis aussi à ma « marraine hydrologique » Julie et aussi à Laure pour les conversations agréables à la cantine ainsi que les pauses café et les soirées Mamanesques. A mon Italienne préférée Gaia Piazzi pour sa joie débordante, à Lila pour ses actions louables de sensibilisation à l'environnement, à Benoît pour m'avoir fait courir jusqu'au bout pour le rattraper aux jeux de frisbee. Merci aussi à François Tilmant et Valentin Mansanarez, deux personnes magnifiques que j'ai rencontrées lors de mon passage à Irstea Lyon et que j'ai agréablement retrouvées lors de mon passage à Antony, et avec qui j'ai pu profiter d'intenses matchs de volley-ball et de frisbee ainsi que de superbes soirées chez Mamane. Je tiens également à remercier les personnes qui m'ont aidé à résoudre avec succès les difficiles processus administratifs français : Nathalie Camus, Valerie Quagliozzi et Laurence Tanton. Un grand merci aux anciens doctorants qui nous ont montré le chemin pour faire une bonne thèse : Léonard, Cédric, Phillipe, Andrea, Angelica et Carine, à ceux qui sont prêts à terminer avec succès cette belle expérience : Daniela, Paul RG (Allez Red Star!) et Anthony et à la nouvelle génération de doctorants qui sont plus que certains d'avoir un grand avenir dans l'équipe : Antoine, Paul A. et Thibaut. Merci aussi aux personnes qui sont passées par ce temps et qui ne sont plus là : les stagiaires, Carina, Alban, Maria, David, Damien, Louise.

Je ne pouvais pas laisser passer l'occasion de remercier les "gens du Lavoisier": Yannick, Sylvain, Mathieu, Timo et Florent pour les conversations toujours drôles, pour m'avoir fait participer à cette fièvre appelée MPG, pour les bons moments en faisant du sport et bien sûr pour les soirées chez Mamane! Un grand merci également aux doctorants de l'équipe Arthemys : Aya et Sami, et surtout à mon grand ami Alexis, une belle personne que j'ai eu le plaisir de rencontrer depuis son stage chez l'équipe Hydro et qui commence maintenant cette aventure doctorale qui sera, j'en suis sûr,un succès!

Pour finir de parler de cette magnifique équipe Hydro, je voudrais remercier ceux avec qui pendant ces 3 années j'ai eu l'honneur de partager le bureau: Arnaud pour sa gentillesse de tous les instants,

Fernando mon grand ami/frère brésilien, unis par notre grande passion pour le football, les motos et les petits poissons mexicains. Et surtout, je tiens à remercier très sincèrement mes deux « co-bureau » de ces deux dernières années: Manon et Morgane. Merci pour tout, les filles. Ces deux dernières années ont été formidables à vos côtés. Toutes les expériences, les joies, les peines, les succès et les échecs que nous avons vécus ensemble (ce n'est pas pour rien que nous appelons notre bureau " le bureau communiste ") sont des choses que je n'oublierai jamais et que je garde profondément en moi. Merci d'avoir partagé tant de choses et de m'avoir enseigné tant de choses. Pour les bons conseils que vous m'avez toujours donnés, pour la tolérance et le soutien dans mes mauvais moments. Pour avoir pris soin de moi et m'avoir considéré comme un ami et confident. Sincèrement et sans crainte de me tromper, je peux dire que j'ai eu les deux meilleures personnes que l'on aurait pu souhaiter pour cette étape qui, hélas, se termine. Je suis sûr que même si dorénavant chacun de nous prendra des chemins différents, notre amitié restera intacte. Et surtout, je suis assuré que, grâce à leurs grandes qualités, je sais que la vie ne leur apportera que succès et bonheur sur le plan personnel et professionnel.

A MI FAMILIA

Quisera finalmente agradecer a mi familia que siempre creyo en mi y me apoyo desde el principio de esta aventura : a mis papis Melquiades y Gina, asi como a mis dos queridas hermanas Gina y Carla. Sin vuestro apoyo nunca hubiera podido llegar tan lejos. Los quiero mucho y siempre serán una parte importante en mi vida.

Abstract

Recent technological advances allow measuring high-frequency chemical concentrations in rivers over long periods. These new data sets, well adapted to the temporal variations of discharge, allows us today to specify the links between hydrological processes in catchments and the water stream chemistry. However, they require the development of adapted methods for data treatment. This thesis tries to answer to the following questions: which models and methods can we use to exploit high-frequency measurements and the way they are transforming our knowledge of the chemical water-quality?

During the course of this thesis, we adapted different methods and methodologies originally designed for low / medium frequency data and applied then to high-frequency dataset of the *River Lab* of the Oracle-Orgeval observatory (France).

For many years, since the size of the C-Q datasets was limited, it was difficult to analyse in much detail the precise shape of the C-Q relationship. In many cases, the power-law relationship appeared adequate, which explains its popularity, although many additions to the basic relation have been proposed to improve it. With the advent of high-frequency measuring devices, all the range of the relationship can now be included in the analysis. As a progressive alternative to the power law relationship and a log-log transformation, we propose to use a two-sided affine power scaling relationship.

Hydrograph separation is perhaps one of the oldest unsolved problems of hydrology. In the thesis we aim to use jointly the Recursive Digital Filter (RDF) and Mass Balance (MB) methods in order to identify the RDF model parameter leading to the most realistic MB parameters. We show that a simple methodology proposed for the hydrograph separation (RDF-MB coupling approach) works, with a specific calibration and with the simple hypothesis of two sources of path flow.

To combine the power-law relationship and the two-component mixing model, we applied the two-side affine power scaling relationship to the so-called base flow and quick flow (C_b and C_q) components, with a multicriterion identification procedure. The new combined model significantly improves, compared to power and mixing models, the simulation of stream river concentrations.

Last, we develop a methodology for identifying and quantifying sources from a purely chemical point of view. The new method developed here, without any preliminary assumption on the composition of the potential sources, allows us analyzing the temporal variability of the end-member sources and their relationship to the different flow regimes.

Résumé

Les progrès technologiques récents permettent de mesurer à haute-fréquence les concentrations en ions dissous des eaux de rivières, sur de longues périodes. Ces nouvelles données, bien adaptées aux variations temporelles des débits, permettent aujourd'hui de préciser les liens entre les processus hydrologiques du bassin versant et la chimie du cours d'eau. Cependant, elles nécessitent le développement de méthodes adaptées. Cette thèse tente de répondre aux nouvelles questions qui se posent aujourd'hui: quels modèles et méthodes pouvons-nous utiliser pour exploiter les données haute-fréquences et comment transforment-elles notre connaissance de la qualité chimique des rivières?

Au cours de cette thèse, nous avons adapté différentes méthodes et méthodologies conçues à l'origine pour les données basse / moyenne fréquence et les avons appliquées au jeu de données haute-fréquence du *River Lab* de l'Observatoire Oracle-Orgeval (France).

Pendant de nombreuses années, la taille des jeux de données concentrations-débits ayant été limitée, il était difficile d'analyser de manière détaillée la forme précise de la relation C-Q. Dans de nombreux cas, l'équation de puissance précédée d'une transformation logarithmique, semblait adequate. Aujourd'hui, toute la gamme des relations C-Q à haute-fréquance peut maintenant être incluse dans l'analyse. De cette dernière, comme alternative à la relation de puissance, nous proposons d'utiliser une transformation affine de puissance bilatérale.

La séparation d'hydrogramme est peut-être l'un des plus anciens problèmes non résolus de l'hydrologie. Dans la thèse, nous avons utilisé conjointement les méthodes de séparation d'hydrogramme de type filtre numérique (RDF) et une équation de mélange à deux composantes basée sur le bilan de masse (MB). Le but etait d'identifier le paramètre du modèle RDF menant aux paramètres de l'équation de mélange les plus réalistes. Nous montrons que cette approche de couplage RDF-MB fonctionne avec un étalonnage spécifique et sur l'hypothèse simple de deux sources d'écoulement.

Pour combiner la relation simple de puissance et le modèle de mélange, nous avons appliqué la transformation affine de puissance bilatérale aux deux composantes de l'équation de mélange, à l'aide d'une procédure d'identification multicritère. Le nouveau modèle combiné améliore considérablement, par rapport aux modèles de puissance et de mélange, la simulation des concentrations dans le cours d'eau.

Enfin, nous avons développé une méthodologie pour identifier et quantifier les sources sur la seule base d'une analyse chimique. La nouvelle méthode développée au cours de la thèse, sans aucune hypothèse préalable sur la composition des sources potentielles, nous permet d'analyser la variabilité temporelle des sources chimiques et leur relation avec les différents régimes d'écoulement.

Sinopsis

Los recientes avances tecnológicos han permitido medir las concentraciones de iones disueltos en el agua de los ríos de manera continua (alta frecuencia) durante largos periodos de tiempo. Estos nuevos datos bien adaptados a las variaciones temporales de los caudales, permiten con una mayor precisión determinar los vínculos entre los procesos hidrológicos de la cuenca hidrográfica y la química del rio. Sin embargo esta tecnología requiere del desarrollo de métodos apropiados. Esta tesis intenta responder a las nuevas preguntas que surgen al día de hoy: ¿qué modelos y métodos podemos utilizar para explotar los datos de alta frecuencia y cómo transforman nuestro conocimiento de la calidad química de los ríos?

Durante esta tesis, adaptamos y/o desarrollamos diferentes métodos y metodologías originalmente diseñados para datos de baja/media frecuencia (medidas puntuales hechas una vez cada semana/día) y las aplicamos al conjunto de datos de alta frecuencia del laboratorio *in-situ River Lab* del observatorio Oracle-Orgeval (Francia).

Durante muchos años, el tamaño de los conjuntos de datos de caudales y concentraciones fue limitado, lo que dificultó el análisis detallado de la relación C-Q. En muchos casos, la ecuación de tipo potencia transformada en su forma logarítmica parecía lo mas adecuado. Ahora gracias a la alta frecuencia toda la gama completa de relaciones C-Q puede incluirse en el análisis y asi, podemos proponer una alterntiva a la ecuación de tipo potencia.

La separación de hidrogramas es quizás uno de los problemas más antiguos no resueltos en hidrología. En la tesis, utilizamos conjuntamente los métodos de separación de hidrogramas del tipo de filtro digital (RDF) y una ecuación de mezcla de dos componentes basada en el balance de masa (MB). El objetivo era identificar el parámetro del modelo RDF que conduce a los parámetros más realistas de la ecuación de mezcla. Demostramos que este enfoque de acoplamiento RDF-MB funciona con una calibración específica y con la simple suposición de dos componentes de caudal.

Para combinar la ecuación de tipo potencia y el modelo de mezcla, aplicamos la transformación afín de potencia bilateral a ambos componentes de la ecuación de mezcla, utilizando un procedimiento de identificación multicriterio. El nuevo modelo combinado mejora significativamente la simulación de las concentraciones en el río. Sin embargo, no distingue las fuentes químicas del río y sus variaciones.

Finalmente, hemos desarrollado una metodología para identificar y cuantificar componentes químicas basadas únicamente en el análisis químico. El nuevo método desarrollado durante la tesis, sin hipótesis previas sobre la composición de las fuentes potenciales, permite analizar la variabilidad temporal de las fuentes químicas y su relación con diferentes regímenes de caudal.

Contents

General introduction

1 Context

Rivers are among the most threatened ecosystems of the world. For more than a century, river science has evolved to define these threatening trends and the mechanisms that cause them. What has emerged, while still incomplete, is a picture of imposing complexity. Rivers increasingly suffer from pollution, water abstraction, channelization, and damming. Fundamental knowledge of ecosystem structure and function is necessary to understand how human activities interfere with natural processes and what interventions are feasible to rectify this.

Today, it is important **to guide management toward a sustainable future of riverine ecosystems**. For more than half a century, management science has striven to base decisions primarily on experiment-driven data. Management based on conventional, tradition-based intuition or opinion has often been the default option when measurement proves difficult. In France, the evaluation of the chemical quality of river water, the mission of the French Water Agencies, focuses on two types of indicators: annual concentrations (arithmetic means, medians, upper quantiles, flow weighted means) that are reported to a quality grid (SEQ-Eau) and annual flows generally reduced to the surface of the producing catchment area. This assessment of concentrations and flows is carried out on the basis of regular monitoring monthly, bi-monthly frequency (in the best of cases). These data are available over relatively long time series, and can be used for mass balances of watershed inputs/outputs, or to detect and measure long-term trends in water quality (Godsey et al., 2009; Moatar et al., 2017). However, these point measurements do not allow to know in detail all the different hydrochemical processes involved in the formation of the stream water chemical concentrations (Kirchner et al., 2004; Neal et al., 2012).

Efforts to measure are often stymied by resource (time, money) limitations and system complexity. However, the long-term environmental observatories combined with the grant project EQUIPEX, allow us today to have access to new high-frequency data, for a number of solutes. These considerable efforts of measurement give us a new picture of the river system. Over the past two and a half decades, continuous chemical concentration measurement studies have been conducted to solve the lack of information and clarify the links between catchment hydrological processes and stream chemistry (e.g Floury et al., 2017; Jarvie et al., 2001; Neal et al., 2013; Robson, 1993). Figure 1 gives an idea of the total loss of information related to the frequency of sampling (sub-hourly, daily, monthly). It shows the different types of sampling (high-frequency, medium frequency and low frequency respectively)

carried out on the Oracle-Orgeval observatory (Avenelles sub-catchment) from June 2015 to July 2017 (example for chloride). Due to the difficulties of implementation and the relatively high costs, the first high-frequency measurements were performed over short periods of a few days or during a storm event (Brick and Moore, 1996; Chapman et al., 1997) and this lead to significantly underestimating the legacy of past hydrological conditions (Kirchner, 2006). However, with the development of in situ chemical measurements for certain parameters such as conductivity and nitrate probes (e.g. Jarvie et al., 2001; Robson, 1993) or for a set of parameters such as River-Lab (Floury et al., 2017), we have long series of high-frequency chemical measurements of the water stream.

It has been shown that refining the time between two samples (C, Q) makes it possible to reveal behaviors (chemical, physical and hydrological) that were not previously apparent at larger time steps, especially in small catchments (Kirchner et al., 2004). In addition, measuring the physico-chemical parameters continuously as in the case of discharge allows direct comparisons between chemistry and streamflow. The new availability of high-frequency monitoring data raised renewed interest in the concentration-discharge (C-Q) relationship (Bieroza et al., 2018; Kirchner, 2019; Zhang et al., 2016). High-frequency analysis with hourly (or sub-hourly) time steps makes it possible to highlight day cycles, but also biological phenomena (Moraetis et al., 2010; Rusjan et al., 2008), to identify the apparition of a point pollution and its origin (Bowes et al., 2015) or to reveal the sources of water stream chemical concentrations (Chapin et al., 2004).

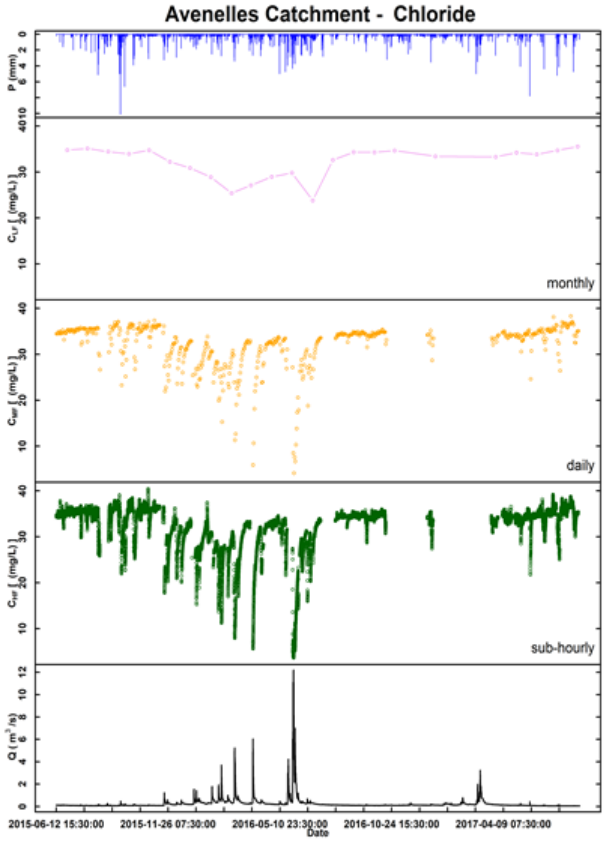


Figure 1 : Precipitation (blue line), Sub-hourly flow (black line) and Chloride concentration measurements: in green high-frequency measurements (sub-hourly.), in orange medium frequency measurements (daily) and in purple, low frequency measurements (monthly) taken in the Avenelles gaging station from June 2015 to July 2017

High frequency observations over a day may be equivalent to the amount of bimonthly data collected over more than two years. Before we could only hear one note every minute or two of a Beethoven symphony (Kirchner et al., 2004). Today we can finally hear each note of this symphony. How is this new symphony of high-frequency measurements transforming our knowledge of chemical water-quality? Will it allow us to hear new chords? To find new harmony between concentrations and flow? If these new data require new processing and storage, can we still analyze it with the tools of yesterday? Can we still play a complete Beethoven symphony on a clavichord, where the same string is used for multiple notes? Will hydrologists, who have been searching for reliable and realistic hydrograph separation methods for a long time, be able to adapt their methods to these new data? Will geochemists improve their equations closely associated fluctuations in concentrations and flow rates? After all who did not dream to know the chemical and hydrological sources of the river to predict the concentration of pollutants or to estimate more realistic nutrients budgets in surface waters?

During the course of this thesis, we adapted different methods and methodologies originally designed for low / medium frequency data and applied to high-frequency dataset of the River-Lab (Floury et al., 2017) of the ORACLE-Orgeval observatory. We explored different methods of hydrograph separation,

different C-Q relationships, and tested new models estimating end-member sources, with the aim of developing different parsimonious models that help us simulate adequately the chemical concentrations as well as the different flow components of our catchment.

2 Scientific questions of this thesis

This thesis represents an attempt to address the three following questions: (i) Can in-stream chemical concentrations be described based on discharge, without any mixing assumption? (Figure 2); (ii) can in-stream chemical concentrations be described based on the dynamics of several sources (end-members) linked to discharge? (Figure 3); and (iii) can in-stream chemical concentrations be described based on the dynamics of several sources (end-members), without any direct relation to discharge (Figure 4).

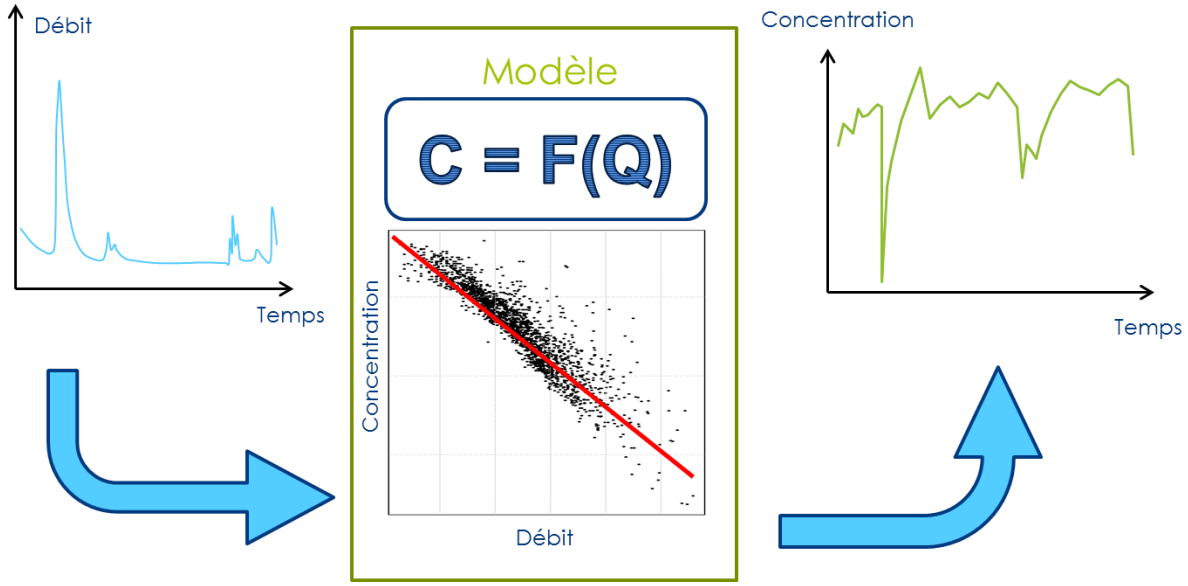


Figure 2. Objective 1: assessing whether in-stream chemical concentrations can be described based on discharge, without any mixing assumption

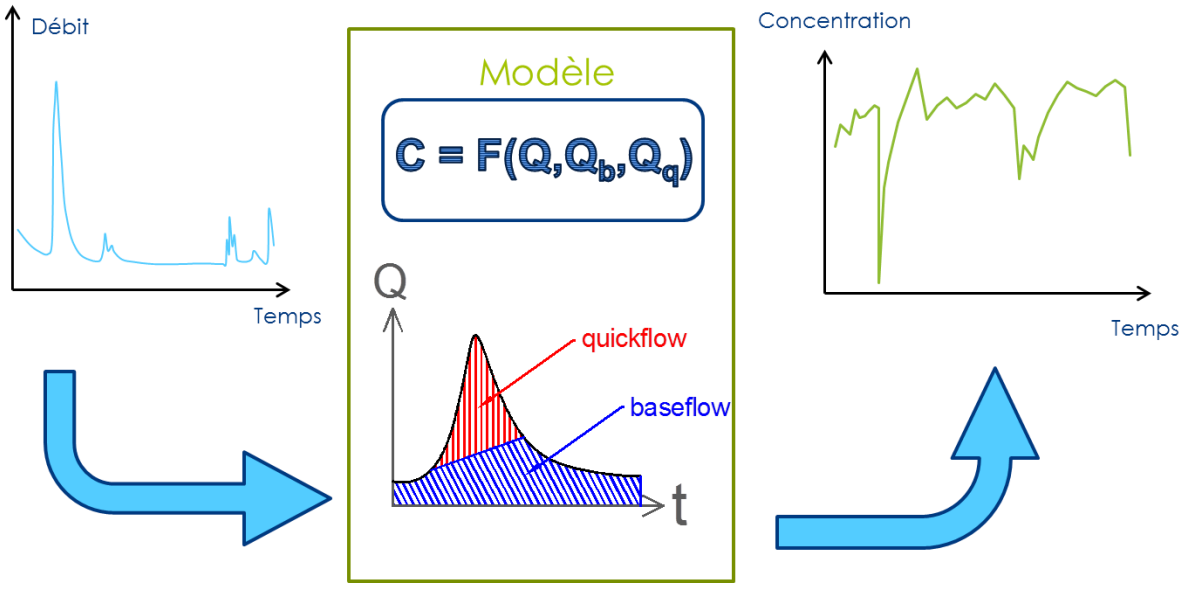


Figure 3. Objective 2: assessing whether in-stream chemical concentrations can be described based on the dynamics of several sources (end-members) linked to discharge

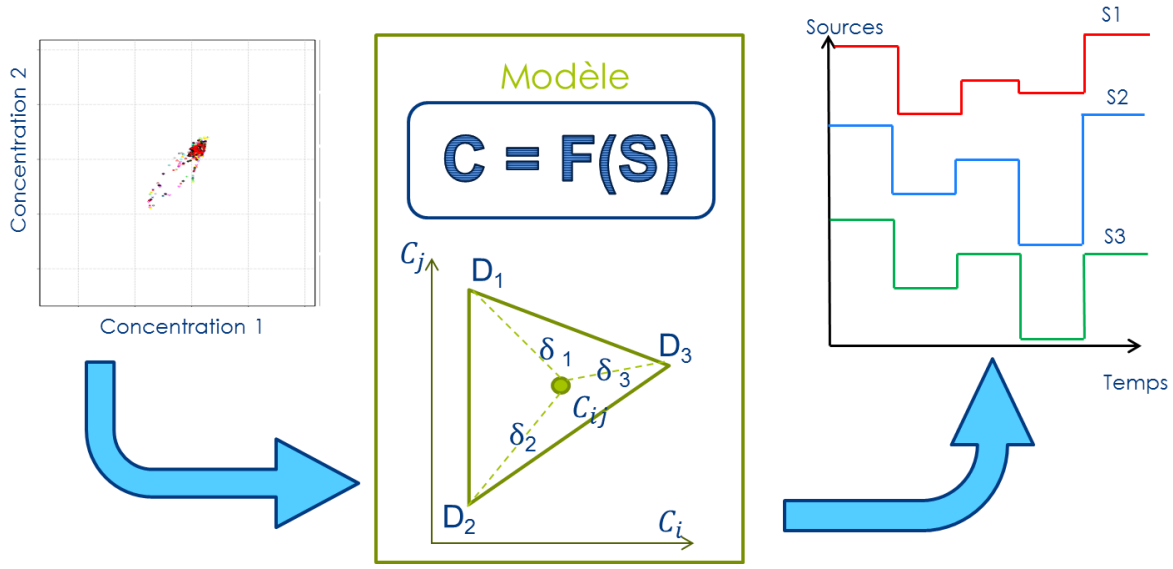


Figure 4. Objective 3: assessing whether in-stream chemical concentrations can be described based on the dynamics of several sources (end-members), without any direct relation to discharge

3 Structure of the thesis

This thesis consists of six main chapters. After a first chapter giving a **brief review of the concentration-discharge (C-Q) relationships (Part I)**, we present the thesis works through four chapters as scientific articles (Part II).

The first article, « A two-sided affine power scaling relationship to represent the concentration–discharge relationship » (Tunqui Neira et al., 2019), **revisiting the one-component model as power-law relationship**, deals with the mathematical representation of concentration-discharge relationships and with the identification of its parameters. For many years, since the size of the C-Q datasets was limited, it was difficult to analyze in much detail the precise shape of the C-Q relationship. In many cases, the power-law relationship appeared visually adequate (and conceptually simple), which explains its popularity, although many additions to the basic relation have been proposed to improve it. With the advent of high-frequency measuring devices, the size of the datasets has exploded, and all the range of the relationship can now be included in the analysis. As a progressive alternative to the power-law relationship, we propose to use a two-sided affine power scaling relationship (2S-APS) (Box and Cox, 1964). Overall, the two-sided affine power scaling relationship with a parameter identification procedure improves fit quality over the entire calibration period dataset (17,500 points) and the validation period dataset (3,200 points).

The second article focuses on “the hydrograph separation issue using high-frequency chemical measurements”, **revisiting the n-component model as mixing model.** Hydrograph separation and the identification of the baseflow contribution to streamflow is perhaps one of the oldest unsolved problems of hydrology (Hall, 1968; Nathan and McMahon, 1990). In this paper we aim to use jointly the Recursive Digital Filter (RDF) and Mass Balance (MB) methods in order to identify the RDF model parameter leading to the most realistic (most stable). MB parameters. Because we are conscious of the limits of the hydrograph separation methods and because we want to propose a generic methodology, we tested two RDF methods and six ions concentrations plus the conductivity. We show that a simple methodology proposed for the hydrograph separation (RDF – MB coupling approach) works for certains ions and EC, with a specific calibration and with the simple hypothesis of two sources of path flow.

The third article, « Combining concentration-discharge relationships with mixing models », discusses a new model, **combining a power-law relationship (2S-APS) and the two-component mixing model based on the mass balance (MB) approach**. With this aim, we developed a unifying hypothesis, that introduces the benefits of the a two-sided affine power scaling relationship (2S-APS) (Tunqui Neira et al., 2019) into the representative chemical components (C_b and C_q) of the base flow (Q_b) and quick flow (Q_q) of the mixing equation using a multicriterion identification procedure. The new reconciliation model significantly improves, compared to power and mixing models, the simulation of stream river concentrations. The new conceptual model proposed in this paper does not contradict the classical streamflow-dependent and mixing approaches but it reconciles them.

The fourth article, **« Identification of potential end members and their apportionment from downstream high-frequency chemical data »,** develops a methodology for identifying and quantifying

sources from a purely chemical point of view. A series of reasonable hypothesis drive to an objective function that was applied to the high-frequency dataset of the Oracle-Orgeval observatory. The new method developed here, without any preliminary assumption on the composition of the potential sources, allows us analyzing the temporal variability of the end-member sources and their relationship to the different flow regimes.

The last chapter (**Conclusions and perspectives**, Part III) provides a global synthesis of the work, highlighting the main contributions and the implications of the results for water quality applications. Finally, some ideas for future studies are suggested.

Part I

A brief review of concentration-discharge (C-Q) relationships

Avant-propos

Concentration-discharge (C-Q) relationships are an old science topic for over 70 years. The subject refers to more than 4000 scientific publications. In this first part, we can only present a brief review of the bibliography. We will also focus more specifically on dissolved ions in rivers.

1 Introduction

The Earth system can be considered as a complex biogeochemical reactor in which water, gases, sediments and solutes interact on time scales that range from a few days to millions of years. The biogeochemical elements are transported through many different pathways or stored in various pools or storage places during short or long periods in continuous cycles.

In the different compartments of the Earth system, the chemical, biological, and geological processes vary in their rates of cycling (Figure 5). How fast substances cycle depends on their chemical reactivity (Meybeck, 1983). A given element will be present in this cycle in various specific forms, gases, ions, molecules, dissolved and particulate organic (Meybeck, 1983). Water acquires its chemical composition, by crossing the different compartments of the Earth system, from the lithosphere to the atmosphere through the biosphere, in a continuous cycle (Figure 5). It interacts with these solid-liquid and gas-liquid interfaces whose compositions and reactivity alter aqueous geochemistry (Figure 5). Biogeochemical cycles are subject to disturbance by human activities. Humans accelerate natural cycles when elements are extracted from their pools, or sources, and deposited back into the environment (Garnier et al., 2014).

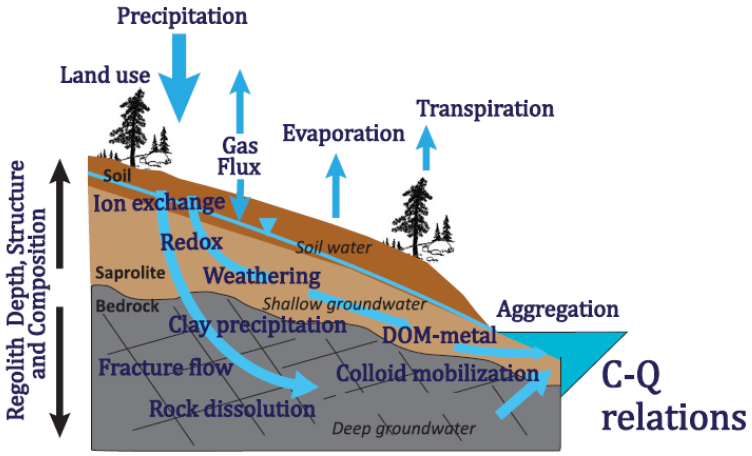


Figure 5 : An array of coupled hydrologic and biogeochemical processes superimposed on a complex structure give rise to concentration-discharge relations observed in surface waters. Source: Chorover et al. (2017)

In a catchment, water flow determines connectivity between stream and catchment, mediating what landscape components contribute particulate materials and solutes to the stream network at what times (Abbott et al., 2018). Hydrology controls the residence time of those particulates and solutes in different components of the stream network (Benettin et al., 2015; Garnier et al., 2014; Rode et al., 2016). The processes determining the concentrations in the river and therefore the quality of the water are multi-faceted and complex. River chemistry results from various origins: atmospheric inputs, soil

leaching and weathering of minerals. These solutes and particulates can reach streams and rivers through various pathways and be controlled by multiple physical, chemical and biological processes (Cohen et al., 2013). There are myriad characteristics that can alter retention or removal capacity of a catchment, most of which are not measured or measurable at catchment scales (Abbott et al., 2018; Rode et al., 2016). Variations in residence times affect the extent to which kinetically limited processes such as weathering and microbial transformation of nutrients (Moatar et al., 2017). They generate a chemical signal in the water moving through the catchment, whereas variations in hydrologic flow paths and their connectivity control the mixing of waters from distinct pools, and their relative contributions to the chemical load in receiving stream water (Moatar et al., 2017). Interactions among these processes are expected to vary across hydrologic events, such as rainfall, thereby producing a range of potential correlations between chemical composition and flux of receiving waters, i.e., **concentration-discharge** (**C-Q) relationships**.

Researchers are trying to understand all of the various pathways and flows of the biogeochemical cycle. One well studied postulate is that hydro-chemical controls can be related quantitatively to solute releases measured down-gradient of reactive flow paths of variable length and residence time, whose mixing is reflected in concentration-discharge (C-Q) relationships (Chorover et al., 2017). (C-Q) relationships have been studied in a multitude of papers, showing their complexity. Studies of **hysteresis loops**, observed in storm-event concentration–discharge plots, show the relative timing and volume of mixing of the hydrological sources (Chanat et al., 2002; Evans and Davies, 1998; Evans et al., 1999). Since the first plots between the concentration of ions and stream discharge were traced (Hem, 1948; Lenz and Sawyer, 1944) scientists are trying to set different **models** and solve this relationship. Because (C-Q) relationships are interesting for a variety of purposes and potential users, and because they employ simple approaches to describe complex hydro-chemical interactions, hydrologists and geochemists have been using and exploring them for over 70 years (Chanat et al., 2002; Durum, 1953; Hem, 1948; Johnson et al., 1969; Kirchner, 2009; Moatar et al., 2017).

2 Non-univocal concentration-discharge (C-Q) relationship (hysteresis)

Non-univocal relationships between river discharge and the concentration of a solute (hysteresis) can provide insight into river catchment functioning. Hysteresis between discharge and suspended sediment or dissolved solids during storm events was first observed by Gunnerson (1967) and since has been noted in many other water quality parameters (Bowes et al., 2009; House and Warwick, 1998; Lawler et al., 2006). Williams (1989) was one of the first to describe the most common shapes of hysteresis loops. Commonly used to examine nutrient concentrations (e.g., Bowes et al. (2015)) or sediment transport (e.g., House and Warwick (1998)) as function of discharge, hysteresis patterns can also provide information on the source(s) of water contributing to river flow (Cartwright et al., 2014; Evans and Davies, 1998; House and Warwick, 1998; Lloyd et al., 2016). The slope, rotational direction (clockwise or counter clockwise), and size of hysteresis loops reflect the source, timing, and relative contributions to rivers (Evans and Davies, 1998).

The C-Q hysteresis could be examined with different approaches and conceptualization of the hydrograph separation; i.e. as a simple early episode flushing of soluble material (a “prevent” before the event) or as a mixing component (Evans and Davies, 1998). Evans and Davies (1998) give a set of solutions to examine hysteresis through a simple three components mixing model (Figure 6).

Hysteresis patterns are generally clockwise or anti-clockwise. With clockwise patterns, solute concentrations on the rising limb of the hydrograph are higher than that on the falling limb. Conversely, with anti-clockwise patterns, the opposite holds true. Groundwater generally has a higher salinity than surface water; hence, bank return flows should result in clockwise hysteresis loops in Total Dissolved Solids (TDS) and/or Electrical Conductivity (EC), although the input of other low salinity waters such as interflow or soil water following the peak in river discharge may produce similar trends (Evans and Davies, 1998).

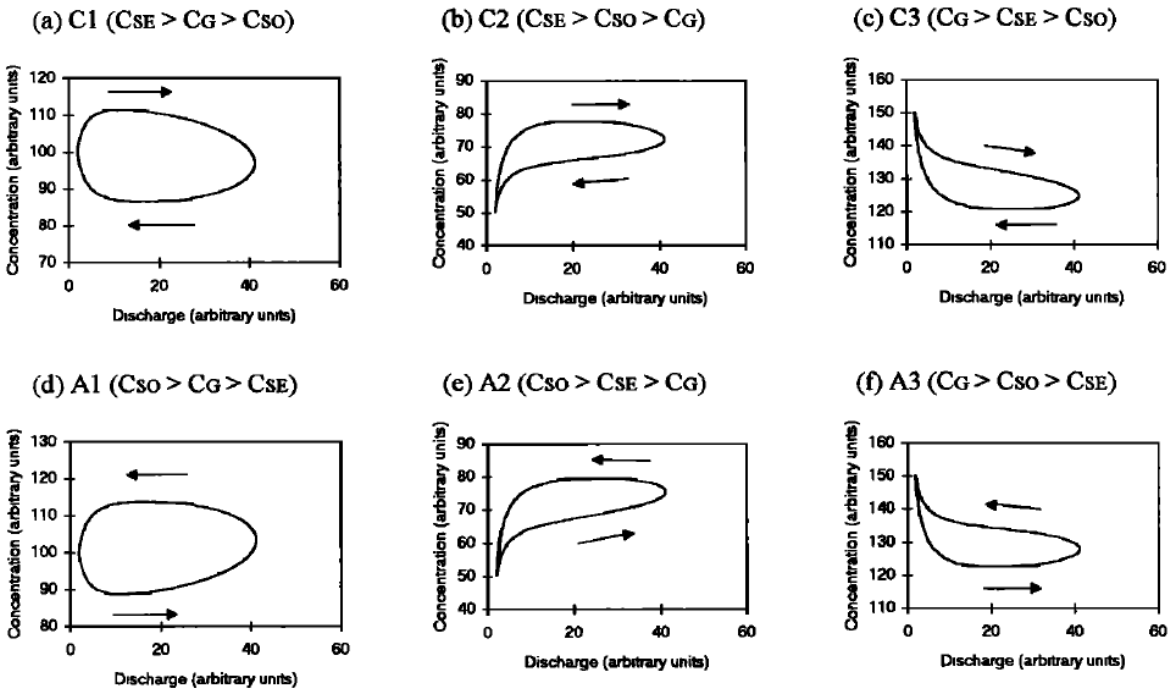


Figure 6 : (C-Q) hysteresis loops for each combination of component concentrations from a simple conceptual three-component hydrograph (Cso: Soils concentrations; CSE: surface event concentrations; CG: groundwater concentrations). (Source: Evans and Davies (1998))

Other interpretations of event-scale (C-Q) hysteresis have focused on the mixing of hydrologic sources (Bowes et al., 2009; Chanat et al., 2002; Evans and Davies, 1998; Ledbetter and Gloyna, 1964), biogeochemical transformations (Bowes et al., 2009; Musolff et al., 2017), and variability of end-member concentrations (Chanat et al., 2002) as drivers of observed patterns.

Similarly, investigations of longer term (C-Q) dynamics incorporating data from multiple storm events over many years have focused on comparing the broad-scale behaviour of solutes across flow regimes within a single catchment and across multiple catchments (Godsey et al., 2009; Howarth, 1984; Musolff et al., 2015; Reimann et al., 2002; Thompson et al., 2011). Some authors showed differences between catchment behaviours even during similar storm conditions or differences between water quality parameters under same events (Campbell and Bauder, 1940; Lloyd et al., 2016). Studying the long-term (C-Q) relationships across multiple storms and years, Rose et al. (2018) showed that (C-Q) dynamics can be explained by the sequence and timing of hydrologic source contributions and connectivity during events. According to Rose et al. (2018), the sequence and timing of hydro-biogeochemical source contributions to the stream during storms are the main drivers of observed (C-Q) patterns at event and inter-annual timescales.

3 Concentration-discharge (C-Q) relationship models

Establishing mathematical relationships between two adjacent processes is not something new in science. It is at best an equation, at worst a system of equations, with two unknowns. Since the first plots between the concentration of ions and stream discharge were traced (Hem, 1948; Lenz and Sawyer, 1944), scientists are trying to set different models and solve this relationship.

These models, empirical or conceptual, have in common to have only two types of variables: a concentration type variable and a hydrological type variable.

Today, the concentration-discharge (C-Q) relationship models can divide into two groups: one-component models and *n*-component models. To describe the C-Q relationship, the one-component models use only one hydrological volume or source (e.g. the total discharge), whereas the n-components models could use several distinct hydrological sources (e.g., baseflow and quickflow discharges).

The first ones try to explain the processes controlling the mobilization and delivery of chemical elements into streams (i.e. export regimes) as well as geochemical transformations in river networks. The *n*-component models, more complex, have additional ambition to characterize the sources of the chemical concentrations measured in the stream water.

3.1 One-component models

To analyze the C-Q relationship, these models use a statistic regression linear and non-linear. The one-component models using linear regression describe the C-Q relationship from a power-law equation. A log-log transformation allows linearizing the C-Q relationship. The one-component models using non-linear regression describe the C-Q relationship from a hyperbolic model, without transformation. For the two types of one-component model, the regression parameters or the indices derived from them, describe the different C-Q relationships and consequently different associated hydro-chemical processes.

3.1.1 Power-law model

The power-law model follows three assumptions whatever its form : (1) mixing is ideal (i.e., complete) within the volume, (2) the dissolved constituents move in the flow system in the same fashion as does the water, and (3) the influence of factors such as biologic uptake or evapotranspiration is negligible (Hall, 1970).

The power-law model presents the general form below:

$$C = aQ^b \quad \textbf{Eq. (1)}$$

Where C is the measured concentration, Q is the total discharge, a and b are dimensionless constants.

The Eq. (1) is solved with a log-log transformation and a simple linear regression:

$$\ln(C) = \ln(a) + b.\ln(Q) \quad \textbf{Eq. (2)}$$

With slope b and y-intercept $\ln(a)$.

This approach was first applied in sediment transport (Campbell and Bauder, 1940). In our knowledge, Durum (1953) was the first to apply this approach to a chemical tracer, specifically to the chloride ions measured at the Russel gauge station, Saline River catchment, Kansas, USA. Because of the high reactivity of the chloride with respect to the discharge in this catchment, Durum (1953) had to set the value of *b*=-1.

According to the b slope, Edwards (1973) established three patterns:

i. the slope is positive ($b > 0$) when the concentration of the solutes increases with the increase of discharge,

ii. the slope is negative ($b < 0$) when the concentration of the solutes decreases with the increase of discharge,

iii. the slope is near of zero ($b \approx 0$) when there is no variation in the concentration of solutes with the increase of discharge.

The three patterns correspond to the three principal mechanisms, resumed by Salmon et al. (2001), to describe the linkages between stream chemistry and hydrology: dilution ($b < 0$), hydrologically constant ($b = 0$) or enhanced hydrological access ($b > 0$) (Figure 7). According to Salmon et al. (2001), dilution occurs when water delivery to the stream is greater than the increase in chemical delivery. Hydrologically constant behavior occurs when there is a balance between changes in concentration and discharge. Enhanced hydrological access refers to an increase in chemical concentration with increasing discharge.

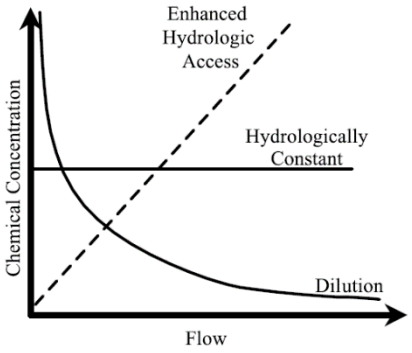


Figure 7: Schematic of chemical responses to hydrological controls (Salmon et al., 2001)

Since Durum (1953), the C-Q relationships from power-law model were studied in a number of catchments under a large panel of locations, climates and temporal scales; from UK (Edwards, 1973) to USA (e.g., Barco et al. (2008), Thompson et al. (2011), Zimmer et al. (2019)), through west of Europe and France (Meybeck and Moatar, 2012; Moatar et al., 2017) or Porto Rico (Shanley et al., 2011), etc.

Number of authors described also the relationship between load as a product of the concentration and discharge, and the stream discharge (i.e. L-Q relationships) with power-law model (Basu et al., 2010; McDonnell et al., 1991; Musolff et al., 2015; von Freyberg et al., 2017; von Freyberg et al., 2018). If some argued that it is a spurious correlation since discharge is a component of load calculation, the L-Q relationships could give good linear correlations with high R^2 coefficient (Basu et al., 2010). However, for some nutrient, this correlation could be very poor, due to the seasonal variability of the concentrations (von Freyberg et al., 2018). Hall (1971) had already noted that the C-Q relationships over long periods of time may be obscured by non-random diurnal or seasonal fluctuations due to biological or chemical processes. Because of changes in point sources, land-use practices, and atmospheric deposition, the variability of the concentrations is also observed over the years (Hirsch et al., 2010).

To solve the temporal variability, Hirsch et al. (2010), Zhang et al. (2016) and Zhang (2018) developed the Weighted Regression on Time, Discharge and Season model (WRTDS model) that uses time, discharge, and season as explanatory variables for the C-Q relationship:

$$\ln(C_i) = \beta_{0,i} + \beta_{1,i}.t_i + \beta_{2,i}.\ln(Q_i) + \beta_{3,i}.\sin(2\pi t_i) + \beta_{4,i}.\cos(2\pi t_i) + \varepsilon_i \qquad \textbf{Eq. (3)}$$

Where t_i is time in decimal years, C_i is the daily concentration at time t_i, $\beta_{0,i} \sim \beta_{4,i}$ are fitted coefficients and ε_i is the error term. The first three terms of Eq. (3) represent respectively, the intercept, time effects, and discharge effects, whereas the fourth and fifth terms collectively represent cyclical seasonal long-term effects (Zhang et al., 2016), which efficiently interpolate low-frequency time series (Minaudo et al., 2019). The main disadvantage of the WRTDS model is that it includes four parameters

and therefore, cannot be considered a parsimonious approach. These coefficients are calibrated at each time step, which makes difficult to interpret what drives the heterogeneity often observed in terms of catchment behavior (Minaudo et al., 2019).

Meybeck and Moatar (2012) and Moatar et al. (2017) proposed a split-hydrograph method to deal with the non-linearity of the long-term log-log transformation of the C-Q relationship (Figure 8). In this method the C-Q relationship is split at the median daily flow (see Figure 8) with individual slopes for low and high flows ($b_{50\,inf} < Q_{50} < b_{50\,sup}$), solving non-linearity and possible changes in hydrological and biogeochemical controls in different hydrological states (Meybeck and Moatar, 2012).

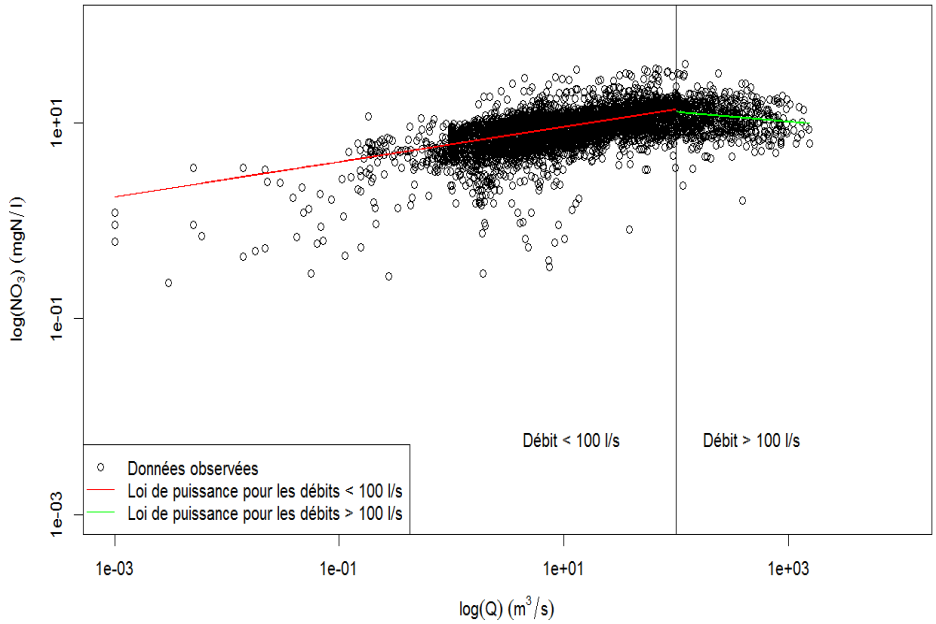


Figure 8: C-Q relationship (log–log transformation) for low frequency nitrates concentrations and discharge of Mélarchez station (Orgeval-Oracle observatory), from 1975 to 2014

With this method, from 293 water quality stations in France, Moatar et al. (2017) determine export regimes due to biogeochemical processes and develop a conceptual diagram of the interplay between biological and hydrological long-term dynamics in the C-Q relationships (Figure 9).

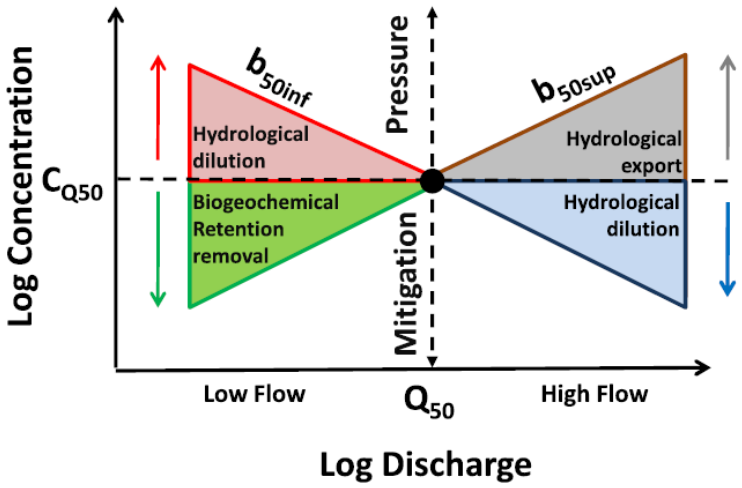


Figure 9 : Diagram of the interplay between biological and hydrological long-term dynamics in the C-Q relationships (Moatar et al., 2017)

Others used a statistical method and have compared the rate (CV) between the coefficient of variation of concentration (CV_c) to that of discharge (CV_q) (Musolff et al., 2017; Musolff et al., 2015; Thompson et al., 2011). To enhance the performance of the log-log transformation between the concentration of solutes and the discharge, relations have been established between the two indices CV and b slope (e.g Musolff et al., 2015; Zimmer et al., 2019).

According to Reimann et al. (2002), with respect to the slope (b in Eq. (2)), the CV rate has been developed to represent a different alternative to determine two types of export regime: chemostatic and chemo-dynamic. In chemostatic export regimes (when the slope is near to zero) the coefficient of variation of concentration is lower than the discharge (when the slope is near to zero and $CV_c/CV_q \leq 0.5$) and conversely for chemo-dynamic export regimes (when the slope is positive or negative and $CV_c/CV_q > 0.5$) (Reimann et al., 2002).

Chemostatic behavior occurs when the contributions of solutes coming from the different pools (groundwater, soil, and runoff) are constant and stable despite the variation of streamflow (Godsey et al., 2009; Moatar et al., 2017; Thompson et al., 2011). Chemostatic behavior was observed generally with solutes from geological weathering (Godsey et al., 2009). The ascending chemo-dynamic behavior (positive slope) is due to the enhanced erosion during high flows (Godsey et al., 2009). It supposes that in some of the hydro-chemical pools of the basin this concentration is higher than in the others. This export regime is considered transport limit because it depends more on the good interconnection of the water pathways between the pools and the stream than of the quantity of solute of the pools (Moatar et al., 2017). Last, the descending chemo-dynamic behavior (negative slope) is observed in the dilution process of the water stream solutes during high flows. This export regime is considered as source limited because its delivery to the stream is subject to the abundance or production capacity of

the solute, rather than the transport capacity in the different hydro-chemical pools of the catchment (Basu et al., 2011; Moatar et al., 2017)

From this basic concept obtained by the power-law model, it has been possible to make other practical applications to explain the hydro-chemical processes, such as hydrological residence time (Duncan et al., 2017; Genereux et al., 1993) or the influence on export dynamics of time scales and catchment characteristics (Botter et al., 2019; Godsey et al., 2009; Moatar et al., 2017).

If the power-law model contributes greatly to the knowledge of the hydro-chemical processes at the catchment scale, it has also shortcomings. It does not take into account biological processes influencing the nitrates and phosphates concentrations (Basu et al., 2010; Moatar et al., 2017; Thompson et al., 2011). Even log-log transformed, the C-Q plots can keep a large dispersion and as any linear regression it is sensitive to outliers. According to Minaudo et al. (2019) this is due to : (i) hysteresis and non-linearity effects due to source and transport limitations; (ii) instream biogeochemical transformations on nutrient concentration without temporal correlation with hydrological variations; (iii) seasonal and long-term variations in C-Q relationships.

Because regression-based methods does not rely on those methods' problematic assumptions about the homoscedasticity of model errors, constancy of seasonal trends in concentration, or constancy of the concentration-discharge relationship, others models have been developed (Chanat et al., 2002; Grunsky, 2010; Hirsch et al., 2010).

3.1.2 Hyperbolic model

To alleviate the problems mainly due to the linear regression approach of the log-log transformation (i.e. sensitivity to outliers, overfitting), Johnson et al. (1969) developed the hyperbolic model. The hyperbolic model uses as the power-law model only one component to describe the C-Q relationship (i.e. the total discharge Q). This model was developed by Johnson et al. (1969) from the mass balance equation form:

$$C_0.V_0 + C_1.V_1 = C(V_0 + V_1) \qquad \textbf{Eq. (4)}$$

Where the concentration of the solute before dilution is C_0, V_0 is the volume before dilution, C_1 is the diluent concentration, V_1 is the volume of dilution and C is the concentration from the mixed solution.

With the Eq. (4), Johnson et al. (1969) stated the following hypotheses:

(i) when discharge (Q) approaches zero, V_1 also approaches zero, thus $(V_0 + V_1) \rightarrow V_0$;

(ii) at any given time if $V_1 > V_0$, V_1 is directly proportional to discharge ($V_1 = \gamma Q$, where γ is the residence time for water moving through the soil without chemical influences);

(iii) every volume before dilution (V_0) has a characteristic initial concentration (C_0) for any given natural ion.

Under these hypotheses, Eq. (4) can be rewritten as:

$$C_0.V_0 + C_1.\gamma Q = C(V_0 + \gamma Q) \quad \textbf{Eq. (5)}$$

Solving C and grouping yields :

$$C = \frac{C_0 + C_1(\beta Q)}{1 + \beta Q} \quad \textbf{Eq. (6)}$$

Where $\beta = \gamma/V_0$ a constant. Defining $\alpha = C_0 - C_1$ and substituting in Eq. (6) yields:

$$C = \frac{\alpha}{1 + \beta Q} + C_0 \quad \textbf{Eq. (7)}$$

The parameters of the Eq. (7) (α , β and C_0) are computed using non-linear least-squares regression (Barco et al., 2008). Johnson et al. (1969) applied the Eq. (7) for major ions concentrations in the stream water of Hubbard Brook Experimental catchment. As the power-law model, the hyperbolic model distinguishes the three major processes depending on groups of ions; a dilution process and a process of concentration with increasing stream discharge, and a chemical buffering mechanisms (Gregory and Walling, 1973; Johnson et al., 1969). Subsequent studies demonstrated that the first two processes could be linked to α parameter of the Eq. (7) (e.g. Barco et al., 2008; Salmon et al., 2001). Thus $\alpha > 0$ corresponds to the dilution, while $\alpha < 0$ corresponds to the concentration process. Johnson et al. (1969) tried to relate the parameters of the hyperbolic model to physical descriptors. They related C_0 to the soil water, C_1 to the rainwater and γ to the ratio between residence time of soil water and field capacity, without chemical influences. However because these physical connections were difficult to corroborate, this model is considered empirical (e.g. Huntington et al., 1988; Lawrence and Driscoll, 1990; Lawrence et al., 1986). Nevertheless, Barco et al. (2008) and Barco et al. (2013) used it to study the key factors controlling geochemical export in catchments at seasonal and inter-annual time scale.

3.2 N-component models

3.2.1 Mixing models

To approximate the physical processes involved, n-component models or mixing equations, still based on mass balance equation, were developed (Dewalle et al., 1988; Pinder and Jones, 1969; Sklash et al., 1976; Stewart et al., 2007; Uhlenbrook and Hoeg, 2003). These models assume that the chemical concentrations of the stream water are the result of the mixing of various chemical components inherent in the different flow components. In this model, the composition or chemical signature of water is constant and different from each other components (Pinder and Jones, 1969). The mixing

models or n-component models, may include several flow and chemical components (e.g. Dewalle et al., 1988; Miller et al., 2017; Uhlenbrook and Hoeg, 2003). The simplest, with two-component was initially applied by La Sala Jr. (1967): "*The water in a stream usually is a mixture of surface runoff and groundwater. The chemical quality of the stream water is therefore, determined both the chemical qualities of overland runoff and groundwater and by the proportion of each"*.

Following this simple definition, the equation of the 2-components mixing model could be written:

$$C = C_b \frac{Q_b}{Q} + C_q \frac{Q_q}{Q} \qquad \textbf{Eq. (8)}$$

Where C is the concentration observed in the stream, Q is the total observed discharge, Q_b is the groundwater discharge assimilated to the baseflow, Q_q is the runoff discharge assimilated to the quickflow, C_b is the characteristic concentration of groundwater and C_q of runoff.

The baseflow represents the longer-slow discharge derived from the groundwater (Hall, 1968; Nathan and McMahon, 1990). The quickflow represents the direct response to a rainfall event including overland flow (runoff) and direct rainfall onto the stream surface (direct precipitation) (Linsley Jr et al., 1958; Pinder and Jones, 1969).

Mixing models are used by geochemist to study hydrochemistry processes and estimate the end-member sources of the stream water quality (e.g Evans and Davies, 1998; Genereux et al., 1993; Probst, 1985; Uhlenbrook and Hoeg, 2003). They are also used by hydrologists and hydrogeologists to separate quickflow from baseflow (e.g. Nathan and McMahon, 1990; Zhang et al., 2017; Zhang et al., 2013). Dependent on discipline (geochemistry, hydrology or hydrogeology) and purpose, lateral movement in the soil profile (interflow) is integrated to the baseflow (e.g. Buttle, 1994; Klaus and McDonnell, 2013; Sklash et al., 1976) or the quickflow (e.g. Brodie et al., 2007; Chapman and Maxwell, 1996; Nathan and McMahon, 1990). Baseflow and quickflow could be also assimilated to the "old water" and "new water" used by geochemists who work with isotopes (Bansah and Ali, 2017; Klaus and McDonnell, 2013).

Whatever the discipline and the components number, the key points of the mixing models are the hydrograph separation and /or the quantification of sources contribution.

4 Hydrograph separation methods

Hydrograph separation is the oldest age topic in hydrology (Hall, 1968). It is a deconstructive method of streamflow as a two-component or multi-component process (Mei and Anagnostou, 2015). The most commonly used scheme is the two-component scenario that considers streamflow consisting of direct flow (i.e. quick surface or subsurface flow) and baseflow (i.e. flow that comes from groundwater storage or other delayed source) (Hall, 1968; Tallaksen, 1995). As shown in Figure 10, classical hydrograph separation is defined in terms of the delay or lag times of the components, without implication of origin (Hall, 1968).

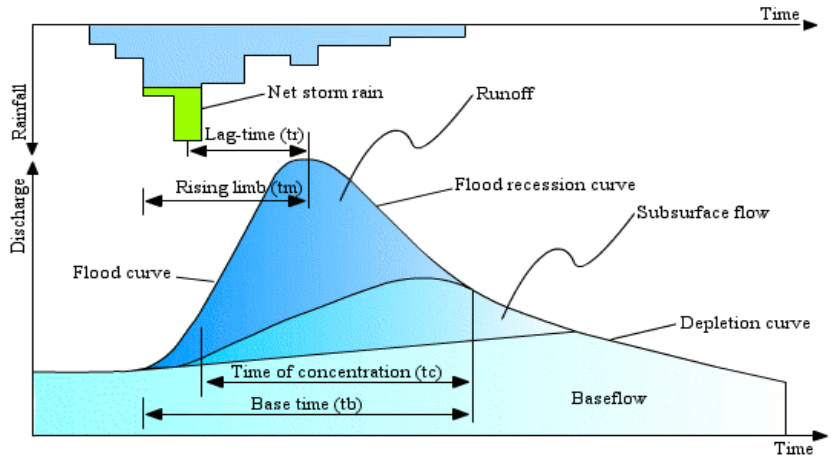


Figure 10 : Classical separation of the different components in hydrograph (Musy, 2001). Two principal methods of hydrograph separation are used, the non-tracers-based methods and the tracers-based methods.

Methods that are not based on tracers are mainly founded on the study of recession curves. The flood recession curve extends from the peak flow rate onward (Figure 10). The end of stormflow (i.e. quickflow or direct runoff) and the return to groundwater-derived flow (baseflow) is often taken as the point of inflection of the recession limb (Figure 10). The recession limb represents the withdrawal of water from the storage built up in the basin during the earlier phases of the hydrograph. From Boussinesq (1877), Maillet (1905) and Horton (1933), the groundwater-derived flow (baseflow) can be expressed by the equation of a single linear reservoir as follow:

$$Q_{(t)} = Q_{0(t)}.exp^{-(\propto t)} \qquad \textbf{Eq. (9)}$$

Where Q_t is the streamflow at time t, Q_0 initial discharge, and $\propto$ a constant. The term $exp^{-(\propto)}$ is normally replaced by K called the recession constant (see Eq. (10)).

$$Q_{(t)} = Q_{0(t)}.K^t \qquad \textbf{Eq. (10)}$$

Some authors (Chapman and Maxwell, 1996; Gustard and Demuth, 2009; Tallaksen, 1995) also use the hydrological recession time constant (τ) of the catchment (in time units) expressed by Eq. (11). Whereas τ has the dimension of time, K is a dimensionless quantity whose value depends on the time unit chosen (Tallaksen, 1995).

$$\tau = -\frac{t}{\ln(K)} \qquad \textbf{Eq. (11)}$$

The recession analysis methods focus primarily on the study of the recession parameter, whatever its form (K, exponential form $exp^{-(\propto)}$ or K= f (τ)). The recession parameter could be associated with a linear reservoir or a non-linear reservoir, each authors trying to be as close as possible to a physical description of this parameter.

Barnes (1939) popularized a first graphical technique to determine the recession constant. The technique, known as a semi-logarithmic plot of flow recession, is originally applied to a single storm event. For analyzing, a set of hydrograph recessions at a particular catchment simultaneously, a master recession curve (MRC) is commonly used (e.g. Nathan and McMahon, 1990). These techniques are inconvenient when separations are to be undertaken on a long continuous record of stream flow, rather than just a few storm period hydrographs, and this has led to the development of numerical algorithms (Chapman, 1999). In addition to the graphical method, analytical approaches of the recession curves have been proposed, based on regression analysis and a non-linear reservoir algorithm (Rutledge, 1998), on wavelet transform (Robson, 1993) or filtering processes (e.g Eckhardt, 2005; Lyne and Hollick, 1979).

To understand the process by which rainfall becomes stream flow, a method was developed used tracers. It allows to determine the time-varying proportion of “old water” and “new water” in the stream hydrograph (Chapman and Maxwell, 1996; Pinder and Jones, 1969). Here “old water” is identified as being water that was already in the catchment before the start of rainfall, while the “new water” has the same quality characteristics as the incoming rainfall (Chapman, 1999). In the precedent methods, recession analysis, the baseflow is generally regarded as being a result of groundwater discharging into the stream, while the direct runoff is considered to result from overland or near-surface flow (Chapman et al., 1997). Table 1 non-exhaustively resumes the main methods of hydrograph separation.

Table 1: Main methods of hydrograph separation.

Approaches		Methods	References
Non tracers-based method	Recession curves analysis methods	– Semi-logarithmic plot technique – Master Recession Curves (MRC)	Nathan and McMahon (1990) Anderson and Burt (1980) Barnes (1939) Snyder (1939) Kulandaiswamy and Seetharaman (1969)
	Filtering methods	– Smoothed Criteria Technique (SMT)	Hydrology (1980) Gustard et al. (1992)
		– Recursive Digital Filter (RDF)	Lyne and Hollick (1979) Chapman and Maxwell (1996) Bentura and Michel (1997) Eckhardt (2005)
Tracers-based method		– Natural tracers – Isotopic tracers	Ladouche et al. (2001) Mul et al. (2008) Pinder and Jones (1969) Uhlenbrook and Hoeg (2003)

4.1 Non tracers-based method

4.1.1 Recession curves analysis methods

A Semi-logarithmic plot technique

Studying the structure of the discharge-recession curves, Barnes (1939) was the first to popularize a graphical method. The technique is known as a semi-logarithmic plot of flow recession used for separating a hydrograph into linear components of surface flow, interflow and baseflow.

The graphical method starts with identifying the points at which the direct runoff starts and ends. The start point is readily identified as the time when the flow starts to increase, while a wide range of graphical techniques is available to define the end-point (Dickinson et al., 1967). Figure 11 presents the most used graphical methods:

(i) The constant discharge method (line A-A'), assumes that the baseflow is constant during the flood hydrograph. The minimum streamflow immediately prior to the rising limb is used as the constant value;

(ii) The constant slope method (line A-B) link the start of the rising limb with the inflection point on the receding limb. This assumes an instant response in baseflow to the rainfall event (Nathan and McMahon, 1990);

(iii) The concave method (line A-C-B) attempts to represent the assumed initial decrease in baseflow during the climbing limb by projecting the declining hydrographic trend evident prior to the rainfall event to directly under the crest of the flood hydrograph. This minima is then connected to the inflection point on the receding limb of storm hydrograph to model the delayed increase in baseflow.

For the constant slope and concave methods the inflection point (point B, Figure 11) can be determined from the peak of the hydrograph using the empirical equation proposed by Linsley Jr et al. (1958):

$$N = 0.827.A^{0.2}$$ **Eq. (12)**

Where N is the number of days between the storm crest and the end of quickflow, and A is the catchment area in square kilometers. The constant 0.2 can vary depending of catchment characteristics such slope, vegetation and geology (Brodie et al., 2007).

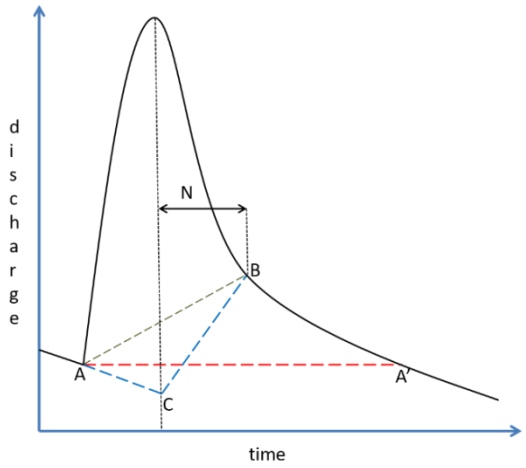


Figure 11: Graphical separation methods including: (A-A') constant discharge method; (A-B) constant slope method and (A-C-B) concave method.

The main advantages of this method is that it is easy to apply and consistent (Nathan and McMahon, 1990). The principal disadvantage of this method is that it cannot be applied in the whole continuous hydrograph (Gonzales et al., 2009; Nathan and McMahon, 1990). Tallaksen (1995) pointed out high variation in recession behavior, not only between catchments but also within a catchment. To provide an average characterization of baseflow in a catchment it is necessary to combine individual baseflow recession (Robson, 1993).

B Master Recession Curve method

The master recession curve (MRC) combines individual recession curve to estimate one recession constant from a catchment. For constructing the MRC, the three principal methods are the correlated

method (Langbein, 1938), the matching strip method (Snyder, 1939) and the BN 77 method (Brutsaert and Nieber, 1977).

The correlation method involves plotting on natural scales discharge at one time against discharge at some arbitrary interval N days later during known recession periods (Figure 12a). All periods during the entire N day period are used and individual points are linked (Nathan and McMahon, 1990). The traces of the individual recession periods generally describe an arc that increasingly becomes steeper as the discharge decreases. An enveloping line can be usually drawn along the perimeter of the region where the curves run together most densely. This enveloping line is defined as the master recession curve, and its slope is used as the recession constant.

The matching strip method involves plotting individual recessions on tracing paper. The recessions are then superimposed and adjusted horizontally until the main recessions overlap to form a set of common lines (Figure 12b). A mean line through the set of common lines represents the master recession curve, and its slope the recession constant (Nathan and McMahon, 1990).

(a) MRC with correlation method (b) MRC with matching strip

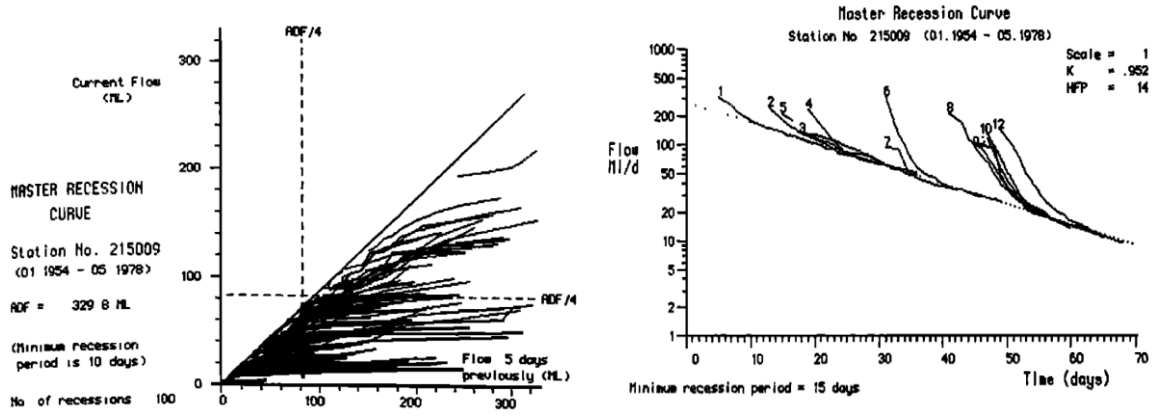


Figure 12 : Example of MRC with (a) correlation method and (b) matching strip (Nathan and McMahon, 1990)

To eliminate the uncertainties involved in defining recession periods from the hydrograph, the BN 77 method expresses the rate of change in flow for a differential dt (Q vs dQ/dt) (Figure 13). The slope of this regression is the lower envelope; i.e. the locus of points for the slowest recession rate by keeping roughly 5% of the data points below it (Cheng et al., 2016).

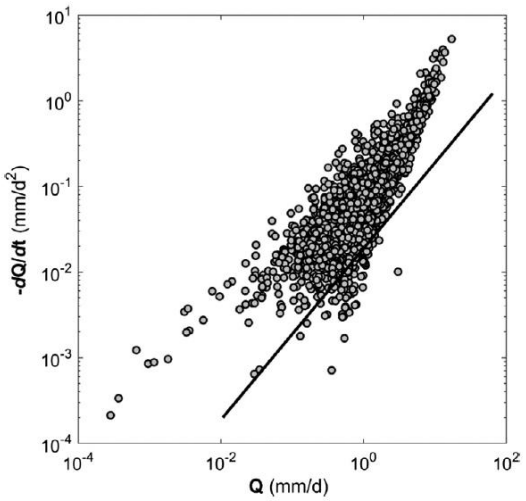


Figure 13: Example of MRC with the BN 77 method (Zhang et al., 2017)

If the correlated method is the easiest to program and the least subjective method, the matching strip method yields more discriminating and accurate information about recession behavior (Nathan and McMahon, 1990). The BN 77 method intended to avoid or minimize the specific difficulties of the other two methods (Cheng et al., 2016). According to Tallaksen (1995) the high time variability in the recession curve argues against the use of a MRC except as an overall approximation which might be applicable for use at the regional scale. These techniques are also inconvenient when separations are to be undertaken on a long continuous record of stream flow, rather than just a few storm period hydrographs (Chapman, 1999).

4.1.2 Filtering methods

These methods were developed in order to standardize graphical methods to be applied in whole continuous hydrograph (Gonzales et al., 2009). Some as the smoothed minima technique (SMT) (Gustard et al., 1992; Hydrology, 1980) stay graphic, whereas others as Recursive Digital Filter (RDF) methods are parametric.

A Smoothed Minima technique

The smoothed minima technique (SMT) corresponds to the simplest of separation rules (Nathan and McMahon, 1990). First, the minima of 5-day non-overlapping periods are found for the entire period of record. Next, this minima series is searched for values that are less than 1.11 times the two outer values; such central values are defined as turning points. The baseflow hydrograph is then constructed by simply connecting all the turning points (see Figure 14) (Nathan and McMahon, 1990). Gustard et al. (1992), have described the detailed operational application of the SMT.

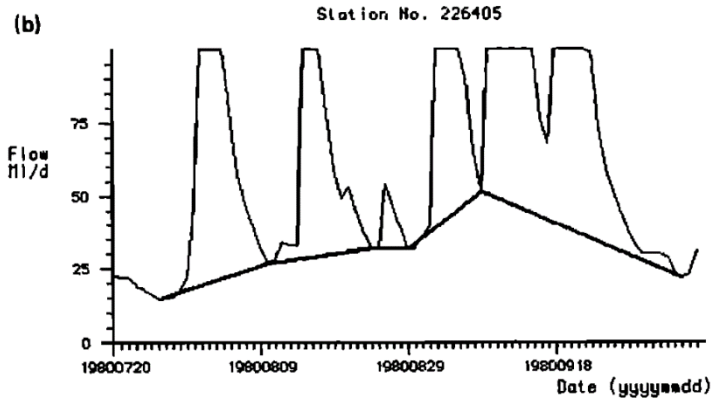


Figure 14: Example of continuous baseflow separation using SMT (Nathan and McMahon, 1990).

The main disadvantage of the method is the determination of the time interval, based on empirical equations and without taking into account the temporal dynamic (Gonzales et al., 2009; Koskelo et al., 2012). Moreover, this technique could lead in some case to a total streamflow less than the baseflow (Nathan and McMahon, 1990).

B Recursive digital filter

The Recursive digital filter (RDF) were proposed at the end of the 70's (Lyne and Hollick, 1979). RDF methods are based on the concept used in signal processing; the quickflow component is assimilated to a high-frequency signal and the baseflow component to a low frequency signal. The RDF method applies a filter in a repetitive process on the set of hydrograph records. Filters have been declined and improved over time. Table 2 summarizes the different RDF methods.

Table 2: Summary Recursive of Digital Filters from literature

Filter equation	Source	Comments
$Q_{q(t)} = \alpha_\tau Q_{q(t-1)} + \frac{1+\alpha_\tau}{2}(Q_{(t)} - Q_{(t-1)})$	Lyne and Hollick (1979)	$Q_{q(t)} \geq 0$ $Q_{b(t)} = Q_{(t)} - Q_{q(t)}$
$Q_{q(t)} = \frac{3\alpha_\tau - 1}{3-\alpha_\tau} Q_{q(t-1)} + \frac{2}{3-\alpha_\tau}(\alpha_\tau Q_{(t)} - Q_{(t-1)})$	Chapman (1991)	Reformulation of the Lyne and Hollick (1979) filter
$Q_{b(t)} = \frac{\alpha_\tau}{2-\alpha_\tau} Q_{b(t-1)} + \frac{1-\alpha_\tau}{2-\alpha_\tau} Q_{(t)}$	Chapman and Maxwell (1996)	
$Q_{b(t)} = \frac{(1 - BFI_{max})\alpha_\tau Q_{b(t-1)} + (1-\alpha_\tau) BFI_{max} Q_{(t)}}{1 - \alpha_\tau BFI_{max}}$	Eckhardt (2005)	filter encompassing the characteristics of Lyne and Hollick (1979) and Chapman (1991) filters

Where:
$Q_{(t)}$: total flow at time *t.*
$Q_{b(t)}$: baseflow at time *t.*
$Q_{q(t)}$: quickflow at time *t.*
α_τ, α_q : filters parameter that can be associated with recession analysis.
C : parameter that allows the shape of the separation to be altered.
BFI_{max} : maximum recharge ratio of the groundwater

The principal advantages of the RDF methods is that they are objectives, computationally efficient, easily automated and seemed to be the most flexible methods to have better performance for a wide

range of climate conditions and catchment properties (Longobardi and Loon, 2018; Nathan and McMahon, 1990). The main disadvantages of this approach is that the calibration of the parameters of these methods are often subjectively determined, resulting in high uncertainties in the baseflow separation estimations (Lott and Stewart, 2016; Zhang et al., 2013). Because these techniques stay arbitrary and physically unrealistic, several authors tried to improve calibration of the RDF method parameter using a recession analysis approach. Table 3 summarizes these approaches.

Table 3: Scientific articles about calibration of the RDF parameter using the recession analysis approach

Authors	Source of the filter equation	Recession analysis approach used
Cartwright et al. (2014)	Lyne and Hollick (1979) Eckhardt (2005)	Correlation method Matching strip method
Rimmer and Hartmann (2014)	Eckhardt (2005)	Correlation method Matching strip method
Mei and Anagnostou (2015)	Eckhardt (2005)	Adaptation of BN 77
Longobardi et al. (2016)	Lyne and Hollick (1979) Eckhardt (2005)	Adaptation of Matching strip method
Zhang et al. (2017)	Lyne and Hollick (1979) Chapman (1991) Eckhardt (2005)	BN 77
Saraiva Okello et al. (2018)	Eckhardt (2005)	Correlation method Matching strip method

4.2 Tracers-based methods

The tracers-based methods are based on the use of isotopes and geochemical tracers to perform the hydrograph separation. This type of method dates back to the late 1960s (Hubert et al., 1969; La Sala Jr., 1967; Pinder and Jones, 1969). It is perceived as a transcendental step in the hydrograph separation. Since unlike the graphical methods based on the arbitrary interpretation of the flood hydrograph (and widely used until then), tracers based approach are measureable, objective and based on components of the water itself (Klaus and McDonnell, 2013). These methods assume that stream water results from the mixture of several sources, and that each source has a constant and unique chemical composition (Pinder and Jones, 1969).

The first studies were performed at the level of a storm event. The hydrograph separation was made into two components: “old water”, which represents the flow from groundwater, and “new water” that generally represented by precipitation and inflow (Buttle, 1994; Klaus and McDonnell, 2013). These terms continue to be used to this day (e.g. Bansah and Ali, 2017).

The tracers based approach is founded on five assumptions (according to Buttle, 1994; Klaus and McDonnell, 2013):

i. the isotopic/tracer content of the event and the pre-event water are significantly different;

ii. the event water maintains a constant isotopic/tracer signature in space and time;

iii. the isotopic/tracer signature of the pre-event water is constant in space and time;

iv. contributions from the vadose zone is negligible, or the isotopic/tracer signature of the soil water is similar to that of groundwater;

v. surface storage contributes minimally to streamflow. The tracer based methods can divided into two groups: (i) isotopic tracers and (ii) geochemical tracers

4.2.1 Isotopic tracers

This approach is generally used for a flood event. Two natural isotopes can be used: oxygen-18 (^{18}O), first applied by Sklash and Farvolden (1979) and deuterium (^{2}H), used for the first time by Sklash et al. (1996). Since they are part of the water molecule, these isotopes of the water molecule are ideal conservative tracers. Added naturally during precipitation events and once free from evaporative exposure, they are only subject to changes due to mixing end-members (Buttle, 1994; Klaus and McDonnell, 2013). Even if weekly or even daily isotope measurements are now becoming available for many catchments (Kirchner, 2019), isotopic analyses can be very laborious and expensive, especially for long term study (Lott and Stewart, 2016; Stewart et al., 2007). Moreover, the conservative character of these istopic tracers can be discussed which increases the uncertainties of the method (Kirchner, 2019). Klaus and McDonnell (2013) illustrated the differences in estimated source fractions due to the effect of evaporated soil water and its differential impact on ^{2}H and ^{18}O signatures and underline the need of a careful assessment of potential end-members (Figure 15).

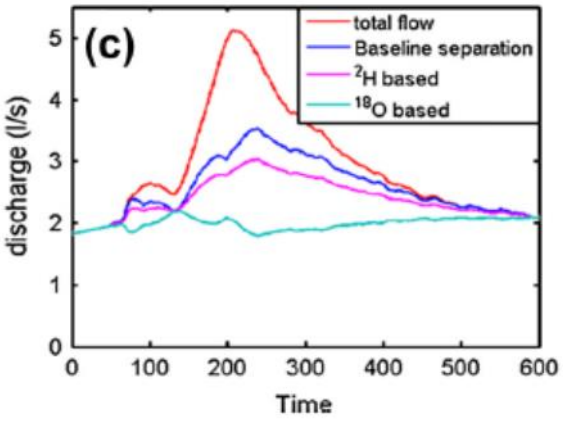


Figure 15: Hydrograph separation results based on a two-component separation using "old water" baseflow (groundwater) and "new water" as end-member. Different results are gained from the use of ^{18}O and ^{2}H that are different to the known "old water" fraction (baseline separation) (Klaus and McDonnell, 2013).

4.2.2 Geochemical tracers

Geochemical tracer methods was first performed by Pinder and Jones (1969) with different solutes and Kunkle (1965) with electro-conductivity (EC). Still on the assumption of the mass balance, the concentrations of the chemical constituents are related to actual physical processes and flowpaths in the basin that generate the different flow components (Stewart et al., 2007).

There are two principal geochemical tracer methods. Either the concentration of each chemical components are measured (e.g. Genereux et al., 1993; Uhlenbrook and Hoeg, 2003) or it is assumed that the chemical component of the baseflow is constant and can be defined as the maximum concentration calculated in the dry period. The chemical component of the quickflow is also considered constant and can be defined as the minimum concentration during high flows (e.g. Stewart et al., 2007; Zhang et al., 2013).

Unlike isotopes, this approach is easier applied to long term studies, especially with EC, that can be inexpensively measured concurrently with stream flow measurements (Lott and Stewart, 2016). However, the assumption of a constant and uniform signature for every component is often fulfilled within short intervals (e.g. within an event) (Gonzales et al., 2009). According to Sherson et al. (2015) the uncertainty of hydrograph separations could be due to laboratory analytical errors, the spatial and temporal variability of component chemical compositions or the non-conservation of the tracer.

5 Quantification of the end-members

In the previous chapter “Hydrological separation methods”, hydrological studies are designed to provide direct information on water pathways and indirect information to explain stream water chemistry. In this chapter, we present the opposite approach, where the chemistry can provide about the hydrological pathways.

The usual situation in the assessment of water quality is the measurement of multiple parameters, taken at different monitoring times, and from many monitoring stations. Therefore a complex data matrix is frequently needed to evaluate water quality (Gaillardet et al., 2018b). For the interpretation of these complex data matrices, several multivariate methods could be used. They allow reducing the data to meaningful terms for identifying sources (also known as end-members) of pollutants and estimating their contributions.

Multivariate methods include the enrichment factors, chemical mass balance, and linear regression methods, as analytical methods, already presented in the previous chapter on Hydrograph separation methods (Pierret et al., 2018). The statistical multivariate methods include eigenvector analysis (also termed principal component analysis (PCA) and factor analysis (FA)), cluster analysis (CA), discriminant analysis (DA), multiple linear regression, neural networks and a number of other multivariate data analysis methods (Fovet et al., 2018; Pierret et al., 2018). Pierret et al. (2018) give clear explanation of each method. However if these statistical methods have been yet used to number of dataset, few of them, resumed in Table 4, have been applied for stream water.

Eigenvector and CA analysis have been widely used to identify the end-members and typology of pollution (Christophersen and Hooper, 1992; Clow and Mast, 2010; Pfister et al., 2018; Simeonov et al., 2003). Several authors used a mixed analysis as PCA with CA, or Multiple regressions on principal components (i.e. EMMA analysis) (Christophersen et al., 1990; Liu et al., 2008; Simeonov et al., 2003). However, if these methods make it possible to identify the end-members, none allows them to be quantified. Understanding of stream water composition requires not only the identification of geochemically distinct water stores but also time-dependent variation in their relative contributions to stream flow. Carrera et al. (2004) and Vázquez-Suñé et al. (2010) have developed other statistical approaches, as MIX method to quantify these end-members and deal with their spatial and temporal variability.

This Chapter presents only some multivariate statistical methods: PCA and FA as descriptive methods, CA and DA as explanatory methods. Two mixed analysis are also presented: EMMA analysis and the MIX method, because they have been developed specifically for the stream water topic.

Table 4 : Statistical multivariate methods used to identify end-members of stream water

Statistical multivariate methods		References
Descriptive methods		
Eigenvector analysis *Analyses eigenvectors of the covariance matrix to describe the covariance*	principal component analysis (PCA)	Clow and Mast (2010) Fovet et al. (2018) Gaillardet et al. (2018a) Simeonov et al. (2003)
	factor analysis (FA) or positive matrix factorization (PMF)	Maher (2011) Fovet et al. (2018) Gaillardet et al. (2018a) Reghunath et al. (2002) Paatero and Tapper (1994)
Explanatory methods		
Classification analysis *Analyses dissimilarity to describe the pattern*	Cluster analysis (CA)	Fovet et al. (2018) Gaillardet et al. (2018a) Simeonov et al. (2003) Eng and Brutsaert (1999)
	Discriminant analysis (DA)	Fovet et al. (2018) Alberto et al., (2001)
	Neural networks	Brutsaert and Lopez (1998) Jasechko (2019) Semkin et al. (1994)
Correlation analysis *Analyses correlation to describe variability of one component associated to several extern variables*	Multiple linear regression	Semkin et al. (1994)

5.1 Eigenvector analysis

Eigenvectors analysis, calculates temporal correlations from a time series of chemical concentrations. Eigenvectors of this correlation matrix are determined and a subset is rotated to maximize and minimize correlations of each factor with each measured species. The factors are interpreted as end-members profiles by comparison of factor loadings with end-member measurements. Several different normalization and rotation schemes have been used, but their physical significances have not been established (Pierret et al., 2018). The most used methods are: Principal Component Analysis (PCA) and the Factor analysis (FA), also called Positive Matrix Factorization (PMF) by Paatero and Tapper (1994).

5.1.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) extracts the eigenvalues and eigenvectors from the covariance matrix of original variables. The PCs are the uncorrelated (orthogonal) variables, obtained by

multiplying the original correlated variables with the eigenvector, which is a list of coefficients (loadings or weightings). Thus, the PCs are weighted linear combinations of the original variables. PC provides information on the most meaningful parameters, which describe whole data set affording data reduction with minimum loss of original information (Singh et al., 2004; Vega et al., 1998). It is a powerful technique for pattern recognition that attempts to explain the variance of a large set of inter-correlated variables and transforming into a smaller set of independent (uncorrelated) variables (principal components)(Singh et al., 2004).

However, application of the method to practical environmental data requires that error estimates for the data be chosen judiciously so that the estimates reflect the quality and reliability of different data points. For a typical environmental problem, with outliers and other problematic data points, Albek (1999) reports this in detail. Moreover, the PCA approach can produce negative values for almost all factors, whereas it is not possible, for example, to have negative amount of a basic constituent in any sample. Despite the transforming ('rotating') factors in order to eliminate the negative entries, there usually remain some negative values (Paatero and Tapper, 1994).

5.1.2 Factor analysis (FA) or positive matrix factorization (PMF)

The term, 'factor analysis' (FA), is ambiguous. FA means principal component analysis (PCA): singular value decomposition (SVD), selection of dimension, and rotations. Statisticians often remark that this is not FA at all according to their definition of FA. In statistics, FA means investigation of correlations of random variables. This leads to a non-linear computation, which cannot be done with SVD. To avoid the ambiguous term 'FA', Paatero and Tapper (1994) called the statistician FA, 'positive matrix factorization' or PMF.

Although highly used in the air pollution domain to perform end-members apportionment of particulate matter (see detailed review of existing methods in Popoola et al., 2018), the PMF method is recent for the domain of hydrochemistry (Capozzi et al., 2018; Haji Gholizadeh et al., 2016; Zanotti et al., 2019). The PMF method utilizes statistical techniques to reduce the data to meaningful terms for identifying the chemical end-members and to estimate the end-member contributions.

The form of the PMF model, most widely used to analyze the chemical network data, is the bilinear model: two matrices, G and F, leading to a reproduction of the dataset variability as a linear combination of a set of constant factors profile and their contribution to each sample (Reff et al., 2007).

The main advantages of this method compared to PCA are that: (i) it takes into account the analytical uncertainties often associated with measurements of environmental samples and (ii) forces all of the values in the solution profiles and contributions to be positive, which can lead to a more realistic representation (Reff et al., 2007; Zanotti et al., 2019).

According to Zanotti et al. (2019), the main limitations of this method, which would lead to its poor development in the field of catchment hydrochemistry, are that PMF requires only data expressed as concentration while some typical measurement of water samples have different units (e.g. pH, Electrical Conductivity, Oxidation Reduction Potential, isotopes analysis, age tracers). Thus, the dataset cannot be directly fed into a PMF model. Moreover, a proper monitoring network, specifically designed to capture the variability of the hydrochemical system, is crucial to obtain a complete representation of the system with a PMF analysis. Under some cases, where a single end-member is present, a multivariate statistical analysis such as PMF might not be appropriate (Zanotti et al., 2019).

5.2 Classification analysis

5.2.1 Cluster analysis (CA)

Cluster analysis (CA), also called unsupervised pattern recognition method, encompasses a wide range of techniques for exploratory data analysis. The main aim of CA is grouping objects (cases) into classes (clusters) so that objects within a class are similar to each other but different from those in other classes. The class characteristics are not known in advance, but may be determined from the data analysis. The results obtained are justified according to their value in interpreting the data and indicating patterns (Gaillardet et al., 2018a; Singh et al., 2004; Singh et al., 2005; Vega et al., 1998).

5.2.2 Discriminant analysis (DA)

Discriminant analysis (DA), also called supervised pattern recognition, provides statistical classification of samples, contrasting with the exploratory features of cluster analysis. DA can be used if the membership of objects to particular groups or clusters (i.e., the temporal or spatial ownership of a water sample as determined from its monitoring time or station) is known in advance. A discriminant function is constructed for each group. Such a function represents a surface dividing data space into regions; then samples sharing common properties will be grouped into the same region with a decision boundary separating two or more groups (Gaillardet et al., 2018a).

All these three MS methods were used to investigate groundwater and surface water chemical data and identify natural phenomena (Blake et al., 2016; Koh et al., 2016) as well as anthropogenic impacts affecting water quality (Barakat et al., 2016; Devic et al., 2014; Phung et al., 2015).

The principal disadvantage of MS methods is that their results (sources compositions and source contributions) can be positive or negative. Since negative compositions and contributions do not actually exist, it is difficult to assess quantitatively the relative contributions from the sources identified (Li et al., 2019; Zanotti et al., 2019). Another important problems seen in MS methods are that they

put the same weight on all the chemical components, although their real contribution to the stream water is negligible (Gaillardet et al., 2018a; Singh et al., 2004) and they are data-sensitive techniques and usually require a previous univariate analysis of the dataset. In fact, solutes stream water datasets are often characterized by skewed distributions and normalization procedures have to be applied (Comero et al., 2011). All these problems have a considerable impact on the results.

5.3 Mixed analysis

5.3.1 End member mixing analysis (EMMA)

End member mixing analysis (EMMA) was developed from works done by Christophersen et al. (1990), Christophersen and Hooper (1992), Hooper (2003) and modified by Liu et al. (2004; 2008) and Barthold et al. (2011). It is a hydrological mixing model which principal aims is identifying the minimum number of end-members required to explain the variability of measured concentrations and co-linearity (Tubau et al., 2014).

EMMA is based on an analytical approach that solves an over-determined set of equations on the basis of a least squares procedure (Barthold et al., 2011). According to Barthold et al. (2011), EMMA uses the following assumptions:

(i) the stream water is a mixture of source solutions with a fixed composition,

(ii) the mixing process is linear and relies completely on hydrodynamic mixing,

(iii) the solutes used as tracers are conservative,

(iv) the source solutions have extreme concentrations.

EMMA allows reducing the dimensionality of the analysis because visualization and interpretation of a large number of chemical analyses of many species is hard. The problem is much easier to handle if, instead of studying a dataset potentially as large as the number of species, it would suffice to analyze their projections on a space of much smaller (say 2 or 3) dimensions (Pelizardi et al., 2017). An EMMA aims at identifying the minimum number of end-members required to explain the variability of measured concentrations in time or space. The explanation of the variance and the contribution of each species to the mixture are obtained from the analysis of information provided by the calculation of the eigenvalues (Tubau et al., 2014).

According to Tubau et al. (2014) the methodological guidelines for the selection of end-members and species are as follows:

(i) Initiate EMMA with all species and take further steps that will eliminate the species with the highest weight in the eigenvalue.

(ii) A common rule is to retain all eigenvectors associated with eigenvalues (n) greater or equal to 1, called the "rule of 1" (Hooper, 2003). This rule consists of selecting the eigenvalues that explain variance greater than [1/number of specie].

(iii) The selection of species according to those eigenvalues that we want to explain. The number of end-members will be n + 1.

(iv) The projection of the observation points in the space defined by the eigenvectors associated with each eigenvalue. With this projection, it is possible to make a preliminary selection of end-members and an analysis of the points that fall outside the area defined by these end-members.

As in the case of the multivariate statistical methods, the main disadvantage of EMMA is that the outcomes (sources compositions) can be positives or negatives. Other disadvantage is that the concentrations of end-members must be different and accurately known (Vázquez-Suñé et al., 2010). Last, EMMA does not allow us to compute mixing contributions of sources.

5.3.2 MIX method

To calculate mixing contributions of sources, Carrera et al. (2004) and Vázquez-Suñé et al. (2010) have developed other statistical approaches, the MIX method to quantify these sources and deal with their spatial and temporal variability.

According to Tubau et al. (2014), the MIX tool consist of

(i) the identification of potential recharge sources,

(ii) the selection of tracers,

(iii) the characterization of the hydro-chemical composition of potential recharge sources and mixed water samples,

(iv) the calculation of mixing ratios and the re-evaluation of end-members.

The main drawbacks of this calculation tool are the uncertainty in end-member definition caused by insufficiently documented spatial and temporal variability, the limited experience in selecting appropriate tracers and co-linearity, which causes the solution to be unidentifiable or unstable.

6 Study site – The Oracle-Orgeval observatory

6.1 Characteristics of Oracle-Orgeval observatory

6.1.1 Location and brief history

The Orgeval catchment is a small sub-catchment of the Oracle-Orgeval observatory located 70 km east of Paris, France (Figure 16). In the department of Seine-et-Marne (77), the Orgeval catchment is a sub-basin of the Grand Morin watershed, main tributary of the Marne River. The Grand Morin watershed has a significant quantitative and qualitative impact on the Seine River and the Parisian conurbation. To understand and monitor the hydrological and chemical processes that lead to floods and eutrophication of the Seine River, Irstea set up an observatory on the Orgeval catchment in 1962. The relatively small extent of the Oracle-Orgeval observatory, 104 km², gives it homogeneous natural conditions to assert a representation at the regional scale.

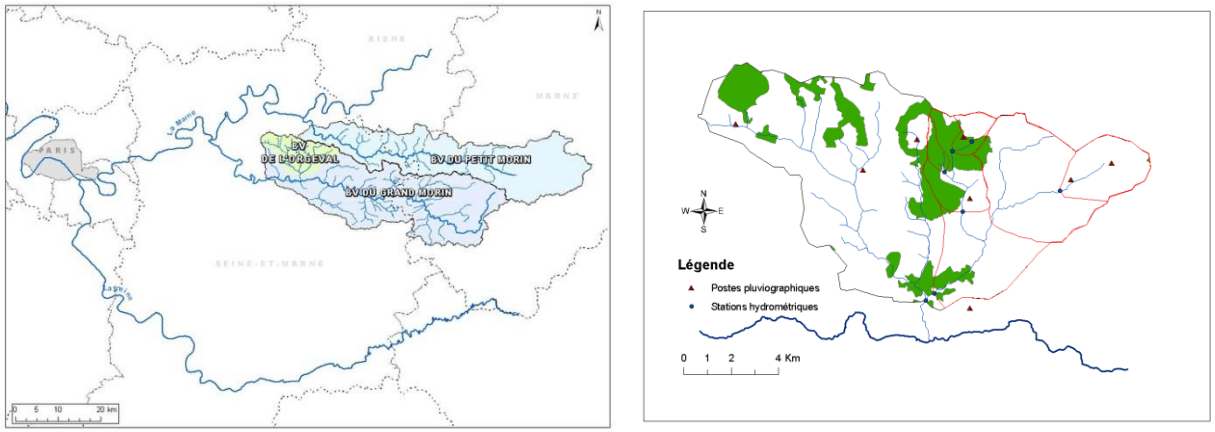


Figure 16 : Location of the ORCALE-Orgeval observatory (France) (left map) and Orgeval catchment with its sub-catchments (red line) (right map)

Since 1962, a complete measurement network monitors the hydrological compartments of the Orgeval catchment and its seven sub-catchments (Figure 16) with perennial equipment: gauging stations at the outlet of each sub-basin and drainage network, piezometers, rain gauges and soil moisture stations. This device is matched by water quality measurements of surface water, rainfall and groundwater. All data measured and validated by Irstea are accessible via the Banque HYDRO database (http://hydro.eaufrance.fr/) and the ORACLE database (https://bdoh.irstea.fr/ORACLE/).

6.1.2 Topography and Climate

The Oracle-Orgeval observatory has a relatively undifferentiated topography (average altitude: 145 m, minimum: 85 m and maximum: 185 m). The only accident is the Doue hilltop, the highest elevation of the catchment.

Climate variables measured in the Orgeval basin are representative of a temperate oceanic climate with cumulative annual rainfall of 675 ± 31 mm. The rainfall events are weak (2.4 ± 0.1 mm) but spread throughout the year (Table 5). Only potential evapotranspiration (PET) (733 ± 23 mm) calculated from temperatures (10±6 °C) and global radiation (1±0.6 kgJ.cm^{-2}) varies significantly during the year, with maximum values in July and minimum in December (Table 5).

Table 5: Monthly and annual averages over the period 1962 - 2018 of the climatic variables of the Orgeval observatory

	Monthly averages												Annual averages
	J	F	M	A	M	J	J	A	S	O	N	D	
PET (mm)	12	18	45	73	101	114	129	113	69	36	14	10	733
Rain (mm)	54	46	54	49	62	59	59	54	54	61	60	63	675
T (°C)	3	4	6	9	13	16	18	17	15	10	6	3	10
Global radiation (kgJ.cm^{-2})	0.3	0.5	0.9	1.4	1.7	1.9	1.7	1.5	1.1	0.6	0.3	0.2	1.0

6.1.3 Geology and hydrogeology

The Oracle-Orgeval observatory is influenced by a multi-layered aquifer system composed of two main geological formations (Flipo et al., 2007; Mouhri et al., 2013a): the Oligocene (Brie limestone) and the Eocene (Champigny limestone). These two aquifer units are separated by an aquitard of green clays and marls (Figure 17).

The Oligocene consists of an alternation of blocks of digging clays and Brie limestone banks (Figure 17). Brie limestone, composed of limestone of lacustrine origin, is located only in the basal part (with a thickness of 1 to 2 m) (Thiry and Simon-Coinçon, 1996). The upper part is essentially composed of a silt-clay horizon containing heterometric pieces of residual limestone and dispersed silicified masses (Figure 17). The heterogeneous and thin (6-10 m of humid height) ensemble forms a free surface aquifer, very sensitive to precipitation variations, on a green clays base almost impermeable (Mouhri et al., 2013b) (Figure 17).

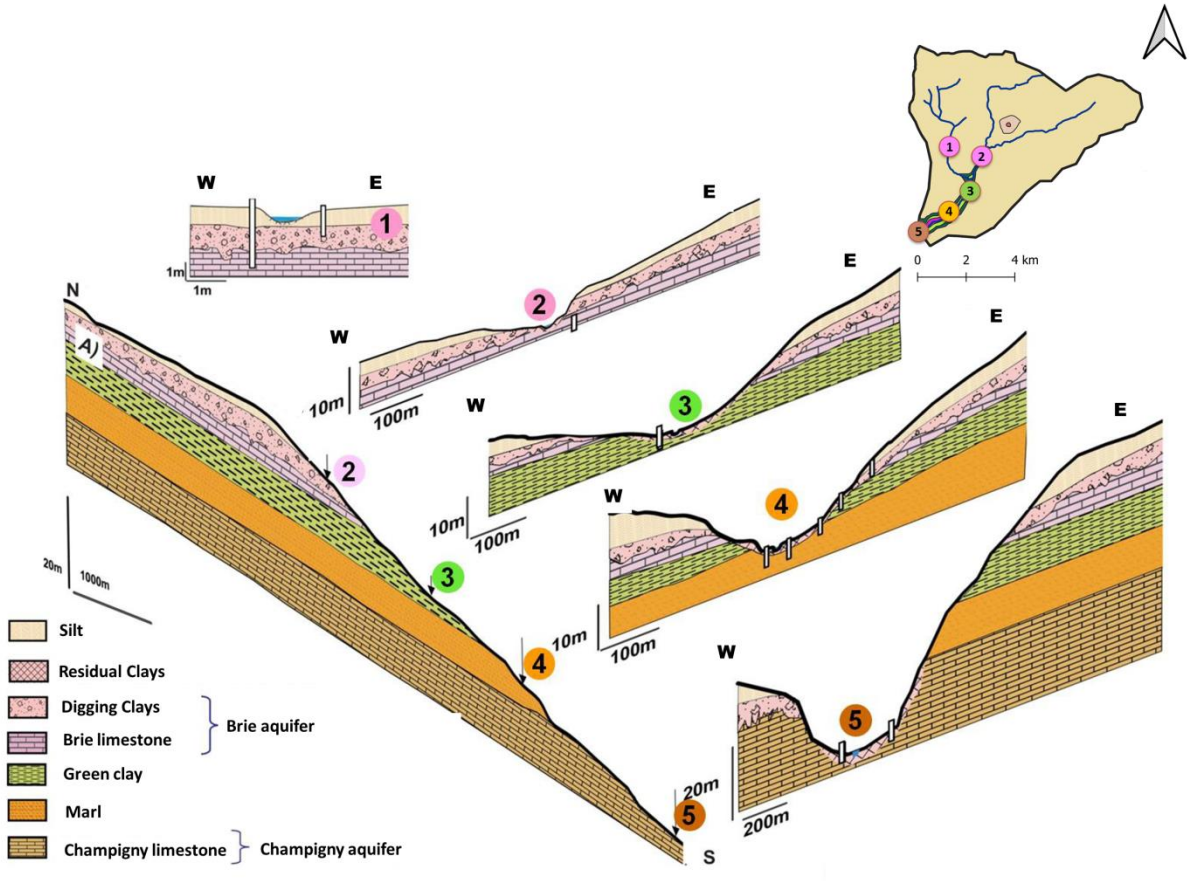


Figure 17 : Simplified geological map of the Avenelles sub-catchment (Oracle-Orgeval observatory). Source: Mouhri et al. (2013b)

The upper and middle Eocene is composed, from top to bottom, of marls (not very permeable) and of the limestone of Champigny (Bartonian) (Figure 17). The Champigny aquifer includes these two formations which are separated by a marl transition. These formations constitute an aquifer of fissure permeability, capped on the plateaus by the layer of green clays. This aquifer is partially free, although topped with a clay layer. It is fed by the valley sides and locally by porous areas (Mouhri et al., 2013b). The strength of the Champigny aquifer can expect about sixty meters with a wet thickness of 30 to 50 m (Mouhri et al., 2012).

6.1.4 Pedology

The catchment is covered by Quaternary wind deposits (maximum a layer of 10 m thick) mainly composed of sand and silt lenses that are not very permeable. The soils on the plateau are brown earth leachate type of silt-sand to silt-clay texture, with temporary hydromorphic characteristics. The presence of more impermeable levels in the silts leads to the local formation of surface aquifers. Over the entire winter wet period, the soil surface (from 5 to 45 cm deep) remained saturated to 55% moisture volume, while the deeper horizon (45 to 155 cm) never exceeded 40% of saturation (Tallec et al., 2013). The soils of the catchment are poor in organic matter with little variation in porosity. This

explains the low structural stability of the soils, which can however be compensated by permeability. Thus, the catchment shows a particular sensitivity to aquifer erosion, even if it remains a function of the vegetation cover (Tallec et al., 2013).

6.1.5 Land use and hydro-agricultural infrastructures

The Oracle-Orgeval observatory is essentially agricultural (81% agricultural, 18% forested, and 1% urban; see Figure 18). Lands under cultivation are dominated by cereal crops (wheat, maize, barley and pea), with conventional practices, mainly based on mineral nitrogen fertilization. Wheat represents up to 40% and maize, depending on the year, 5 to 20% of the farmland (Tallec et al., 2013). Fertilizers are mainly sprayed from February to March, through one to three applications (Vilain et al., 2010). They can be source of nitrogen, potassium, phosphorus, chlorides and calcium in the catchment (Garnier et al., 2016; Garnier et al., 2014; Vilain et al., 2010). Note that soil lime is also practiced on the Orgeval catchment.

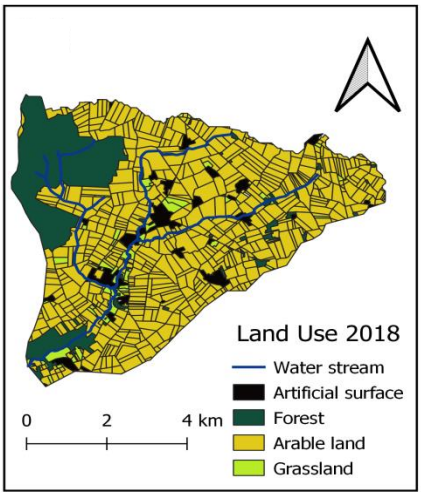


Figure 18 : Land use (2018) of the Avenelles sub-catchment, Oracle-Orgeval observatory

The different types of hydro-agricultural infrastructures concern three categories: infrastructures of protection, earthworks and research infrastructures, homogeneously distributed over the catchment. There are both so-called "flood protection" structures (low walls, dikes and thresholds) and structures related to the use of rivers (bridges, old washhouses or mills). The large-scale earthworks were conducted for the recalibration or dewatering of rivers. The research infrastructures correspond to the Irstea limnimetric stations (stream bed concreting and thresholds).

A large part of the catchment is also drained by underground pipes. According to available information, 41.5% of the total surface area of the basin is drained, 25% un-drained and 33.5% not reported (Tallec et al., 2013).

6.2 Hydrological measurements of Oracle-Orgeval observatory

From 1962 to 1981, the hydrology of the catchment was studied in detail, using a dense network of measuring devices (twenty-three rain gauges stations, five limnimetric gauging stations and a complete meteorological station). The data obtained provided access to a very good knowledge of the hydrological behavior of the catchment. Since 1982, rainfall monitoring has been simplified, while preserving a minimum basic network sufficient to fully understand the variable.

The Orgeval River, tributary of Grand Morin, is constituted by the meeting of two main streams: the Rognon and the Avenelles (see Figure 16 above). The Avenelles stream drains a basin of 46 km² (Figure 19). Because of its high monitoring device, it is our sub-catchment of interest (Figure 19).

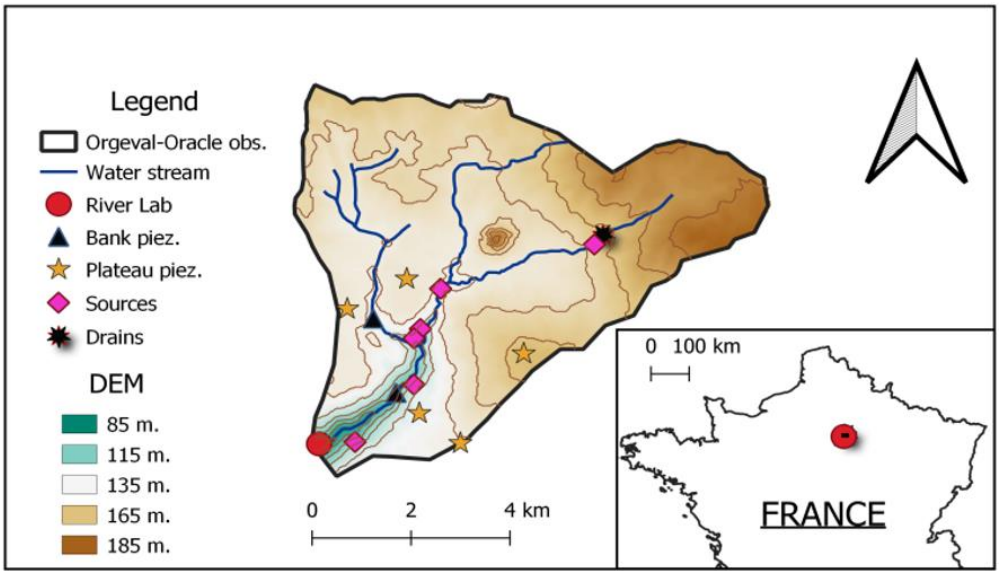


Figure 19 : The Avenelles sub-catchment (Oracle-Orgeval observatory) and its monitoring device.

The average daily inter-annual discharge (1974-2018) of the Avenelles sub-catchment is 0.3 m^3s^{-1} with a maximum of 10.4 m^3s^{-1} and a minimum of 0.01 m^3s^{-1} (Figure 20). Discharges present typical seasonal fluctuations of the rivers of the central-eastern Paris basin. High water levels are observed in winter, from December to March inclusive (with a maximum in January), and low water levels in summer, from July to September (minimum in August and September) (Figure 21).

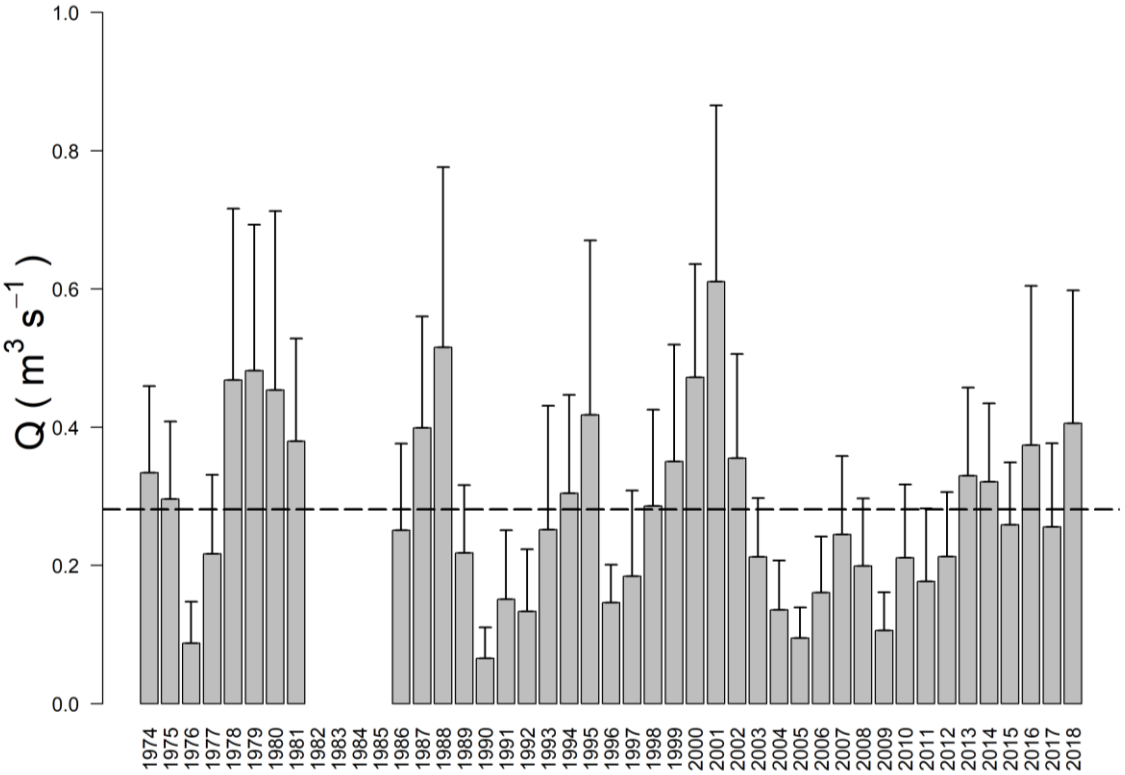


Figure 20 : Average daily discharge at the outlet of the Avenelles sub-catchment (46 km^2, Oracle-Orgeval observatory) and average over the long-term period of observation from 1974 to 2018 (0.3 m^{3}s^{-1}, dotted black line)

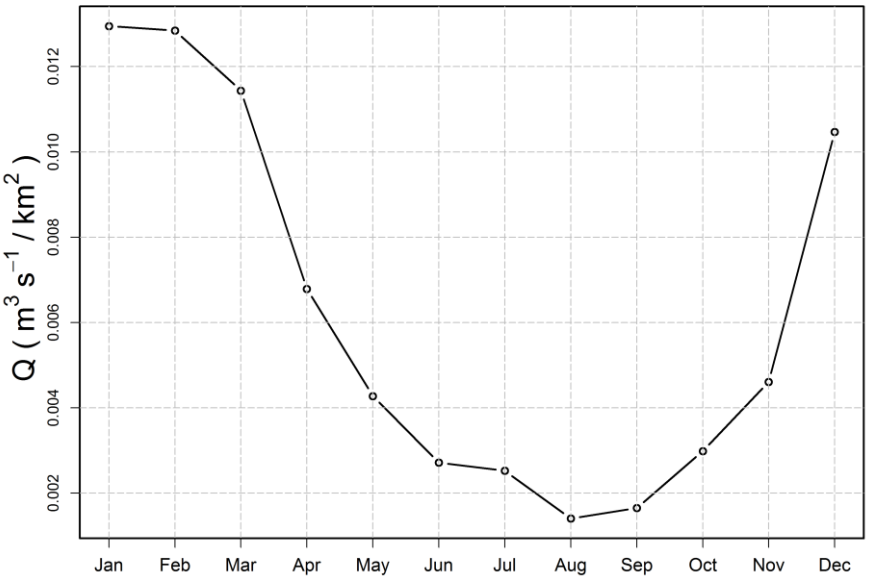


Figure 21 : Curves of the hydrological regime of the Avenelles sub-catchment (Oracle-Orgeval observatory), from1974 to2018.

6.3 Chemical measurements of Oracle-Orgeval observatory

6.3.1 Long-term monitoring

The first water quality station was introduced to monitor nitrates concentrations in 1975 at the outlet of the Mélarchez sub-catchment (tributary of the Avenelles River). This was the first station to show and warn of the effects of nitrate pollution in the stream water. Since 1975, the Mélarchez station, despite some interruptions, has been acquiring nitrate chronicles (one measurement every month or even twice a month), showing the strong impacts of the seasonal nature of hydrological operations linked to the presence of agricultural drainage (Tournebize et al., 2013). Over the period 1975-2006, an increase of 17 $mgNL^{-1}$ was observed then they have stabilized, demonstrating the importance of the contribution of intensive agriculture to the watershed (Tournebize et al., 2013) (Figure 22).

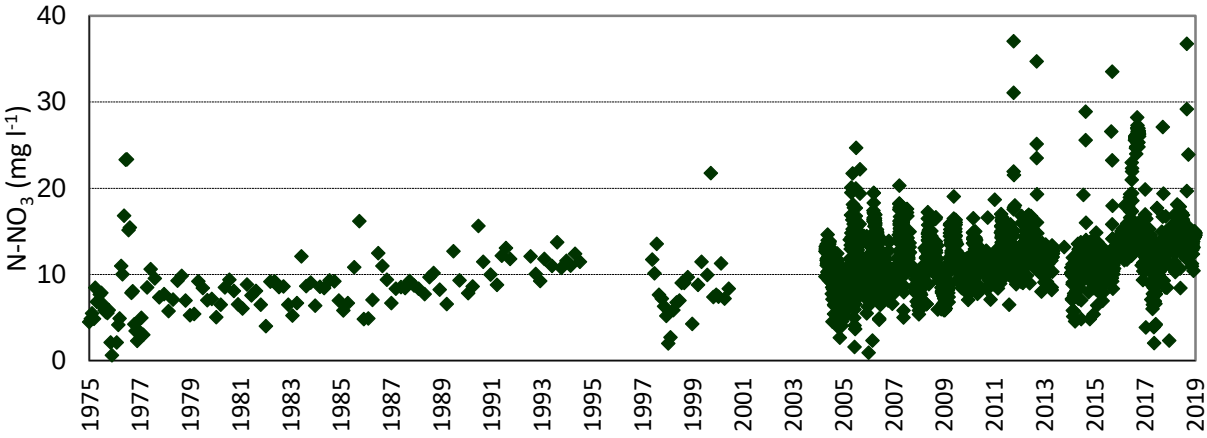


Figure 22 : Nitrate concentrations measured at the Mélarchez station (Oracle-Orgeval observatory), since 1975.

In the middle of 2000, punctual monthly/bimonthly measurements were changed to daily measurements, and new stations of water quality were installed, including the Avenelles station. To date, there are 14 stations (streams, piezometers, rain gauges, drains and sources) that measure water quality in the observatory (Tallec et al., 2013).

Apart from nitrates, other chemical concentrations are measured daily at the Oracle-Orgeval observatory, such as nitrites [NO_2^-], ammonia [NH_4^+], phosphates [PO_4^{3-}], chlorides [Cl^-], dissolved inorganic (DIC) and organic (DOC) carbon (see Figure 23). Electro-conductivity (EC) and pH are also measured. All of these chemical databases are available in the BDOH database (https://bdoh.irstea.fr/ORACLE/). Others data are collected during recurring campaigns as suspended matter flow or pesticides concentrations.

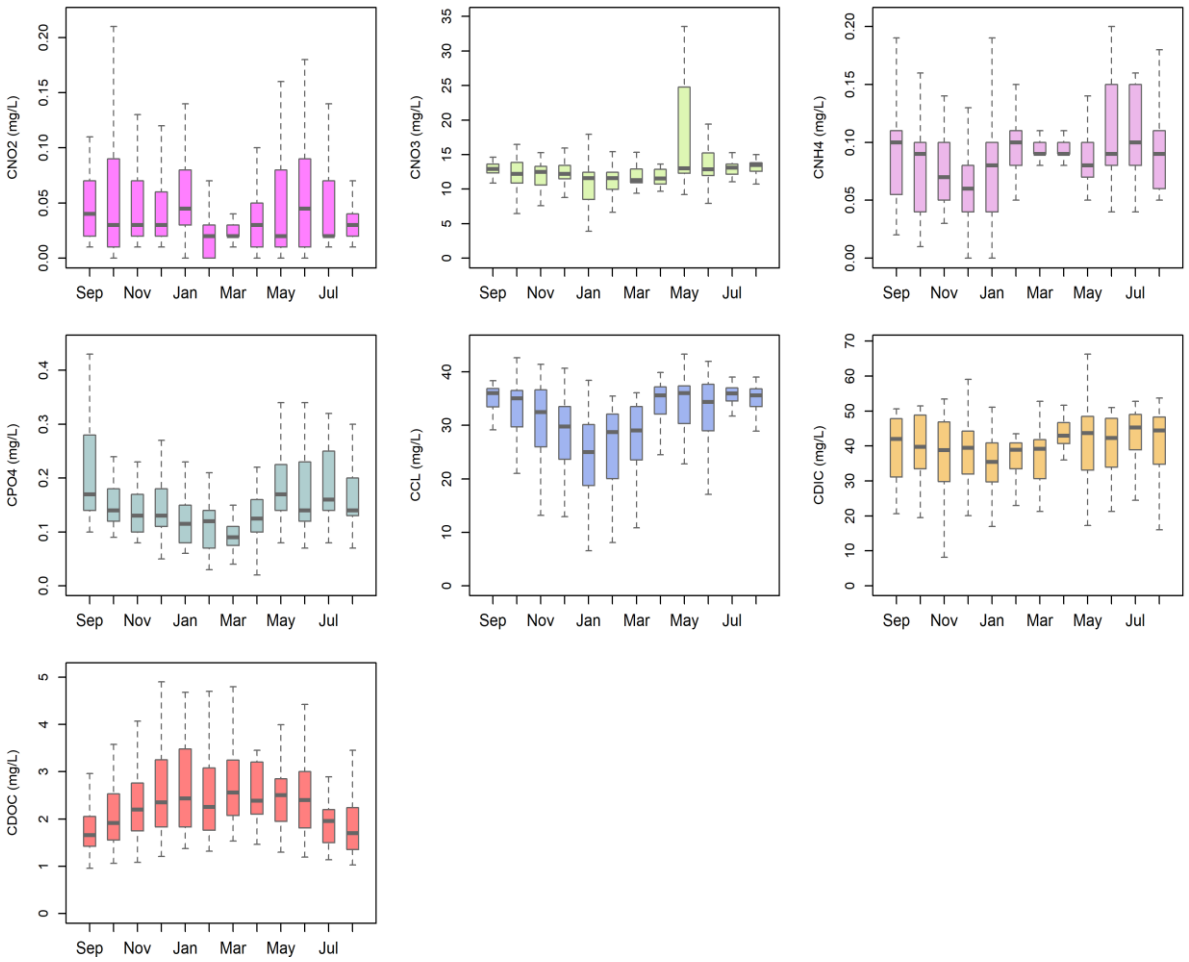


Figure 23 : Monthly variations of measured concentrations (one measurement per day since 2000) at the Avenelles station (Oracle-Orgeval observatory): nitrites (CNO2), nitrates (CNO3), ammonia (CNH4), phosphates (CPO4), chlorides (CCL), dissolved inorganic carbon (CDIC) and dissolved organic carbon (CDOC).

6.3.2 River Lab. and high-frequency measurements

Since June 2015, the "River-Lab", developed by Irstea and IPGP, has been installed at the Avenelles station to measure the concentration of all major dissolved species with high-frequency (Floury et al., 2017). A set of laboratory instruments have been deployed continuously in the field in a confined bungalow next to the river. The River Lab performs a complete analysis of dissolved ions, every 30 minutes, using two ionic chromatographs (Dionex® ICS-2100) by continuous sampling and filtration of stream water (Figure 24). In addition to the major ions (Mg^{2+}, K^{+}, Ca^{2+}, Na^{+}, Sr^{2+}, F^{-}, SO_4^{2-}, NO_3^{-}, Cl^{-}, PO_4^{3}), a set of physico-chemical probes measure every minutes, pH, electro-conductivity (EC), dissolved oxygen (O_2), dissolved organic carbon (DOC), turbidity and temperature.

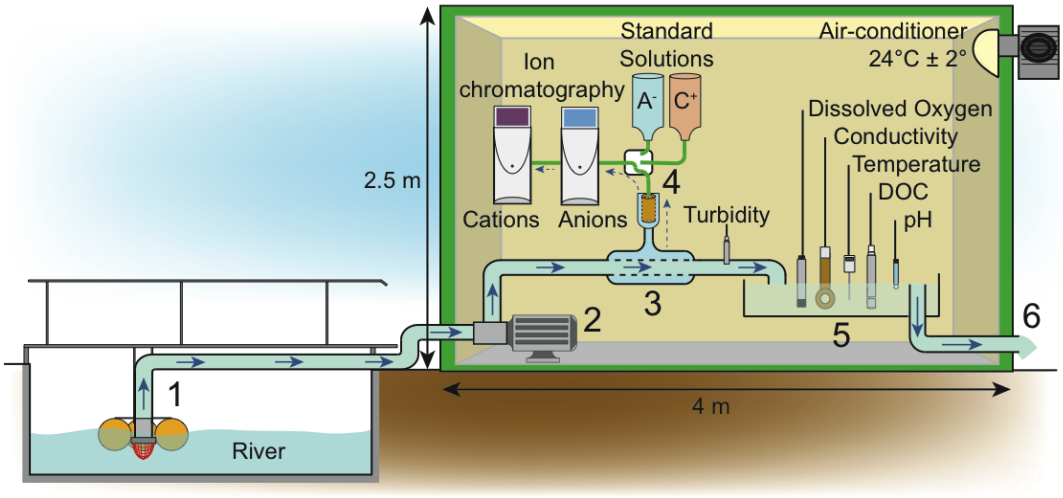


Figure 24 : Diagram of the River-Lab. The blue arrows in bold indicate the primary circuit of unfiltered water. The dotted arrows indicate the filtered water supplied to the ICs. (1) The entry of the primary circuit samples the river at a constant depth of 20 cm maintained by buoys. The water is filtered to 2 mm using a strainer. The distance between the inlet and the pump is 6 m. The entire primary circuit is almost entirely made of PVC pipes. (2) The electric pump runs continuously at constant power, giving a flow rate of 700 liters per hour. (3) Almost all the water taken from the river flows through the pipe and remains unfiltered. Only a fraction is filtered at 2 µm through a stainless steel tangential filtration unit, then filtered through a 0.2 µm cellulose acetate front filter and finally delivered to the ICs at a flow rate of 1 liter per hour. (4) A multiport valve before introduction into the ICs allows switching between filtered river water and standard or blank solutions. (5) All probes are deployed in an overflow tank of 5 liters of unfiltered stream water. (6) The outlet of the primary circuit is downstream in the river. Source: Floury et al. (2017)

The River-Lab is able to capture the abrupt changes in dissolved species concentrations during a typical flood event, as well as low-flow period of summer drought (see Figure 25). All the technical qualities, calibration of the equipment, comparison with laboratory measurements, degree of accuracy, etc. have been well described in Floury et al. (2017).

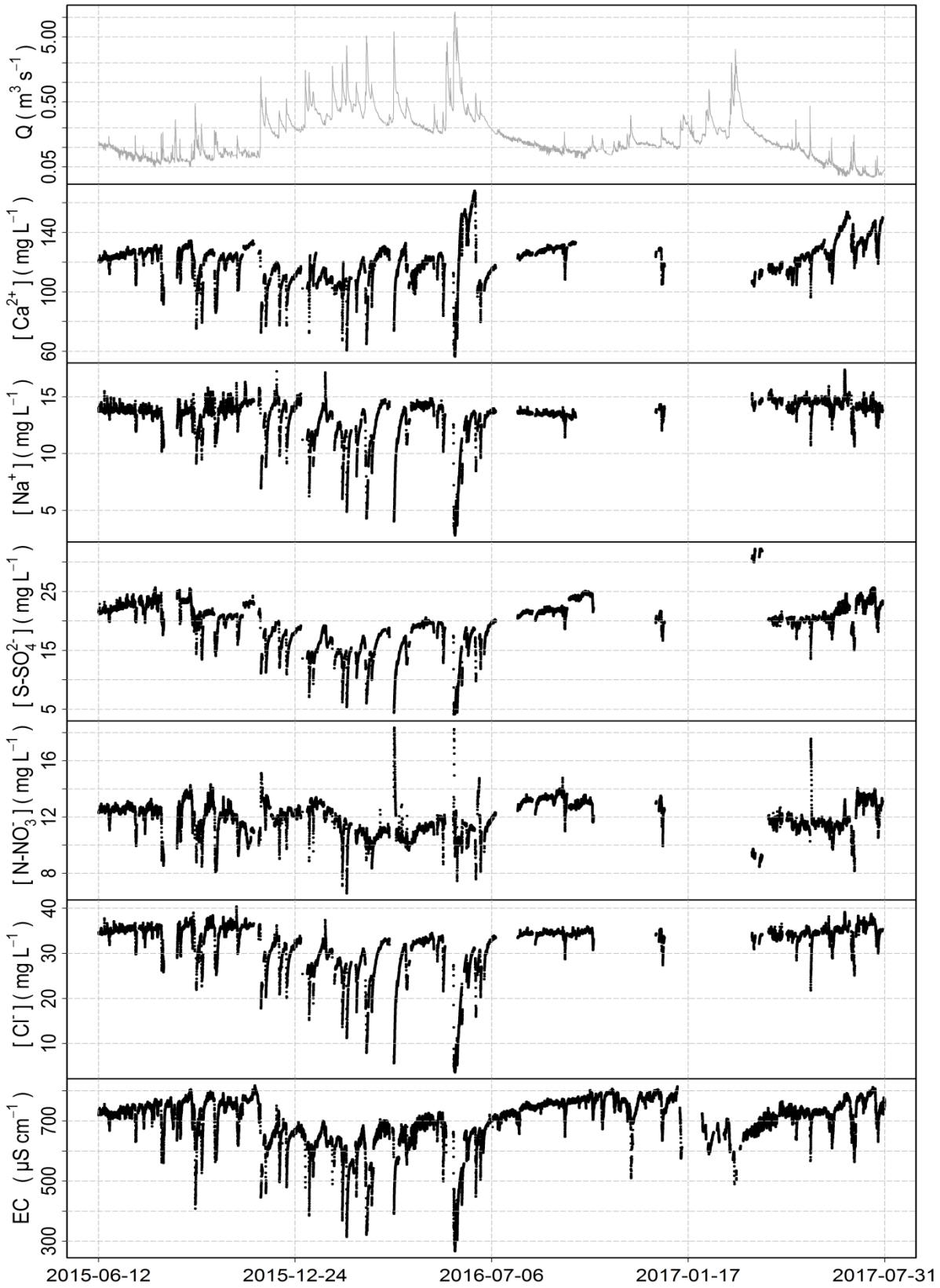
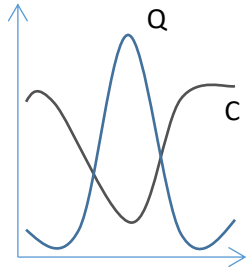


Figure 25 : High-frequency time series of flow, [Ca^{2+}], [Na^{+}], [SO_4^{2-}], [Cl^{-}], [NO_3^{-}] and EC, measured at the Avenelles station, Oracle-Orgeval observatory, from June 2015 to July 2017.

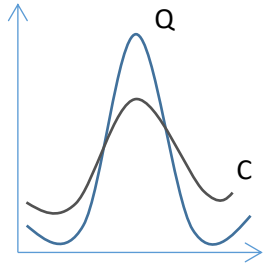
6.3.3 Typology of high-frequency measurements of ions concentrations during flow events

According to Williams (1989), during a flood event, the typology of variations of the concentrations can be divided into three typical behaviors:



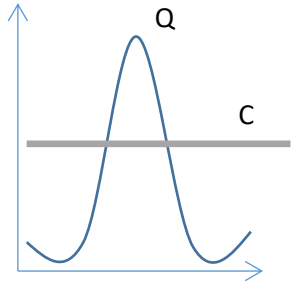
✓ Dilution

During the flood event, the flow increases and chemical concentrations decrease. Once the event is over, the concentration returns to its usual value before the event.



✓ Concentration

During the flood event, the flow increases, and chemical concentrations also increase. Once the event is over, the concentration returns to its usual value before the event.



✓ Chemostasis

During the climatic event, the flow rate increases, but the chemical concentrations remain unchanged; there is no change in the initial values.

It is possible to find combinations of these three behaviors during the same flood event. From June 2015 to August 2016, high-frequency measurements carried out by the River Lab at the Avenelles station, account for about fifty events. Because there is less available data of concentrations than discharge, especially for strontium and phosphates, we have selected only 20 events. All of them clearly distinguish the three typologies. The following table presents the number of incidences of the 3 typologies observed during the 20 flood events (Table 6). It shows that the dilution process is the most widespread behavior for the majority of ions and events. Magnesium, calcium and sulfate ions, only present dilution processes (100% of the flood events, see Table 6). The concentration process only concerns potassium (100% of the flood events), and for some events, fluorides, nitrates and phosphates (Table 6). Phosphates ions are the only ones to present a chemostasis process (Table 6). There is a small percentage combination of the two processes (dilution and concentration) during a same event, but it can concern all ions, except magnesium, potassium, calcium and sulfates (Table 6).

Table 6 : Number of incidences of processes observed on twenty flood events measured at the Avenelles station from June 2015 to August 2016 (Oracle-Orgeval observatory).

Number of incidences of typologies from 20 flood events						
ions	main typologies			Combinations		Total
	Dilution (D)	Concentration (C)	Chemostasis (CH)	D-C	C-D	
Mg^{2+}	20	---	---	---	---	20
K^{+}	---	20	---	---	---	20
Ca^{2+}	20	---	---	---	---	20
Na^{+}	17	---	---	2	1	20
Sr^{2+}	11	---	---	1	1	13
F^{-}	6	10	---	3	1	20
S-SO_4^{2-}	20	---	---	---	---	20
N-NO_3^{-}	12	2	---	3	3	20
Cl^{-}	16	---	---	2	1	20
P-PO_4^{3-}	2	8	3	5	---	18
Total	123	41	4	16	7	191

Figure 26 presents three of the twenty flood events, for magnesium, potassium, nitrates and chlorides ions. The first event (02/12-15/2016) represents a classical flood of the Avenelles sub-catchment. The second event (03/03-06/2016) presents a double peak of flood. The third event (03/30 to 04/02/2016) is an exceptional flood. The three flood events show a six-hour delay between the maximum peak of rainfall and the peak of the flood. The reaction time of the Avenelles sub-catchment is fast (sub-day scale) typical of small catchments (Kirchner et al., 2004). Magnesium and Chlorides ions exhibit a dilution process whatever the flood event, whereas nitrate ions show different processes according to the flood event (Figure 26).

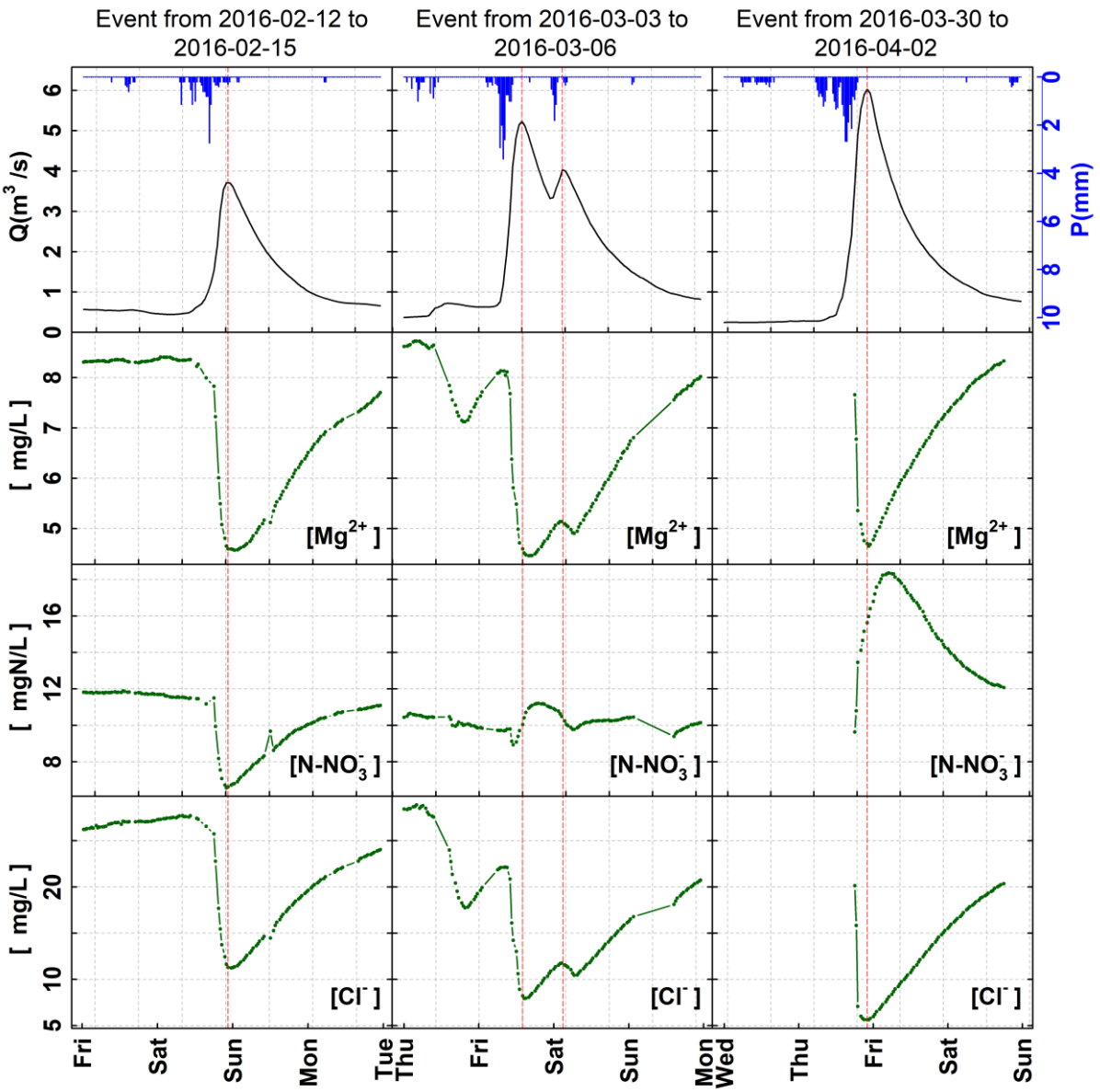
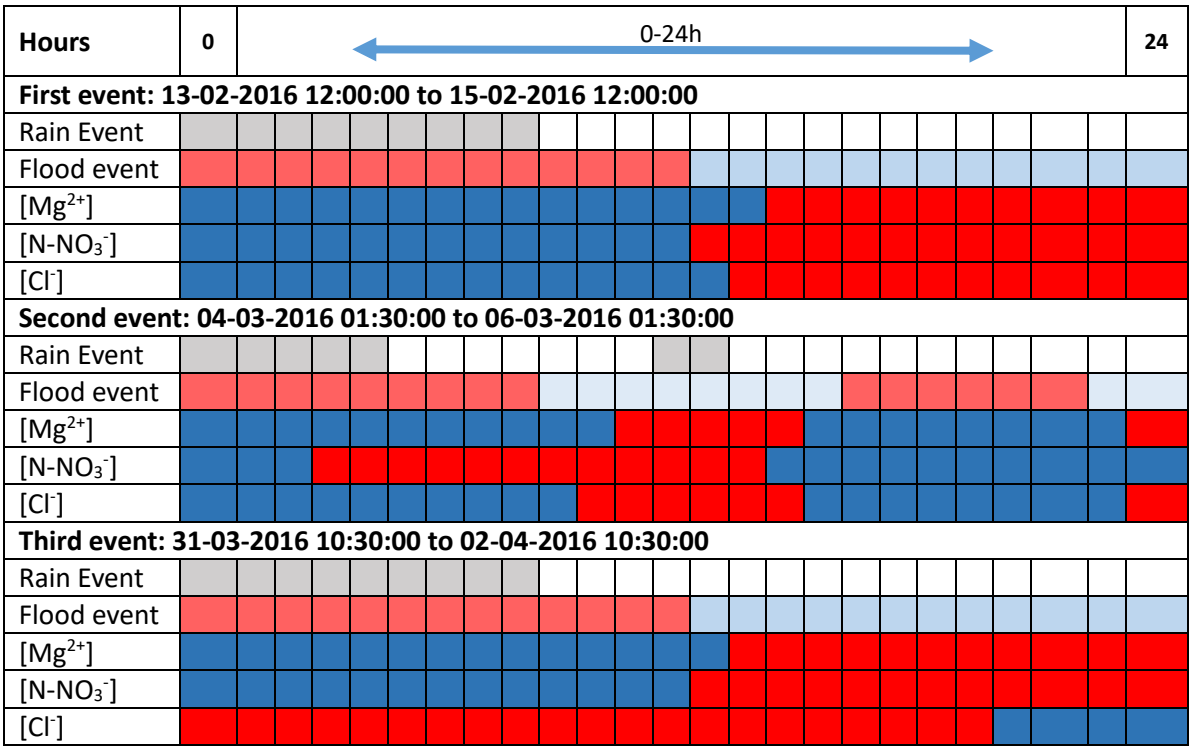


Figure 26 : Study of the typologies of concentrations for Magnesium, Potassium, Nitrates and Chlorides for three flood events, at the Avenelles station (River Lab, Oracle-Orgeval observatory). The red straight line represents the peak of the flood.

According to the ion and the flood event, the process and its duration could be variable. Table 7 resumes the processes and duration of each event and for each ion. It mainly shows the time-lag between hydrological processes and the ion dilution-concentration processes. This time-lag depends on the ion considered (Table 7). We also note that some water bodies can be mobilized during exceptional events (see chlorides during the third events (03/30 to 04/02/2016), Table 7).

This shows the essential interest of the high-frequency measurements to understand hydro-chemical processes, especially on small catchments with sub-daily functioning.

Table 7 : Processes and duration of each events for each ion (For precipitation: grey=rainfall. For discharge: light red = rising limb, light blue = falling limb. For ions: blue=dilution process, red=concentration process)



6.4 Hydrological and chemical behavior of the catchment

Observed and studied since 1962, a large number of studies (more than 500 references) have been developed in the Orgeval catchment and especially in the Avenelles sub-catchment.

Various processes of the hydrological cycle of the Orgeval catchment have been the subject of many research works based on soil measurements, spatial observations and modelling. Thus, some authors have been able to link runoff and/or infiltration to the physical characteristics of the soil (moisture, texture, surface condition, occupation) (Aubert et al., 2011; Augeard et al., 2005; Boureghda, 1988; Loumagne, 1984; Loumagne et al., 2013), to precipitation via rainfall simulations and remote sensing techniques (Jarry, 1987). Others have also evaluated these processes with the implementation of experimental *in situ* monitoring (Augeard, 2006), or the use of chemical and isotopic tracers such as oxygen 18 (Filiz, 1973). In the field of hydrological model, studies were made on the re-analysis of rainfall data of the catchment for the improvement of rainfall-runoff models (Lobligeois et al., 2013), flood forecasting in the different sub-catchments of the Oracle-Orgeval observatory (Furusho et al., 2013) and drought forecasting (Perrin et al., 2013). The analysis of the impact on different hydro-climatic data (rainfall, evapotranspiration, discharge, temperature) under different climate change scenarios was also conducted (Thirel and Perrin, 2013).

In the case of the hydrological interaction with tile-drains, Trincal (1994) has demonstrated that the increase in the proportion of drained areas did not have a significant influence on the catchment hydrological behavior. Henine (2010) also showed that drainage had the effect of reducing high floods, through a transitory storage role.

The Oracle-Orgeval observatory is the subject of an extensive study aimed at better understanding the exchange processes between groundwater and stream network of the catchment. These studies focused on the complex relationship between low-frequency hydrological processes developing at the scale of aquifers of large spatial extent and high-frequency hydrological processes developing in stream waters. For this reason an experimental system for monitoring exchanges (MoLoNaRi station) has thus been developed in the basin (Berrhouma, 2018; Mouhri et al., 2013a). The data collected are used as a basis for hydrological and hydrogeological modelling of the catchment around the groundwater-river interface. At these MoLoNaRi stations, the water levels in the river, the piezometric level in the groundwater and the temperatures have been measured continuously since the end of 2012. These measures have made it possible to establish a first water balance for the catchment (Mouhri et al., 2013a). To validate these first estimates, research is now focusing on the use of chemical tracers. The first campaigns based on major ions, and nutrients, as well as the subsequent campaigns based on isotopes and temperature, demonstrated a good agreement between hydrogeophysical interpretations (Berrhouma, 2018), geochemical monitoring (Mouchel et al., 2016), isotopic monitoring (Guillon et al., 2017) and groundwater temperature (Rivière et al., 2018).

In case of the chemical behavior of the observatory, most studies have focused on the behavior of nitrates. The interactions of the nitrogen cycle with the hydrological cycle in the catchment were investigated in detail (e.g. Garnier et al., 2016; Garnier et al., 2014; Vilain et al., 2010; Vilain et al., 2012), as well as the denitrification processes in the catchment (Billy, 2008; Billy et al., 2010; Flipo et al., 2007) and hydro-chemical modelling of nitrates (Hiver, 2015; MA, 1991).

Studies were also conducted on the source of the different chemical concentrations measured in the stream of the catchment (Floury et al., 2018; Le Cloarec et al., 2007) and continuous monitoring of chemical concentrations and potential applications (Floury, 2017; Floury et al., 2017).

7 References

Abbott, B.W., Gruau, G., Zarnetske, J.P., Moatar, F., Barbe, L., Thomas, Z., Fovet, O., Kolbe, T., Gu, S., Pierson-Wickmann, A.C., Davy, P., Pinay, G., 2018. Unexpected spatial stability of water chemistry in headwater stream networks. Ecology Letters, 21(2): 296-308. DOI:10.1111/ele.12897

Albek, E., 1999. Identification of the Different Sources of Chlorides in Streams by Regression Analysis Using Chloride-Discharge Relationships. Water Environment Research, 71(7): 1310-1319. DOI:10.2175/106143096x122384

Anderson, M., Burt, T., 1980. Interpretation of recession flow. Journal of Hydrology, 46(1-2): 89-101.

Aubert, M., Baghdadi, N., Zribi, M., Douaoui, A., Loumagne, C., Baup, F., El Hajj, M., Garrigues, S., 2011. Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust. Remote Sensing of Environment, 115(8): 1801-1810. DOI:https://doi.org/10.1016/j.rse.2011.02.021

Augeard, B., 2006. Mécanismes de genèse du ruissellement sur sol agricole drainé sensible à la battance : études expérimentales et modélisation, 1 vol. (236 p.) pp.

Augeard, B., Kao, C., Chaumont, C., Vauclin, M., 2005. Mechanisms of surface runoff genesis on a subsurface drained soil affected by surface crusting: A field investigation. Physics and Chemistry of the Earth, Parts A/B/C, 30(8): 598-610. DOI:https://doi.org/10.1016/j.pce.2005.07.014

Bansah, S., Ali, G., 2017. Evaluating the Effects of Tracer Choice and End-Member Definitions on Hydrograph Separation Results Across Nested, Seasonally Cold Watersheds. Water Resources Research, 53(11): 8851-8871.

Barakat, A., El Baghdadi, M., Rais, J., Aghezzaf, B., Slassi, M., 2016. Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. International Soil and Water Conservation Research, 4(4): 284-292. DOI:https://doi.org/10.1016/j.iswcr.2016.11.002

Barco, J., Gunawan, S., Hogue, T.S., 2013. Seasonal controls on stream chemical export across diverse coastal watersheds in the USA. Hydrological Processes, 27(10): 1440-1453. DOI:10.1002/hyp.9294

Barco, J., Hogue, T.S., Curto, V., Rademacher, L., 2008. Linking hydrology and stream geochemistry in urban fringe watersheds. Journal of Hydrology, 360(1): 31-47. DOI:https://doi.org/10.1016/j.jhydrol.2008.07.011

Barnes, B.S., 1939. The structure of discharge-recession curves. Eos, Transactions American Geophysical Union, 20(4): 721-725.

Barthold, F.K., Tyralla, C., Schneider, K., Vaché, K.B., Frede, H.-G., Breuer, L., 2011. How many tracers do we need for end member mixing analysis (EMMA)? A sensitivity analysis. Water Resources Research, 47(8). DOI:10.1029/2011wr010604

Basu, N.B., Destouni, G., Jawitz, J.W., Thompson, S.E., Loukinova, N.V., Darracq, A., Zanardo, S., Yaeger, M., Sivapalan, M., Rinaldo, A., 2010. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. Geophysical Research Letters, 37(23).

Basu, N.B., Thompson, S.E., Rao, P.S.C., 2011. Hydrologic and biogeochemical functioning of intensively managed catchments: A synthesis of top-down analyses. Water Resources Research, 47(10).

Benettin, P., Kirchner, J.W., Rinaldo, A., Botter, G., 2015. Modeling chloride transport using travel time distributions at Plynlimon, Wales. Water Resources Research, 51(5): 3259-3276. DOI:10.1002/2014wr016600

Bentura, P.L., Michel, C., 1997. Flood routing in a wide channel with a quadratic lag-and-route method. Hydrological Sciences Journal, 42(2): 169-189.

Berrhouma, A., 2018. Fonctionnement hydrothermique de l'interface nappe-rivière du bassin des Avenelles, PSL Research University.

Bieroza, M.Z., Heathwaite, A.L., Bechmann, M., Kyllmar, K., Jordan, P., 2018. The concentration-discharge slope as a tool for water quality management. Science of The Total Environment, 630: 738-749. DOI:https://doi.org/10.1016/j.scitotenv.2018.02.256

Billy, C., 2008. Transfert et rétention d'azote à l'échelle d'un bassin versant agricole artificiellement drainé, Université Paris VI, 1 vol. (222 p.) pp.

Billy, C., Billen, G., Sebilo, M., Birgand, F., Tournebize, J., 2010. Nitrogen isotopic composition of leached nitrate and soil organic matter as an indicator of denitrification in a sloping drained agricultural plot and adjacent uncultivated riparian buffer strips. Soil Biology and Biochemistry, 42(1): 108-117. DOI:https://doi.org/10.1016/j.soilbio.2009.09.026

Blake, S., Henry, T., Murray, J., Flood, R., Muller, M.R., Jones, A.G., Rath, V., 2016. Compositional multivariate statistical analysis of thermal groundwater provenance: A hydrogeochemical case study from Ireland. Applied Geochemistry, 75: 171-188. DOI:https://doi.org/10.1016/j.apgeochem.2016.05.008

Botter, M., Burlando, P., Fatichi, S., 2019. Anthropogenic and catchment characteristic signatures in the water quality of Swiss rivers: a quantitative assessment. Hydrol. Earth Syst. Sci., 23(4): 1885-1904. DOI:10.5194/hess-23-1885-2019

Boureghda, S., 1988. Influence des caracteristiques physiques (texture, etat de surface) de divers sols a vegetation naturelle ou cultives sur leur susceptibilite au ruissellement et a l'erosion: etude experimentale au champ sous pluies simulees, Orléans, 251 pp.

Boussinesq, J., 1877. Essai sur la théorie des eaux courantes, Memoires presentés par divers savants à l'Academic des Sciences de l'Institut National de France, XXIII. Impr. nationale, Paris, 63 pp.

Bowes, M., Jarvie, H., Halliday, S., Skeffington, R., Wade, A., Loewenthal, M., Gozzard, E., Newman, J., Palmer-Felgate, E., 2015. Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration–flow relationships. Science of the Total Environment, 511: 608-620.

Bowes, M.J., Smith, J.T., Neal, C., 2009. The value of high-resolution nutrient monitoring: A case study of the River Frome, Dorset, UK. Journal of Hydrology, 378(1–2): 82-96. DOI:http://dx.doi.org/10.1016/j.jhydrol.2009.09.015

Box, G.E., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2): 211-243.

Brick, C.M., Moore, J.N., 1996. Diel variation of trace metals in the upper Clark Fork River, Montana. Environmental Science & Technology, 30(6): 1953-1960. DOI:10.1021/es9506465

Brodie, R., Sundaram, B., Tottenham, R., Hostetler, S., Ransley, T., 2007. An overview of tools for assessing groundwater-surface water connectivity. Bureau of Rural Sciences, Canberra, Australia: 57-70.

Brutsaert, W., Lopez, J.P., 1998. Basin-scale geohydrologic drought flow features of riparian aquifers in the Southern Great Plains. Water Resources Research, 34(2): 233-240. DOI:10.1029/97wr03068

Brutsaert, W., Nieber, J.L., 1977. Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resources Research, 13(3): 637-643.

Buttle, J.M., 1994. Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins. Progress in Physical Geography: Earth and Environment, 18(1): 16-41.

Campbell, F.B., Bauder, H.A., 1940. A rating-curve method for determining silt-discharge of streams. Eos, Transactions American Geophysical Union, 21(2): 603-607. DOI:10.1029/TR021i002p00603

Capozzi, S.L., Rodenburg, L.A., Krumins, V., Fennell, D.E., Mack, E.E., 2018. Using positive matrix factorization to investigate microbial dehalogenation of chlorinated benzenes in groundwater at a historically contaminated site. Chemosphere, 211: 515-523. DOI:https://doi.org/10.1016/j.chemosphere.2018.07.180

Carrera, J., Vázquez-Suñé, E., Castillo, O., Sánchez-Vila, X., 2004. A methodology to compute mixing ratios with uncertain end-members. Water Resources Research, 40(12). DOI:10.1029/2003wr002263

Cartwright, I., Gilfedder, B., Hofmann, H., 2014. Contrasts between estimates of baseflow help discern multiple sources of water contributing to rivers. Hydrology and Earth System Sciences, 18(1): 15-30.

Chanat, J.G., Rice, K.C., Hornberger, G.M., 2002. Consistency of patterns in concentration-discharge plots. Water Resources Research, 38(8): 22-1-22-10. DOI:10.1029/2001wr000971

Chapin, T.P., Caffrey, J.M., Jannasch, H.W., Coletti, L.J., Haskins, J.C., Johnson, K.S., 2004. Nitrate sources and sinks in Elkhorn Slough, California: results from long-term continuous in situ nitrate analyzers. Estuaries, 27(5): 882-894.

Chapman, P.J., Reynolds, B., Wheater, H.S., 1997. Sources and controls of calcium and magnesium in storm runoff: the role of groundwater and ion exchange reactions along water flowpaths. Hydrology and Earth System Sciences, 1(3): 671-685.

Chapman, T., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrological Processes, 13(5): 701-714.

Chapman, T., Maxwell, A., 1996. Baseflow separation-comparison of numerical methods with tracer experiments, Hydrology and Water Resources Symposium 1996: Water and the Environment; Preprints of Papers. Institution of Engineers, Australia, pp. 539.

Chapman, T.G., 1991. Comment on “Evaluation of automated techniques for base flow and recession analyses” by R. J. Nathan and T. A. McMahon. Water Resources Research, 27(7): 1783-1784.

Cheng, L., Zhang, L., Brutsaert, W., 2016. Automated selection of pure base flows from regular daily streamflow data: objective algorithm. Journal of Hydrologic Engineering, 21(11): 1-7.

Chorover, J., Derry, L.A., McDowell, W.H., 2017. Concentration-Discharge Relations in the Critical Zone: Implications for Resolving Critical Zone Structure, Function, and Evolution. Water Resources Research, 53(11): 8654-8659. DOI:10.1002/2017wr021111

Christophersen, N., Hooper, R.P., 1992. Multivariate analysis of stream water chemical data: The use of principal components analysis for the end-member mixing problem. Water Resources Research, 28(1): 99-107. DOI:10.1029/91WR02518

Christophersen, N., Neal, C., Hooper, R.P., Vogt, R.D., Andersen, S., 1990. Modelling streamwater chemistry as a mixture of soilwater end-members—a step towards second-generation acidification models. Journal of Hydrology, 116(1-4): 307-320.

Clow, D.W., Mast, M.A., 2010. Mechanisms for chemostatic behavior in catchments: Implications for CO2 consumption by mineral weathering. Chemical Geology, 269(1): 40-51. DOI:https://doi.org/10.1016/j.chemgeo.2009.09.014

Cohen, M.J., Kurz, M.J., Heffernan, J.B., Martin, J.B., Douglass, R.L., Foster, C.R., Thomas, R.G., 2013. Diel phosphorus variation and the stoichiometry of ecosystem metabolism in a large spring-fed river. Ecological Monographs, 83(2): 155-176. DOI:10.1890/12-1497.1

Comero, S., Locoro, G., Free, G., Vaccaro, S., De Capitani, L., Gawlik, B.M., 2011. Characterisation of Alpine lake sediments using multivariate statistical techniques. Chemometrics and Intelligent Laboratory Systems, 107(1): 24-30. DOI:https://doi.org/10.1016/j.chemolab.2011.01.002

Devic, G., Djordjevic, D., Sakan, S., 2014. Natural and anthropogenic factors affecting the groundwater quality in Serbia. Science of The Total Environment, 468-469: 933-942. DOI:https://doi.org/10.1016/j.scitotenv.2013.09.011

Dewalle, D.R., Swistock, B.R., Sharpe, W.E., 1988. Three-component tracer model for stormflow on a small Appalachian forested catchment. Journal of Hydrology, 104(1-4): 301-310.

Dickinson, W., Holland, M., Smith, G.L., 1967. An experimental rainfall-runoff facility. Hydrology papers (Colorado State University); no. 25.

Duncan, J.M., Band, L.E., Groffman, P.M., 2017. Variable nitrate concentration–discharge relationships in a forested watershed. Hydrological Processes, 31(9): 1817-1824. DOI:10.1002/hyp.11136

Durum, W.H., 1953. Relationship of the mineral constituents in solution to stream flow, Saline River near Russell, Kansas. Eos, Transactions American Geophysical Union, 34(3): 435-442. DOI:10.1029/TR034i003p00435

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes, 19(2): 507-515.

Edwards, A.M.C., 1973. The variation of dissolved constituents with discharge in some Norfolk rivers. Journal of Hydrology, 18(3): 219-242. DOI:https://doi.org/10.1016/0022-1694(73)90049-8

Eng, K., Brutsaert, W., 1999. Generality of drought flow characteristics within the Arkansas River Basin. Journal of Geophysical Research: Atmospheres, 104(D16): 19435-19441. DOI:10.1029/1999jd900087

Evans, C., Davies, T.D., 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. Water Resources Research, 34(1): 129-137. DOI:10.1029/97wr01881

Evans, C., Davies, T.D., Murdoch, P.S., 1999. Component flow processes at four streams in the Catskill Mountains, New York, analysed using episodic concentration/discharge relationships. Hydrological Processes, 13(4): 563-575. DOI:10.1002/(sici)1099-1085(199903)13:4<563::aid-hyp711>3.0.co;2-n

Filiz, S., 1973. Etude du ruissellement et de l'infiltration sur le bassin versant experimental de l'Orgeval à l'aide de l'oxygène 18, Université Paris VI, 95 pp.

Flipo, N., Even, S., Poulin, M., Théry, S., Ledoux, E., 2007. Modeling nitrate fluxes at the catchment scale using the integrated tool CAWAQS. Science of The Total Environment, 375(1): 69-79. DOI:https://doi.org/10.1016/j.scitotenv.2006.12.016

Floury, P., 2017. La dynamique des bassins versants sous l'angle de l'analyse chimique à haute fréquences des rivières, Sorbonne Paris 259 pp.

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., Roubaty, J.L., 2017. The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data. Hydrol. Earth Syst. Sci., 21(12): 6153-6165.

Floury, P., Gaillardet, J., Tallec, G., Ansart, P., Bouchez, J., Louvat, P., Gorge, C., 2018. Chemical weathering and CO2 consumption rate in a multilayered-aquifer dominated watershed under intensive farming: The Orgeval Critical Zone Observatory, France. Hydrological Processes: 1-19.

Fovet, O., Ruiz, L., Gruau, G., Akkal, N., Aquilina, L., Busnot, S., Dupas, R., Durand, P., Faucheux, M., Fauvel, Y., Fléchard, C., Gilliet, N., Grimaldi, C., Hamon, Y., Jaffrezic, A., Jeanneau, L., Labasque, T., Le Henaff, G., Mérot, P., Molénat, J., Petitjean, P., Pierson-Wickmann, A.-C., Squividant, H., Viaud, V., Walter, C., Gascuel-Odoux, C., 2018. AgrHyS: An Observatory of Response Times in Agro-Hydro Systems. Vadose Zone Journal, 17(1). DOI:10.2136/vzj2018.04.0066

Furusho, C., Lilas, D., Perrin, C., Andréassian, V., Coron, L., Peschard, J., Berthet, L., Ansart, P., Loumagne, C., 2013. Prévision des crues par modélisation hydrologique In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 63-74.

Gaillardet, J., Braud, I., Hankard, F., Anquetin, S., Bour, O., Dorfliger, N., de Dreuzy, J.R., Galle, S., Galy, C., Gogo, S., Gourcy, L., Habets, F., Laggoun, F., Longuevergne, L., Le Borgne, T., Naaim-Bouvet, F., Nord, G., Simonneaux, V., Six, D., Tallec, T., Valentin, C., Abril, G., Allemand, P., Arènes, A., Arfib, B., Arnaud, L., Arnaud, N., Arnaud, P., Audry, S., Comte, V.B., Batiot, C., Battais, A., Bellot, H., Bernard, E., Bertrand, C., Bessière, H., Binet, S., Bodin, J., Bodin, X., Boithias, L., Bouchez, J., Boudevillain, B., Moussa, I.B., Branger, F., Braun, J.J., Brunet, P., Caceres, B., Calmels, D., Cappelaere, B., Celle-Jeanton, H., Chabaux, F., Chalikakis, K., Champollion, C., Copard, Y., Cotel, C., Davy, P., Deline, P., Delrieu, G., Demarty, J., Dessert, C., Dumont, M., Emblanch, C., Ezzahar, J., Estèves, M., Favier, V., Faucheux, M., Filizola, N., Flammarion, P., Floury, P., Fovet, O., Fournier, M., Francez, A.J., Gandois, L., Gascuel, C., Gayer, E., Genthon, C., Gérard, M.F., Gilbert, D., Gouttevin, I., Grippa, M., Gruau, G., Jardani, A., Jeanneau, L., Join, J.L., Jourde, H., Karbou, F., Labat, D., Lagadeuc, Y., Lajeunesse, E., Lastennet, R., Lavado, W., Lawin, E., Lebel, T., Le Bouteiller, C., Legout, C., Lejeune, Y., Le Meur, E., Le Moigne, N., Lions, J., Lucas, A., Malet, J.P., Marais-Sicre, C., Maréchal, J.C., Marlin, C., Martin, P., Martins, J., Martinez, J.M., Massei, N., Mauclerc, A., Mazzilli, N., Molénat, J., Moreira-Turcq, P., Mougin, E., Morin, S., Ngoupayou, J.N., Panthou, G., Peugeot, C., Picard, G., Pierret, M.C., Porel, G.,

Probst, A., Probst, J.L., Rabatel, A., Raclot, D., Ravanel, L., Rejiba, F., René, P., Ribolzi, O., Riotte, J., Rivière, A., Robain, H., Ruiz, L., Sanchez-Perez, J.M., Santini, W., Sauvage, S., Schoeneich, P., Seidel, J.L., Sekhar, M., Sengtaheuanghoung, O., Silvera, N., Steinmann, M., Soruco, A., Tallec, G., Thibert, E., Lao, D.V., Vincent, C., Viville, D., Wagnon, P., Zitouna, R., 2018a. OZCAR: The French Network of Critical Zone Observatories. Vadose Zone Journal, 17(1). DOI:10.2136/vzj2018.04.0067

Gaillardet, J., Calmels, D., Romero-Mujalli, G., Zakharova, E., Hartmann, J., 2018b. Global climate control on carbonate weathering intensity. Chemical Geology. DOI:https://doi.org/10.1016/j.chemgeo.2018.05.009

Garnier, J., Anglade, J., Benoit, M., Billen, G., Puech, T., Ramarson, A., Passy, P., Silvestre, M., Lassaletta, L., Trommenschlager, J.M., Schott, C., Tallec, G., 2016. Reconnecting crop and cattle farming to reduce nitrogen losses to river water of an intensive agricultural catchment (Seine basin, France): past, present and future. Environmental Science & Policy, 63: 76-90.

Garnier, J., Billen, G., Vilain, G., Benoit, M., Passy, P., Tallec, G., Tournebize, J., Anglade, J., Billy, C., Mercier, B., Ansart, P., Azougui, A., Sebilo, M., Kao, C., 2014. Curative vs. preventive management of nitrogen transfers in rural areas: Lessons from the case of the Orgeval watershed (Seine River basin, France). Journal of Environmental Management, 144: 125-134.

Genereux, D.P., Hemond, H.F., Mulholland, P.J., 1993. Use of radon-222 and calcium as tracers in a three-end-member mixing model for streamflow generation on the West Fork of Walker Branch Watershed. Journal of Hydrology, 142(1): 167-211. DOI:https://doi.org/10.1016/0022-1694(93)90010-7

Godsey, S.E., Kirchner, J.W., Clow, D.W., 2009. Concentration-discharge relationships reflect chemostatic characteristics of US catchments. Hydrological Processes, 23(13): 1844-1864. DOI:10.1002/hyp.7315

Gonzales, A., Nonner, J., Heijkers, J., Uhlenbrook, S., 2009. Comparison of different base flow separation methods in a lowland catchment. Hydrology and Earth System Sciences, 13(11): 2055-2068.

Gregory, K.J., Walling, D.E., 1973. Quantitative evaluation of drainage basin process, Drainage basin form and process. Fletcher & Son Ltd, United Kingdom, pp. 183-233.

Grunsky, E.C., 2010. The interpretation of geochemical survey data. Geochemistry: Exploration, Environment, Analysis, 10(1): 27. DOI:10.1144/1467-7873/09-210

Guillon, S., Rivière, A., Flipo, N., 2017. Premiers retours sur la faisabilité du traçage des écoulements à l'aide des isotopes stables de l'eau et du radon. Tech. rep. PIREN Seine, France.

Gunnerson, C.G., 1967. Streamflow and quality in the Columbia River basin. Journal of the Sanitary Engineering Division, 93(6): 1-16.

Gustard, A., Bullock, A., Dixon, J.M., 1992. Low flow estimation in the United Kingdom. Report 108. Institute of Hydrology, UK, 88 pp.

Gustard, A., Demuth, S., 2009. Manual on low-flow estimation and prediction. Opera.

Haji Gholizadeh, M., Melesse, A.M., Reddi, L., 2016. Water quality assessment and apportionment of pollution sources using APCS-MLR and PMF receptor modeling techniques in three major rivers of South Florida. Science of The Total Environment, 566-567: 1552-1567. DOI:https://doi.org/10.1016/j.scitotenv.2016.06.046

Hall, F.R., 1968. Base-flow recessions-a review. Water Resources Research, 4(5): 973-983.

Hall, F.R., 1970. Dissolved solids-discharge relationships .1. Mixing models. Water Resources Research, 6(3): 845-&. DOI:10.1029/WR006i003p00845

Hall, F.R., 1971. Dissolved solids-discharge relationships .2. Applications to field data. Water Resources Research, 7(3): 591-&. DOI:10.1029/WR007i003p00591

Hem, J.D., 1948. Fluctuations in concentration of dissolved solids of some southwestern streams. Eos, Transactions American Geophysical Union, 29(1): 80-84. DOI:10.1029/TR029i001p00080

Henine, H., 2010. Couplage des processus hydrologiques reliant parcelles agricoles drainées, collecteurs enterrés et émissaire à surface libre : intégration à l'échelle du bassin versant, Paris VI, 1 vol. (210 p.) pp.

Hirsch, R.M., Moyer, D.L., Archfield, S.A., 2010. Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs1. JAWRA Journal of the American Water Resources Association, 46(5): 857-880. DOI:10.1111/j.1752-1688.2010.00482.x

Hiver, M., 2015. Modélisation d'un bassin versant artificiellement drainé et desflux de nitrates associés: Les bassins de l'Orgeval et de Porijõgi, Paris VI, 50 pp.

Hooper, R.P., 2003. Diagnostic tools for mixing models of stream water chemistry. Water Resources Research, 39(3).

Horton, R.E., 1933. The role of infiltration in the hydrologic cycle. Eos, Transactions American Geophysical Union, 14(1): 446-460.

House, W.A., Warwick, M.S., 1998. Hysteresis of the solute concentration/discharge relationship in rivers during storms. Water Research, 32(8): 2279-2290. DOI:10.1016/s0043-1354(97)00473-9

Howarth, R.J., 1984. Statistical applications in geochemical prospecting: A survey of recent developments. Journal of Geochemical Exploration, 21(1): 41-61. DOI:https://doi.org/10.1016/0375-6742(84)90033-5

Hubert, P., Martin, E., Meybeck, M., Oliver, P., Siwertz, E., 1969. Aspects hydrologique, géochimique et sédimentologique de la crue exceptionelle de la Dranse du Chablais du 22 sept. 1968. Arch. Sci. Soc. Phys. (Genève), 22(3): 581-603.

Huntington, T.G., Ryan, D.F., Hamburg, S.P., 1988. Estimating Soil Nitrogen and Carbon Pools in a Northern Hardwood Forest Ecosystem. Soil Science Society of America Journal, 52(4): 1162-1167. DOI:10.2136/sssaj1988.03615995005200040049x

Hydrology, I.o., 1980. Low Flow Studies report, Wallingford, UK.

Jarry, F., 1987. Le ruissellement sur les terres agricoles. Approche par simulation de pluie et par teledectetion, Université Paris 7, 276 pp.

Jarvie, H.P., Neal, C., Smart, R., Owen, R., Fraser, D., Forbes, I., Wade, A., 2001. Use of continuous water quality records for hydrograph separation and to assess short-term variability and extremes in acidity and dissolved carbon dioxide for the River Dee, Scotland. Science of the Total Environment, 265(1-3): 85-98. DOI:10.1016/s0048-9697(00)00651-3

Jasechko, S., 2019. Global Isotope Hydrogeology—Review. Reviews of Geophysics, 0(0). DOI:10.1029/2018rg000627

Johnson, N.M., Likens, G.E., Bormann, F.H., Fisher, D.W., Pierce, R.S., 1969. A working model for variation in stream water chemistry at Hubbard-Brook-experimental-forest, new-hampshire. Water Resources Research, 5(6): 1353-&. DOI:10.1029/WR005i006p01353

Kirchner, J.W., 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research, 42(3).

Kirchner, J.W., 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. Water Resources Research, 45(2). DOI:10.1029/2008wr006912

Kirchner, J.W., 2019. Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: theory and benchmark tests. Hydrology & Earth System Sciences, 23(1).

Kirchner, J.W., Feng, X., Neal, C., Robson, A.J., 2004. The fine structure of water-quality dynamics: the (high-frequency) wave of the future. Hydrological Processes, 18(7): 1353-1359.

Klaus, J., McDonnell, J., 2013. Hydrograph separation using stable isotopes: Review and evaluation. Journal of Hydrology, 505: 47-64.

Koh, D.-C., Chae, G.-T., Ryu, J.-S., Lee, S.-G., Ko, K.-S., 2016. Occurrence and mobility of major and trace elements in groundwater from pristine volcanic aquifers in Jeju Island, Korea. Applied Geochemistry, 65: 87-102. DOI:https://doi.org/10.1016/j.apgeochem.2015.11.004

Koskelo, A.I., Fisher, T.R., Utz, R.M., Jordan, T.E., 2012. A new precipitation-based method of baseflow separation and event identification for small watersheds (< 50 km2). Journal of hydrology, 450: 267-278.

Kulandaiswamy, V.C., Seetharaman, S., 1969. A note on Barnes' method of hydrograph separation. Journal of Hydrology, 9(2): 222-229. DOI:https://doi.org/10.1016/0022-1694(69)90080-8

Kunkle, G.R., 1965. Computation of ground-water discharge to streams during floods, or to individual reaches during base flow by use of specific conductance. US Geol. Surv., Prof. Pap: 207-210.

La Sala Jr., A.M., 1967. New Approaches to Water-Resources Investigations in Upstate New Yorkaa. Groundwater, 5(4): 6-11. DOI:10.1111/j.1745-6584.1967.tb01619.x

Ladouche, B., Probst, A., Viville, D., Idir, S., Baqué, D., Loubet, M., Probst, J.-L., Bariac, T., 2001. Hydrograph separation using isotopic, chemical and hydrological approaches (Strengbach catchment, France). Journal of Hydrology, 242(3): 255-274.

Langbein, W., 1938. Some channel-storage studies and their application to the determination of infiltration. Eos, Transactions American Geophysical Union, 19(1): 435-447.

Lawler, D.M., Petts, G.E., Foster, I.D.L., Harper, S., 2006. Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK. Science of The Total Environment, 360(1): 109-126. DOI:https://doi.org/10.1016/j.scitotenv.2005.08.032

Lawrence, G.B., Driscoll, C.T., 1990. Longitudinal patterns of concentration-discharge relationships in stream water draining the Hubbard Brook Experimental Forest, New Hampshire. Journal of Hydrology, 116(1): 147-165. DOI:https://doi.org/10.1016/0022-1694(90)90120-M

Lawrence, G.B., Fuller, R.D., Driscoll, C.T., 1986. Spatial relationships of aluminum chemistry in the streams of the Hubbard Brook Experimental Forest, New Hampshire. Biogeochemistry, 2(2): 115-135. DOI:10.1007/bf02180190

Le Cloarec, M.-F., Bonté, P., Lefèvre, I., Mouchel, J.-M., Colbert, S., 2007. Distribution of 7Be, 210Pb and 137Cs in watersheds of different scales in the Seine River basin: Inventories and residence times. Science of The Total Environment, 375(1): 125-139. DOI:https://doi.org/10.1016/j.scitotenv.2006.12.020

Ledbetter, J.E., Gloyna, E.F., 1964. Predictive Techniques for Water Quality-Inorganics. Journal of the Sanitary Engineering Division, 90(1): 127-154.

Lenz, A., Sawyer, C.N., 1944. Estimation of stream-flow from alkalinity-determinations. Eos, Transactions American Geophysical Union, 25(6): 1005-1011. DOI:10.1029/TR025i006p01005

Li, T., Sun, G., Yang, C., Liang, K., Ma, S., Huang, L., Luo, W., 2019. Source apportionment and source-to-sink transport of major and trace elements in coastal sediments: Combining positive matrix factorization and sediment trend analysis. Science of The Total Environment, 651: 344-356. DOI:https://doi.org/10.1016/j.scitotenv.2018.09.198

Linsley Jr, R.K., Kohler, M.A., Paulhus, J.L., 1958. Hydrology for engineers. McGraw-Hill Civil Engineering Series. McGraw-Hill Book Company, USA, 340 pp.

Liu, F., Bales, R.C., Conklin, M.H., Conrad, M.E., 2008. Streamflow generation from snowmelt in semi-arid, seasonally snow-covered, forested catchments, Valles Caldera, New Mexico. Water Resources Research, 44(12).

Lloyd, C.E.M., Freer, J.E., Johnes, P.J., Collins, A.L., 2016. Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments. Science of the Total Environment, 543: 388-404. DOI:10.1016/j.scitotenv.2015.11.028

Lobligeois, F., Andréassian, V., Perrin, C., Loumagne, C., 2013. Réanalyse des larmes d'eau radar pour la modélisation hydrologique pluie-débit. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 39-62.

Longobardi, A., Loon, A.F.V., 2018. Assessing baseflow index vulnerability to variation in dry spell length for a range of catchment and climate properties. Hydrological Processes, 32(16): 2496-2509. DOI:10.1002/hyp.13147

Longobardi, A., Villani, P., Guida, D., Cuomo, A., 2016. Hydro-geo-chemical streamflow analysis as a support for digital hydrograph filtering in a small, rainfall dominated, sandstone watershed. Journal of Hydrology, 539: 177-187.

Lott, D.A., Stewart, M.T., 2016. Base flow separation: A comparison of analytical and mass balance methods. Journal of Hydrology, 535: 525-533.

Loumagne, C., 1984. Prédétermination du coefficient d'écoulement, Université Paris-Sud, 114 pp.

Loumagne, C., Ottlé, C., Zribi, M., 2013. Assimilation de données d'état hydrique des sols pour la modélisation des débits. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 269-284.

Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modelling, Institute of Engineers Australia National Conference, pp. 89-93.

MA, Z.C., 1991. Modélisation du transfert des nitrates: du bassin de recherche au grand bassin (exemples des bassins de Mélarchez et de la Charente), Strasbourg 1.

Maher, K., 2011. The role of fluid residence time and topographic scales in determining chemical fluxes from landscapes. Earth and Planetary Science Letters, 312(1): 48-58. DOI:https://doi.org/10.1016/j.epsl.2011.09.040

Maillet, E.T., 1905. Essais d'hydraulique souterraine & fluviale. A. Hermann, Paris, 286 pp.

McDonnell, J.J., Stewart, M.K., Owens, I.F., 1991. Effect of Catchment-Scale Subsurface Mixing on Stream Isotopic Response. Water Resources Research, 27(12): 3065-3073. DOI:10.1029/91wr02025

Mei, Y., Anagnostou, E.N., 2015. A hydrograph separation method based on information from rainfall and runoff records. Journal of Hydrology, 523: 636-649.

Meybeck, M., 1983. Atmospheric inputs and river transport of dissolved substances. Dissolved Loads of Rivers and Surface Water Quantity/Quality Relationships(141): 173-192.

Meybeck, M., Moatar, F., 2012. Daily variability of river concentrations and fluxes: indicators based on the segmentation of the rating curve. Hydrological Processes, 26(8): 1188-1207.

Miller, M.P., Tesoriero, A.J., Hood, K., Terziotti, S., Wolock, D.M., 2017. Estimating Discharge and Nonpoint Source Nitrate Loading to Streams From Three End-Member Pathways Using High-Frequency Water Quality Data. Water Resources Research.

Minaudo, C., Dupas, R., Gascuel-Odoux, C., Roubeix, V., Danis, P.-A., Moatar, F., 2019. Seasonal and event-based concentration-discharge relationships to identify catchment controls on nutrient export regimes. Advances in Water Resources, 131: 103379. DOI:https://doi.org/10.1016/j.advwatres.2019.103379

Moatar, F., Abbott, B., Minaudo, C., Curie, F., Pinay, G., 2017. Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. Water Resources Research, 53(2): 1270-1287.

Moraetis, D., Efstathiou, D., Stamati, F., Tzoraki, O., Nikolaidis, N.P., Schnoor, J.L., Vozinakis, K., 2010. High-frequency monitoring for the identification of hydrological and bio-geochemical processes in a Mediterranean river basin. Journal of Hydrology, 389(1-2): 127-136. DOI:10.1016/j.jhydrol.2010.05.037

Mouchel, J.-M., Rocha, S., Rivière, A., Tallec, G., 2016. Caractérisation de la géochimie des interfaces nappe-rivière du bassin des Avenelles. Tech. rep. PIREN Seine, France, 27 pp.

Mouhri, A., Flipo, N., Réjiba, F., De Fouquet, C., Bodet, L., Kurtulus, B., Tallec, G., Durand, V., Jost, A., Ansart, P., 2013a. Designing a multi-scale sampling system of stream–aquifer interfaces in a sedimentary basin. Journal of Hydrology, 504: 194-206.

Mouhri, A., Flipo, N., Rejiba, F., de Fouquet, C., Tallec, G., Bodet, L., Durand, V., Jost, A., Guérin, R., Ansart, P., 2012. Stratégie d'échantillonnage des échanges nappe-rivière du bassin agricole de l'Orgeval. Tech. rep. PIREN Seine, France, 27 pp.

Mouhri, A., Flipo, N., Vitale, Q., Bodet, L., Tallec, G., Ansart, P., Rejiba, F., 2013b. Influence du contexte hydrogéologique sur la connectivité nappe - rivière. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 89-98.

Mul, M.L., Mutiibwa, R.K., Uhlenbrook, S., Savenije, H.H., 2008. Hydrograph separation using hydrochemical tracers in the Makanya catchment, Tanzania. Physics and Chemistry of the Earth, Parts A/B/C, 33(1): 151-156.

Musolff, A., Fleckenstein, J.H., Rao, P.S.C., Jawitz, J.W., 2017. Emergent archetype patterns of coupled hydrologic and biogeochemical responses in catchments. Geophysical Research Letters, 44(9): 4143-4151. DOI:10.1002/2017gl072630

Musolff, A., Schmidt, C., Selle, B., Fleckenstein, J.H., 2015. Catchment controls on solute export. Advances in Water Resources, 86: 133-146. DOI:10.1016/j.advwatres.2015.09.026

Musy, A., 2001. E-drologie. Lausanne, Switzerland, Ecole Polytechnique Fédérale.

Nathan, R., McMahon, T., 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research, 26(7): 1465-1473.

Neal, C., Reynolds, B., Kirchner, J.W., Rowland, P., Norris, D., Sleep, D., Lawlor, A., Woods, C., Thacker, S., Guyatt, H., Vincent, C., Lehto, K., Grant, S., Williams, J., Neal, M., Wickham, H., Harman, S., Armstrong, L., 2013. High-frequency precipitation and stream water quality time series from Plynlimon, Wales: an openly accessible data resource spanning the periodic table. Hydrological Processes, 27(17): 2531-2539.

Neal, C., Reynolds, B., Rowland, P., Norris, D., Kirchner, J.W., Neal, M., Sleep, D., Lawlor, A., Woods, C., Thacker, S., Guyatt, H., Vincent, C., Hockenhull, K., Wickham, H., Harman, S., Armstrong, L., 2012. High-frequency water quality time series in precipitation and streamflow: From fragmentary signals to scientific challenge. Science of the Total Environment, 434: 3-12.

Paatero, P., Tapper, U., 1994. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. Environmetrics, 5(2): 111-126. DOI:doi:10.1002/env.3170050203

Pelizardi, F., Bea, S.A., Carrera, J., Vives, L., 2017. Identifying geochemical processes using End Member Mixing Analysis to decouple chemical components for mixing ratio calculations. Journal of Hydrology, 550: 144-156. DOI:https://doi.org/10.1016/j.jhydrol.2017.04.010

Perrin, C., Andréassian, V., Nicolle, P., 2013. Prévision des étiages. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 75-88.

Pfister, L., Thielen, F., Deloule, E., Valle, N., Lentzen, E., Grave, C., Beisel, J.-N., McDonnell, J.J., 2018. Freshwater pearl mussels as a stream water stable isotope recorder. Ecohydrology, 11(7): e2007. DOI:10.1002/eco.2007

Phung, D., Huang, C., Rutherford, S., Dwirahmadi, F., Chu, C., Wang, X., Nguyen, M., Nguyen, N.H., Do, C.M., Nguyen, T.H., Dinh, T.A.D., 2015. Temporal and spatial assessment of river surface water quality using multivariate statistical techniques: a study in Can Tho City, a Mekong Delta area, Vietnam. Environmental Monitoring and Assessment, 187(5): 229. DOI:10.1007/s10661-015-4474-x

Pierret, M.-C., Cotel, S., Ackerer, P., Beaulieu, E., Benarioumlil, S., Boucher, M., Boutin, R., Chabaux, F., Delay, F., Fourtet, C., Friedmann, P., Fritz, B., Gangloff, S., Girard, J.-F., Legtchenko, A., Viville, D., Weill, S., Probst, A., 2018. The Strengbach Catchment: A Multidisciplinary Environmental Sentry for 30 Years. Vadose Zone Journal, 17(1). DOI:10.2136/vzj2018.04.0090

Pinder, G.F., Jones, J.F., 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. Water Resources Research, 5(2): 438-445.

Probst, J.L., 1985. Nitrogen and Phosphorus exportation in the Garonne basin (France). Journal of Hydrology, 76(3-4): 281-305. DOI:10.1016/0022-1694(85)90138-6

Reff, A., Eberly, S.I., Bhave, P.V., 2007. Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods. Journal of the Air & Waste Management Association, 57(2): 146-154. DOI:10.1080/10473289.2007.10465319

Reimann, C., Filzmoser, P., Garrett, R.G., 2002. Factor analysis applied to regional geochemical data: problems and possibilities. Applied Geochemistry, 17(3): 185-206. DOI:https://doi.org/10.1016/S0883-2927(01)00066-X

Rimmer, A., Hartmann, A., 2014. Optimal hydrograph separation filter to evaluate transport routines of hydrological models. Journal of Hydrology, 514: 249-257.

Rivière, A., Flipo, N., Ansart, P., Baudin, A., Marlot, L., 2018. Revue des données de température du bassin des Avenelles. Tech. rep. PIREN Seine, France.

Robson, A.J., 1993. The use of continuous measurement in understanding and modelling the hydrochemistry of the uplands. Phd Thesis, University of Lancaster.

Rode, M., Wade, A.J., Cohen, M.J., Hensley, R.T., Bowes, M.J., Kirchner, J.W., Arhonditsis, G.B., Jordan, P., Kronvang, B., Halliday, S.J., Skeffington, R.A., Rozemeijer, J.C., Aubert, A.H., Rinke, K., Jomaa, S., 2016. Sensors in the Stream: The High-Frequency Wave of the Present. Environmental Science & Technology, 50(19): 10297-10307. DOI:10.1021/acs.est.6b02155

Rose, L.A., Karwan, D.L., Godsey, S.E., 2018. Concentration–discharge relationships describe solute and sediment mobilization, reaction, and transport at event and longer timescales. Hydrological Processes, 32(18): 2829-2844. DOI:10.1002/hyp.13235

Rusjan, S., Brilly, M., Mikoš, M., 2008. Flushing of nitrate from a forested watershed: an insight into hydrological nitrate mobilization mechanisms through seasonal high-frequency stream nitrate dynamics. Journal of Hydrology, 354(1): 187-202.

Rutledge, A., 1998. Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records: Update. US Geological Survey, 98: 4148.

Salmon, C.D., Walter, M.T., Hedin, L.O., Brown, M.G., 2001. Hydrological controls on chemical export from an undisturbed old-growth Chilean forest. Journal of Hydrology, 253(1): 69-80. DOI:https://doi.org/10.1016/S0022-1694(01)00447-4

Saraiva Okello, A.M.L., Uhlenbrook, S., Jewitt, G.P., Masih, I., Riddell, E.S., Van der Zaag, P., 2018. Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid mesoscale catchment. Hydrological Processes, 32(10): 1334-1350.

Semkin, R.G., Jeffries, D.S., Clair, T.A., 1994. Hydrochemical methods and relationships for study of stream output from small catchments. SCOPE-SCIENTIFIC COMMITTEE ON PROBLEMS OF THE ENVIRONMENT INTERNATIONAL COUNCIL OF SCIENTIFIC UNIONS, 51: 163-163.

Shanley, J.B., McDowell, W.H., Stallard, R.F., 2011. Long-term patterns and short-term dynamics of stream solutes and suspended sediment in a rapidly weathering tropical watershed. Water Resources Research, 47(7).

Sherson, L.R., Van Horn, D.J., Gomez-Velez, J.D., Crossey, L.J., Dahm, C.N., 2015. Nutrient dynamics in an alpine headwater stream: use of continuous water quality sensors to examine responses to wildfire and precipitation events. Hydrological Processes, 29(14): 3193-3207. DOI:10.1002/hyp.10426

Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M., Kouimtzis, T., 2003. Assessment of the surface water quality in Northern Greece. Water Research, 37(17): 4119-4124. DOI:https://doi.org/10.1016/S0043-1354(03)00398-1

Singh, K.P., Malik, A., Mohan, D., Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water Research, 38(18): 3980-3992. DOI:https://doi.org/10.1016/j.watres.2004.06.011

Singh, K.P., Malik, A., Sinha, S., 2005. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. Analytica Chimica Acta, 538(1): 355-374. DOI:https://doi.org/10.1016/j.aca.2005.02.006

Sklash, M., Farvolden, R., Fritz, P., 1976. A conceptual model of watershed response to rainfall, developed through the use of oxygen-18 as a natural tracer. Canadian Journal of Earth Sciences, 13(2): 271-283.

Sklash, M.G., Beven, K.J., Gilman, K., Darling, W.G., 1996. Isotope studies of pipeflow at Plynlimon, Wales, UK. Hydrological Processes, 10(7): 921-944. DOI:10.1002/(sici)1099-1085(199607)10:7<921::aid-hyp347>3.0.co;2-b

Sklash, M.G., Farvolden, R.N., 1979. The role of groundwater in storm runoff. Journal of Hydrology, 43(1): 45-65. DOI:https://doi.org/10.1016/0022-1694(79)90164-1

Snyder, F.F., 1939. A conception of runoff-phenomena. Eos, Transactions American Geophysical Union, 20(4): 725-738.

Stewart, M., Cimino, J., Ross, M., 2007. Calibration of base flow separation methods with streamflow conductivity. Groundwater, 45(1): 17-27.

Tallaksen, L., 1995. A review of baseflow recession analysis. Journal of Hydrology, 165(1-4): 349-370.

Tallec, G., Ansart, P., Guérin, A., Derlet, N., Pourette, N., Guenne, A., Delaigue, O., Boudhraa, H., Loumagne, C., 2013. Introduction. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 11-33.

Thirel, G., Perrin, C., 2013. Impact du changement climatique sur le bassin de l'Orgeval à l'horizon 2050. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 209-308.

Thiry, M., Simon-Coinçon, R., 1996. Tertiary paleoweatherings and silcretes in the southern Paris Basin. CATENA, 26(1): 1-26. DOI:https://doi.org/10.1016/0341-8162(95)00044-5

Thompson, S., Basu, N., Lascurain, J., Aubeneau, A., Rao, P., 2011. Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients. Water Resources Research, 47(10).

Tournebize, J., Garnier, J., Billen, G., Vilain, G., Passy, P., Tallec, G., Billy, C., Mercier, B., Ansart, P., Sebilo, M., Kao, C., 2013. Transfert de nitrates et transformations dans un petit bassin versnat de la seine. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 145-157.

Trincal, L., 1994. Recensement des superficies drainées d'un bassin versant agricole à l'aide d'un SIG, 60 pp.

Tubau, I., Vàzquez-Suñé, E., Jurado, A., Carrera, J., 2014. Using EMMA and MIX analysis to assess mixing ratios and to identify hydrochemical reactions in groundwater. Science of The Total Environment, 470-471: 1120-1131. DOI:https://doi.org/10.1016/j.scitotenv.2013.10.121

Tunqui Neira, J.M., Andréassian, V., Tallec, G., Mouchel, J.M., 2019. A two-sided affine power scaling relationship to represent the concentration–discharge relationship. Hydrol. Earth Syst. Sci. Discuss., 2019: 1-15. DOI:10.5194/hess-2019-550

Uhlenbrook, S., Hoeg, S., 2003. Quantifying uncertainties in tracer-based hydrograph separations: a case study for two-, three- and five-component hydrograph separations in a mountainous catchment. Hydrological Processes, 17(2): 431-453.

Vázquez-Suñé, E., Carrera, J., Tubau, I., Sánchez-Vila, X., Soler, A., 2010. An approach to identify urban groundwater recharge. Hydrol. Earth Syst. Sci., 14(10): 2085-2097. DOI:10.5194/hess-14-2085-2010

Vega, M., Pardo, R., Barrado, E., Debán, L., 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Research, 32(12): 3581-3592. DOI:https://doi.org/10.1016/S0043-1354(98)00138-9

Vilain, G., Garnier, J., Tallec, G., Cellier, P., 2010. Effect of slope position and land use on nitrous oxide (N2O) emissions (Seine Basin, France). Agricultural and Forest Meteorology, 150(9): 1192-1202.

Vilain, G., Garnier, J., Tallec, G., Tournebize, J., 2012. Indirect N2O emissions from shallow groundwater in an agricultural catchment (Seine Basin, France). Biogeochemistry, 111(1-3): 253-271. DOI:10.1007/s10533-011-9642-7

von Freyberg, J., Studer, B., Kirchner, J.W., 2017. A lab in the field: high-frequency analysis of water quality and stable isotopes in stream water and precipitation. Hydrology and Earth System Sciences, 21: 1721-1739.

von Freyberg, J., Studer, B., Rinderer, M., Kirchner, J.W., 2018. Studying catchment storm response using event- and pre-event-water volumes as fractions of precipitation rather than discharge. Hydrol. Earth Syst. Sci., 22(11): 5847-5865. DOI:10.5194/hess-22-5847-2018

Williams, G.P., 1989. Sediment concentration versus water discharge during single hydrologic events in rivers. Journal of Hydrology, 111(1): 89-106.

Zanotti, C., Rotiroti, M., Fumagalli, L., Stefania, G.A., Canonaco, F., Stefenelli, G., Prévôt, A.S.H., Leoni, B., Bonomi, T., 2019. Groundwater and surface water quality characterization through positive matrix factorization combined with GIS approach. Water Research, 159: 122-134. DOI:https://doi.org/10.1016/j.watres.2019.04.058

Zhang, J., Zhang, Y., Song, J., Cheng, L., 2017. Evaluating relative merits of four baseflow separation methods in Eastern Australia. Journal of Hydrology, 549: 252-263.

Zhang, Q., 2018. Synthesis of nutrient and sediment export patterns in the Chesapeake Bay watershed: Complex and non-stationary concentration-discharge relationships. Science of The Total Environment, 618: 1268-1283. DOI:https://doi.org/10.1016/j.scitotenv.2017.09.221

Zhang, Q., Harman, C.J., Ball, W.P., 2016. An improved method for interpretation of riverine concentration-discharge relationships indicates long-term shifts in reservoir sediment trapping. Geophysical Research Letters, 43(19): 10215-10224. DOI:10.1002/2016gl069945

Zhang, R., Li, Q., Chow, T.L., Li, S., Danielescu, S., 2013. Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method. Hydrological Processes, 27(18): 2659-2665.

Zimmer, M.A., Pellerin, B., Burns, D.A., Petrochenkov, G., 2019. Temporal Variability in Nitrate-Discharge Relationships in Large Rivers as Revealed by High-Frequency Data. Water Resources Research, 55(2): 973-989. DOI:10.1029/2018wr023478

Part II

Revisiting the concentration-discharge (C-Q) relationships

Chapter 1: Revisiting the one-component models – power law model

Avant-propos

In this chapter, we analyse one of the most used C-Q relationships: the one-sided power scaling relationship, commonly known as “power-law”, and its performance using high-frequency data. We noticed that this relationship can be improved using a new approach developed especially for this continuous measurements. We named this new relationship a “two-sided affine power scaling” relationship.

.

Technical Note: A two-sided affine power scaling relationship to represent the concentration–discharge relationship

José Manuel Tunqui Neira[1, 2], Vazken Andréassian[1], Gaëlle Tallec[1] & Jean-Marie Mouchel[2]

(1) Irstea, HYCAR Research Unit, Antony, France

(2) Sorbonne Université, CNRS, EPHE, UMR Metis 7619, Paris, France

Under review to the HESS journal: https://doi.org/10.5194/hess-2019-550, in review, 2019.

Abstract

This technical note deals with the mathematical representation of concentration–discharge relationships. We propose a two-sided affine power scaling relationship (2S-APS) as an alternative to the classic one-sided power scaling relationship (commonly known as “power law”). We also discuss the identification of the parameters of the proposed relationship, using an appropriate numerical criterion. The application of 2S-APS to the high-frequency chemical time series of the Oracle-Orgeval observatory is presented (in calibration and validation mode): It yields better results for several solutes and for electrical conductivity in comparison with the power law relationship.

Keywords

Concentration–discharge relationships; log–log space; power law, high-frequency chemical data

1 Introduction

The relationship between solute concentrations and river discharge (from now on “C-Q relationship”) is an age-old topic in hydrology (see among others Durum, 1953; Hem, 1948; Lenz and Sawyer, 1944). It would be impossible to list here all the articles that have addressed this subject, and we refer our readers to the most recent reviews (e.g. Bieroza et al., 2018; Botter et al., 2019; Moatar et al., 2017) for an updated view of the ongoing research on C-Q relationships.

Many complex models have been proposed to represent C-Q relationships, from the tracer mass balance (e.g. Minaudo et al., 2019) to the multiple regression methods (e.g. Hirsch et al., 2010). Nonetheless, for the past 50 years the simple mathematical formalism known as “power law” has

enjoyed lasting popularity among hydrologists and hydrochemists (see e.g. Edwards, 1973; Gunnerson, 1967; Hall, 1970; Hall, 1971). Over the years, however, some shortcomings of this relationship have become apparent: Recently, Minaudo et al. (2019) mentioned that, "fitting a single linear regression on C-Q plots is sometimes questionable due to large dispersion in C-Q plots (even log transformed)". Also, Moatar et al. (2017) present an extensive typology of shapes (in log–log space) for the French national water quality database, which shows that the power law must be modified to represent the C-Q relationship for dissolved components as well as for particulate-bound elements.

This technical note presents a two-sided affine power scaling relationship (named "2S-APS") that can be seen as a generalization of the power law. And although we do not wish to claim that it can be universally applicable, we argue here that it allows for a better description and modeling of the C-Q relationship of some solutes as a natural extension of the power law.

2 Tested dataset

We used the half-hourly (every 30 min) hydrochemical dataset collected by the in situ *River Lab* laboratory at the Oracle-Orgeval observatory (Floury et al., 2017; Tallec et al., 2015). A short description of the study site is given in Appendix 1. We used dissolved concentrations of three ions – sodium [Na^+], sulfate [$S\text{-}SO_4^{2-}$], and chloride [Cl^-] – as well as electrical conductivity (EC). This dataset was collected from June 2015 to March 2018, averaging 20,700 measurement points.

As our main objective in this note is to compare the performance of two relationships (the new 2S-APS and the classic power law), we divided our dataset into two parts to perform a split-sample test (Klemeš, 1986): We used June 2015 to July 2017 for calibration (of both relationships), and August 2017 to March 2018 for validation. Table 8 presents the main characteristics of both periods.

Table 8: Summary of high-frequency dissolved concentrations and electrical conductivity (EC; average, minimum, maximum values and ratio between quantiles 90 and 10 divided by the mean) from the River Lab at the Oracle-Orgeval observatory, divided into two groups: June 2015 to July 2017 (calibration period) and August 2017 to March 2018 (validation period).

Solute	Unit	Calibration period (June 2015 to July 2017)			
		Mean (μ)	*Min*	*Max*	*$(q_{90}-q_{10})/\mu$*
Sodium	$mg.L^{-1}$	13	2	17	0.22
Sulfate	$Smg.L^{-1}$	19	2	32	0.44
Chloride	$mg.L^{-1}$	30	4	40	0.28
EC	$\mu S.cm^{-1}$	704	267	1015	0.23
		Validation period (August 2017 to March 2018)			
Sodium	$mg.L^{-1}$	13	3	17	0.59
Sulfate	$Smg.L^{-1}$	18	3	26	0.70
Chloride	$mg.L^{-1}$	29	4	40	0.71
EC	$\mu S.cm^{-1}$	576	171	813	0.65

3 Mathematical formulations

3.1 Classic one-sided power scaling relationship (power law)

Since at least 50 years ago, a one-sided power scaling relationship (commonly known as power law) has been used to represent and model the relationship between solute concentration (C) and discharge (Q) (Eq. (13)).

$$C = aQ^b \quad \textbf{Eq. (13)}$$

From a numerical point of view, the relationship presented in Eq. (13) is generally adjusted by first transforming the dependent (C) and independent (Q) variables using a logarithmic transformation, and then adjusting a linear model (Eq. (14)).

$$\ln(C) = \ln(a) + b.\ln(Q) \quad \textbf{Eq. (14)}$$

Graphically, this is equivalent to plotting concentration and discharge in a log–log space, where parameters a and b can be identified either graphically or numerically, under the assumptions of linear regression.

3.2 Limits of the power law

In many cases, the power law appears visually adequate (and conceptually simple), which explains its lasting popularity. With the advent of high-frequency measuring devices in recent years, the size of the

datasets has exploded, and the C-Q relationship can now be analyzed on a wider span (Kirchner et al., 2004). Figure 27 shows an example from our own high-frequency dataset: the 17,500 data points (which correspond to the calibration period of Table 8) represent half-hourly measurements collected over a 2-year period, during which the catchment was exposed to a variety of high- and low-flow events, thus providing a great opportunity for exploring the shape of the C-Q relationship. This being said, we do not wish to imply that a similar behavior could not been identified in medium- and low-frequency datasets, which remain essential tools with which to analyze and understand long-term hydrochemical processes (e.g. Godsey et al., 2009; Moatar et al., 2017).

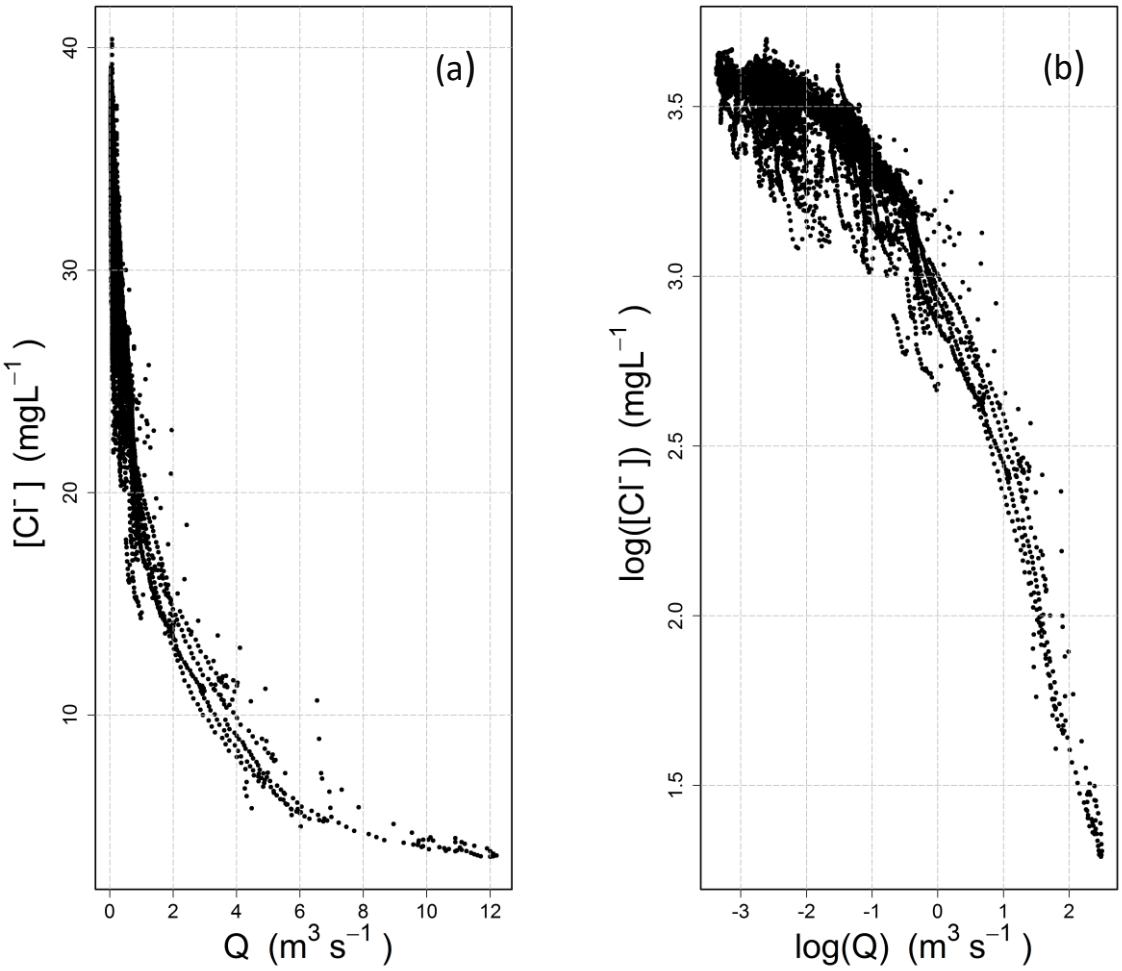


Figure 27: Concentration–discharge relationship observed at the Oracle-Orgeval observatory (measurements from the River Lab) for chloride ions [Cl^-]: (a) standard axes, (b) logarithmic axes.

Figure 27 illustrates the inadequateness of the power law for this dataset: The C-Q relationship evolves from a well-defined concave shape on the left to a slightly convex shape on the right in the log–log space. From the point of view of a modeler wishing to adjust a linear model, one has gone beyond the straight shape that was aimed at. Note that this is true for our dataset, and that it does not need to

always be the case: The log–log space can be well adapted in some situations (see examples in the paper by Moatar et al., 2017).

3.3 A two-sided affine power scaling relationship as a progressive alternative to the power law

As a progressive alternative to the one-sided power scaling relationship (power law), we propose to use a two-sided affine power scaling (2S-APS) relationship as shown in Eq. (15) (Box and Cox, 1964; Howarth and Earle, 1979).

$$C^{\frac{1}{n}} = a + bQ^{\frac{1}{n}} \qquad \textbf{Eq. (15)}$$

From a numerical point of view, the relationship presented in Eq. (15) is equivalent to first transforming the dependent (C) and independent (Q) variables using a so-called Box–Cox transformation (Box and Cox, 1964), and then adjusting a linear model. In comparison with the logarithmic transformation, the additional degree of freedom offered by *n* allows for a range of transformations, from the untransformed variable (*n* = 1) to the logarithmic transformation (*n* → ∞). This "progressive" property was underlined long ago by Box and Cox (1964): When *n* takes high values, Eq. (15) converges toward the one-sided power scaling relationship (power law) (Eq. (13)). The reason is simple:

$C^{\frac{1}{n}} = e^{\frac{1}{n}lnC} \approx 1 + \frac{1}{n}lnC$ when *n* is large.

Thus, for large values of *n*, Eq. (15) can be written as:

$$1 + \frac{1}{n}lnC \approx a + b + \frac{b}{n}lnQ$$

That is equivalent to:

$\ln C \approx A + b.\ln Q$ (with $A = n(a + b - 1)$)

The progressive behavior and the convergence toward the log–log space are clearly evident in Figure 28.

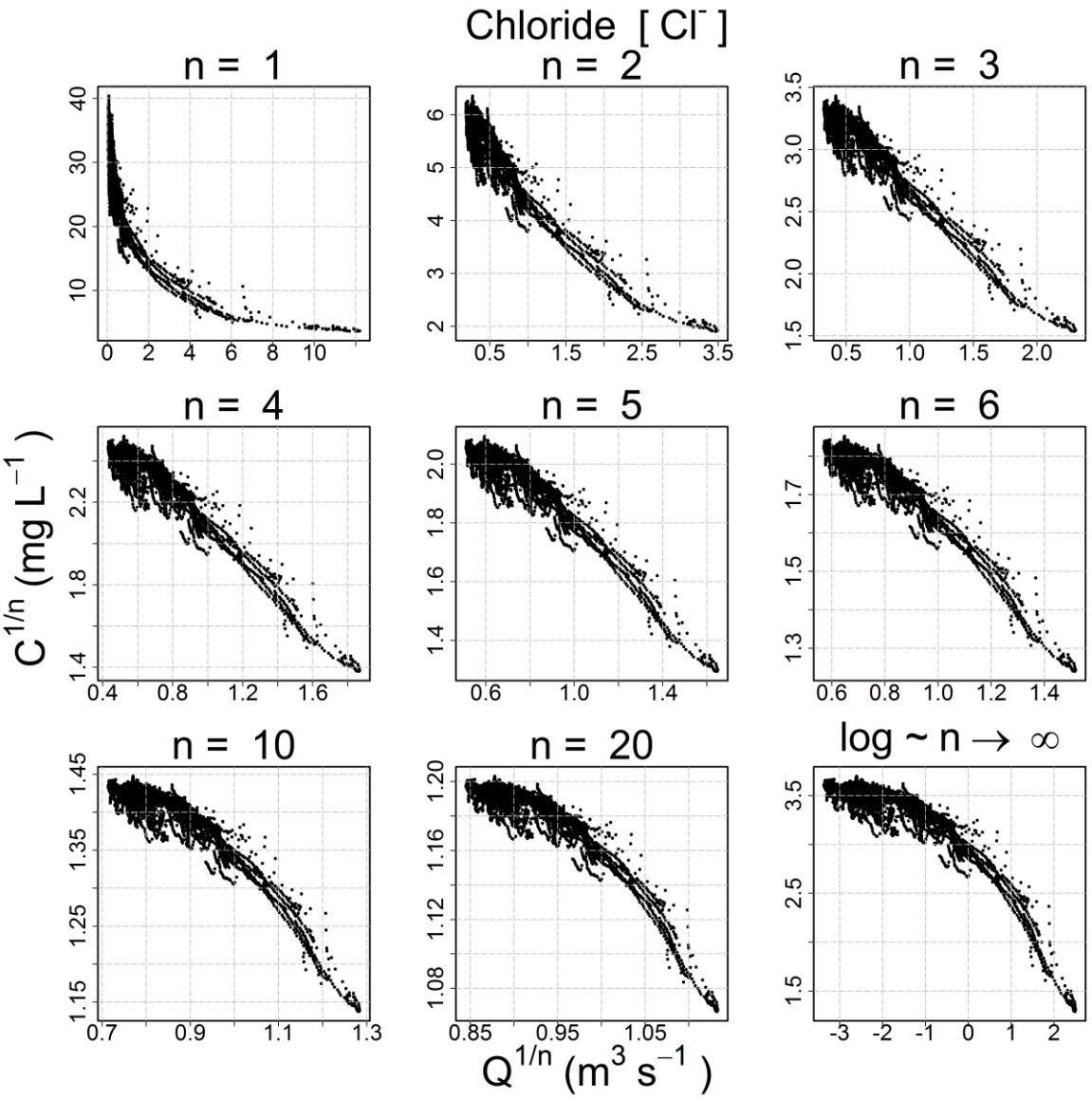


Figure 28: Evolution of the shape of the concentration–discharge scatterplot for chloride ion with two-sided affine power scaling (2S-APS) and an increasing value of parameter *n*.

3.4 Choosing an appropriate transformation for different ion species (calibration mode)

Because the hydro-biogeochemical processes that control the transport and reaction of ions are different, different ionic species may have a C-Q relationship of distinct shape (Moatar et al., 2017). In Figure 29, we show the behavior of three ions and the EC from the same catchment and the same dataset (all four from the Oracle-Orgeval observatory) with different transformations (*n* =1, 3, 5 and logarithmic transformation). The optimal shape was chosen numerically: We transformed our data

series of C and Q using different values of n (i.e., $C^* = C^{1/n}$ and $Q^* = Q^{1/n}$) and logarithmic transformation (i.e., $C^{**} = \log(C)$ and $Q^{**} = \log(Q)$). With these transformed values, we performed a linear regression and computed parameter a and b and the coefficient of determination (R^2) (see Table 9). The n considered as optimal has the highest R^2 value (see Table 9). However, we could also have followed the advice of Box et al. (2016, p. 331) and done it visually (Figure 29).

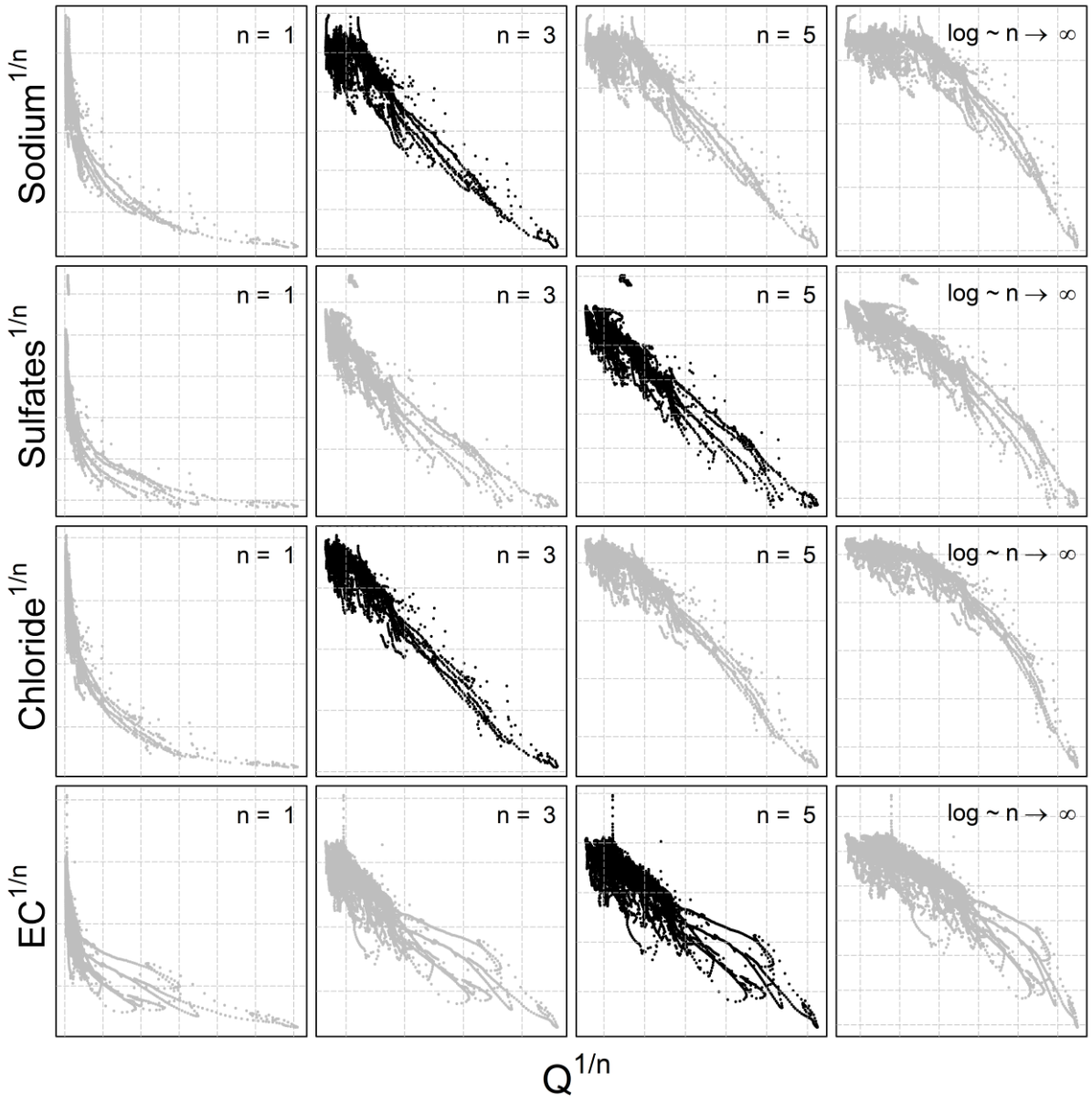


Figure 29: C-Q behavior of three different chemical species and the electrical conductivity with different 2S-APS transformations (*n* =1, 3, 5, and log). The optimal power parameter (black dots) was chosen based on the R^2 criterion. Note that we have removed the scale on the axes to focus only on the change in shape in the C-Q relationship.

Table 9: Coefficient of determination (R^2) calculated for n =1 (no transformation), n = optimal value for two-sided affine power scaling relationship (Figure 29) and n → ∞ (log–log space) for each ion and for electrical conductivity (EC). Note that the R^2 is computed from transformed values.

Solute	n	R^2
Sodium	n = 1 (no transformation)	0.53
	n = 3 (optimal)	0.73
	$n \rightarrow \infty$ (log–log)	0.53
Sulfate	n = 1 (no transformation)	0.32
	n = 5 (optimal)	0.81
	$n \rightarrow \infty$ (log–log)	0.77
Chloride	n = 1 (no transformation)	0.52
	n = 3 (optimal)	0.88
	$n \rightarrow \infty$ (log–log)	0.69
EC	n = 1 (no transformation)	0.38
	n = 5 (optimal)	0.79
	$n \rightarrow \infty$ (log–log)	0.74

The results given in Table 9 show the better quality of the fit obtained with the optimal value of n.

4 Numerical identification of the parameters for the 2S-APS relationship

The extremely large number of values in this high-frequency dataset may cause problems for a robust identification over the full range of discharges using a simple linear regression. Indeed, the largest discharge values are in small numbers (in our dataset only 1% of discharges are in the range [2.6 m^3s^{-1}, 12.2 m^3s^{-1}], and they correspond to the lowest concentrations (see Figure 27)).

To address this question, we successively tested a large number of (a,b) pairs (n remaining fixed at the optimal value given in Table 9). Each pair yields a series of simulated concentrations (C_{sim}) that can be compared with the observed concentrations (C_{obs}). Among the many numerical criteria that could be used, we chose the bounded version of the Nash and Sutcliffe (1970) efficiency criterion *NSEB* (Mathevet et al., 2006), which is commonly used in hydrological modeling. *NSEB* can be computed on concentrations or on discharge-weighted concentrations (which corresponds to the load). We chose the average of both, because we found that it allows more weight to be given to the extremely low concentrations and thus to avoid the issue of under-representation of high-discharge/low-concentration measurement points. Table 10 presents the formula for these numerical criteria.

We retained as optimal the pair of (a,b) that yielded the highest $NSEB_{comb}$ value (we explored in a systematic fashion the range [1–5] for a and [-1.2–1.2] for b).

Table 10: Numerical criteria used for optimization (C_{obs} – observed concentration, C_{sim} – simulated concentration, Q – observed discharge). The Nash and Sutcliffe (1970) efficiency (NSE) criterion is well known and widely used in the field of hydrology. The rescaling proposed by Mathevet et al. (2006) transforms NSE into NSEB, which varies between -1 and 1 (its optimal value). The advantage of this rescaled version is to avoid the occurrence of large negative values (the original NSE criterion varies in the range [-∞, 1]).

$NSE_{conc} = 1 - \frac{\sum_t \left(C_{obs}^t - C_{sim}^t\right)^2}{\sum_t (C_{obs}^t - \overline{C_{obs}})^2}$	**Eq. (16)**
$NSEB_{conc} = \frac{NSE_{conc}}{2 - NSE_{conc}}$	**Eq. (17)**
$NSE_{load} = 1 - \frac{\sum_t \left(Q^t C_{obs}^t - Q^t C_{sim}^t\right)^2}{\sum_t (Q^t C_{obs}^t - \overline{QC_{obs}})^2}$	**Eq. (18)**
$NSEB_{load} = \frac{NSE_{load}}{2 - NSE_{load}}$	**Eq. (19)**
$NSEB_{comb} = \frac{1}{2}(NSEB_{conc} + NSEB_{load})$	**Eq. (20)**

5 Results

5.1 Results in calibration mode

The optimal values of a and b corresponding to the simulation of each ion and EC with the highest $NSEB_{comb}$ criterion and the *n* value identified in Figure 29 and Table 9 are presented in Table 11.

Table 11: Summary of values a, b, and n used to obtain the optimal $NSEB_{comb}$ criterion.

Ion	*n*	*a*	*b*	$NSEB_{comb}$
Sodium	3	2.70	-0.60	0.68
Sulfate	5	2.20	-0.55	0.69
Chloride	3	3.70	-1.00	0.83
EC	5	4.20	-0.70	0.77

For comparing the two relationships, we used the RMSE criterion. The results are shown in Table 12; they illustrate (for our catchment) the better performance (i.e., lower RMSE value) of the proposed 2S-APS relationship for the three ions (sodium, sulfate, and chloride) over the power law relationship. For EC, there is a slight advantage over the power law. A test of the equality of variance (F-test) was

performed between the RMSE obtained for the two relationships: Because of the very large number of points in our dataset, all differences were highly significant (*p*-value <0.001).

Table 12: Summary of values of RMSE criterion calculated for the three ions and EC.

Solute	*2S-APS*	*Power law*
	RMSE	***RMSE***
Sodium	1.00 mgL^{-1}	1.22 mgL^{-1}
Sulfate	2.17 mgL^{-1}	2.22 mgL^{-1}
Chloride	2.00 mgL^{-1}	2.91 mgL^{-1}
EC	42.0 $\mu S.cm^{-1}$	41.3 $\mu S.cm^{-1}$

Figure 30 illustrates the comparison of the quality of simulation over the entire calibration dataset between the power law and 2S-APS relationships. In general, the two-sided affine power scaling relationship yields better simulated concentrations than the classic power law relationship for the three ions (according to the results of Table 12). This is particularly evident over the low concentrations. This better performance is more apparent in the case of sodium and chloride ions.

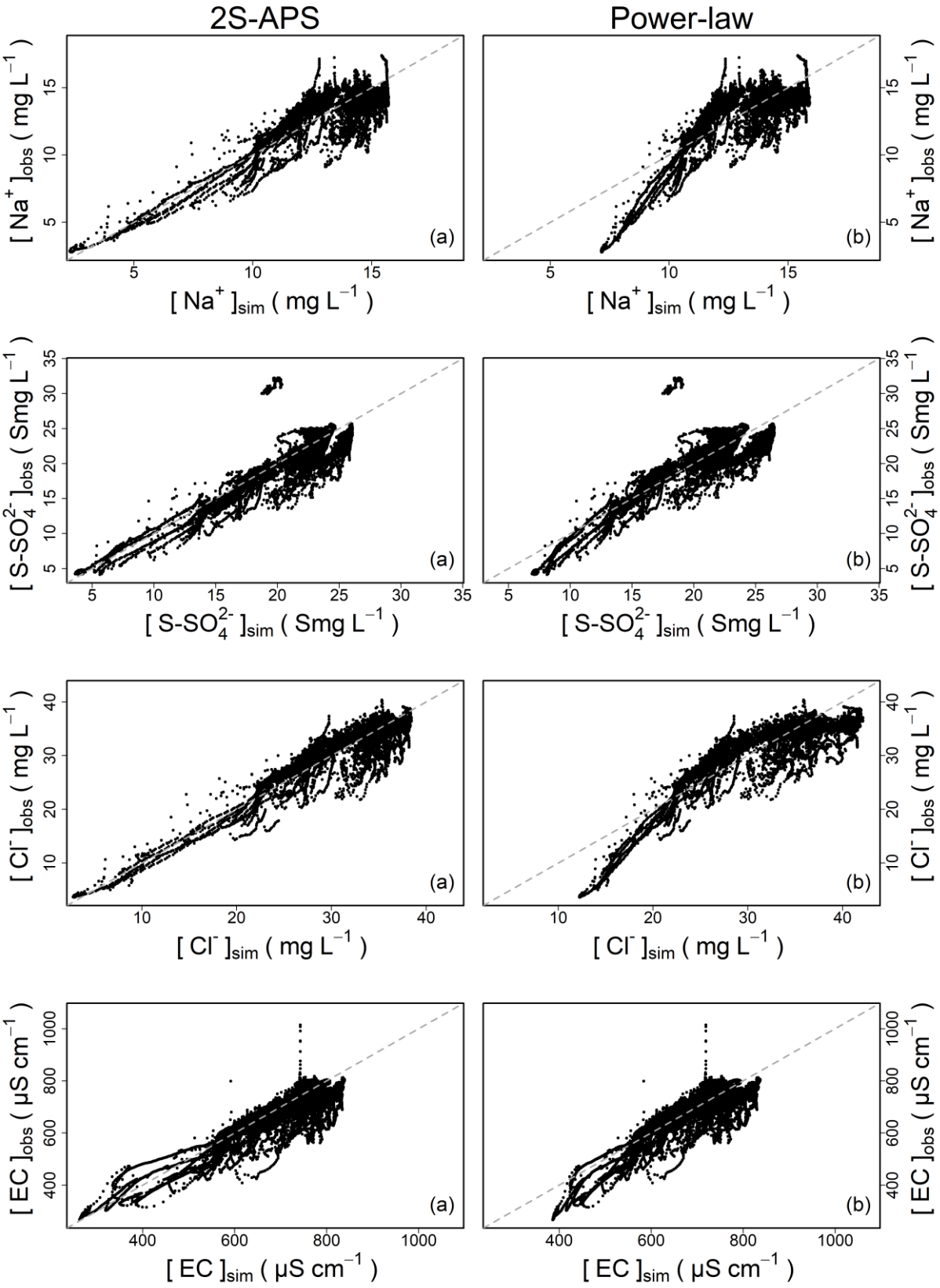


Figure 30: Comparison of simulated concentrations with observed concentrations for: (a) two-sided affine power scaling (2S-APS) relationship, (b) power law (calibration mode).

5.2 Results in validation mode

For the validation mode, we applied the above-calibrated relationships to a different time period (August 2017 to March 2018). The results are shown in Table 13. The RMSE criterion illustrates (for our catchment) the better performance of the proposed 2S-APS relationship over the power law relationship for all the solutes. Unlike the calibration case, the quality of the simulation of EC using the 2S-APS relationship has a much better performance than the one simulated by the power law relationship.

Table 13: Summary of values of RMSE criterion calculated for the three ions and EC with the validation dataset.

Solute	*2S-APS*	*Power law*
	RMSE	***RMSE***
Sodium	1.48 mgL^{-1}	1.90 mgL^{-1}
Sulfate	1.65 mgL^{-1}	2.33 mgL^{-1}
Chloride	3.69 mgL^{-1}	4.34 mgL^{-1}
EC	62.3 $\mu S.cm^{-1}$	78.8 $\mu S.cm^{-1}$

6 Conclusion

In this technical note, we tested and validated a three-parameter relationship (2S-APS) as an alternative to the classic two-parameter one-sided power scaling relationship (commonly known as “power law”), to represent the concentration–discharge relationship. We also proposed a way to calibrate the 2S-APS relationship.

Our results (in calibration and validation mode) show that the 2S-APS relationship can be a valid alternative to the power law: In our dataset, the concentrations simulated for sodium, sulfate, and chloride as well as the EC are significantly better in validation mode, with a reduction in RMSE ranging between 15 and 26%.

Data availability. Data will be available in a dedicated database website after a contract accepted on behalf of all institutes.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. The first author acknowledges the Peruvian Scholarship Cienciactiva of CONCYTEC for supporting his PhD study at Irstea and the Sorbonne University. The authors acknowledge the EQUIPEX CRITEX program (grant no. ANR-11-EQPX-0011) for the data availability. We thank François Bourgin for his kind review.

7 Appendix 1 - Description of the River Lab

In June 2015, the “River Lab” was deployed on the bank of the Avenelles River (within the limits of the Oracle-Orgeval observatory, see Figure 31) to measure the concentration of all major dissolved species at high-frequency (Floury et al., 2017). The River Lab's concept is to "permanently” install a series of laboratory instruments in the field in a confined bungalow next to the river. River Lab performs a complete analysis every 30 min using two Dionex® ICS-2100 ionic chromatography (IC) systems by continuous sampling and filtration of stream water. River Lab measures the concentration of all major dissolved species ([Mg^{2+}], [K^{+}], [Ca^{2+}], [Na^{+}], [Sr^{2+}], [F^{-}], [SO_4^{2-}] [NO_3^{-}], [Cl^{-}], [PO_4^{3-}]). In addition, a set of physico-chemical probes is deployed to measure pH, conductivity, dissolved O_2, dissolved organic carbon (DOC), turbidity, and temperature. The discharge is measured continuously via a gauging station located at the River Lab site.

All the technical qualities, calibration of the equipment, comparison with laboratory measurements, degree of accuracy, etc. have been well described in a publication by Floury et al. (2017).

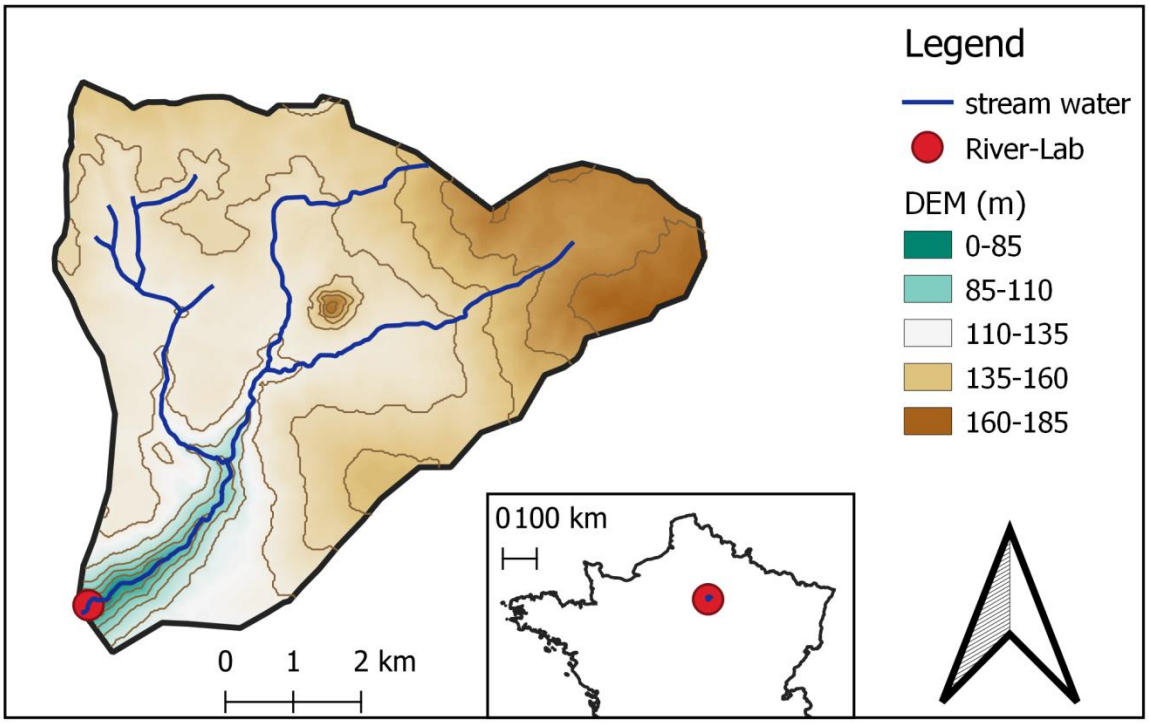


Figure 31: Location of the River Lab (red dot) on the Avenelles river, Oracle-Orgeval observatory.

8 Appendix 2 - Graphical representation of the numerical identification of parameters

The identification of the parameters of our model 2S-APS can be illustrated graphically through a Pareto front (Figure 32). Pareto front serves to show graphically the best compromise between two criteria: the criterion focusing on concentration ($NSEB_{conc}$) and the criterion focusing on load ($NSEB_{load}$)

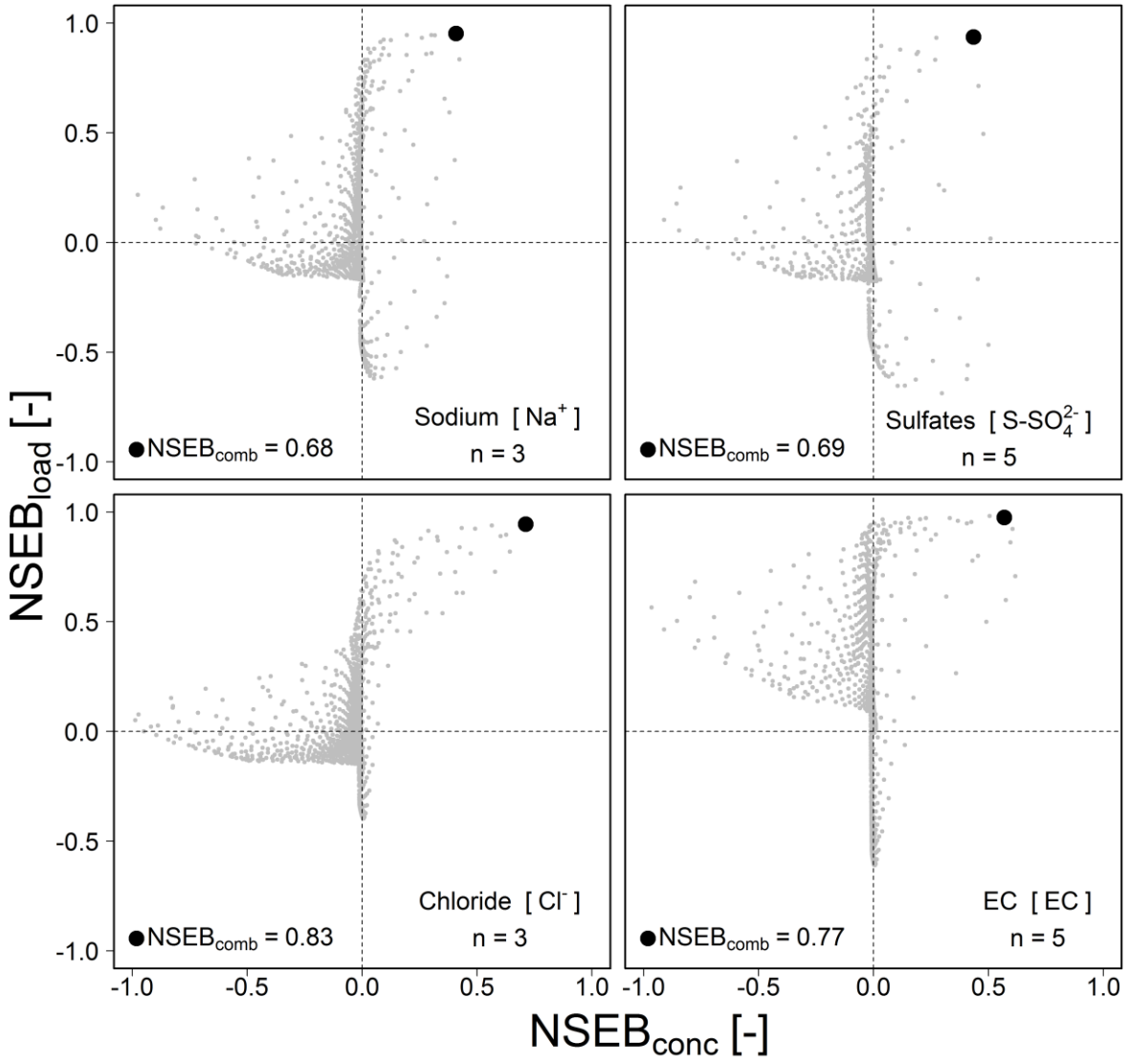


Figure 32 : Pareto plot obtained from 3 chemical species (Sodium, Sulfates, Chloride) and Electrical conductivity (EC) from the Oracle-Orgeval observatory. The black points represent the *NSEB$_{comb}$* criterion.

Figure 32 shows that not all ion species and EC have the same behavior: Chloride has a clear common optimum for the concentration and the load, while the others do not.

9 References

Bieroza, M.Z., Heathwaite, A.L., Bechmann, M., Kyllmar, K., Jordan, P., 2018. The concentration-discharge slope as a tool for water quality management. Science of The Total Environment, 630: 738-749. DOI:https://doi.org/10.1016/j.scitotenv.2018.02.256

Botter, M., Burlando, P., Fatichi, S., 2019. Anthropogenic and catchment characteristic signatures in the water quality of Swiss rivers: a quantitative assessment. Hydrol. Earth Syst. Sci., 23(4): 1885-1904. DOI:10.5194/hess-23-1885-2019

Box, G.E., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2): 211-243.

Durum, W.H., 1953. Relationship of the mineral constituents in solution to stream flow, Saline River near Russell, Kansas. Eos, Transactions American Geophysical Union, 34(3): 435-442. DOI:10.1029/TR034i003p00435

Edwards, A.M.C., 1973. The variation of dissolved constituents with discharge in some Norfolk rivers. Journal of Hydrology, 18(3): 219-242. DOI:https://doi.org/10.1016/0022-1694(73)90049-8

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., Roubaty, J.L., 2017. The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data. Hydrol. Earth Syst. Sci., 21(12): 6153-6165.

Godsey, S.E., Kirchner, J.W., Clow, D.W., 2009. Concentration-discharge relationships reflect chemostatic characteristics of US catchments. Hydrological Processes, 23(13): 1844-1864. DOI:10.1002/hyp.7315

Gunnerson, C.G., 1967. Streamflow and quality in the Columbia River basin. Journal of the Sanitary Engineering Division, 93(6): 1-16.

Hall, F.R., 1970. Dissolved solids-discharge relationships .1. Mixing models. Water Resources Research, 6(3): 845-&. DOI:10.1029/WR006i003p00845

Hall, F.R., 1971. Dissolved solids-discharge relationships .2. Applications to field data. Water Resources Research, 7(3): 591-&. DOI:10.1029/WR007i003p00591

Hem, J.D., 1948. Fluctuations in concentration of dissolved solids of some southwestern streams. Eos, Transactions American Geophysical Union, 29(1): 80-84. DOI:10.1029/TR029i001p00080

Hirsch, R.M., Moyer, D.L., Archfield, S.A., 2010. Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs1. JAWRA Journal of the American Water Resources Association, 46(5): 857-880. DOI:10.1111/j.1752-1688.2010.00482.x

Howarth, R., Earle, S., 1979. Application of a generalized power transformation to geochemical data. Journal of the International Association for Mathematical Geology, 11(1): 45-62.

Kirchner, J.W., Feng, X., Neal, C., Robson, A.J., 2004. The fine structure of water-quality dynamics: the (high-frequency) wave of the future. Hydrological Processes, 18(7): 1353-1359.

Klemeš, V., 1986. Dilettantism in Hydrology: transition or destiny? Water Resour. Res., 22(9): 177S-188S.

Lenz, A., Sawyer, C.N., 1944. Estimation of stream-flow from alkalinity-determinations. Eos, Transactions American Geophysical Union, 25(6): 1005-1011. DOI:10.1029/TR025i006p01005

Mathevet, T., Michel, C., Andreassian, V., Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. IAHS PUBLICATION, 307: 211.

Minaudo, C., Dupas, R., Gascuel-Odoux, C., Roubeix, V., Danis, P.-A., Moatar, F., 2019. Seasonal and event-based concentration-discharge relationships to identify catchment controls on nutrient export regimes. Advances in Water Resources, 131: 103379. DOI:https://doi.org/10.1016/j.advwatres.2019.103379

Moatar, F., Abbott, B., Minaudo, C., Curie, F., Pinay, G., 2017. Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. Water Resources Research, 53(2): 1270-1287.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology, 10(3): 282-290.

Tallec, G., Ansard, P., Guérin, A., Delaigue, O., Blanchouin, A., 2015. Observatoire Oracle [Data set]. DOI:https://dx.doi.org/10.17180/obs.oracle

Chapter 2: Revisiting the n-component models – mixing model

Avant-propos

The hydrograph separation, a widely used method in hydrology, is considered as poorly justified. Can the new availability of high-frequency measurements on a large number of dissolved ions change this situation? In this chapter, we try to calibrate two hydrological methods using the concept of mass balance, with the help of several ions concentrations.

Hydrograph separation issue using high-frequency chemical measurements

José Manuel Tunqui Neira1, 2, Vazken Andréassian1, Gaëlle Tallec1 & Jean-Marie Mouchel2

(1) Irstea, HYCAR Research Unit, Antony, France

(2) Sorbonne Université, CNRS, EPHE, UMR Metis 7619, Paris, France

Works presented to :
European Geosciences Union (EGU) general assembly 2018 (Austria) in poster format.
Computational Methods in Water Resources (CMWR) conference 2018 (France)in oral presentation.
TERENO International Conference 2018 (Germany) in poster format.

1 Introduction

1.1 Hydrograph separation: an age-old issue

Hydrograph separation and the identification of the baseflow contribution to streamflow is perhaps one of the oldest unsolved problems of hydrology (Hall, 1968; Nathan and McMahon, 1990). Discussions on hydrograph separation can be traced back to early studies of groundwater flow (Boussinesq, 1877; Horton, 1933; Maillet, 1905), and since then, many methods have been proposed to quantify the baseflow contributions to streamflow (see among others Dickinson et al., 1967; Hall, 1968; Nathan and McMahon, 1990; Smakhtin, 2001; Tallaksen, 1995).

Despite a quest spanning over more than a century, the existing hydrograph separation methods still have many limitations, which have been studied in detail since at least the 1950s (e.g. Chow, 1964a section 1.10). In **Table 14** we have collected a few statements from prominent hydrologists concerning the inconsistencies of hydrograph separation and/or of recession analysis methods. We believe that these citations will convince the reader that the only certitude concerning hydrograph separation is … that there is no good separation procedure!

Table 14: A collection of statements by prominent hydrologists concerning the inconsistences of hydrograph separation/recession analysis

Sentences	Source	Commentary
"Since there is no real basis for distinguish between the direct and groundwater flow in stream at any instant and since the definitions of these two components are relatively arbitrary, the method of [baseflow] separation is usually equally arbitrary"	Linsley Jr et al. (1958). Hydrology for engineers. p. 156	A famous textbook by the forefathers of hydrological modelling
"The separation of the different components of the hydrograph (runoff and baseflow) can only be approximate since there is no simple experimental method to identify the origin and the previous path of flows that reach the outlet", [thus the design and identification of components (runoff and baseflow) in the hydrograph should be] "left to the discretion of the operator"	Réménérias (1960). L'Hydrologie de l'Ingénieur. pp. 333-337	Classical textbook in French
[The different methods of recession analysis]: "It is practical enough to treat certain problems, but the justifications given for the form of these laws are not demonstrable and are based on too simplistic physical hypotheses"	Roche (1963). Hydrologie de Surface. p. 268	Classical textbook in French
"In practical hydrograph analysis the baseflow separation is made usually in an arbitrary manner, and it is not significant even to consider the exact amount (which is unknown anyway)"	Chow (1964b). Handbook of Applied Hydrology. S. 14.9 & 14.10	The most famous of all hydrologists, once considered the "bible" of the hydrologists
"The diversity of recession analysis formulas show that they are not the expression of rigorous laws, but mathematical expressions of the empirical data"	Castany (1967). Traité pratique des eaux souterraines. p. 529	Classical textbook in French
[Among baseflow separation methods] "there is no way of determining at present which method is most applicable"	Dickinson et al. (1967). An experimental rainfall-runoff facility. p. 26	A study of almost all graphical hydrograph separation methods existing up to that date
[The hydrograph separation is] "one of the most desperate analysis technique in use in hydrology	Hewlett and Hibbert (1967). Factors affecting the response of small watersheds to precipitations in humid areas. p. 276	A classical paper of the literature

"Although much of the basic mathematics and many of the fundamental concepts for base-flow recessions have been available for a long time, in actual practice the status of the art does not seem very far advanced"	Hall (1968). Base-flow recession: a review. p. 980	A classical review of recession analysis
[The separation of the baseflow from the streamflow was a] "fascinating arena of fancy and speculation"	Appleby (1970). Recession and the baseflow problem. p. 1398	
"Standard analytical procedures of hydrograph separation […] are nothing more than hydrologic epicycles […] with hydrologic names"	Klemeš (1986). Dilettantism in hydrology. p. 177	A famous paper on hydrologic misconceptions
"the term baseflow itself is more of a conceptual convenience than a precise description of the nature of the source, or degree of attenuation, of subsurface flow"	Nathan and McMahon (1990). Evaluation of automated techniques for baseflow and recession analysis. p. 1472	An famous paper on hydrograph separation
[Hydrograph separation methods] "are poor tools for estimating ground water discharge or discharge when the major assumptions of the methods are commonly and grossly violated" [Baseflow] "frequently is not equivalent to groundwater discharge because other hydrologic phenomena can significantly affect stream-discharge during recession periods."	Halford and Mayer (2000). Problems associated with estimating ground water discharge and recharge from stream-discharge records. p. 340	An article on the inconsistencies of the hydrograph separation and recession analysis
" the best method of dealing with hydrograph separation is to avoid it all together"	Beven (2001). Rainfall – Runoff Modeling. The Primer. p. 32	
"The [baseflow] separations methods all suffer the disadvantage of being arbitrary and somewhat inaccurate. At present, baseflow separation is more an art than a science"	Bedient et al. (2012). Hydrology and floodplain analysis. p. 67	

1.2 A variety of solutions proposed for hydrograph separation

Although all hydrograph separation methods are criticizable, this procedure is a necessary part of many hydrologic studies: determination of catchment water budget (Stewart et al., 2007), calibration of hydrological models (Yu and Schwartz, 1999; Zhang et al., 2013), management of water resources (Koskelo et al., 2012), etc. This is the reason why hydrologists have been actively looking for alternatives to improve it. In the late 1970s, Lyne and Hollick (1979) proposed the Recursive Digital Filter method (RDF), adapted from the signal-processing theory. In the RDF method, baseflow is considered as a low-

frequency signal, and surface runoff is considered as a high-frequency signal: by filtering out the high-frequency signal, the low-frequency signal (baseflow) can be revealed (Nathan and McMahon, 1990; Zhang et al., 2013). Since 1979, several RDF methods have been developed (a detailed review of existing RDF methods for baseflow analysis can be found in Brodie et al., 2007). Presently, the RDF is one of the most used methods for hydrograph separation because it is objective, computationally efficient, easily automated and applied to long continuous streamflow records (Chapman, 1991; Eckhardt, 2005). Nevertheless, the RDF method requires the calibration of at least one parameter (called the filter parameter), which main objective is to define the speed of streamflow recession, and it is common in practice to adapt the constant of the filter to the hydrological recession time constant of the catchment, a process which also carries its share of uncertainties (Brutsaert, 2008; Cheng et al., 2016).

According to Nathan and McMahon (1990) recession analysis consists in "combining individual flow recessions in such a way as to provide an average characterization of baseflow response": the resulting construction is usually termed Master Recession Curve (MRC). Early methods (that are still in use) to calculate the MRC were proposed by Langbein (1938) and Snyder (1939). These methods present difficulties because of problems in interpreting the hydrograph, particularly in regard to determining initial discharge and eliminating effects of occasional rises due to rainfall during the recession period (Hall, 1968). Despite multiple attempts to improve these methods (e.g. Brutsaert and Nieber, 1977; Cheng et al., 2016; Lamb and Beven, 1997) their application still presents many speculative and arbitrary aspects (Brutsaert and Nieber, 1977; Cheng et al., 2016).

Some authors have investigated the hydrograph separation from chemistry (La Sala Jr., 1967; Pinder and Jones, 1969). They assume that stream water results from the mixture of several sources, and that each source has a constant and unique chemical composition (La Sala Jr., 1967; Longobardi et al., 2016; Uhlenbrook and Hoeg, 2003). Generally, the Mass Balance method (MB) uses an approach based on a *n*-component mixing model with chemical conservative tracers (Dewalle et al., 1988; Ladouche et al., 2001; Uhlenbrook and Hoeg, 2003). These tracers, as environmental isotopes (such as tritium, Oxygen-18 and deuterium), ions (such as Chloride, Silicate, Sodium) or electro conductivity (EC), can be used separately or together (Klaus and McDonnell, 2013).

The MB method is based on five assumptions (according to Buttle, 1994; Klaus and McDonnell, 2013): (i) the isotopic/tracer content of the event and the pre-event water are significantly different, (ii) the event water maintains a constant isotopic/tracer signature in space and time, (iii) the isotopic/tracer signature of the pre-event water is constant in space and time, (iv) contributions from the vadose zone is negligible, or the isotopic/tracer signature of the soil water is similar to that of groundwater, (v) surface storage contributes minimally to streamflow. To our knowledge, Pinder and Jones (1969) and Hubert et al. (1969) were the first to apply the concept of MB method. Pinder and Jones (1969) used

the total sum of various solutes (bicarbonate, chloride, nitrate, sulfate, silicate, calcium, iron, magnesium, potassium and sodium) to separate the storm hydrograph into a direct flow and groundwater flow component. Hubert et al. (1969) did a similar separation using only one isotopic tracer, the tritium. Since 1969, more than 200 papers have been written about the hydrograph separation based on the MB method (Klaus and McDonnell, 2013). Because this method relies on *in situ* measurements, it brings a physical bond to the hydrograph separation (Lott and Stewart, 2016; Stewart et al., 2007). Nevertheless, the MB method frequently violates some of the five underlying assumptions, and more particularly that of conservation even with isotopes (Buttle, 1994; Klaus and McDonnell, 2013). We should also mention a lack of detailed studies of the uncertainties (Bansah and Ali, 2017; Genereux, 1998). The others disadvantages of the MB method are its high cost in terms of fieldwork and the difficulty of its application in large catchments or for long-term studies (Stewart et al., 2007).

1.3 Joining the strengths of hydrological and chemical approaches for hydrograph separation

Several recent studies have attempted to combine the RDF and MB hydrograph separation approaches. Both methods have been combined by calibrating the RDF method with chemical data through the MB principle (Longobardi et al., 2016). All of these studies show promising results for this coupled method, which combines the RDF method, easy to implement and applicable over the long term, and the MB method which provides information on physical processes at catchment scale (Saraiva Okello et al., 2018; Stewart et al., 2007; Zhang et al., 2013). In almost all this studies the electro-conductivity (EC) was the only chemical descriptor used, because it can be inexpensively measured concurrently with streamflow measurements (Lott and Stewart, 2016; Stewart et al., 2007)

1.4 Using high-frequency chemical measurements to revisit the hydrograph separation issue

For this chapter, we considered that the simplest way to solve the hydrograph separation issue is to work with the RDF-MB coupling. Unlike the previously discussed studies, we had at our disposal high-frequency chemical measurements for several stream water ions over more than two years (Floury et al., 2017). The great advantage of this rather long series of high-frequency chemical measurements is that it captures a wide range of hydrological conditions.

In this chapter we aim to use jointly the RDF and MB methods in order to identify the RDF model parameter leading to the most realistic (most stable) MB parameters. Because we are conscious of the

limits of the hydrograph separation methods (see **Table 14**) and because we want to propose a rather generic methodology, we will test two RDF methods and six ions concentrations plus the electro-conductivity.

2 Methodology

2.1 Study site and data set

The Avenelles catchment is a small low-relief catchment of 46 km^2, located 70 km east of Paris, France. It is a sub-catchment of the ORACLE-Orgeval observatory (https://gisoracle.irstea.fr/). The Orgeval is affected by a temperate oceanic climate (annual average temperature is about 11 ± 1 °C and mean annual rainfall about 674 ± 31 mm) (Tallec et al., 2013). The average streamflow at the Avenelles outlet is about 0.2 m^3/s, with minimum flows in the summer (~0.1 m^3/s) and floods up to 10 m^3/s in the spring (1962-2017, Tallec et al., 2015). With respect to geology, the catchment is formed entirely by limestone rocks, with two aquifers: the shallower aquifer of the Brie limestone, and the deeper Champigny limestone aquifer (Mouhri et al., 2013). The upper layer is constituted by loess deposit (up to 10 m thick), enriched in clay and sand. These silty leached brown soils are characterized by low permeability resulting in temporary waterlogged soil in winter. Nearly 60% of the surface of the catchment is drained with man-made buried drains. Land use is agricultural with intensive farming practices, mainly based on mineral nitrogen fertilization (Garnier et al., 2016).

Since 1962, streamflow is measured continuously by a gauging station at the Avenelles outlet and stored in the French national database Banque Hydro (http://hydro.eaufrance.fr/). We used two hydrological datasets: one was at daily time step (flow from January 1991 and December 2017) for the recession analysis part. For RDF-MB method application, the second data set covers the period and time step of the chemical measurement, i.e. between June 2015 and July 2017 at half-hourly time step. The chemical concentrations are *in situ* measured, continuously, at the Avenelles outlet, through the "River Lab" (Floury et al., 2017) (Table 15). We also used concentrations measured in the framework of the PIREN Seine program in bank piezometers, drains, deep piezometers (plateau), and springs from the ORACLE-Orgeval observatory (Table 15). These measurements were made from January 2016 to April 2018 (see Mouchel et al., 2016; Mouhri et al., 2013).

Table 15: Chemical concentrations (median and standard deviation) measured in hydrological compartments in the framework of the PIREN Seine program (January 2016-April 2018) and at the Avenelles outlet by the River Lab (June 2015 - July 2017), from the ORACLE-Orgeval observatory. Rainfall concentrations were taken from (Floury et al., 2018).

Solute	unit	stream (River Lab)	bank	drains	plateau	springs	Rainfall (Floury et al., 2018)
		median ± sd					
Chloride	mg L^{-1}	30 ± 5	33 ± 12	32 ± 7	35 ± 8	34 ± 5	1.60
Sulfate	mgS L^{-1}	19 ± 4	11 ± 26	10 ± 2	16 ± 50	18 ± 8	0.38
Magnesium	mg L^{-1}	9 ± 1	7 ± 6	6 ± 1	6 ± 2	8 ± 1	0.19
Sodium	mg L^{-1}	13 ± 2	12 ± 3	15 ± 3	18 ± 5	12 ± 4	0.75
Nitrate	mgNL^{-1}	12 ± 1	4 ± 5	14 ± 3	13 ± 4	15 ± 3	0.75
Calcium	mg L^{-1}	124 ± 13	133 ± 31	67 ± 7	129 ± 60	145 ± 11	1.56
EC	µS cm^{-1}	704 ± 76	711 ± 152	465 ± 77	742 ± 204	794 ± 62	NA

2.2 Hydrograph separation methods

RDF methods are used to partition the streamflow into two components, direct runoff and baseflow. Long-slow waves are associated with baseflow while the high-frequency variability of the streamflow is caused by direct runoff. We will use a common notation which may differ from the initial notation: Q_b is the slow response (baseflow), Q_q the quick response (quickflow or direct runoff) and Q the total flow ($Q = Q_b + Q_q$).

We compare two RDF methods: the Lyne-Hollick (LH) and Eckhardt filters (ECK). The two methods share a common parameter (α_τ) which can be calibrated by adapting it to the hydrological recession time constant of the catchment (τ).

2.2.1 Lyne – Hollick method (LH)

In this method proposed by Lyne and Hollick (1979), the quickflow $Q_{q(t)}$ is defined as follows:

$$Q_{q(t)} = \alpha_\tau Q_{q(t-1)} + \frac{1+\alpha_\tau}{2}\left(Q_{(t)} - Q_{(t-1)}\right) \quad \textbf{Eq. (21)}$$

An iterative passage of the filter allows smoothing data and nullifying phase distortion (Longobardi et al., 2016; Nathan and McMahon, 1990). Different number of passages can be found in the literature (e.g. Ladson et al., 2013; Longobardi et al., 2016; Zhang et al., 2017). We use the 3 passages approach (forward – backward – forward) proposed by Nathan and McMahon (1990).

The Lyne – Hollick algorithm (Eq. (21)) can be rewritten according to Q_b :

$$Q_{b(t+1)} = \min\left(\alpha_\tau\, Q_{b(t)} + \frac{1-\alpha_\tau}{2}\left(Q_{(t+1)} + Q_{(t)}\right), Q_{(t)}\right) \quad \textbf{Eq. (22)}$$

Note that, during long lasting dry periods, when the quickflow can be assumed zero, the filter reduces to a constant baseflow as $Q_{b(t+1)} = Q_{b(t)}$. Then, for the LH method, the physical meaning of the parameter α_τ is only a recession of the quickflow and the steady-state solution is $Q_q = 0$. From the LH method, Chapman and Maxwell (1996), who pointed out that the LH method incorrectly provides a constant flow or baseflow, respectively when runoff has ceased, proposed a new method, with $Q_b = Q/2$ as the steady-state limit. However, Eckhardt (2005) showed that filters are special cases of a two-parameter filter. In addition to the recession constant, he introduced a second parameter, the maximum value of the baseflow index (BFI_{max}), which refers to the mean groundwater recharge in the catchment. Eckhardt (2005) proposed a two-parameter method, better describing the fraction of water discharged by the slow flow over long periods (Eckhardt, 2005; Longobardi and Loon, 2018).

2.2.2 Eckhardt method (ECK)

Eckhardt algorithm (2005) assumes a linear relation between the outflow from the aquifer and its storage, but with an exponential baseflow recession during periods without groundwater recharge. Parameters A and B in Eq. (24) are expressed as functions of the filter parameter α_τ and a second parameter BFI_{max} (the long-term ratio of baseflow to the total streamflow).

The baseflow is written as:

$$Q_{b(t)} = \min\left(A.\, Q_{b(t-1)} + B.\, Q_{(t)}, Q_{(t)}\right) \quad \textbf{Eq. (23)}$$

$$\begin{cases} A = \left(\dfrac{1 - BFI_{max}}{1 - \alpha_\tau BFI_{max}}\right).\alpha_\tau \\ B = \dfrac{(1-\alpha_\tau).\, BFI_{max}}{1 - \alpha_\tau BFI_{max}} \end{cases} \quad \textbf{Eq. (24)}$$

It can be rewritten as:

$$Q_{b(t+1)} = \min\left(\alpha_\tau Q_{b(t)} + \frac{BFI_{max} - \alpha_\tau\, BFI_{max}}{1 - BFI_{max}}.\, Q_{q(t+1)}, Q_{(t)}\right) \quad \textbf{Eq. (25)}$$

As in LH method, α_τ is defined as the filter parameter of the baseflow during periods when the quickflow is zero. With the Eckhardt two-parameters filter the steady-state limit situation is $Q_b = BFI$, whereas it was $Q_b = 0$ in LH method.

Eckhardt recommends using a BFI_{max} value according to stream and aquifer properties. He suggests a BFI_{max} of 0.80 for a perennial stream with porous aquifers, 0.50 for an ephemeral stream with porous aquifers, and 0.25 for a perennial stream with hard rock aquifers. The BFI_{max}, corresponding to the situation of the Avenelles catchment, is about 0.65. This value was chosen because our catchment, fed

by two aquifers, does not adapt to the Eckhardt recommendations; thus an arbitrary value between 0.5 and 0.8 (suggested for a perennial aquifer, see Eckhardt, 2005; Eckhardt, 2008) was chosen.

2.2.3 Hydrological recession time constant of the catchment

During seasons without groundwater recharge, stream flow may recess exponentially and follows the form:

$$Q_{(t)} = Q_0 . exp^{-\left(\frac{t - t_o}{\tau}\right)} \quad \textbf{Eq. (26)}$$

Where Q_0 is the streamflow at time $t_o$assuming no significant recharge, and $\boldsymbol{\tau}$ is the hydrological recession time constant of the catchment (in time units). When fixed time step are used. The exponential term of Eq. (26) can be replaced by:

$$Q_{(t+\Delta t)} = Q_{(t)} . exp^{(-\frac{\Delta t}{\tau})} = Q_{(t)} . K \quad \textbf{Eq. (27)}$$

Where K is the so-called recession constant of the catchment.

It is a common practice to adapt the RDF parameter (α_τ) to the recession constant of the catchment (K) which depends on the hydrological recession time constant of the catchment (τ). In our study, τ is the parameter of interest, which allows us to compare hydrological methods between them and to evaluate the RDF-MB method. To provide the hydrological recession time constant of the catchment (τ), the Master Recession Curve (MRC) approach is commonly used. It is graphical method applied on the daily discharge time series. Because we use high-frequency measurements (0.5h), we calculate the τ value for an unusual half-hour time step.

2.2.4 MB method

The mass balance method is based on the assumption that baseflow has different chemical characteristics compared with surface runoff due to the different flow paths of these two types of flows (Longobardi et al., 2016). Consequently, total streamflow hydrograph can be separated into different components on the base of the single component concentration (Longobardi et al., 2016). To apply a hydrograph separation with a chemical approach, we used a classical two-component mixing model (Pinder and Jones, 1969). The relationship between quickflow and baseflow components can be expressed in the form of a mass-balance equation as follows:

$$C_{(t)} = C_{bj} \frac{Q_{b(t)}}{Q_{(t)}} + C_{qj} \frac{Q_{q(t)}}{Q_{(t)}} \quad \textbf{Eq. (28)}$$

$C_{(t)}$: total chemical streamflow concentration at time t

$Q_{(t)}$: total streamflow at time t

$Q_{b(t)}$: baseflow at time t (computed using a RDF method with a parameter ατ)

$Q_{q(t)}$: quickflow at time t ($Q_{q(t)}$ = $Q_{(t)}$ - $Q_{b(t)}$)

C_{bj} : Representative daily parameter of the concentration from the baseflow (mg/L)

C_{qj} : Representative daily parameter of the concentration from the quickflow (mg/L).

Eq. (28) can be further simplified as follow:

$$C(t) = \left(C_{bj} - C_{qj}\right)\frac{Q_{b(t)}}{Q_{(t)}} + C_{qj} \quad \textbf{Eq. (29)}$$

2.3 Comparison of the two RDF methods through an hydrological and a chemical calibration

2.3.1 The hydrological recession time constant (τ) from hydrological MRC approach

We first evaluate τ through a classical recession analysis with the hydrological MRC approach. This approach was applied on daily flow data of the Avenelles station from January 1, 2000 to December 31, 2017 as follows: the daily stream flow data of several seasons were overlapped according to the day of the year, starting from the beginning of June, and assuming that the stream flow is under a continuous process of recession for the period June-September (day of the year 153–275). The master recession curve was created by averaging the daily flow during all the seasons, and optimization procedure (exponential regression) was used to find the hydrological recession time constant of the catchment (τ). Because our measurements are at a sub-hourly scale, the values of τ are given in hours and not in days as usually done in the literature.

2.3.2 Sensibility of the RDF methods to the hydrological recession time constant of the catchment (τ)

Thus, we have evaluated the sensibility of the RDF methods to the τ parameter. We tested 200 different τ values; ranging between 24 and 4800 hours and compared the BFI (Base Flow Index) obtained from the two RDF methods. The BFI is the ratio between the sums of the base flows with respect to the total flows (Gustard et al., 1992)

2.3.3 The hydrological recession time constant of the catchment (τ) from MB approach

With the 200 different τ values used for sensibility assessment, we solved the mass balance equation (i.e. we calculated the unknown C_{bj} and C_{qj}, see Eq. (29)). We used a simple linear relationship at daily time step. Because of the high-frequency measurements dataset, we have between 25 to 48 points per day, which should be statistically enough. To considerate the daily linear regression as valid, it must

meet two essentials requirements: (i) the linear regression should be statistically significant (ρ-value < 0.05), and (ii) the value of C_{bj} should be always positive.

When C_{qj} was identified as negative, the linear regression was forced to go through the origin (thus C_{qj} = 0, only if the p-value < 0.05). It can also happen that in all points of the daily record $Q_{b(t)}$ = $Q_{(t)}$ then C_{qj} could not be identified and C_{bj} is the average of $C_{(t)}$.

At the end of this process, we obtained around 250 daily pairs of values of C_{bj} and C_{qj}, for each τ values, for each RDF method, for each ions and EC tested.

Because the MB method is based on the mass conservation principle, we made the hypothesis that the optimal hydrograph separation parameter would be the one minimizing the variability of the series of concentrations. To express the variability of the C_{bj} and C_{qj} time series we used the standard deviation (sd).

Because the separate analysis of C_{bj} and C_{qj} does not necessary yield the same minimum τ value, we used the multi-criterion selection (Pareto front) between the sd values of C_{bj} and C_{qj} to find the optimal τ, for each tested ions and EC and for the two RDF method.

2.4 Comparison between the components C_{bj} and C_{qj} obtained from the optimal hydrograph separation and the field chemical dataset.

To corroborate the validity of our hydrograph separation, we compared the values obtained of C_{bj} and C_{qj} from Eq. (29) to the field data collected at different sample sites of the Oracle-Orgeval observatory (Mouchel et al., 2016). To make this comparison we follow the hypotheses formulated by Pinder and Jones (1969): The component C_{bj} is the representation of the groundwater concentration and C_{qj} is the representation of the inflow and precipitation concentration. Therefore C_{bj} was compared to the field data coming from the plateau and spring zones that were measured from the Brie groundwater (see Table 15) and C_{qj} was compared to the to the field data coming from the drains and banks zone, two representative areas of the inflow.

3 Results and Discussion

3.1 Hydrological approach to calculate the hydrological recession time constant of the catchment (τ)

First, we have calculated the hydrological recession time constant of the catchment (τ) of the Avenelles catchment using the MRC approach. The τ value of our catchment is 768 hours (32 days). According to

the literature, the τ value for catchments less than 100 km^2 should be in the 10 to 45 days range (Brutsaert, 2008; Cheng et al., 2016; Vogel and Kroll, 1992), thus a value of 32 days seems reasonable. Second, the Avenelles τ of 768 hours was applied to the LH and ECK RDF methods. For each method, we calculated a Base Flow Index (BFI), which is the ratio of the base flow to the total flow (Gustard et al., 1992). The two RDF methods (LH and ECK) give comparable BFI estimates, 0.59 and 0.55 for LH and ECK respectively, having only a difference of only 0.04 between both methods, thus we can point out that the computed baseflow by both RDF methods for the whole flow hydrograph (high and low flows between June or 2015 to July 2017) is statistically similar.

According to Gustard et al. (1992), the BFI can be thought of as measuring the proportion of the river's runoff that derives from aquifer storage. Values of BFI range from 0.1 for a very flashy river to nearly unity for a very stable river with a high base flow proportion (Gustard et al., 1992). The BFI calculated with both RDF methods shows us that our stream is mostly fed by groundwater, but it is influenced by seasonal hydrological variations. This can be seen in the record of flows, where we have marked regimes of high flows (January to April) and low flows (June-September)

However, it should be remembered that we cannot observe baseflow itself. From a hydrological point of view, without more quantitative knowledge on the water pathways, the two methods seem to give plausible results.

Although both LH and ECK yield similar baseflow series, there are differences when plotting the baseflow on a flood event (Figure 33). These differences have been observed in other studies where both RDF methods have been analyzed.(e.g. Cartwright et al., 2014; Longobardi et al., 2016; Zhang et al., 2017). This is mainly due to the number of times the filter is passed in the data to smooth the baseflow plot and to minimize the distortion effected produced by the filter itself (Nathan and McMahon, 1990). In the case of the LH method the filter is passed 3 times, unlike the ECK method which is passed only once.

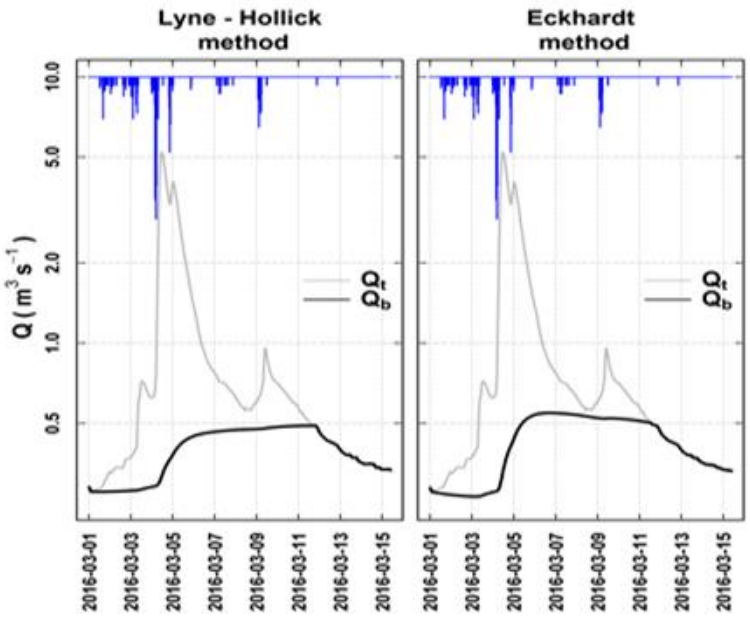


Figure 33 : Comparison of baseflow behavior (black line, Qb) from the 2 RDF methods during a flow peak (grey line, Qt) in the Avenelles catchment using the same τ =768 h.

3.2 Baseflow exploration from RDF methods with hydrological recession time constant (τ) values

To evaluate the sensitivity of the RDF methods to hydrological recession time constant of the catchment, we have tested 200 different τ values, ranging between 24 (1 day) and 4800 hours (200 days). The different BFI values obtained are shown in the Figure 34.

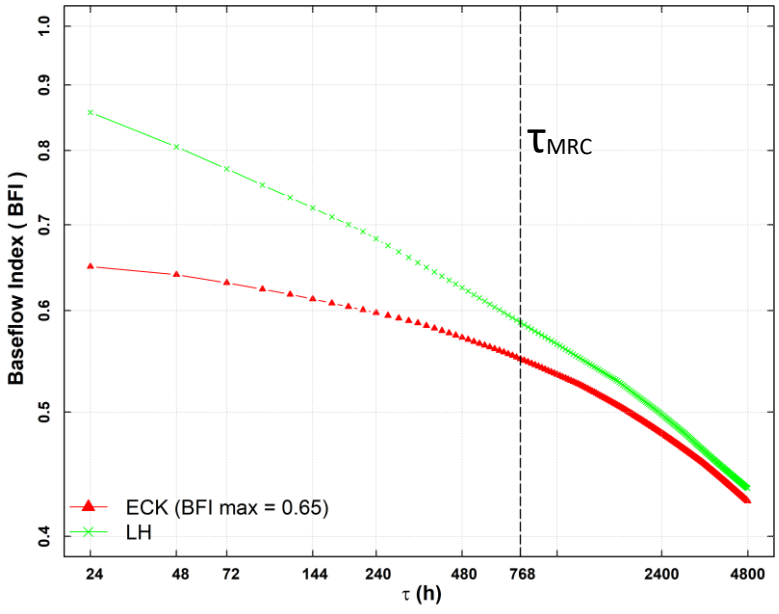


Figure 34 : BFI calculated from 200 different values of τ (ranging from 24 to 4800 hours) for the 2 RDF methods studied. Vertical dotted line represents the τ calculated with MRC approach (768h).

Figure 34 shows that a slight change in the selection of τ can lead to significant differences in calculated baseflow. From 24 to 4800 h, the BFI decreases continuously for both methods, from 0.86 to 0.43 for the LH method and from 0.65 to 0.42 for ECK (Figure 34). Figure 34 also shows that the LH method is more sensitive to the selection of τ than ECK method, but note that ECK was used with its second parameter fixed. This demonstrates that the parameter of the LH method should prioritize the calibration of its parameter instead of using the generic value of 0.925 proposed by Nathan and McMahon (1990), which we must insist was set for Australian catchments and should not be generalized for use for all the catchments where the LH method would be applied.

In τ = 24 h (start of τ sample) we can observe that the difference of BFI between the two methods is significant (BFI of 0.85 for LH against 0.65 for ECK method). As τ increases, this BFI difference tends to decrease. When we arrive at the tau value obtained through the RMC approach (768 h, dotted black line in Figure 34) this difference is already slight (0.04) and continues to decrease until the limit of tau analyzed (τ =4800 h). However they never become equal.

Under a τ of 48 h the BFI for ECK method, is equal to its BFI_{max} parameter (see Figure 34), i.e. 0.65. BFI_{max} represents the maximum groundwater recharge in the catchment (Eckhardt, 2005; Eckhardt, 2008). ECK BFI values vary slightly less than LH method, due to the influence of BFI_{max} that acts as a stabilization parameter. Several studies (e.g. Cartwright et al., 2014; Zhang et al., 2017) have shown that the BFI_{max} parameter has a more decisive influence than the filter parameter (linked to the recession constant; Eq. (27)) on computing the baseflow. Therefore BFI_{max} parameter should also be calibrated (Eckhardt, 2008).

The fundamental problem is that the true value of BFI is unknown (Zhang et al., 2013). Therefore, one cannot say, which one of the methods approximates reality best. Tracer measurements would provide an independent estimate of the baseflow contribution to streamflow.

3.3 MB approach to find the optimal hydrological recession time constant

In this section, we estimate a τ value with the MB method (Eq. (29)) from chemical dataset at the sub-hourly time step (30 min). We use different chemical dataset corresponding to different ions and EC. Table 16 presents the more stable concentrations (median ± sd) obtained for the quickflow and the baseflow and the corresponding optimal τ value. The corresponding BFI calculated from the new calibration of the LH and ECK methods are presented in Table 16.

Table 16 : Results of the Pareto front for LH-MB and ECK-MB methods for each ions and EC: optimal τ values, Median and standard deviation values of C_{bj} and C_{qj} and corresponding BFI values.

RDF-MB methods	Ions and EC	Best compromise (Pareto front)			BFI
		Optimal τ (h)	Median ± sd (%) C_{bj} (mgL^{-1})	Median ± sd (%) C_{qj} (mgL^{-1})	
LH-MB	Chloride	888	34 ± 14 %	30 ± 42 %	0.58
	Sulfate	912	21 ± 16 %	18 ± 42 %	0.57
	Magnesium	888	9 ± 16 %	8 ± 40 %	0.58
	Sodium	888	14 ± 15 %	13 ± 47 %	0.58
	Nitrate	408	12 ± 12 %	12 ± 36 %	0.64
	Calcium	840	125 ± 13 %	123 ± 32 %	0.58
	EC	960	732 ± 10 %	723 ± 38 %	0.57
ECK-MB (BFI_{max} =0.65)	Chloride	648	33 ± 17 %	31 ± 40 %	0.56
	Sulfate	600	21 ± 19 %	18 ± 39 %	0.56
	Magnesium	384	9 ± 13 %	8 ± 38 %	0.58
	Sodium	744	14 ± 19 %	13 ± 44 %	0.55
	Nitrate	408	12 ± 17 %	12 ± 36 %	0.58
	Calcium	528	124 ± 13 %	124 ± 37 %	0.57
	EC	624	720 ± 13 %	725 ± 37 %	0.56

Among the 200 τ values tested, those that give less variability in C_{bj} and C_{qj} averaged 562 ± 23% h and 826 ± 23% h for ECK-MB and LH-MB methods respectively. The values of τ obtained for LH method are in the same range for almost all solutes (except for nitrates) (888 h to 960 h or 37 to 40 days, thus only 3 days of difference). The BFI value is ~0.58 for almost all ions (except for nitrates with a BFI of 0.64) and EC.

Unlike LH method, in the ECK method the τ values obtained are more variable (360 hours or 15 days difference between the τ obtained for magnesium and the τ obtained for sodium ion; see Table 16). However, if we check the BFI values calculated with these τ obtained, they are quite similar (~0.57) for all solutes (see Table 16). This is due to the effect of the second parameter (BFI_{max}) which serves as a stabilizing parameter of the RDF method.

If we compare the calculated average BFI of the two RDF methods (0.58 for LH and 0.57 for ECK respectively, see Table 16) from the MB approach, the difference between the two RDF methods is further shortened than the purely hydrological MRC approach (BFIs of 0.59 for LH and 0.55 for ECK).

Hydrologically, the base flow is defined as the hydrogeological water that feeds the river without precipitation. Should not we have a single τ value representative of a single baseflow rate? Only LH

method seems to allow approximating a single τ value for all ions. The average of τ values for LH method is 896 ± 4 h without nitrate ions (see Table 16).

3.4 Comparison between the components C_{bj} and C_{qj} obtained from the optimal hydrograph separation and the field chemical dataset.

Figure 35 illustrates the cumulative frequency of concentration of the field chemical data representing the groundwater (i.e. measurements coming from the plateau and springs sample zones of the catchment) and the C_{bj} components obtained using the baseflow given (adjusted with the optimal τ from Table 16) by the LH-MB and ECK-MB methods for each ion and EC.

Note that in the case of sulfate and calcium ions measured in the plateau zone, concentrations are extremely variable. This is due to the seasonal phenomena of gypsum dissolution which contains significant concentrations of sulfate and calcium ions (Mouchel et al., 2016).(see Figure 35)

We can observe the cumulative frequencies of the C_{bj} components of both RDF-MB methods (LH and ECK) produce identical performances (without counting the extremes), although the optimal τ (and consequently baseflows) for each RDF method are different (except for the case of nitrates).

From Figure 35, we can also point out the case of chlorides ions: we observe in general a good fit between the computed C_{bj} components for both RDF-MB methods (LH and ECK) and the field concentrations measured in the springs. The same occurs in the case of the EC with the concentration measurements of the plateau zone (except in the range of [0.9 to 1] of the cumulative frequency). For the other ions i.e. magnesium, nitrates and sodium, we do not obtain a good fit as in the case of chloride ions and EC, but we can observe certain approximations between these ions and one of the field concentrations measurements: (i) for magnesium, sodium, sulfates and calcium ions, in both RDF-MB methods, the results obtained for C_{bj} values fit better to the concentrations measured in the springs than those measured in the plateau zone (see Figure 35); (ii) for nitrates, the C_{bj} results of both methods fit better to the concentrations measured in the plateau zone.

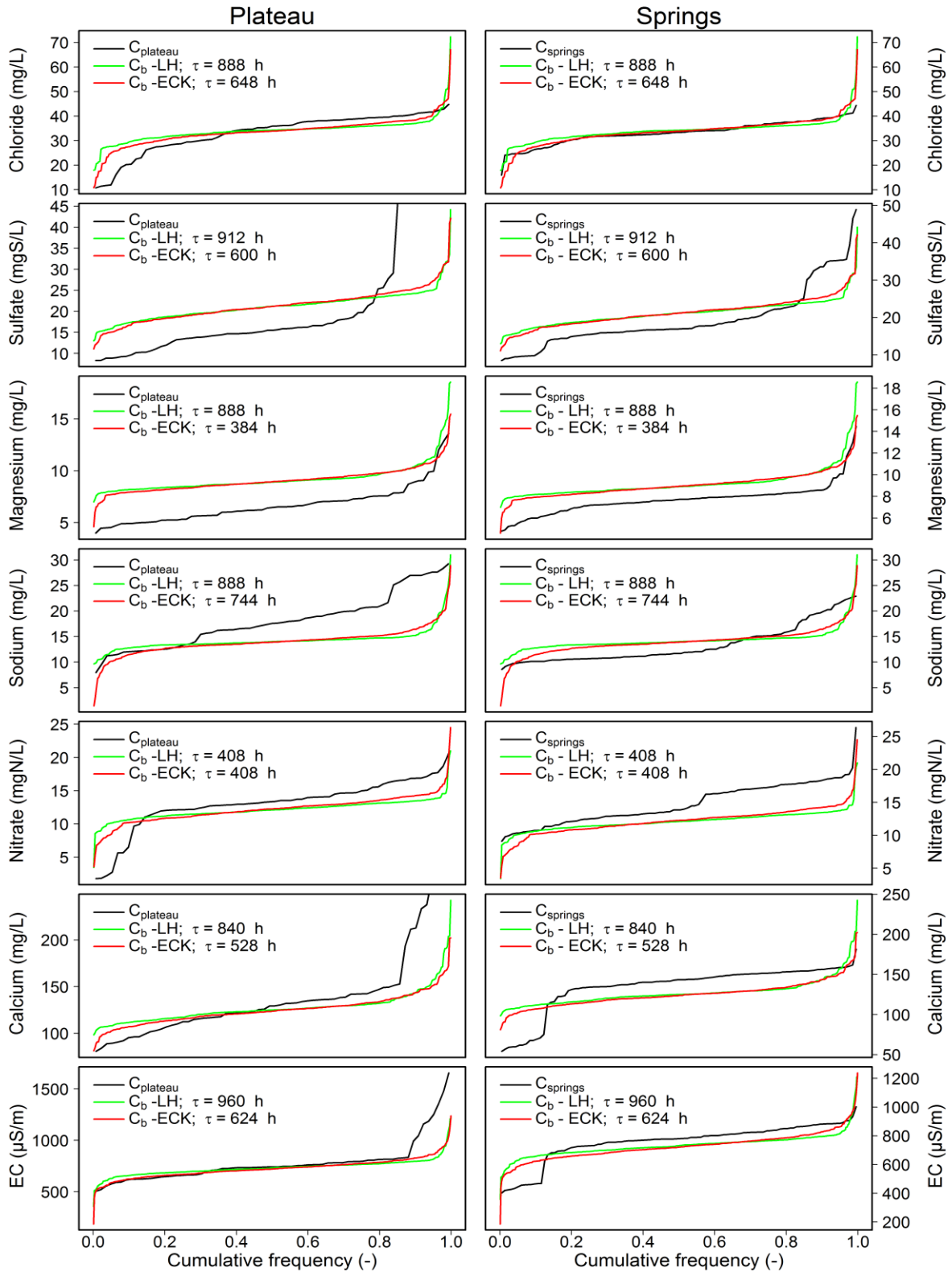


Figure 35 : Cumulative frequencies of concentrations coming from field measurements (i.e. plateau and springs, black line) and the C_{bj} components computed using the optimal baseflow of the LH (green line) and ECK (red line) RDF methods.

Figure 36 illustrates the cumulative frequency of concentration of the field chemical data representing the inflow (i.e. measurements coming from the drains and banks sample zones of the catchment) and the C_{qj} components obtained using the baseflow given (adjusted with the optimal τ from Table 16) by the LH-MB and ECK-MB methods for each ion and EC.

As in the case of C_{bj}, the values of C_{qj} obtained for both methods have almost the same performance (without counting the extreme values indeed).

Once again we can observe that between all the solutes compared, chloride ions and EC remain the two elements that most approximate the performances of the field chemical data (i.e. drains and banks): In the case of chloride ions the C_{qj} values are better fitted to the measurements of concentration of the drains (see Figure 36); and for the EC we can observe an adequate fit between the values of C_{qj} obtained for the two RDF-MB methods (LH and ECK) and the measurements done in the banks.

For the other ions tested, sulfates, magnesium and nitrate ions the ions seem to have no correlation with the data measured in the field (i.e. drains and banks, see Figure 36). The calculated C_{qj} values for the calcium ion have certain adjustment similarities with the concentrations measured in the banks (see Figure 36). Finally for the sodium ion it is difficult to determine with which of the two field measurements it could have a better fit, although in our opinion there is a better correspondence with the concentration measurements of the banks (see Figure 36).

From this cumulative frequency analysis we can indicate that of all the ions and EC, only in the case of the EC and the chloride ion has it been possible to determine an acceptable correlation with the field measurements made in the catchment: (i) for chloride ion C_{bj} with the springs and C_{qj} with the drains; (ii) for EC C_{bj} with the plateau measurements and C_{qj} with the banks. This confirms what other studies indicate (using other approaches e.g. McHale et al., 2002; Rice and Hornberger, 1998; Stewart et al., 2007; Zhang et al., 2013): that chlorides and EC are good tracers for hydrograph separation. For the other ions we can set some similarities with the field chemical data, but these are not robust enough to ensure an acceptable correlation.

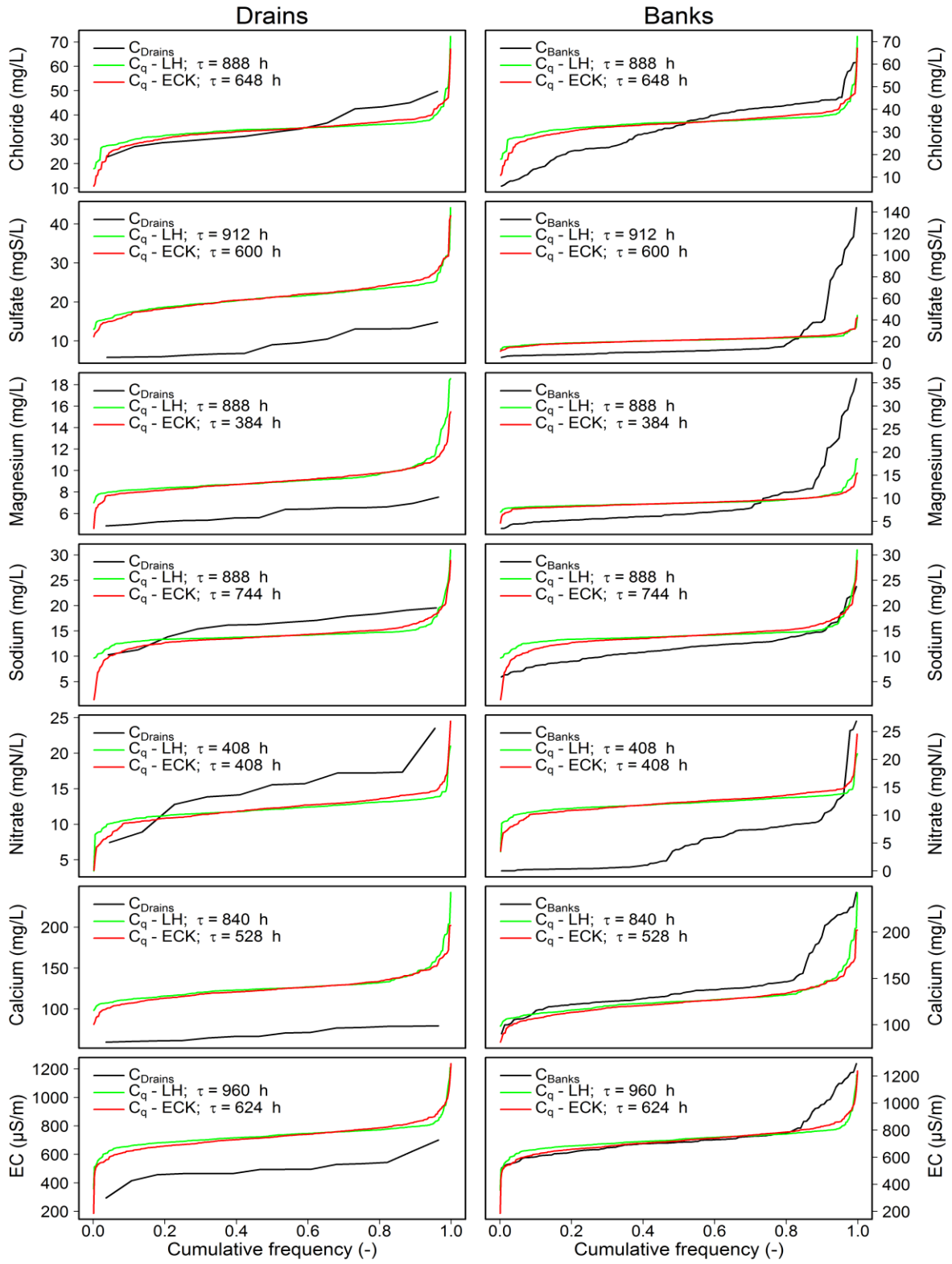


Figure 36 : Cumulative frequencies of concentrations coming from field measurements (i.e. drains and banks, black line) and the C_{qj} components computed using the optimal baseflow of the LH (green line) and ECK (red line) RDF methods.

4 Conclusions

In this chapter we have studied the hydrograph separation, a methodology that has always been present since the establishment of hydrology as a science (Hall, 1968). Although objective, this methodology is based on concepts that are difficult to verify in a physical reality. It is for this reason that we have proposed a new approach, trying to calibrate a hydrological method of hydrograph separation (RDF method) with the help of chemical concentrations, combining them via a mixing equation based on the concept of mass balance (MB).

For this purpose, first we have used two well-known RDF methods: Lyne-Hollick (LH) and Eckhardt (ECK) whose parameter (α_τ) can be adapted from the hydrological recession time constant (τ) of the catchment given by the recession analysis (Master Recession Curve-MRC). The results show that between the two RDF methods the baseflows are similar (BFI of 0.59 for LH and 0.55 for ECK). Variations between the two methods are observed in flood events (see Figure 33). This is produced by the number of times the method parameter is passed over the flow data (3 times for the LH method and one time for the ECK method). RDF methods, for even arbitrary hydrograph separation, work perfectly well for hydrological modeling purposes.

Then we have studied the sensitivity of the two RDF methods with respect to the tau parameter. We have verified that the LH method is more sensitive to this parameter than the ECK method. This is due to the second parameter (BFI_{max}) of the ECK method which serves as a stabilizing parameter, thus ECK is also sensitive to the BFI_{max} value This leads us to conclude that these two RDF methods, to be as efficient as possible, require a calibration on the BFI_{max} parameter and / or the τ value.

Finally we have shown that our simple methodology proposed for the hydrograph separation (RDF – MB coupling approach) could be optimally applied for the chloride ion and EC (after comparison with field chemical data of the catchment, see Figure 35 and Figure 36) for both baseflow methods (LH and ECK). Based on these two solutes, the optimal τ value for our catchment (and therefore the optimal hydrograph separation) would be 900 h for the LH method and 648 h for the ECK method (with a BFI_{max} of 0.65).

The RDF-MB method allows having a good hydrograph separation, with two years of high-frequency chemical measurement, against 26 years for the RMC methods (without possibility of automatization of the method). However, this hydrograph separation occurs with an annual time step. However, our results show that the calibration of the τ value is time step dependent, which have to be taken into account during the RDF-MB coupling method using.

5 References

Appleby, F.V., 1970. Recession and the baseflow problem. Water Resources Research, 6(5): 1398-1403.

Bansah, S., Ali, G., 2017. Evaluating the Effects of Tracer Choice and End-Member Definitions on Hydrograph Separation Results Across Nested, Seasonally Cold Watersheds. Water Resources Research, 53(11): 8851-8871.

Bedient, P.B., Huber, W.C., Vieux, B.E., 2012. Hydrology and floodplain analysis. Pearson, USA, 816 pp.

Beven, K.J., 2001. Rainfall-runoff modeling. The Primer. Willey - Blackwell, UK, 457 pp.

Boussinesq, J., 1877. Essai sur la théorie des eaux courantes, Memoires presentés par divers savants à l'Academic des Sciences de l'Institut National de France, XXIII. Impr. nationale, Paris, 63 pp.

Brodie, R., Sundaram, B., Tottenham, R., Hostetler, S., Ransley, T., 2007. An overview of tools for assessing groundwater-surface water connectivity. Bureau of Rural Sciences, Canberra, Australia: 57-70.

Brutsaert, W., 2008. Long-term groundwater storage trends estimated from streamflow records: Climatic perspective. Water Resources Research, 44(2).

Brutsaert, W., Nieber, J.L., 1977. Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resources Research, 13(3): 637-643.

Buttle, J.M., 1994. Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins. Progress in Physical Geography: Earth and Environment, 18(1): 16-41.

Cartwright, I., Gilfedder, B., Hofmann, H., 2014. Contrasts between estimates of baseflow help discern multiple sources of water contributing to rivers. Hydrology and Earth System Sciences, 18(1): 15-30.

Castany, G., 1967. Traité pratique des eaux souterraines. Dunod, Paris, 661 pp.

Chapman, T., Maxwell, A., 1996. Baseflow separation-comparison of numerical methods with tracer experiments, Hydrology and Water Resources Symposium 1996: Water and the Environment; Preprints of Papers. Institution of Engineers, Australia, pp. 539.

Chapman, T.G., 1991. Comment on “Evaluation of automated techniques for base flow and recession analyses” by R. J. Nathan and T. A. McMahon. Water Resources Research, 27(7): 1783-1784.

Cheng, L., Zhang, L., Brutsaert, W., 2016. Automated selection of pure base flows from regular daily streamflow data: objective algorithm. Journal of Hydrologic Engineering, 21(11): 1-7.

Chow, V.T., 1964a. Hydrology and its development. In: Chow, V.T. (Ed.), Handbook of applied hydrology. McGraw - Hill Book Company, USA, pp. 1.1-1.22.

Chow, V.T., 1964b. Runoff. In: Chow, V.T. (Ed.), Handbook of applied hydrology. McGraw - Hill Book Company, USA, pp. 14.1-14.54.

Dewalle, D.R., Swistock, B.R., Sharpe, W.E., 1988. Three-component tracer model for stormflow on a small Appalachian forested catchment. Journal of Hydrology, 104(1-4): 301-310.

Dickinson, W., Holland, M., Smith, G.L., 1967. An experimental rainfall-runoff facility. Hydrology papers (Colorado State University); no. 25.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes, 19(2): 507-515.

Eckhardt, K., 2008. A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. Journal of Hydrology, 352(1-2): 168-173.

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., Roubaty, J.L., 2017. The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data. Hydrol. Earth Syst. Sci., 21(12): 6153-6165.

Floury, P., Gaillardet, J., Tallec, G., Ansart, P., Bouchez, J., Louvat, P., Gorge, C., 2018. Chemical weathering and CO2 consumption rate in a multilayered-aquifer dominated watershed under intensive farming: The Orgeval Critical Zone Observatory, France. Hydrological Processes: 1-19.

Garnier, J., Anglade, J., Benoit, M., Billen, G., Puech, T., Ramarson, A., Passy, P., Silvestre, M., Lassaletta, L., Trommenschlager, J.M., Schott, C., Tallec, G., 2016. Reconnecting crop and cattle farming to reduce nitrogen losses to river water of an intensive agricultural catchment (Seine basin, France): past, present and future. Environmental Science & Policy, 63: 76-90.

Genereux, D., 1998. Quantifying uncertainty in tracer-based hydrograph separations. Water Resources Research, 34(4): 915-919.

Gustard, A., Bullock, A., Dixon, J.M., 1992. Low flow estimation in the United Kingdom. Report 108. Institute of Hydrology, UK, 88 pp.

Halford, K.J., Mayer, G.C., 2000. Problems associated with estimating ground water discharge and recharge from stream-discharge records. Groundwater, 38(3): 331-342.

Hall, F.R., 1968. Base-flow recessions-a review. Water Resources Research, 4(5): 973-983.

Hewlett, J.D., Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: Sooper, W., Lull, H. (Eds.), Forest hydrology. Pergamon Press, New York, pp. 275-290.

Horton, R.E., 1933. The role of infiltration in the hydrologic cycle. Eos, Transactions American Geophysical Union, 14(1): 446-460.

Hubert, P., Martin, E., Meybeck, M., Oliver, P., Siwertz, E., 1969. Aspects hydrologique, géochimique et sédimentologique de la crue exceptionelle de la Dranse du Chablais du 22 sept. 1968. Arch. Sci. Soc. Phys. (Genève), 22(3): 581-603.

Klaus, J., McDonnell, J., 2013. Hydrograph separation using stable isotopes: Review and evaluation. Journal of Hydrology, 505: 47-64.

Klemeš, V., 1986. Dilettantism in hydrology: Transition or destiny? Water Resources Research, 22(9S): 177-188.

Koskelo, A.I., Fisher, T.R., Utz, R.M., Jordan, T.E., 2012. A new precipitation-based method of baseflow separation and event identification for small watersheds (< 50 km2). Journal of hydrology, 450: 267-278.

La Sala Jr., A.M., 1967. New Approaches to Water-Resources Investigations in Upstate New Yorkaa. Groundwater, 5(4): 6-11. DOI:10.1111/j.1745-6584.1967.tb01619.x

Ladouche, B., Probst, A., Viville, D., Idir, S., Baqué, D., Loubet, M., Probst, J.-L., Bariac, T., 2001. Hydrograph separation using isotopic, chemical and hydrological approaches (Strengbach catchment, France). Journal of Hydrology, 242(3): 255-274.

Ladson, A., Brown, R., Neal, B., Nathan, R., 2013. A standard approach to baseflow separation using the Lyne and Hollick filter. Australasian Journal of Water Resources, 17(1): 25-34.

Lamb, R., Beven, K., 1997. Using interactive recession curve analysis to specify a general catchment storage model. Hydrology and Earth System Sciences, 1(1): 101-113.

Langbein, W., 1938. Some channel-storage studies and their application to the determination of infiltration. Eos, Transactions American Geophysical Union, 19(1): 435-447.

Linsley Jr, R.K., Kohler, M.A., Paulhus, J.L., 1958. Hydrology for engineers. McGraw-Hill Civil Engineering Series. McGraw-Hill Book Company, USA, 340 pp.

Longobardi, A., Loon, A.F.V., 2018. Assessing baseflow index vulnerability to variation in dry spell length for a range of catchment and climate properties. Hydrological Processes, 32(16): 2496-2509. DOI:10.1002/hyp.13147

Longobardi, A., Villani, P., Guida, D., Cuomo, A., 2016. Hydro-geo-chemical streamflow analysis as a support for digital hydrograph filtering in a small, rainfall dominated, sandstone watershed. Journal of Hydrology, 539: 177-187.

Lott, D.A., Stewart, M.T., 2016. Base flow separation: A comparison of analytical and mass balance methods. Journal of Hydrology, 535: 525-533.

Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modelling, Institute of Engineers Australia National Conference, pp. 89-93.

Maillet, E.T., 1905. Essais d'hydraulique souterraine & fluviale. A. Hermann, Paris, 286 pp.

McHale, M.R., McDonnell, J.J., Mitchell, M.J., Cirmo, C.P., 2002. A field-based study of soil water and groundwater nitrate release in an Adirondack forested watershed. Water Resources Research, 38(4): 2-1-2-16. DOI:10.1029/2000wr000102

Mouchel, J.-M., Rocha, S., Rivière, A., Tallec, G., 2016. Caractérisation de la géochimie des interfaces nappe-rivière du bassin des Avenelles. Tech. rep. PIREN Seine, France, 27 pp.

Mouhri, A., Flipo, N., Vitale, Q., Bodet, L., Tallec, G., Ansart, P., Rejiba, F., 2013. Influence du contexte hydrogéologique sur la connectivité nappe - rivière. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 89-98.

Nathan, R., McMahon, T., 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research, 26(7): 1465-1473.

Pinder, G.F., Jones, J.F., 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. Water Resources Research, 5(2): 438-445.

Réménérias, G., 1960. L'hydrologie de l'ingénieur. Collection du Laboratoire National d'Hydraulique. Eyrolles, Paris, 413 pp.

Rice, K.C., Hornberger, G.M., 1998. Comparison of hydrochemical tracers to estimate source contributions to peak flow in a small, forested, headwater catchment. Water Resources Research, 34(7): 1755-1766.

Roche, M., 1963. Hydrologie de surface. Gauthier-Villars Editeur, Paris, 430 pp.

Saraiva Okello, A.M.L., Uhlenbrook, S., Jewitt, G.P., Masih, I., Riddell, E.S., Van der Zaag, P., 2018. Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid mesoscale catchment. Hydrological Processes, 32(10): 1334-1350.

Smakhtin, V.U., 2001. Low flow hydrology: a review. Journal of Hydrology, 240(3-4): 147-186.

Snyder, F.F., 1939. A conception of runoff-phenomena. Eos, Transactions American Geophysical Union, 20(4): 725-738.

Stewart, M., Cimino, J., Ross, M., 2007. Calibration of base flow separation methods with streamflow conductivity. Groundwater, 45(1): 17-27.

Tallaksen, L., 1995. A review of baseflow recession analysis. Journal of Hydrology, 165(1-4): 349-370.

Tallec, G., Ansard, P., Guérin, A., Delaigue, O., Blanchouin, A., 2015. Observatoire Oracle [Data set]. DOI:https://dx.doi.org/10.17180/obs.oracle

Tallec, G., Ansart, P., Guérin, A., Derlet, N., Pourette, N., Guenne, A., Delaigue, O., Boudhraa, H., Loumagne, C., 2013. Introduction. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 11-33.

Uhlenbrook, S., Hoeg, S., 2003. Quantifying uncertainties in tracer-based hydrograph separations: a case study for two-, three- and five-component hydrograph separations in a mountainous catchment. Hydrological Processes, 17(2): 431-453.

Vogel, R.M., Kroll, C.N., 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. Water Resources Research, 28(9): 2451-2458.

Yu, Z., Schwartz, F.W., 1999. Automated calibration applied to watershed-scale flow simulations. Hydrological Processes, 13(2): 191-209.

Zhang, J., Zhang, Y., Song, J., Cheng, L., 2017. Evaluating relative merits of four baseflow separation methods in Eastern Australia. Journal of Hydrology, 549: 252-263.

Zhang, R., Li, Q., Chow, T.L., Li, S., Danielescu, S., 2013. Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method. Hydrological Processes, 27(18): 2659-2665.

Chapter 3: Combining the one- and n-component models

Avant-propos

We have seen in chapter one and two, that the one-component model (i.e. power-law) in terms of quality of fitting, and the mixing approach, in terms of physical interpretability, are efficient to describe the C-Q relationship. We try in this chapter, to combine them in order to obtain an efficient model able to describe the variations of the C-Q relationships as well as the functioning of the catchment.

Combining concentration-discharge relationships with mixing models

José Manuel Tunqui Neira[1, 2], Gaëlle Tallec[1], Vazken Andréassian[1] & Jean-Marie Mouchel[2]
[(1)] Irstea HYCAR Research Unit, Antony, France
[(2)] Sorbonne Université, CNRS, EPHE, UMR METIS 7619, Paris, France

Draft submission to Hydrological Processes journal

Abstract

Discharge is one of the major factors influencing the evolution of solute concentration in river water. Different modelling approaches exist to characterize the dependency of concentration to discharge: the simplest require calibration, they are based on measurable quantities (stream discharge and stream water concentration) but do not allow for an explicit, physical, flow-path interpretation; the more complex are based on mixing assumptions with different end-members sources, but require the knowledge of (unmeasurable) flow components. We present here a combination between the simple concentration-discharge (C-Q) models with the Mass Balance (MB) mixing approach, which we apply to a new high-frequency series on the Oracle-Orgeval observatory (France).

Keywords

Concentration-discharge relationships; power-law relationship; mixing approach

1 Introduction

Because concentration-discharge (C-Q) relationships are interesting for a variety of purposes and potential users, and because they employ simple approaches to describe complex hydro-chemical interactions, hydrologists and geochemists have been using and exploring them for over 70 years (see Chanat et al., 2002; Durum, 1953; Hem, 1948; Johnson et al., 1969; Kirchner, 2019; Moatar et al., 2017). The established relations can be purely empirical, and used to complete lacking concentration data and compute fluxes over long time periods, or to characterize the behavior of chemical components in a number of catchements and empirically derive euristic patterns on a regional and global scale (e.g. Bieroza et al., 2018; Meybeck and Moatar, 2012; Moatar et al., 2017). In some cases, the relations have been derived from a simplified model, describing a physical representation of water circulation and

composition, from various kinds of inputs, with the objective to derive relevant information on its internal functionning (Johnson et al., 1969).

Depending on the type of solute, on the catchment and on the hydro-meteorological conditions, C-Q relationships can either exhibit a dilution effect, a flow-enhancement effect, or no-effect at all (which is referred to chemostasis) (see e.g. Salmon et al., 2001). To model C-Q relationships, two approaches have been developed. The simple one-component models that try to explain the processes controlling the mobilization and delivery of chemical elements into streams (i.e. export regimes) as well as biogeochemical transformations in river networks (Minaudo et al., 2019). The more complex *n*-component models, that besides to explain the previously mentioned processes, have the aim of quantifying the sources of the chemical concentrations measured in the river (Barthold et al., 2011). The simplest C-Q models, that can be of linear or power-law type (see Eq. (30), Eq. (31) and Eq. (32) in Table 17), are based on measurable quantities (discharge and chemical concentration of the stream) (Durum, 1953; Hem, 1948; Tunqui Neira et al., 2019). More complex models, introduced at the end of the 1960s (Hubert et al., 1969; Johnson et al., 1969; Pinder and Jones, 1969), are based on the partition of hydrological sources, assuming constant concentrations in each of the sources and assuming the existence of a methodology to compute the flow issued from each of the sources. This mixing approach is founded on the mass balance equation (see its simplest expression, Eq. (33), Table 17). Hall (1970) gives an exhaustive presentation of these models, their different expressions and assumptions.

The simple one-component model is, first and foremost, based on observations of concentrations relatively stable over time, as electro-conductivity (EC) or chlorides (Durum, 1953; Hem, 1970). The main intention was to identify the contribution of the ground-water to surface waters (Durum, 1953). Despite the first attempts, various new concepts on catchment functioning have enriched the C-Q relations development. Based on some physical description of water circulation and concentration generation processes, the mixing model was developed (Hall, 1970; Johnson et al., 1969; Pinder and Jones, 1969). It was applied at different time scale from flood and event scale (Hubert et al., 1969; Pinder and Jones, 1969) or annual scale (Johnson et al., 1969) and in different purposes as hydrograph separation (Pinder and Jones, 1969), understanding of flowpaths in catchment (Johnson et al., 1969), or the dynamics of the components of storm flow (Hubert et al., 1969). If the simple one-component model considered only one phenomena of dilution, the mixing equation proposes a mixing of different types of waters, each characterized by one concentration. The n-components could designed different source waters: event and pre-event water (Evans and Davies, 1998); old and new water (Godsey et al., 2009); or soil water, groundwater and event water (Johnson et al., 1969) or in a simplified way, as used by hydrologists in hydrograph separation, the base flow and the quick flow (Lyne and Hollick, 1979).

Table 17: Classical C-Q relationships used in literature

Approach	Equation	Number of parameters	Formula	Eq. n°
One-component	Inverse equation	1	$C = \frac{a}{Q}$	**Eq. (30)**
	One-sided power-law equation	2	$C = a\mathrm{Q}^{b}$	**Eq. (31)**
	Two-sided power-law equation	3	$C^{\frac{1}{\mathrm{n}}} = a + b\mathrm{Q}^{\frac{1}{\mathrm{n}}}$	**Eq. (32)**
Many-components	mixing model (*with at least two sources which can be simplified with base flow and quick flow*)	at least 2	$C = C_b \frac{Q_b}{Q} + C_q \frac{Q_q}{Q}$	**Eq. (33)**

With

C : *total chemical streamflow concentration at time t (mgL^{-1})*

Q : *total streamflow (m^{3}s^{-1}) at time t*

Q_b : *baseflow (m^{3}s^{-1})*

Q_q : *quickflow (m^{3}s^{-1}, $Q_q = Q - Q_b$)*

C_b : *Concentration of the baseflow (mgL^{-1})*

C_q : *Concentration of the quickflow (mgL^{-1})*

Despite their simplicity, the one-component models can yield excellent fits, explaining up to 90 % of the variance of the concentrations. For this reason, they are widely used today (Barco et al., 2008; Godsey et al., 2009; Moatar and Meybeck, 2007; Probst and Bazerbachi, 1986). However, several authors have underlined that the one-component models lump different hydro-chemical processes and dynamics, and do not allow for an explicit physical flow-path interpretation (Moatar et al., 2017; Rose et al., 2018). Godsey et al. (2009) compared different C-Q models, and concluded that it is difficult to find simple generalizable models that accurately represent the typical form of the C-Q relationship, which are internally consistent, and make plausible assumptions about catchment behavior.

C-Q relationships have been studied in a multitude of papers, showing their complexity. Studies of hysteresis loops, observed in storm-event concentration–discharge plots, show the relative timing and volume of mixing of the hydrological sources (Chanat et al., 2002; Evans and Davies, 1998; Evans et al., 1999). Legacy effects from antecedent moisture conditions (Biron et al., 1999) or time lags of precedent events (Biron et al., 1999; Lloyd et al., 2016) strongly affect C-Q relationships and can make them ambiguous. Some authors have proposed a classification of hysteresis loops within C-Q relationships (Evans and Davies, 1998; Williams, 1989). Stream water chemistry is viewed as the result of a variable mixture of soil water, groundwater and event water end-members. However, resulting loops from the assumption that the catchment behaves as a simple well-mixed bucket, hide the true complexity of flow paths and export behaviors (Godsey et al., 2009).

The variability of C-Q relationships could be due to a large number of processes, varying in space, time and with the characteristics of the catchment. The exact mechanisms leading to C-Q relations remain an open question. Both natural and anthropogenic factors affect the biogeochemical response of streams, and, while the majority of solutes show identifiable behaviors in individual catchments, only a minority of behaviors can be generalized (Botter et al., 2019). Some recent studies differentiate within this variability two distinct behaviors: chemo-dynamic and chemostasic export. Chemostasic export is defined as relatively small variations in concentrations compared to discharge (Musolff et al., 2015; Thompson et al., 2011). Chemostasis could be associated with constant rate of chemical weathering (Godsey et al., 2009), "a legacy storage" of anthropogenic nutriment which buffer the variability in concentrations (Basu et al., 2010; Clow and Mast, 2010), or with a significant hydraulic residence time compared to weathering kinetic (Ameli et al., 2017; Maher, 2011). Chemostasic export processes could represent the long-term trend of basin chemistry and can therefore be identified as an average of concentrations (Musolff et al., 2015). Chemodynamic patterns are characterized as high variations in concentrations compared to discharge that are, flushing, enrichment behavior, or dilution behavior (Musolff et al., 2015). It can be controlled: (i) by activation of solute sources heterogeneously distributed in space; (ii) by threshold-driven transport of constituents or (iii) high reactivity of constituents (Jones et al., 2017; Musolff et al., 2015; Vaughan et al., 2017; Zhang et al., 2016).

To deal with the non-linearity of the C-Q relationship, some authors have proposed to increase the number of hydrological components (e.g. Evans and Davies, 1998; Probst, 1985), make specific catchment calibration (Godsey et al., 2009), segment the flow in low and high flow (Meybeck and Moatar, 2012; Moatar et al., 2017) or integrate significant time steps (Kirchner, 2019). To gain insight into the linkages between chemical and hydrologic processes that yield the observed concentration-discharge relationship, research has often focused on only one of the modeling solutions (i.e. on either a one-component model or an *n*-component mixing model).

2 Procedure for combining mixing models and C-Q relationships

We used the two-sided affine power scaling relationship (2S-APS), which we recently proposed to fit the C-Q relationship using high-frequency concentrations measurements in the *River Lab* of the Orgeval-Oracle observatory, as a natural extension of the linear and log-log relations (see Tunqui Neira et al., 2019). The power-law relationship (Eq. (31), Table 17) could be advantageously replaced by the 2S-APS relationship (Eq. (32), Table 17).

To combine this 2S-APS relationship with the two-component mixing equation of Eq. (33), we propose to write the concentrations of each of components in Eq. (33) (i.e. C_b and C_q) as a function of total discharge, using the same transformation:

$$C_b = \left(a_b + b_b Q^{\frac{1}{n}}\right)^n$$

$$C_q = \left(a_q + b_q Q^{\frac{1}{n}}\right)^n$$

The total discharge and not the discharge of each of the two hydraulic components was used for seek of simplification, not increasing too much the number of parameters, and to propose a model which can be reduced to the two situations described below. Then the mixing equation becomes:

$$C = \left(a_b + b_b Q^{\frac{1}{n}}\right)^n \frac{Q_b}{Q} + \left(a_q + b_q Q^{\frac{1}{n}}\right)^n \frac{Q_q}{Q} \quad \textbf{Eq. (34)}$$

Depending on the values taken by parameters a and b in Eq. (34) reduced situations can be identified as briefly described below.

2.1 Case 1: chemostatic components ($b_b = b_q = 0$)

In this case, C_b and C_q are constant and independent of river discharge. The case 1 represents the chemostatic export characterizing a catchment controlled by large legacy stores (see Musolff et al., 2015). It corresponds to the assumption of the classical mixing equation approach, where the C_b and C_q values are set according to the average of concentrations, measured respectively during the dry and wet seasons (i.e. Saraiva Okello et al., 2018; Stewart et al., 2007; Zhang et al., 2013).

2.2 Case 2: single 2S-APS relationship ($a_b = a_q = a$ and $b_b = b_q = b$)

This case reduces to $C_b = C_q$ and to the simple C-Q relation described in Tunqui Neira et al. (2019):

$$C = \left(a + bQ^{\frac{1}{n}}\right)^n$$

The observed concentration is only a function of discharge.

2.3 Case 3: General case (a and b are different)

This is the general case of the transformed mixing equation (Eq. (34)). Unlike the classical mixing equation (chemostatic components), the C_b and C_q values are not constant but vary individually as a function of the stream flow. The general case allows accounting for complex chemo-dynamic export processes, or in other words, it considers the temporal variation of the chemical concentrations, and may be able to disconnect concentration and total discharge, taking advantage of the changing Q_b to Q_q ratio.

The “combining” procedure proposed in this paper consists in a single mathematical formulation (Eq. (34)) which merges the two most used equations of the hydro-chemical literature (i.e. Eq. (31) and Eq. (33)).

3 Application of the combining model

3.1 Study site and datasets

The combining model (Eq. (34)) was applied to the high-frequency hydro-chemical dataset measured in the Oracle-Orgeval observatory (Tallec et al., 2015) by the *River Lab* (Floury et al., 2017).

The Oracle-Orgeval observatory is a small catchment located 70 km east of Paris, France. It is subject to a temperate and oceanic climate, with annual average temperature of 11 ± 1 °C and a mean annual rainfall of 674 $\pm$ 31 mm (Tallec et al., 2013). The average measured streamflow at the Avenelles outlet (sub-catchment of 46 km^2 and location of the *River Lab* on the Oracle-Orgeval observatory) is about 0.2 m^3/s (1962 - 2017), with minimum flows in summer (< 0.1 m^3/s) and floods up to 10 m^3/s in winter and spring. With respect to geology, the catchment is underlay entirely by limestone rocks, with two aquifers: the shallower aquifer of the Brie limestone and the deeper Champigny limestone aquifer (Mouhri et al., 2013). Land use is mostly agricultural with few villages, and with intensive farming practices, mainly based on mineral nitrogen fertilization (Garnier et al., 2016). Nearly 60% of the surface of the catchment is drained with tile drains.

Among all ions measured every 30 minutes by the *River Lab* laboratory, we used 5 ions (see Table 18) whose behavior exhibit notable differences in the Orgeval catchment: sodium, chloride, sulfate, calcium and nitrates. Sodium and chloride mainly come from rain inputs and to a reduced extent (chloride) from fertilizers. Sulfate comes from the chemical weathering of gypsum, which makes it highly variable, depending on the season and the leaching of concerned localized underground layers. Calcium comes from the chemical weathering of limestones, largely distributed in the Orgeval catchment, with a specific variability, calcium is also on soils as agricultural lime. Nitrates mostly come from agricultural activities and fertilizers inputs, with specific seasonal leaching rates (Floury et al., 2018; Mouchel et al., 2016). Finally electro-conductivity reflects the presence of all these ions in stream water (EC, see Table 18). From all these, sodium, sulfate, chloride and EC show a great reactivity with the discharge (Tunqui Neira et al., 2019). The main data set (flow rates and chemical concentrations) covers the period between June 2015 and July 2017 (Figure 37), i.e. 17,500 measurements over 22 months.

Table 18: High-frequency measurements of chemical concentrations (average, minimum, maximum values and difference between quantiles 90 and 10 divided by the mean) from the *River Lab* at the Avenelles outlet, Oracle-Orgeval observatory (June 2015 - July 2017)

item	Unit	River Lab station (Avenelles outlet)			
		Mean (μ)	Min	Max	$(q_{90}-q_{10})/\mu$
calcium	$mg.L^{-1}$	124	37	202	0.23
sodium	$mg.L^{-1}$	13	2	17	0.22
sulfate	$mgS.L^{-1}$	19	2	32	0.44
nitrates	$mgN.L^{-1}$	12	3	18	0.24
chloride	$mg.L^{-1}$	30	4	40	0.28
EC	$\mu S.cm^{-1}$	704	267	1015	0.23

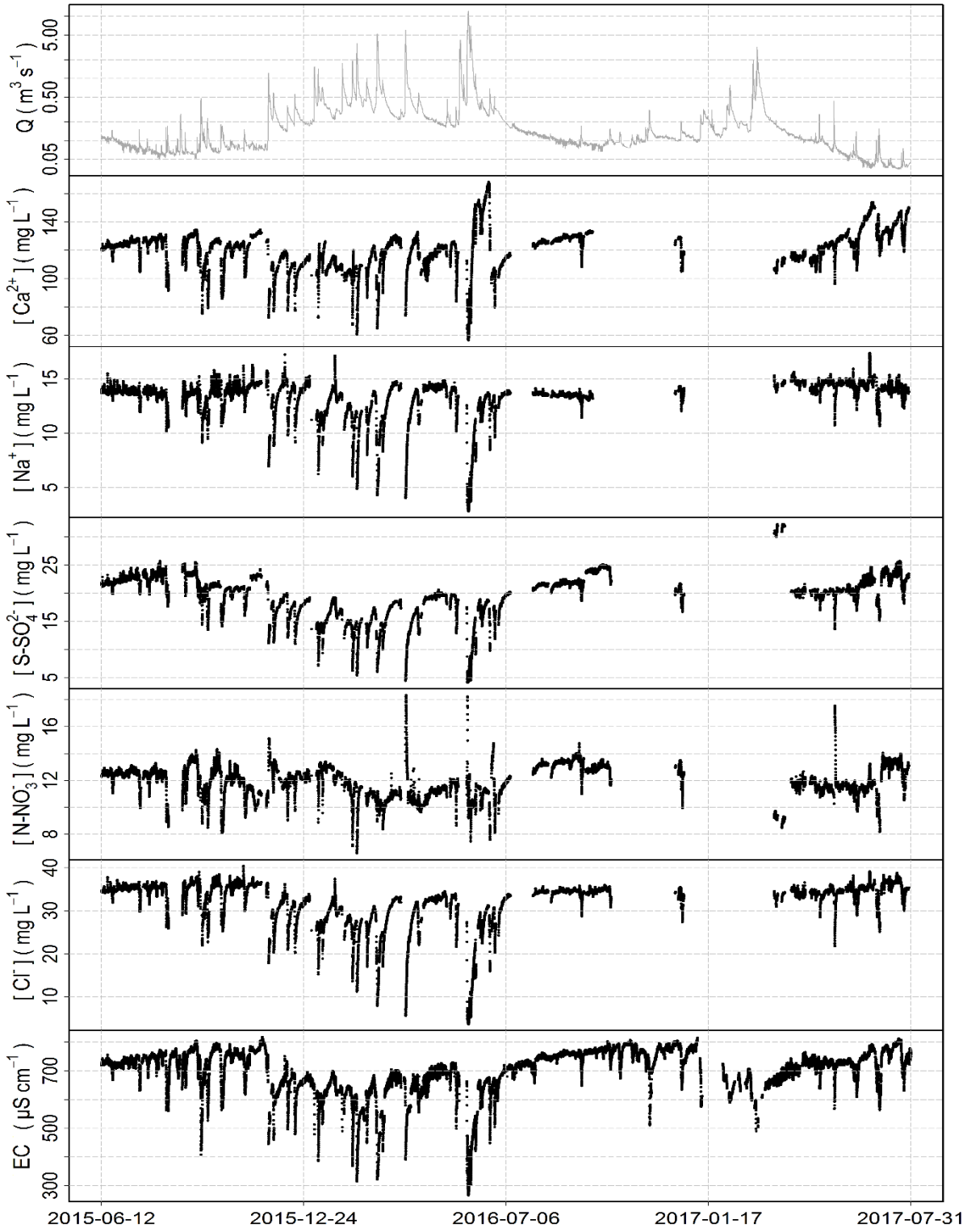


Figure 37 : High-frequency time series of flow, [Ca^{2+}], [Na^{+}], [$S\text{-}SO_4^{2-}$], [Cl^{-}], [$N\text{-}NO_3^{-}$] and EC, measured at the *River Lab* Avenelles station, Oracle-Orgeval observatory, from June 2015 to July 2017.

3.2 Methodology

To apply the combining model (Eq. (34)), we have to define the values of the baseflow (Q_b) and the quickflow (Q_q). The Recursive digital filter (RDF) hydrograph separation approach was used for this purpose. The RDF approach, adapted in the late 1970s from the signal-processing theory, is widely used for hydrograph separation. Indeed, RDF methods are computationally efficient, easily automated and applied to long continuous streamflow records (Chapman, 1991; Eckhardt, 2005). Among all the RDF-methods existing in literature (Brodie et al., 2007 p.62), we used the well-known Lyne-Hollick method (LH-RDF method) (Lyne and Hollick, 1979; Nathan and McMahon, 1990). Baseflow is considered here as a low-frequency signal, and surface runoff as a high-frequency signal. By filtering out the high-frequency signal, the low-frequency signal (i.e. baseflow) can be revealed (Longobardi and Loon, 2018; Nathan and McMahon, 1990).

The LH-RDF method is defined as follows:

$$Q_{b(t+1)} = \min\left(\alpha_\tau\, Q_{b(t)} + \frac{1-\alpha_\tau}{2}\left(Q_{(t+1)} + Q_{(t)}\right), Q_{(t)}\right) \quad \textbf{Eq. (35)}$$

And

$$Q_{q(t)} = Q_{(t)} - Q_{b(t)} \quad \textbf{Eq. (36)}$$

Iterative application of the filter allows smoothing data and nullifying phase distortion. We have used the 3 applications (forward-backward-forward) proposed by Nathan and McMahon (1990).

The LH-RDF method is characterized by one parameter (α_τ) with the main objective of defining the speed of convergence of the filter.

During seasons without significant recharge, stream flow may recess exponentially and follows the form:

$$\mathrm{Q}_{(t)} = Q_0 . exp^{-\left(\frac{\mathrm{t}-t_o}{\tau}\right)} \quad \textbf{Eq. (37)}$$

Where Q_0 is the streamflow at time t_o and $\boldsymbol{\tau}$ is the hydrological recession time constant of the catchment (in time units). When fixed time step are used, the exponential term of Eq. (37) can be replaced by:

$$\mathrm{Q}_{(t+\Delta\, t)} = Q_{(t)} . \exp^{\left(-\frac{\Delta \mathrm{t}}{\tau}\right)} = Q_{(t)} . K \quad \textbf{Eq. (38)}$$

Where K is the so-called recession constant of the catchment.

It is a common practice (e.g Longobardi et al., 2016; Zhang et al., 2017) to adapt the filter parameter (α_τ) to the recession constant of the catchment (K). Otherwise, either Q_b or Q_q would have an unwanted behavior on a seasonal time scale (to slow convergence of Q_b or too fast decrease of Q_q during flood events). Another important reason for this adaptation, is that the default value of α_τ

(0.925, proposed by Nathan and McMahon, 1990) applied in small catchments controlled by the regional scale factors such as slope and shape, has shown poor performance when computing Q_b (Ladson et al., 2013; Zhang et al., 2017)

According to Nathan and McMahon (1990), the recession analysis consists in "*combining individual flow recessions periods in such a way as to provide an average characterization of baseflow response"*: the resulting construction is usually termed Master Recession Curve (MRC). We used the *matching strip* MRC approach (Nathan and McMahon, 1990; Snyder, 1939). The analysis was applied on daily flow data of the Avenelles station from January 1, 2000 to December 31, 2017 as follows: the daily stream flow data of several seasons were overlapped according to the day of the year, starting from the beginning of June, and assuming that the stream flow decreases under a continuous recession process for the period June-September. The master recession curve was created by averaging the daily flow of all considered years, and an exponential regression was used to compute the hydrological recession time constant of the catchment (τ).

The combining model (Eq. (34)) was applied with the baseflow calculated with the LH-RDF method. For each ion and the EC, we used the parameter n previously determined by Tunqui Neira et al. (2019) on the same data set, but without flow separation. Model calibration relied, on the bounded Nash-Sutcliffe criterion (*NSEB*, Mathevet et al., 2006) is described in Tunqui Neira et al. (2019). The NSE criterion (Nash and Sutcliffe, 1970) and NSEB were computed according to Table 19. The Pareto plot between the two *NSEB* criteria (on loads and concentrations) allows us to visualize the best compromise between the criterion focusing on concentration ($NSEB_{conc}$) and the criterion focusing on load ($NSEB_{load}$). Since the scales $NSEB_{conc}$ and $NSEB_{load}$ are the same, the more natural and best compromise maximizes the average of $NSEB_{conc}$ and $NSEB_{load}$, and we name $NSEB_{comb}$ (see Eq. (43) in Table 19). The criteria were calculated using 420000 tuples, covering variations of each of the parameters (a_b,b_b) and (a_q,b_q) for each ion and EC, in the range of [1.0 to 5.0] for the a and [-1.2 to 1.2] for the b. For a linear model, the NSE criterion has a direct relationship to the commonly used correlation coefficient, which is more accurately called the Pearson product–moment correlation coefficient. However, the NSE criterion is more sensitive to data bias than R^2, which only reflects the covariance between observed and simulated values. The NSE criterion is a quadratic error (between simulated and observed data) divided by the variance of a so-called "naive" unbiased model (variance of the observed data compared to the average of the observed data) (see NSE equations, Table 19) (Nash and Sutcliffe, 1970). In addition to the $NSEB_{comb}$ criterion, we also used the root mean square error (RMSE, see Eq. (44) in Table 19).

Table 19: Numerical criteria used for optimization (C_{obs} – observed concentration, C_{sim} – simulated concentration, Q – observed discharge)

$NSE_{conc} = 1 - \frac{\sum_t \left(C_{obs}^t - C_{sim}^t\right)^2}{\sum_t (C_{obs}^t - \overline{C_{obs}})^2}$	**Eq. (39)**
$NSEB_{conc} = \frac{NSE_{conc}}{2 - NSE_{conc}}$	**Eq. (40)**
$NSE_{load} = 1 - \frac{\sum_t \left(Q^t C_{obs}^t - Q^t C_{sim}^t\right)^2}{\sum_t (Q^t C_{obs}^t - \overline{QC_{obs}})^2}$	**Eq. (41)**
$NSEB_{load} = \frac{NSE_{load}}{2 - NSE_{load}}$	**Eq. (42)**
$NSEB_{comb} = \frac{1}{2}(NSEB_{conc} + NSEB_{load})$	**Eq. (43)**
$RMSE = \sqrt{\frac{1}{N}\sum_t (C_{obs}^t - C_{cal}^t)^2}$	**Eq. (44)**

4 Results and discussion

4.1 Identification of the parameters and overall performance of the models

The τ value obtained for the Avenelles sub-catchment is about 768 hours or 32 days. This τ value is coherent with published values between 10 and 45 days (Brutsaert, 2008; Brutsaert and Lopez, 1998; Cheng et al., 2016; Vogel and Kroll, 1992). This calibration allowed us to compute the values of the baseflow (Q_b) and the quickflow (Q_q) on the whole time series (from June 2015 to July 2017).

Figure 38 shows the Pareto fronts plotted for each of the ions and EC, as well as the best chosen compromises (*NSEB$_{comb}$* designated with points of different shape and color for each case studied). In Table 20, all the parameters (a_b,b_b) and (a_q,b_q) of the best compromise found in the Pareto front are presented, as well as the value of *NSEB$_{comb}$* and the *RMSE.*

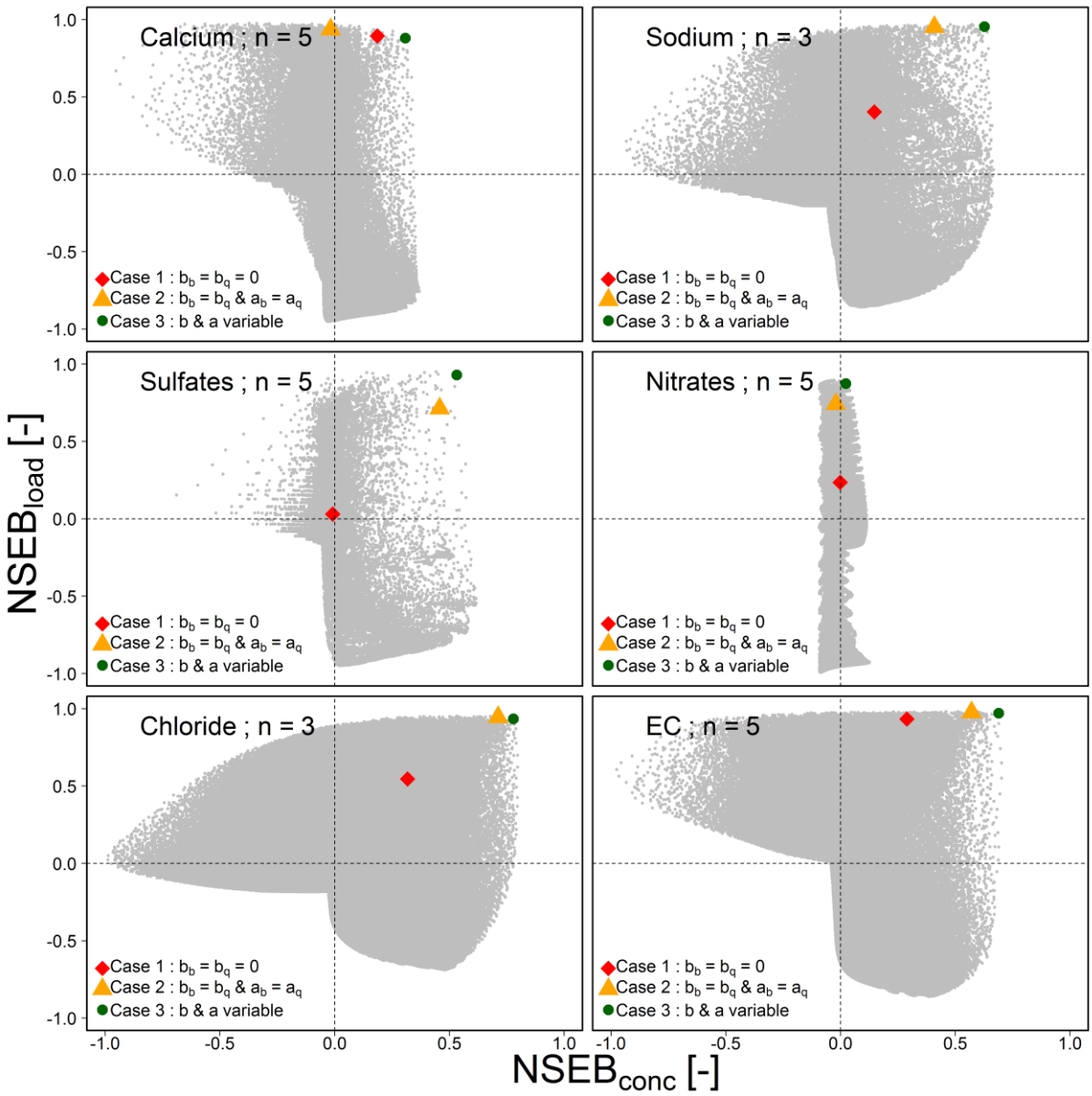


Figure 38: Pareto plot of $NSEB_{conc}$ vs $NSEB_{load}$ (grey dots) obtained from the combining model with 5 ions and EC. Red diamonds correspond to the best-combined criterion NSE_{comb} for chemostatic components (case 1), orange triangles correspond to the single 2S-APS relationship (case 2) and green circles correspond to the general case of the model (case 3).

Table 20: Values used for n in the combining model and obtained for optimal NSE_{comb} criterion, (a_b,b_b) and (a_q,b_q) parameters and *RMSE* for each case and solute. Note that case 1 corresponds to chemostatic components, case 2 to the single 2S-APS relationship and case 3 to general case of the combining model.

Solute	*n*	Case	a_b	b_b	a_q	b_q	$NSEB_{comb}$	*RMSE*
Calcium	5	1	2.7	0	2.2	0	0.54	15.55 mg.L^{-1}
		2	2.9	-0.4	2.9	-0.4	0.46	12.24 mg.L^{-1}
		3	2.8	-0.2	2.0	0.2	0.59	12.92 mg.L^{-1}
Sodium	3	1	2.5	0	1.4	0	0.66	1.49 mg.L^{-1}
		2	2.7	-0.6	2.7	-0.6	0.68	1.00 mg.L^{-1}
		3	2.6	-0.3	2.1	-0.3	0.79	0.90 mg.L^{-1}
Sulfate	5	1	1.9	0	1.3	0	0.46	3.66 mgS.L^{-1}
		2	2.2	-0.55	2.2	-0.55	0.69	2.20 mgS.L^{-1}
		3	2.2	-0.55	2	-0.4	0.74	1.9 0mgS.L^{-1}
Nitrate	5	1	1.8	0	1.6	0	0.39	5.86 mgN.L^{-1}
		2	1.8	-0.1	1.8	-0.1	0.41	3.81 mgN.L^{-1}
		3	2.3	-1.2	2.1	-0.3	0.45	3.31 mgN.L^{-1}
Chloride	3	1	3.3	0	1.6	0	0.52	4.45 mg.L^{-1}
		2	3.7	-1	3.7	-1	0.83	2.00 mg.L^{-1}
		3	3.6	-0.7	3.3	-0.8	0.86	1.72 mg.L^{-1}
EC	5	1	3.8	0	3.1	0	0.61	72.30 µS.cm^{-1}
		2	4.2	-0.7	4.2	-0.7	0.77	42.00 µS.cm^{-1}
		3	4.2	-0.7	3.9	-0.5	0.83	36.90 µS.cm^{-1}

As expected, whatever the solute species considered, the optimal $NSEB_{comb}$ values are obtained for the general case (Case 3, see Figure 38 and Table 20). Except for calcium, the Case 2 presents better $NSEB_{comb}$ than the Case 1 (see yellow vs red points in Figure 38). These results show that a single 2S-APS relationship (Case 2) better explains the variations of the stream water concentrations than a mass balance equation with constant concentration components (Case 1). The general case (Case 3) based on a mixing equation and including a C-Q relationship for each of the flow components, improves all the performances as expected. The same can be observed for the *RMSE* criterion, which minimum value has always been obtained for the general case for almost the ions studied (except calcium, see Table 20).

This means that concentrations cannot be considered constant across time in each of the flow components (chemostatic components), in as much as the hydrograph separation can be considered ad physically relevant. The streamwater quality of the Avenelles sub-catchment appears strongly influenced by chemo-dynamic processes (modeled by the single 2S-APS relationship and the general case).

The most evident improvement of the full case 3 model (for both $NSEB_{comb}$ and *RMSE* criteria see Table 20) is observed for chloride, sodium and EC. According to Floury et al. (2018), sodium and chloride

would come from the Brie aquifer with mainly external input from rainfall during the wet season and additional sources from mineral fertilizers for chloride.

For calcium, unlike other ions, the 2S-APS model with one hydrological component only is the one that reaches the best performance. However, it should be reminded that the *RMSE* criterion only relies on the errors regarding concentrations, while $NSEB_{comb}$, besides the correlation, the concentration budget and variability between, simulated and observed concentrations, are also considered (Gupta et al., 2009). For calcium, we obtained modest results with the $NSEB_{comb}$ criterion for the three cases. The study site, mainly located on limestone, would lead to expect a stable inter-annual behavior of this ion. However, studies made by Mouchel et al. (2016) and Floury et al. (2018) have pointed out the considerable seasonal variability of calcium. External contributions of calcium, such as rainfall or local atmospheric dust are negligible (Floury et al., 2018; Mouchel et al., 2016). Calcium concentrations would come essentially from the both aquifers in relation to the river (Berrhouma, 2018; Mouchel et al., 2016). They can also come from anthropogenic contributions such as agricultural lime, a popular soil additive used in agricultural catchments such as the Oracle-Orgeval observatory (West and McBride, 2005).

The lowest value of the $NSEB_{comb}$ criterion is obtained for nitrate, it remains below 0.50 (see Table 20). Nitrates present a large seasonal variability, strongly related to farming practices (Garnier et al., 2016) and chemo-dynamic processes especially in the hyporheic-zone (Floury et al., 2018). This larger variability, either environmental or anthropogenic, cannot be simulated, whatever the case. The hydrological model does not explicitly separate between several potentially slow hydrological components affecting soil water and aquifers, whereas nitrate is present in contrasted concentrations in these compartments.

Sulfate $NSEB_{comb}$ is much improved when variable concentrations are introduced in the quick and slow compartments (i.e., $NSEB_{comb}$ from 0.46 to 0.73 and *RMSE* from 3.66 to 2.19 mgS.L^{-1}, Table 20). Sulfate probably come from the dissolution of the gypsum, present in lenticular layers in gypsum marl separating the two aquifers (Mouchel et al., 2016; Mouhri et al., 2013), with an additional fertilizer end-member (Floury et al., 2018). The variability of sulfate concentrations is related to the spatial heterogeneity of these gypsum lenses and to the temporary variability of the water table heights and farming practices. The spatial and temporal variability of sulfate concentrations is thus due to chemo-dynamical processes, well modeled for the sulfate by the general case (case 3).

A significant feature, common to almost all ions (except nitrate) and cases of the model, is the lower C_q concentration, compared to C_b, for any Q (see Figure 39). This very clearly shows that dilution, from rainwater and runoff water is a major process in the Avenelles catchment. More in details, in the case 1 (constant concentration in each of the hydrological components) the overall shape of the simulated C-Q relationship is controlled by the distribution of the Q_b/Q and Q_q/Q ratios with changing Q.

Regarding sodium and chloride, which are the better simulated ions, Figure 39 shows that the chemostatic hypothesis (case 1) fails to account for the distribution of sodium concentrations over the whole range of Q values. The case 2 (single component 2S-APS relationship, n = 3) accounts for an acceptable general relation, while the combining general case (case 3) improves this relationship with a relevant variability around the relation simulated in case 2 (Figure 39).

Similar conclusions can be drawn for sulfate. The observed concentration patterns for nitrate are much more complicated. The general case simulates much more variables nitrate concentrations, which seem more appropriate that what was obtained in cases 1 and 2, but the model appears unable to simulate in detail the complex variability of nitrate concentrations during runoff events. As suggested by the weak $NSEB_{comb}$ value, calcium is not well simulated. Although less complex than for nitrate, specific dilution or mixing patterns appear during high flow periods, they all seem different from each other. Since much better results were obtained for sodium and chloride with the same hydrological sub-model, it seems that the assumed constancy of the C-Q relation for all events should be reconsidered.

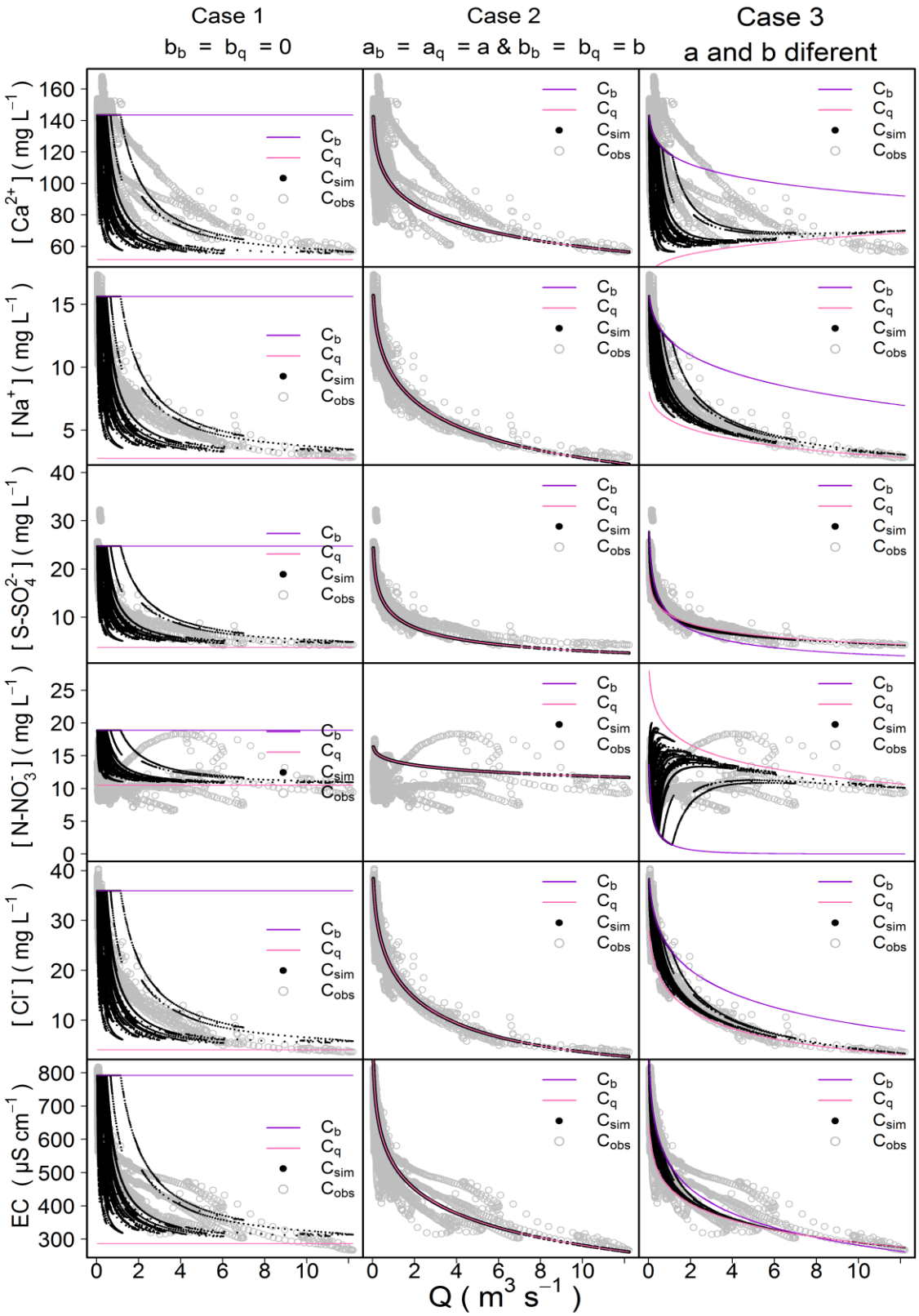


Figure 39: Comparison (C-Q relationships) between the different simulated concentration (C_{sim}) and the observed concentration (C_{obs}) for each ion and EC and the 3 cases. Comparison of C_b and C_q also obtained for each ion and EC and each case. Note that case 1 corresponds to chemostatic components, case 2 to the single 2S-APS relationship and case 3 to general case of the combining model.

4.2 Performances for selected storm events

Compared to the single 2S-APS relationship (case 2), a potential advantage of the combining model (general case) is to decouple rising and decreasing flow periods with changing Q_q/Q and Q_b/Q ratios, which provides an opportunity to simulate hysteresis phenomena, that cannot be simulated by a single C-Q relation. The recursive filter is supposed to separate the base flow to the quick flow, and to simulate a quick response of the quick flow at the beginning of flood periods, and a progressive increase of the base flow during the second part of flood periods, if they are along enough with regards to the time constant of the filter. Accordingly, the two components model should improve the simulation of complex concentration changes during flood events.

Figure 40 presents a representative short duration flood event during a wet period (February 2016, high groundwater level, high soil moisture and drain) and Figure 41 a similar event during a dry period (November 2015, low groundwater level, low soil moisture and no contribution of tile drains). The two Figures present data and simulations with the combining model as a function of time (letter a in Figure 40a and Figure 41a), and as a function of discharge to figure out hysteresis patterns for observed data (letter b in Figure 40 and Figure 41) and simulated data (letters c and d in Figure 40 and Figure 41), for the solute and the EC.

During the wet and dry hydrological seasons, almost all studied hysteresis loops for the ions (except for nitrates in dry season, see Figure 41c) and for EC, show a dilution step with water of lower concentration followed by an increasing concentrations step (see dotted line C_{obs}, Figure 40a and Figure 41a). A common pattern with 3 or 4 stages (see Figure 40b and Figure 41b) can be observed in most situations with a clockwise hysteresis.

During a first stage, the concentration remains almost constant during the first increase of the discharge, then, in a second stage, generally short, the concentrations strongly decrease while the discharges quickly increase. In a third stage, not always visible (i.e. Figure 40b), the concentrations continue to decrease while the discharges start to decrease again. In the fourth and final stage, both concentrations increase again while discharges continue to decrease and reach again a lower flow and higher concentration (see Figure 40b and Figure 41b).

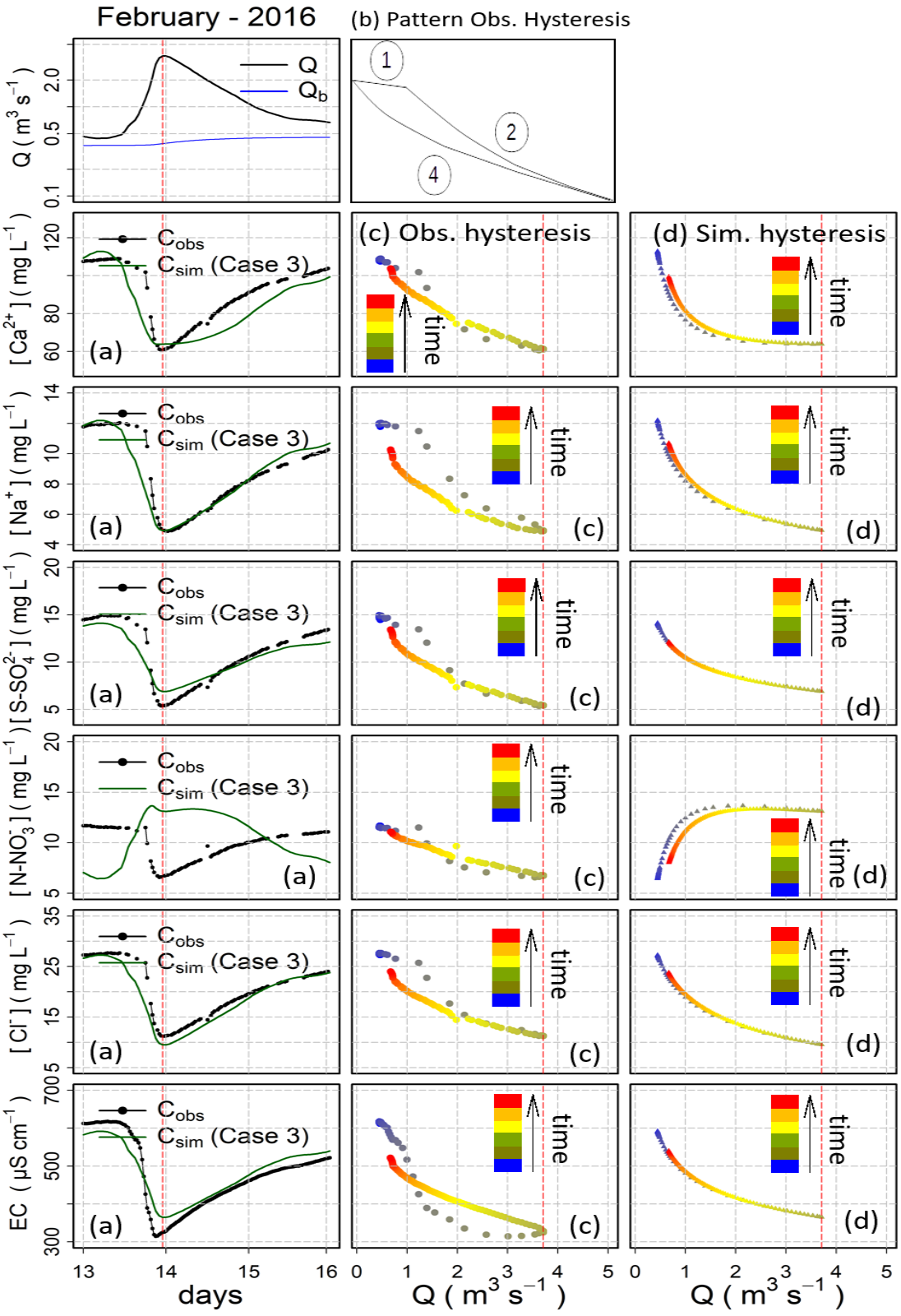


Figure 40: Storm event during the wet season: (a) comparison of observed (black line) and simulated (green line) concentrations from the general case (Case 3); (b) Common pattern description of the observed hysteresis; (c) plots of the corresponded observed hysteresis loop and (d) plots of the corresponded simulated hysteresis loop for the solutes and EC. The red dot line represents the flow peak of the event.

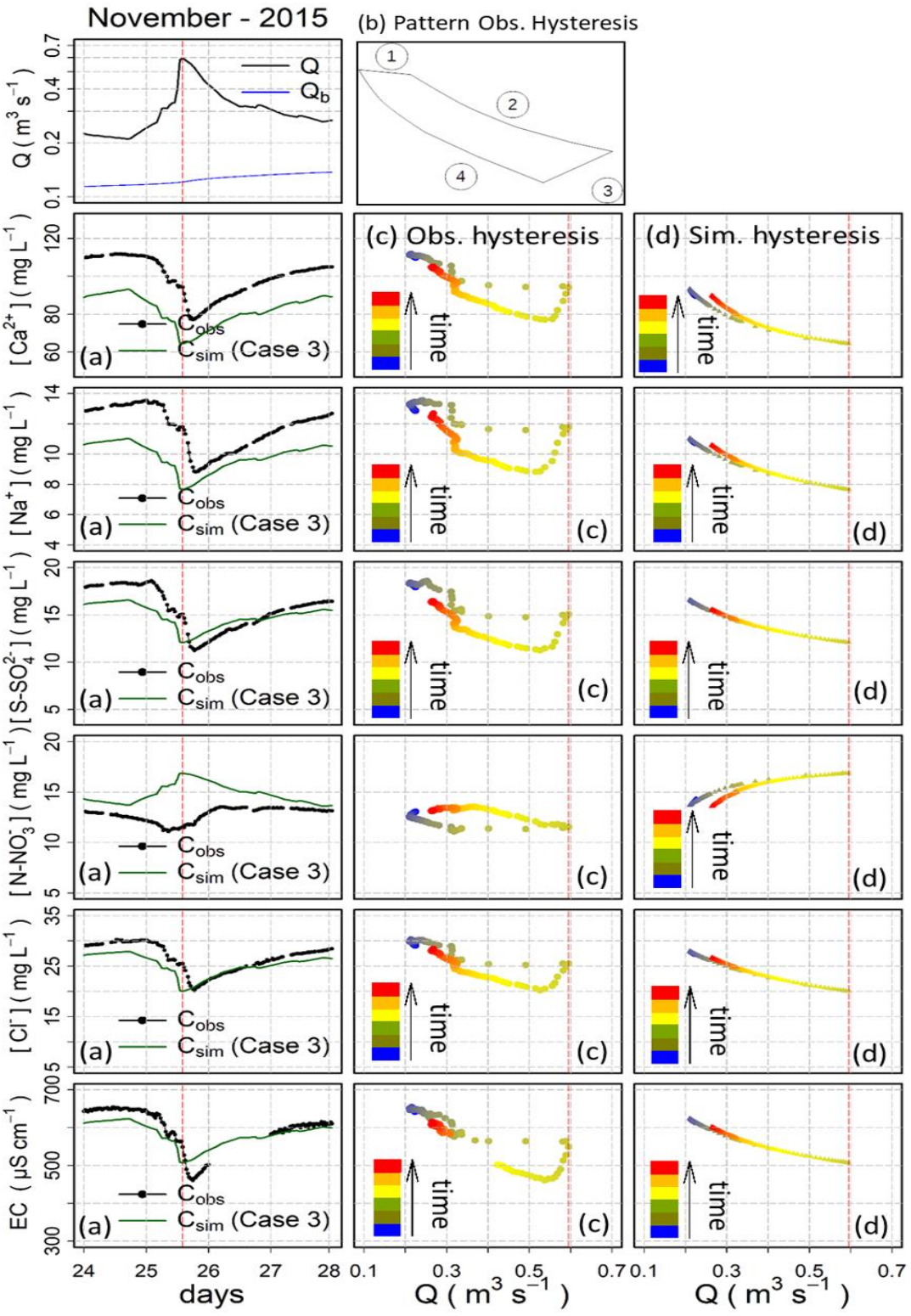


Figure 41: Storm event during the dry season: (a) comparison of observed (black line) and simulated (green line) concentrations from the Case 3; (b) Common pattern description of the observed hysteresis (except for nitrates); (c) plots of the corresponded observed hysteresis loop and (d) plots of the corresponded simulated hysteresis loop for the solutes and EC. The red dot line represents the flow peak of the event.

The first stage would correspond to a pre-event pattern already described in the literature (Chanat et al., 2002; Evans and Davies, 1998; Evans et al., 1999; Rose et al., 2018), during which the concentrations come mostly from the groundwater pool. Short time lags between the beginning of the increase of discharge and the decrease of concentration (stage 1) or the decrease of discharge and increase of concentration (when stage 3 occurs) could be explained by hydraulic shifts that could be explained by the lower in-stream velocity of water as compared to that of the flood wave, and by some piston effect in near stream pore-water. The following stages correspond to the soil waters and / or runoff mixing in varying proportions during the event. The contribution of these waters (and the dilution that goes with it) persists for some time after the peak flow, but in weaker and weaker in proportion compared to those of the groundwater. Note that in the dry season, for all ions and EC, the soft dilution stage (stage 1), is interrupted by rapid and abrupt dilution (Figure 41c). This last dilution would correspond to a larger proportion of rain water, it is indistinguishable during the wet season, because rain water is mixes with water from tile drains during this period of the year.

Only nitrate ions show a specific pattern with anticlockwise hysteresis during the dry season (Figure 41c). Unlike other ions, nitrates appear more concentrated in soils and drain water than in groundwater or rainwater. According to Garnier et al. (2014), in the Avenelles sub-catchment, sub-root nitrate concentrations average 22 $mgN.L^{-1}$ close to the average concentration observed in drains in the same area (26 $mgN.L^{-1}$). Nitrate concentrations in the Brie aquifer are only around 13.2 $mgN.L^{-1}$ (Garnier et al., 2014), whereas in rainfall they are about 0.75 $mgN.L^{-1}$ (Floury et al., 2018).

The clockwise behavior of the hysteresis loop is counter intuitive, except for nitrates which paradoxically show an anti-clockwise hysteresis (Figure 41c). It is expected from the hydrological RDF model we adopted, that the fraction of quickflow should be higher during the rising stage and lower during the descending stage. Then, since the concentration of the quickflow are lower than the concentrations of the baseflow, as simulated by the combining model, the hysteresis loop should be counter clockwise, with lower concentrations during the rising stage and higher concentrations during the descending stage. However, as suggested by Evans and Davies (1998), a clockwise hysteresis can be explained when the pre-event concentration is higher than the event concentration, by taking into account a time lag between the arrival flow and concentration. This may transform anti-clockwise hysteresis into clockwise hysteresis as shown in Figure 42, the stages that appear are not unlike those which were obtained for the November 2015 and February 2016 floods. The average velocity in the Avenelles river is not well known in the reaches upstream of the *River Lab* station. Rough estimates, based on the theoretical ratio between flood wave velocity and water velocity in 1D-flows, show that the lag time between flow and concentrations should lay in the range about 1 to 4 hours (the latter for lower river discharges). The model, which does not include any time shift between concentrations and discharges data, very poorly fits to the results, with a very narrow anti-clockwise hysteresis (see Figure

40c and Figure 41c). The very narrow shape is explained by the high dominance of Q_q over Q_b, during the whole event, except at the very beginning of the February 2016 event. Then, during most of the time, for each of the events, the simulated concentration is controlled by the C_q-Q relation and no significant hysteresis can develop. In a flood with a longer duration, the bias due to the time lag would likely be less significant and a wider distribution of the Q_q/Q_b ratio, as simulated by the RDF method, may be expected.

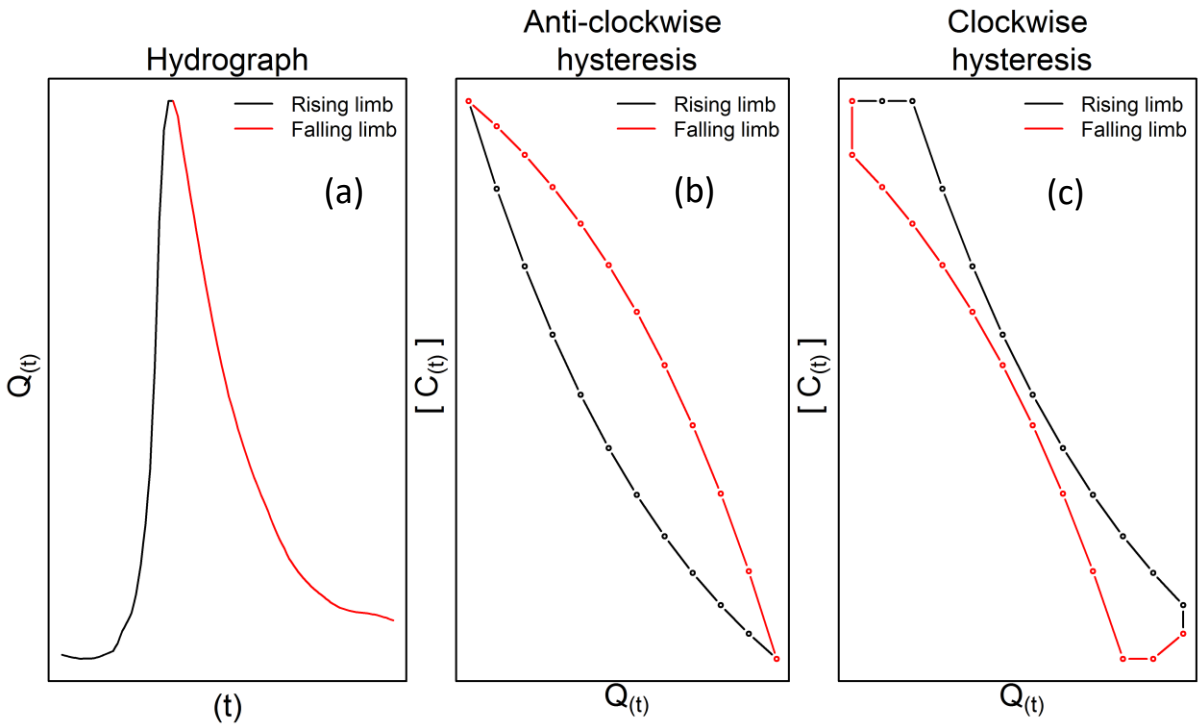


Figure 42: Schematic hysteresis patterns and hydrograph: (a) flow hydrograph behavior during a flood event; (b) a typical C-Q anti-clockwise hysteresis, with a higher fraction of lower concentration quick flow water during the rising limb; (c) the same concentration data form a clockwise hysteresis if concentration data are shifted forward, here by two time steps.

The longer wet season flood of March 2016 (Figure 43a), well covered by concentration data set, was selected, although three short data gaps appear on days 6, 7 and 8. Its duration was about 15 days, but unfortunately, as most long flood events, it structured by a succession of several flood events (three events here) of varying importance which may perturb the analysis of data. A first peak flow (~0.7 m^3/s) on day 4 is quickly followed by the main flood peak (5m^3/s) occurring at the end of day 4, and a third, lower peak (1 m^3/s) occurs during day 9. As shown in Figure 43, because of the longer duration of the flood, the recursive filter simulates important variations of the Q_q/Q_b ratio during the event: during most of peak flow period, most of the discharge is due to the quick flow component, while the quick flow component turns back to zero after day 10.

As expected from the wider range of Q_q/Q_b ratio, the model simulates more difference between concentrations during the rising part of the event than the descending part (Figure 43c). The simulated

rising part gets controlled by the C_q-Q relationship, while the end of descending part is controlled by the C_b-Q relationship.

Because sodium and chloride are better simulated than the other ions (Table 20), these ions will be used to analyze the behavior of the model and its ability to simulate the hysteresis loop. Despite the longer time scale of the event, sharp changes of the discharge data still appears, immediately before each of the peak flow. During these short periods the simulated concentrations are slightly shift forward compared to the observed concentration. Clockwise loops are highly visible in Figure 43b for each of the flood peaks, with a higher effect for the main peak. This is typically the clockwise loop described by Evans and Davis and already observed during the shorter events studied, it is still active at specific moments inside the longer flood period. Once this is identified the analysis should be focused on the remaining observations during the longer periods between sharp discharge changes. Concentration data during these periods are not unlike simulated concentrations for sodium, chloride and also for sulfate, but with very similar C-Q relations for Q_q and Q_b. For calcium the fitted relation fails to properly simulate the main flow descending period, but the general pattern is nevertheless observed. A shift is however observed during the falling limb after the main peak between observed and simulated data. The seasonal or event-scale variability of C-Q relations for calcium for high discharge sequences was already discussed in a previous paper.

The model does not reproduce the interesting daily fluctuations for calcium and EC (and also likely for alkalinity as the main missing term to reconstruct EC from ion concentration data). They are clearly out of the scope of what a C-Q based model can reproduce.

In the case of nitrate, unlike the two other flood events analyzed, we can observe an apparent chemostatic behavior of the observed concentration (see nitrates in Figure 43a). This is due to fertilization and to a lower consumption of nitrogen by the vegetation at this period of (Billy et al., 2013; Garnier et al., 2016), therefore the soil and drains hydro-chemical pools show high concentrations that likely balanced a possible dilution by rain water, therefore leading to apparently constant concentrations.

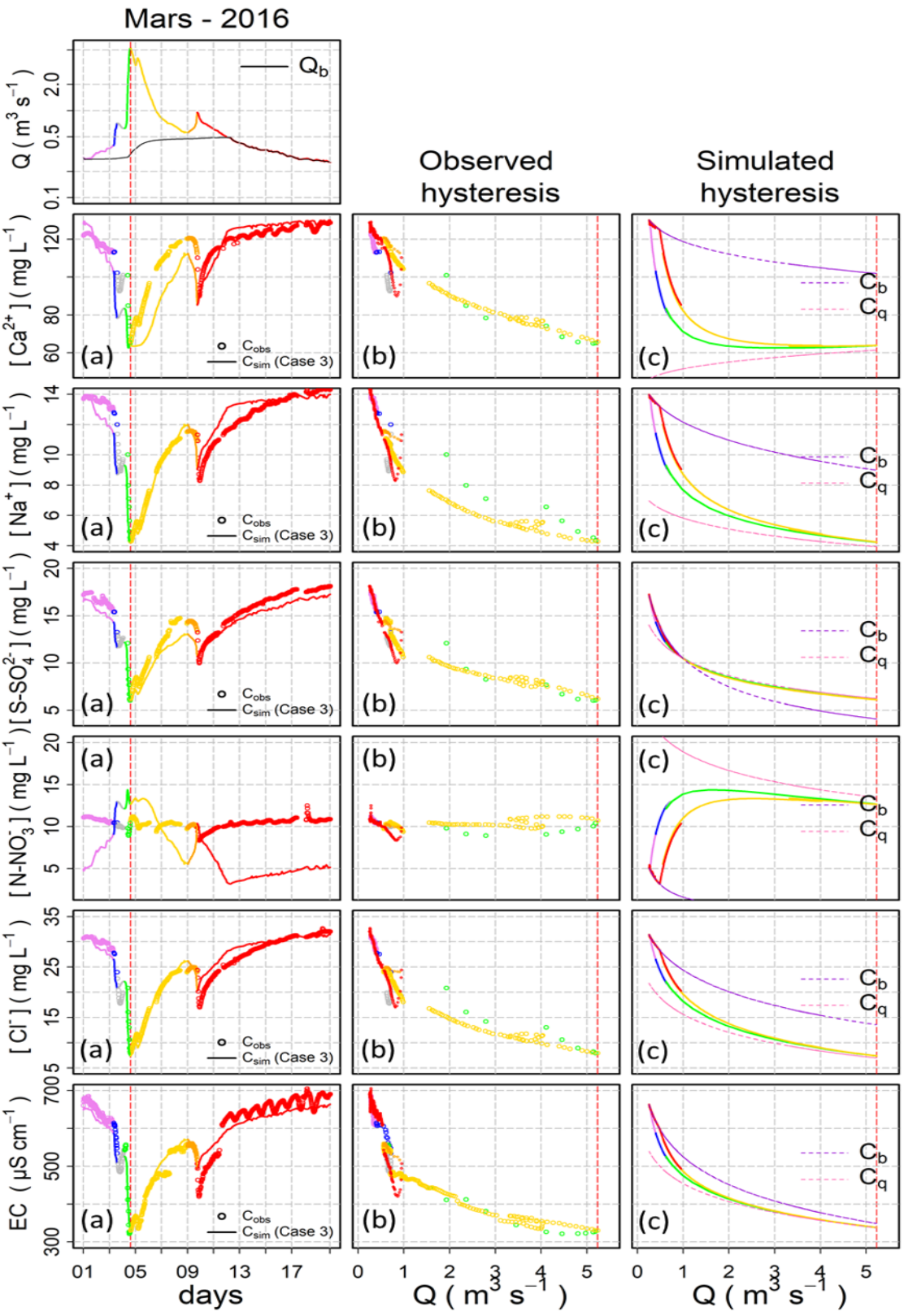


Figure 43: Long storm event during mars 2016 (wet season): (a) comparison of observed and simulated concentrations from the general case (Case 3) vs. time; (b) plots of the corresponded observed hysteresis loop and (c) plots of the corresponded simulated hysteresis loop and C_q and C_b vs. Q. The red dot line represents the flow peak of the event. The different colors of the plot correspond to the different stages.

Last, to quantify the simulation error, Table 21 shows the results of the RMSE criterion calculated between the simulated (using the general case) and observed concentrations of the three flood events for each ion and EC. For most ions (except for nitrates) and EC, the simulation of concentration by the combining model (general case) in the wet season floods (i.e. short and long hysteresis illustrate in Figure 40 and Figure 43) seems much improved compared to the short dry season flood.

Table 21: RMSE criterion (simulated with combining model (general case) vs. observed concentrations) calculated for the three flood events analyzed for each ion and EC.

Solute	*Short flood event, wet season* (Figure 40)	*Short flood event, dry season* (Figure 41)	*Longer flood event, wet season* (Figure 43)
	RMSE	***RMSE***	***RMSE***
Calcium	8.89 mgL^{-1}	18.1 mgL^{-1}	9.97 mgL^{-1}
Sodium	0.80 mgL^{-1}	2.14 mgL^{-1}	0.80 mgL^{-1}
Sulfate	1.04 $mgSL^{-1}$	1.51 $mgSL^{-1}$	1.10 $mgSL^{-1}$
Nitrate	4.12 $mgNL^{-1}$	2.54 $mgNL^{-1}$	4.82 $mgNL^{-1}$
Chloride	1.82 mgL^{-1}	2.07 mgL^{-1}	1.58 mgL^{-1}
EC	29.6 μScm^{-1}	29.6 μScm^{-1}	27.2 μScm^{-1}

5 Conclusions

The new conceptual model, coupling flow separation and C-Q relations for each of the components, proposed in this paper allows better estimations of concentrations in streamflow for most ions at inter-annual scale. Even for nitrate (which do not exhibit a clear C-Q relationship), the new model improves significantly the estimates (based on the $NSEB_{comb}$ and *RMSE* criteria). Also we analyzed the performances of the model during three floods events. We show that the observed C-Q hysteresis are in the majority of the clockwise type (Evans and Davies, 1998) and mainly due to a lag time between flow and concentrations. After removing this specific effect, due to a short hydrodynamic process in the longer flood event that cannot be simulated by a mixing model, the anti-clockwise behavior is revealed and it is fairly consistent with the simulations. However, hysteresis and its behavior have always been a powerful factor in the restriction of hydrochemical modeling (Fovet et al., 2015).

Our proposed model shows the advantage of coupling a time dynamic hydrological model with a C-Q relationship for each of the flow components, makes it possible to simulate time dependent relations that do not only rely on $Q\ (t)$ to simulate this dependence. This is an attempt to connect hydrological models that conceptually distributed flows with non-constant chemical composition. This made possible by the very dense set of made available on the Avenelles catchment by the *River Lab*. This approach remains conceptual and there is no possible physical identification of the water masses with varying concentrations that produce the quick and base flow.

The two-member hydrograph separation tested here may be a limiting factor: for nitrate and calcium ions the hydrograph separation seems to be insufficient, while for sodium, sulfate, chloride ions and

EC the two components hydrograph separation (C_q and C_b) seems adequate. The baseflow separation method used in this article, shares uncertainties due to the arbitrary and speculative hypotheses used in its conception (Brutsaert, 2008; Cheng et al., 2016). The "simple" separation into two components cannot explain the complexity of the behavior of the nitrate and calcium ions. Indeed, several studies (Miller et al., 2017; Probst, 1985) have shown that the introduction of a third component (which would come to represent the soil pool component) improves the modeling performance with respect to a two-component model. Adding complexity to a model always improves simulations.

Seasonality is also a factor of variation. Several studies talked about the seasonal variation of calcium (Barco et al., 2008; Markewitz et al., 2011) and nitrate (Aubert et al., 2013; Barco et al., 2008; Botter et al., 2019; Rusjan et al., 2008). In the Orgeval-Oracle observatory the seasonality of calcium ion might be due to the interaction of the two existing limestone aquifers (Brie and Champigny) that occurs at different times of the year (Berrhouma, 2018; Mouhri et al., 2013). This interaction happens because of heterogeneity in the structure of the geological layers and particular circulations between the two aquifers (Mouchel et al., 2016; Mouhri et al., 2013). In the case of nitrates the strong seasonality is due to an anthropogenic origin linked to agricultural technical sequences and plant-nitrogen interactions (Garnier et al., 2016).

The decoupling of discharge and concentration patterns, due to a short time lag explained by hydraulic mechanisms, is not simulated by the model; doing so would be an interesting extension of this work. Other priorities would be to test the implementation of similar models which will be made possible by the extension of the *River Labs* in other catchments in France and other implementation of high-frequency water quality stations in other countries (Kirchner et al., 2004). Different types of chemical signatures, reflecting different types of hydrological and hydro-chemical functions underlying the transfer processes, could be studied in details, including different flow decomposition concepts and may be other C-Q relations possibly including specific seasonal features.

6 References

Ameli, A.A., Beven, K., Erlandsson, M., Creed, I.F., McDonnell, J.J., Bishop, K., 2017. Primary weathering rates, water transit times, and concentration-discharge relations: A theoretical analysis for the critical zone. Water Resources Research, 53(1): 942-960. DOI:10.1002/2016wr019448

Aubert, A.H., Gascuel-Odoux, C., Gruau, G., Akkal, N., Faucheux, M., Fauvel, Y., Grimaldi, C., Hamon, Y., Jaffrezic, A., Lecoz-Boutnik, M., Molenat, J., Petitjean, P., Ruiz, L., Merot, P., 2013. Solute transport dynamics in small, shallow groundwater-dominated agricultural catchments: insights from a high-frequency, multisolute 10 yr-long monitoring study. Hydrology and Earth System Sciences, 17(4): 1379-1391. DOI:10.5194/hess-17-1379-2013

Barco, J., Hogue, T.S., Curto, V., Rademacher, L., 2008. Linking hydrology and stream geochemistry in urban fringe watersheds. Journal of Hydrology, 360(1): 31-47. DOI:https://doi.org/10.1016/j.jhydrol.2008.07.011

Barthold, F.K., Tyralla, C., Schneider, K., Vaché, K.B., Frede, H.-G., Breuer, L., 2011. How many tracers do we need for end member mixing analysis (EMMA)? A sensitivity analysis. Water Resources Research, 47(8). DOI:10.1029/2011wr010604

Basu, N.B., Destouni, G., Jawitz, J.W., Thompson, S.E., Loukinova, N.V., Darracq, A., Zanardo, S., Yaeger, M., Sivapalan, M., Rinaldo, A., 2010. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. Geophysical Research Letters, 37(23).

Berrhouma, A., 2018. Fonctionnement hydrothermique de l'interface nappe-rivière du bassin des Avenelles, PSL Research University.

Bieroza, M.Z., Heathwaite, A.L., Bechmann, M., Kyllmar, K., Jordan, P., 2018. The concentration-discharge slope as a tool for water quality management. Science of The Total Environment, 630: 738-749. DOI:https://doi.org/10.1016/j.scitotenv.2018.02.256

Billy, C., Birgand, F., Ansart, P., Peschard, J., Sebilo, M., Tournebize, J., 2013. Factors controlling nitrate concentrations in surface waters of an artificially drained agricultural watershed. Landscape Ecology, 28(4): 665-684. DOI:10.1007/s10980-013-9872-2

Biron, P.M., Roy, A.G., Courschesne, F., Hendershot, W.H., Côté, B., Fyles, J., 1999. The effects of antecedent moisture conditions on the relationship of hydrology to hydrochemistry in a small forested watershed. Hydrological Processes, 13(11): 1541-1555. DOI:10.1002/(sici)1099-1085(19990815)13:11<1541::aid-hyp832>3.0.co;2-j

Botter, M., Burlando, P., Fatichi, S., 2019. Anthropogenic and catchment characteristic signatures in the water quality of Swiss rivers: a quantitative assessment. Hydrol. Earth Syst. Sci., 23(4): 1885-1904. DOI:10.5194/hess-23-1885-2019

Brodie, R., Sundaram, B., Tottenham, R., Hostetler, S., Ransley, T., 2007. An overview of tools for assessing groundwater-surface water connectivity. Bureau of Rural Sciences, Canberra, Australia: 57-70.

Brutsaert, W., 2008. Long-term groundwater storage trends estimated from streamflow records: Climatic perspective. Water Resources Research, 44(2).

Brutsaert, W., Lopez, J.P., 1998. Basin-scale geohydrologic drought flow features of riparian aquifers in the Southern Great Plains. Water Resources Research, 34(2): 233-240. DOI:10.1029/97wr03068

Chanat, J.G., Rice, K.C., Hornberger, G.M., 2002. Consistency of patterns in concentration-discharge plots. Water Resources Research, 38(8): 22-1-22-10. DOI:10.1029/2001wr000971

Chapman, T.G., 1991. Comment on "Evaluation of automated techniques for base flow and recession analyses" by R. J. Nathan and T. A. McMahon. Water Resources Research, 27(7): 1783-1784.

Cheng, L., Zhang, L., Brutsaert, W., 2016. Automated selection of pure base flows from regular daily streamflow data: objective algorithm. Journal of Hydrologic Engineering, 21(11): 1-7.

Clow, D.W., Mast, M.A., 2010. Mechanisms for chemostatic behavior in catchments: Implications for CO2 consumption by mineral weathering. Chemical Geology, 269(1): 40-51. DOI:https://doi.org/10.1016/j.chemgeo.2009.09.014

Durum, W.H., 1953. Relationship of the mineral constituents in solution to stream flow, Saline River near Russell, Kansas. Eos, Transactions American Geophysical Union, 34(3): 435-442. DOI:10.1029/TR034i003p00435

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes, 19(2): 507-515.

Evans, C., Davies, T.D., 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. Water Resources Research, 34(1): 129-137.

Evans, C., Davies, T.D., Murdoch, P.S., 1999. Component flow processes at four streams in the Catskill Mountains, New York, analysed using episodic concentration/discharge relationships. Hydrological Processes, 13(4): 563-575. DOI:10.1002/(sici)1099-1085(199903)13:4<563::aid-hyp711>3.0.co;2-n

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., Roubaty, J.L., 2017. The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data. Hydrol. Earth Syst. Sci., 21(12): 6153-6165.

Floury, P., Gaillardet, J., Tallec, G., Ansart, P., Bouchez, J., Louvat, P., Gorge, C., 2018. Chemical weathering and CO2 consumption rate in a multilayered-aquifer dominated watershed under intensive farming: The Orgeval Critical Zone Observatory, France. Hydrological Processes: 1-19.

Fovet, O., Ruiz, L., Hrachowitz, M., Faucheux, M., Gascuel-Odoux, C., 2015. Hydrological hysteresis and its value for assessing process consistency in catchment conceptual models. Hydrology and Earth System Sciences, 19(1): 105-123.

Garnier, J., Anglade, J., Benoit, M., Billen, G., Puech, T., Ramarson, A., Passy, P., Silvestre, M., Lassaletta, L., Trommenschlager, J.M., Schott, C., Tallec, G., 2016. Reconnecting crop and cattle farming to reduce nitrogen losses to river water of an intensive agricultural catchment (Seine basin, France): past, present and future. Environmental Science & Policy, 63: 76-90.

Garnier, J., Billen, G., Vilain, G., Benoit, M., Passy, P., Tallec, G., Tournebize, J., Anglade, J., Billy, C., Mercier, B., Ansart, P., Azougui, A., Sebilo, M., Kao, C., 2014. Curative vs. preventive management of nitrogen transfers in rural areas: Lessons from the case of the Orgeval watershed (Seine River basin, France). Journal of Environmental Management, 144: 125-134.

Godsey, S.E., Kirchner, J.W., Clow, D.W., 2009. Concentration-discharge relationships reflect chemostatic characteristics of US catchments. Hydrological Processes, 23(13): 1844-1864. DOI:10.1002/hyp.7315

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1): 80-91. DOI:https://doi.org/10.1016/j.jhydrol.2009.08.003

Hall, F.R., 1970. Dissolved solids-discharge relationships .1. Mixing models. Water Resources Research, 6(3): 845-&. DOI:10.1029/WR006i003p00845

Hem, J.D., 1948. Fluctuations in concentration of dissolved solids of some southwestern streams. Eos, Transactions American Geophysical Union, 29(1): 80-84. DOI:10.1029/TR029i001p00080

Hem, J.D., 1970. Study and interpretation of the chemical characteristics of natural water. US Government Printing Office, 363 pp.

Hubert, P., Martin, E., Meybeck, M., Oliver, P., Siwertz, E., 1969. Aspects hydrologique, géochimique et sédimentologique de la crue exceptionelle de la Dranse du Chablais du 22 sept. 1968. Arch. Sci. Soc. Phys. (Genève), 22(3): 581-603.

Johnson, N.M., Likens, G.E., Bormann, F.H., Fisher, D.W., Pierce, R.S., 1969. A working model for variation in stream water chemistry at Hubbard-Brook-experimental-forest, new-hampshire. Water Resources Research, 5(6): 1353-&. DOI:10.1029/WR005i006p01353

Jones, C.S., Wang, B., Schilling, K.E., Chan, K.-s., 2017. Nitrate transport and supply limitations quantified using high-frequency stream monitoring and turning point analysis. Journal of Hydrology, 549: 581-591. DOI:https://doi.org/10.1016/j.jhydrol.2017.04.041

Kirchner, J.W., 2019. Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: theory and benchmark tests. Hydrology & Earth System Sciences, 23(1).

Kirchner, J.W., Feng, X., Neal, C., Robson, A.J., 2004. The fine structure of water-quality dynamics: the (high-frequency) wave of the future. Hydrological Processes, 18(7): 1353-1359.

Ladson, A., Brown, R., Neal, B., Nathan, R., 2013. A standard approach to baseflow separation using the Lyne and Hollick filter. Australasian Journal of Water Resources, 17(1): 25-34.

Lloyd, C.E.M., Freer, J.E., Johnes, P.J., Collins, A.L., 2016. Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments. Science of the Total Environment, 543: 388-404. DOI:10.1016/j.scitotenv.2015.11.028

Longobardi, A., Loon, A.F.V., 2018. Assessing baseflow index vulnerability to variation in dry spell length for a range of catchment and climate properties. Hydrological Processes, 32(16): 2496-2509. DOI:10.1002/hyp.13147

Longobardi, A., Villani, P., Guida, D., Cuomo, A., 2016. Hydro-geo-chemical streamflow analysis as a support for digital hydrograph filtering in a small, rainfall dominated, sandstone watershed. Journal of Hydrology, 539: 177-187.

Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modelling, Institute of Engineers Australia National Conference, pp. 89-93.

Maher, K., 2011. The role of fluid residence time and topographic scales in determining chemical fluxes from landscapes. Earth and Planetary Science Letters, 312(1): 48-58. DOI:https://doi.org/10.1016/j.epsl.2011.09.040

Markewitz, D., Lamon, E.C., Bustamante, M.C., Chaves, J., Figueiredo, R.O., Johnson, M.S., Krusche, A., Neill, C., Silva, J.S.O., 2011. Discharge–calcium concentration relationships in streams of the Amazon and Cerrado of Brazil: soil or land use controlled. Biogeochemistry, 105(1): 19-35. DOI:10.1007/s10533-011-9574-2

Mathevet, T., Michel, C., Andreassian, V., Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. IAHS PUBLICATION, 307: 211.

Meybeck, M., Moatar, F., 2012. Daily variability of river concentrations and fluxes: indicators based on the segmentation of the rating curve. Hydrological Processes, 26(8): 1188-1207.

Miller, M.P., Tesoriero, A.J., Hood, K., Terziotti, S., Wolock, D.M., 2017. Estimating Discharge and Nonpoint Source Nitrate Loading to Streams From Three End-Member Pathways Using High-Frequency Water Quality Data. Water Resources Research.

Minaudo, C., Dupas, R., Gascuel-Odoux, C., Roubeix, V., Danis, P.-A., Moatar, F., 2019. Seasonal and event-based concentration-discharge relationships to identify catchment controls on nutrient export regimes. Advances in Water Resources, 131: 103379. DOI:https://doi.org/10.1016/j.advwatres.2019.103379

Moatar, F., Abbott, B., Minaudo, C., Curie, F., Pinay, G., 2017. Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. Water Resources Research, 53(2): 1270-1287.

Moatar, F., Meybeck, M., 2007. Riverine fluxes of pollutants: Towards predictions of uncertainties by flux duration indicators. Comptes Rendus Geoscience, 339(6): 367-382. DOI:https://doi.org/10.1016/j.crte.2007.05.001

Mouchel, J.-M., Rocha, S., Rivière, A., Tallec, G., 2016. Caractérisation de la géochimie des interfaces nappe-rivière du bassin des Avenelles. Tech. rep. PIREN Seine, France, 27 pp.

Mouhri, A., Flipo, N., Réjiba, F., De Fouquet, C., Bodet, L., Kurtulus, B., Tallec, G., Durand, V., Jost, A., Ansart, P., 2013. Designing a multi-scale sampling system of stream–aquifer interfaces in a sedimentary basin. Journal of Hydrology, 504: 194-206.

Musolff, A., Schmidt, C., Selle, B., Fleckenstein, J.H., 2015. Catchment controls on solute export. Advances in Water Resources, 86: 133-146. DOI:10.1016/j.advwatres.2015.09.026

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology, 10(3): 282-290.

Nathan, R., McMahon, T., 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research, 26(7): 1465-1473.

Pinder, G.F., Jones, J.F., 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. Water Resources Research, 5(2): 438-445.

Probst, J., Bazerbachi, A., 1986. Solute and particulate transports by the upstream part of the Garonne river. Sciences Géologiques Bulletin: 79-98.

Probst, J.L., 1985. Nitrogen and Phosphorus exportation in the Garonne basin (France). Journal of Hydrology, 76(3-4): 281-305. DOI:10.1016/0022-1694(85)90138-6

Rose, L.A., Karwan, D.L., Godsey, S.E., 2018. Concentration–discharge relationships describe solute and sediment mobilization, reaction, and transport at event and longer timescales. Hydrological Processes, 32(18): 2829-2844. DOI:10.1002/hyp.13235

Rusjan, S., Brilly, M., Mikoš, M., 2008. Flushing of nitrate from a forested watershed: an insight into hydrological nitrate mobilization mechanisms through seasonal high-frequency stream nitrate dynamics. Journal of Hydrology, 354(1): 187-202.

Salmon, C.D., Walter, M.T., Hedin, L.O., Brown, M.G., 2001. Hydrological controls on chemical export from an undisturbed old-growth Chilean forest. Journal of Hydrology, 253(1): 69-80. DOI:https://doi.org/10.1016/S0022-1694(01)00447-4

Saraiva Okello, A.M.L., Uhlenbrook, S., Jewitt, G.P., Masih, I., Riddell, E.S., Van der Zaag, P., 2018. Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid mesoscale catchment. Hydrological Processes, 32(10): 1334-1350.

Snyder, F.F., 1939. A conception of runoff-phenomena. Eos, Transactions American Geophysical Union, 20(4): 725-738.

Stewart, M., Cimino, J., Ross, M., 2007. Calibration of base flow separation methods with streamflow conductivity. Groundwater, 45(1): 17-27.

Tallec, G., Ansard, P., Guérin, A., Delaigue, O., Blanchouin, A., 2015. Observatoire Oracle [Data set]. DOI:https://dx.doi.org/10.17180/obs.oracle

Tallec, G., Ansart, P., Guérin, A., Derlet, N., Pourette, N., Guenne, A., Delaigue, O., Boudhraa, H., Loumagne, C., 2013. Introduction. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 11-33.

Thompson, S., Basu, N., Lascurain, J., Aubeneau, A., Rao, P., 2011. Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients. Water Resources Research, 47(10).

Tunqui Neira, J.M., Andréassian, V., Tallec, G., Mouchel, J.M., 2019. A two-sided affine power scaling relationship to represent the concentration–discharge relationship. Hydrol. Earth Syst. Sci. Discuss., 2019: 1-15. DOI:10.5194/hess-2019-550

Vaughan, M.C.H., Bowden, W.B., Shanley, J.B., Vermilyea, A., Sleeper, R., Gold, A.J., Pradhanang, S.M., Inamdar, S.P., Levia, D.F., Andres, A.S., Birgand, F., Schroth, A.W., 2017. High-frequency dissolved organic carbon and nitrate measurements reveal differences in storm hysteresis and loading in relation to land cover and seasonality. Water Resources Research, 53(7): 5345-5363. DOI:10.1002/2017WR020491

Vogel, R.M., Kroll, C.N., 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. Water Resources Research, 28(9): 2451-2458.

West, T.O., McBride, A.C., 2005. The contribution of agricultural lime to carbon dioxide emissions in the United States: dissolution, transport, and net emissions. Agriculture, Ecosystems & Environment, 108(2): 145-154. DOI:https://doi.org/10.1016/j.agee.2005.01.002

Williams, G.P., 1989. Sediment concentration versus water discharge during single hydrologic events in rivers. Journal of Hydrology, 111(1): 89-106.

Zhang, J., Zhang, Y., Song, J., Cheng, L., 2017. Evaluating relative merits of four baseflow separation methods in Eastern Australia. Journal of Hydrology, 549: 252-263.

Zhang, Q., Harman, C.J., Ball, W.P., 2016. An improved method for interpretation of riverine concentration-discharge relationships indicates long-term shifts in reservoir sediment trapping. Geophysical Research Letters, 43(19): 10215-10224. DOI:10.1002/2016gl069945

Zhang, R., Li, Q., Chow, T.L., Li, S., Danielescu, S., 2013. Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method. Hydrological Processes, 27(18): 2659-2665.

Chapter 4: Identification and quantication of the end-members

Avant-propos

High-frequency measurements give us an accurate and detailed record of hydrochemical interactions in the catchment. Therefore, it is not absurd to consider that with this information we could identify and quantify the chemical end-members that produce them. The principal aim in this chapter is to develop a methodology for identifying and quantifying end-members, analyze their temporal variability and their relationship to the different flow regimes.

.

Identification of potential end members and their apportionment from downstream high- frequency chemical data

José Manuel Tunqui Neira[1, 2], Jean-Marie Mouchel [2], Gaëlle Tallec[1] & Vazken Andréassian[1]

(1) Irstea, HYCAR Research Unit, Antony, France

(2) Sorbonne Université, CNRS, EPHE, UMR Metis 7619, Paris, France

Draft submission to Water Resources Research journal

Abstract

The high frequency chemical measures measured in the stream give us detailed information on the different hydrochemical processes in the basin. Within this information can also be found the sources (end-members) of whose mixtures are produced the concentrations of the stream. For this reason we propose here the development of a model, called IQEA (Identification and Quantification of End-members and their Apportionment). This model, only using downstream high-frequency chemical data and without any previous hypothesis on the end-members, allow us to identify them, quantify them and determine their contribution to the formation of chemical concentrations of the stream. In this article we present the proposed methodology of the model, the sensitivity analysis of the model and the comparison between the IQEA calculated end-members and pre-identified possible end-members (field data).

Keywords

End-member, high-frequency data, cluster analysis, end-member quantification, nitrates, calcium,

1 Introduction

Where do the measured streamwater chemical concentrations come from? This is the question has both intrigued and fascinated researchers on this topic since it has been possible to measure the chemical concentrations in streams (Lenz and Sawyer, 1944). Knowledge of water sources (from now on “end-members”) is especially important because it allows us to know in detail the hydrochemical behavior of a catchment and consequently, all related and/or dependent hydrochemical processes such as pollution, biochemical processes, hydrochemical modeling, habitat conservation, etc. can be

studied, understood and managed in an efficient manner(Genereux et al., 1993; Liu et al., 2008; Miller et al., 2017). The different chemical signatures of the streamwater are chemical contributions from different end-members that are collected from pools (Chanat et al., 2002; Evans and Davies, 1998; Hrachowitz et al., 2016; Miller et al., 2017; Probst, 1985) that are spatially distributed across the catchment. These pools include near surface (runoff) and deep groundwater pools, such as soils and fractured bedrock aquifers (Dwivedi et al., 2019). These pools are generally connected to each other. These connections allow the chemical solutes collected from the end-members to interact and mix with each other(Genereux et al., 1993). These interactions and mixing process are subject to various factors: (i) the precipitation that provides exogenous chemical solutes to the runoff and soil pool (Neal and Kirchner, 2000; Pearce et al., 2015); (ii) bedrock geology that provides the endogenous chemical signatures of the groundwater and soil pool (Frisbee et al., 2013; Gaillardet et al., 1999); (iii) the hydrological regimes (wet, dry season) that determine whether chemical contributions to the stream water are produced by interaction and mixing of all the pools (Chanat et al., 2002; Rose et al., 2018) or is a function of merely selected number of them (Barco et al., 2008; Muñoz-Villers and McDonnell, 2012; Zhang et al., 2013); (iv) the catchment size which determines the residence time of the solutes in the pools (Ameli et al., 2017; Frisbee et al., 2011; McGuire and McDonnell, 2006) and (v) the human activities, which modify and/or disturb the chemical contributions of different end members (Botter et al., 2019; Garnier et al., 2016).

High frequency measurements provide a detailed record of the interactions between the different pools and both natural and external factors (Bowes et al., 2009; Duncan et al., 2017; Floury et al., 2017; Neal et al., 2012).

The identification and contribution of end members can be divided into four groups: the tracer mass balance (TMB), the multivariate statistical, PCA and end member analysis (EMMA) and Positive Matrix Factorization (PMF).

1.1 Formulation of the problem and main resolution techniques

Assuming the strong basic hypothesis that the number of end-members is finite, that their composition is constant and that chemicals are conservative inside the hydrological system after coming out from the end-member waters, the mixing equation states that the composition of any so-called mixed sample in the hydrological system studied results from the mixing of an unknown proportion of the end-members. It can be written as conservation equations for all chemicals in all samples:

Eq. (45)

$$C_{ij} = \sum (D_{ik}.\delta_{kj})$$

Where the matrix D stands for the composition of end-member waters (*i* is an index for the chemical components, *k* an index for the potential end-members or sources), matrix δ stands for the mixing ratios of mixed samples (*j* is an index for the mixed samples), and matrix C stands for the composition mixed samples. An additional equation is the conservation of water, it states than the sum of mixing ratios for any sample must be 1. Moreover, all contributions of any end-member to any mixed sample, namely each of the terms of δ must be inside the [0,1] interval, which results in a set of inequalities. These results to simple inequalities (inside [0,∞]) if the sum of mixing ratio is set to 1. It is also expected that all terms in matrix C and D are positive.

The set of unknown depends on the question posed, partial information may be available for some of the unknowns and errors terms can be added to any available value. A first question, which is related to the way the system should be represented, is simply "what is the number of end-member waters (n_e) to consider". Since the basic hypothesis is wrong in most situations, the mixing model itself is only a, possibly poor, representation of the real system, and depending on the objective of the study, the available data and the data themselves the optimal or chosen n_e may change. Once the number of end-members is known and once their composition has been defined by additional analysis or considerations, a second problem, where all mixed samples can be treated independently, is to derive the values of the δ for each of them. This problem is a linear problem which can be over-defined or over-defined when the number of end-members (n_{EM}) is strictly larger than the number of species (n_S) plus 1 but under-defined in the opposite situation, with an additional complexity due to the need to respect inequalities. A third more complex and nonlinear problem can be derived when the end-member composition itself is unknown or only partly known. Since end-member waters and their composition are generally poorly defined this situation would be very common in practical application.

1.2 Identification and contribution of end members methods

1.2.1 Tracer mass balance methods (TMB)

Of all the methods cited, the TMB method is the one that has been most used. The TMB approach is an example of the second problem. It uses an approach based on a *n*-end-member mixing model (see detailed review of existing methods in Klaus and McDonnell, 2013), but in most cases a two-end-members mixing model has been used: an end-member representing the internal pools of the catchment (e.g. groundwater and soil water), often referred to as "old water", the other end-member representing the chemical concentrations from precipitation (runoff) or "new water"(Klaus and

McDonnell, 2013; Stewart et al., 2007). The TMB method has been applied with two types of chemical elements: environmental isotopes (such as tritium - ^{3}H, in disuse these days, Oxygen-18 - ^{18}O and deuterium - ^{2}H) and hydrochemical tracers (such as Chloride - Cl^{-}, Silicate - SiO_2, Sodium - Na^{+}, conductivity - EC, etc.), which can be used separately or together. The TMB method is based on five assumptions (according to Buttle, 1994; Klaus and McDonnell, 2013): (i) the isotopic/tracer content of the event and the pre-event water are significantly different, (ii) the event water end-members maintains a constant isotopic/tracer signature in space and time, (iii) the isotopic/tracer signature of the pre-event water is constant in space and time, (iv) contributions from the vadose zone is negligible, or the isotopic/tracer signature of the soil water is similar to that of groundwater, (v) surface storage contributes minimally to streamflow. Although objective, these assumptions are very problematic to corroborate physically (Kirchner, 2019). Pinder and Jones (1969) were the first to apply the total sum of various solutes (bicarbonate, chloride, nitrate, sulfate, silicate, calcium, iron, magnesium, potassium and sodium) to separate the storm hydrograph into two end-members: groundwater and runoff. Hubert et al. (1969) did a similar separation using only one isotopic tracer, the tritium (^{3}H).

The TMB method was much used in storm events, where thanks to the chemical tracers, discharge can be defined as the contribution of 2 end members: the groundwater (generally called "old water") and recent precipitation (or “new water”). A most common assumption is that no end member other than the two described above can contribute to storm discharge; however, several studies (e.g. Chanat et al., 2002; Rose et al., 2018)have shown that soil water is an important contributor of storm discharge. Another typical assumption is that the samples of the end-members are representative (e.g. the sampled precipitation accurately reflects all precipitation) and the chemical tracers (generally isotopic) of the end members are constant through time. This assumption is erroneous, since even in small catchments, the isotope composition of rainfall and groundwater varies considerably spatially and temporally depending on the magnitude of the storm event. (e.g. McDonnell et al., 1991; von Freyberg et al., 2017)

The concentrations of the chemical constituents are related to actual physical processes and flowpaths in the catchment that generate the different flow components (Stewart et al., 2007). Nevertheless, the TMB method frequently violates some of the five underlying assumptions, and more particularly that of conservation even with isotopes (Buttle, 1994; Klaus and McDonnell, 2013) We should also mention, a lack of detailed studies of the uncertainties (Bansah and Ali, 2017; Genereux, 1998) and last the TMB methods are very expensive in terms of fieldwork and it is very difficult their application in large basins or long-term studies (Lott and Stewart, 2016; Stewart et al., 2007).

1.2.2 Multivariate statistical methods (MS)

Several data analysis methods can be used to better assess the end-members from all the concentrations measured in the streamwater. The most used are: Cluster Analysis (CA) and discriminant analysis (DA).

Cluster analysis (CA), also called unsupervised pattern recognition method, encompasses a wide range of techniques for exploratory data analysis. The main aim of CA is grouping objects (cases) into classes (clusters) so that objects within a class are similar to each other but different from those in other classes. The class characteristics are not known in advance, but may be determined from the data analysis. The results obtained are justified according to their value in interpreting the data and indicating patterns (Alberto et al., 2001; Singh et al., 2004; Singh et al., 2005; Vega et al., 1998).

Discriminant analysis (DA, also called supervised pattern recognition) provides statistical classification of samples, contrasting with the exploratory features of cluster analysis. We can use DA if we know in advance the membership of objects to particular groups or clusters (i.e., the temporal or spatial ownership of a water sample as determined from its monitoring time or station). We can construct a discriminant function for each group. Such a function represents a surface dividing our data space into regions (hopefully we have only one group in each region); then samples sharing common properties will be grouped into the same region with a decision boundary separating two or more groups (Alberto et al., 2001).

All these three MS methods were used to investigate groundwater and surface water chemical data and identify natural phenomena (Blake et al., 2016; Koh et al., 2016) as well as anthropogenic impacts affecting water quality (Barakat et al., 2016; Devic et al., 2014; Phung et al., 2015)

The principal disadvantage of MS methods is that their results (end-members compositions and their contributions) can be positive or negative. Since negative compositions and contributions do not actually exist, it is difficult to quantitatively assess the relative contributions from the end-members identified (Li et al., 2019; Zanotti et al., 2019). Another important problems seen in MS methods are that they put the same weight on all the chemical components, although their real contribution to the streamwater is negligible (Alberto et al., 2001; Singh et al., 2004) and they are data-sensitive techniques and usually require a previous univariate analysis of the dataset. In fact, solutes streamwater datasets are often characterized by skewed distributions and normalization procedures have to be applied (Comero et al., 2011). All these problems have a considerable impact on the results.

1.2.3 Principal Component Analysis (PCA) and End member mixing analysis (EMMA)

Principal Component Analysis (PCA) extracts the eigenvalues and eigenvectors from the covariance matrix of original variables. The Principal Components (PC) is the uncorrelated (orthogonal) variables,

obtained by multiplying the original correlated variables with the eigenvector, which is a list of coefficients (loadings or weightings). Thus, the PCs are weighted linear combinations of the original variables. PC provides information on the most meaningful parameters which describe whole data set affording data reduction with minimum loss of original information (Singh et al., 2004; Vega et al., 1998). It is a powerful technique for pattern recognition that attempts to explain the variance of a large set of inter-correlated variables and transforming into a smaller set of independent (uncorrelated) variables (principal components)(Singh et al., 2004).

EMMA (Christophersen and Hooper, 1992; Christophersen et al., 1990; Hooper, 2003; Liu et al., 2008; Liu et al., 2004) is a PCA hydrological mixing model which principal aims is identifying the minimum number of end-members required to explain the variability of measured concentrations and collinearity (Tubau et al., 2014). EMMA is based on an analytical approach that solves an overdetermined set of equations on the basis of a least squares procedure (Barthold et al., 2011). EMMA uses the following assumptions (according to Barthold et al., 2011): (i) the stream water is a mixture of end-member with a fixed composition, (ii) the mixing process is linear and relies completely on hydrodynamic mixing, (iii) the solutes used as tracers are conservative, and (iv) the end-member have extreme concentrations. EMMA allows reducing the dimensionality of the analysis because visualization and interpretation of a large number of chemical analyses of many species is particularly difficult. The problem is much easier to handle if, instead of studying a dataset potentially as large as the number of species, it would suffice to analyze their projections on a space of much smaller (say 2 or 3) dimensions (Pelizardi et al., 2017). An EMMA aims at identifying the minimum number of end-members required to explain the variability of measured concentrations in time or space. The explanation of the variance and the contribution of each species to the mixture are obtained from the analysis of information provided by the calculation of the eigenvalues (Tubau et al., 2014). According to Tubau et al. (2014) the methodological guidelines for the selection of end-members and species are as follows: (i) initiate EMMA with all species and take further steps that will eliminate the species with the highest weight in the eigenvalue. (ii) a common rule is to retain all eigenvectors associated with eigenvalues (n) greater or equal to 1, called the “rule of 1”(Hooper, 2003). This rule consists of selecting the eigenvalues that explain variance greater than [1/number of species]. (iii) the selection of species according to those eigenvalues that we want to explain. The number of end-members will be n + 1. (iv) the projection of the observation points in the space defined by the eigenvectors associated with each eigenvalue. With this projection, it is possible to make a preliminary selection of end-members and an analysis of the points that fall outside the area defined by these end-members.

As in the case of the MS method, the main problem with EMMA is that the outcomes (end-members compositions) can be positive or negative. Other disadvantage is that the concentrations of end-members must be different and accurately known (Vázquez-Suñé et al., 2010). Last EMMA does not

explicitly compute mixing contributions of end-members. Thus, we need to use complementary tools to calculate them, most used is the MIX tool (Carrera et al., 2004; Vázquez-Suñé et al., 2010) described in more detail below.

1.2.4 Positive Matrix Factorization method (PMF)

Positive Matrix Factorization (PMF) was first introduced by Paatero and co-workers (Paatero, 1999; Paatero and Tapper, 1994) who justified the importance of positive factorization for physical applications. Several publications by this group (Anttila et al., 1995; Hopke and Paatero, 1994; Polissar et al., 1998, and additional publications in following years) were devoted the composition of atmospheric aerosols, and this technique was then considerably used in the air pollution domain to perform end-member apportionment of particulate matter (see detailed review by Reff et al., 2007). The US-EPA distributes today a 5th version (PMF-5, Norris et al., 2014) of the software-package based on work by Paatero and co-workers. Only few and much more recent applications were done in the domain of hydrochemistry (Capozzi et al., 2018; Haji Gholizadeh et al., 2016; Zanotti et al., 2019). PMF answers to the third problem, namely find positive D and δ to explain the composition of mixed water samples, with an as low as possible number of end-members, D and δ are the unknown, C are measured with errors that can be estimated. Whatever the algorithm chosen to solve the problem, an objective function has to be minimized, it is generally constructed as the sum of square of deviations between model results and measured mixed water compositions, and preferably weighted with the measurement variances to assuming Gaussian errors to construct a maximum likelihood estimator. Renewed applications of the PMF method (also called NMF, for non-negative matrix factorization) occur in the large domain of artificial intelligence with another most-cited paper by Lee and Seung (1999) who introduced NMF for image recognition and text analysis also introducing other objective function. Recent applications also concern bioinformatics (e.g. Gaujoux and Seoighe, 2010). Whatever the algorithm chosen, an objective function has to be minimized, Paatero and co-workers suggest that it should be constructed as the sum of square of deviations between model results and measured mixed water compositions, and preferably weighted with the measurement variances to construct a maximum likelihood estimator assuming Gaussian errors. Developments are still on-going, motivated by new very large data sets in bio-informatics and deep learning to reduce the dimension of data sets to the effect of a small number of factors (Gaujoux and Seoighe, 2010).

A major problem with PMF is that the solution is not unique. In the case of end-member identification this problem can easily be illustrated in Figure 44 which shows 2 groups of end-members that can equally be used to decompose the set of mixed water samples. More generally, in the illustrated case of 3 end-members, it is clear that any set of end-members (S'_1, S'_2, S'_3) which can be used to positively factorize a primary set (S_1, S_2, S_3) is another equally good solution. An additional problem, which is

common to most minimization problems and generally resolved by several restarts of seeding of the procedure, is that the function to optimize is not convex.

To solve the non-unicity problem, it is necessary to provide additional constraints to better define the solution (Berry et al., 2007). Various types on constraints can be added, typically quadratic functions of *D* and *δ* (Paatero and Tapper, 1994), other types of constraints are added when sparsity of the solution is expected (Berry et al., 2007).

In atmospheric sciences, the PMF method utilize statistical techniques to reduces the complexity of a set of data to meaningful terms to identify chemical end-members and estimate the contributions .Its goal is to resolve the mixture of end-members that contributes to particulate matter samples (Reff et al., 2007). It has been used to resolve primary end-members as well as secondary end-members (namely aerosols formed by reactions in the atmosphere)

The main advantages of this method compared to MS method are: (i) it takes into account the analytical uncertainties often associated with measurements of environmental samples and (ii) forces all of the values in the solution profiles and contributions to be positive, which can lead to a more realistic representation (Reff et al., 2007; Zanotti et al., 2019).

According to Zanotti et al. (2019) the main limitations of this method, which would lead to its poor development in the field of catchment hydrochemistry are: (i) PMF requires only data expressed as concentration while some typical measurement of water samples have different units (e.g. pH, Electrical Conductivity, Oxidation Reduction Potential, isotopes analysis, age tracers) thus they cannot be directly fed into a PMF model, (ii) a proper monitoring network, specifically designed to capture the variability of the hydrochemical system, is crucial to obtain a complete representation of the system with a PMF analysis and (iii) in cases where a single end-member is present a multivariate statistical analysis such as PMF might not be appropriate.

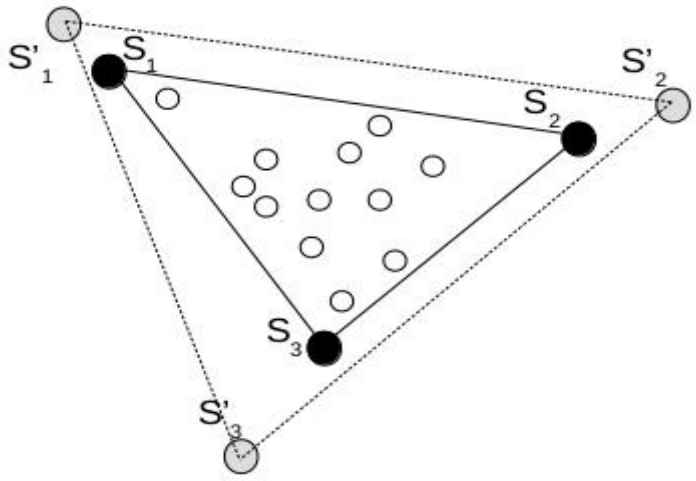


Figure 44 : Graphic description of the "non-unicity" problem of the end-members.

1.2.5 MIX method

Carrera et al. (2004) developed the MIX method to solve the above problem (third type) including positiveness constraints. The method MIX was developed for hydrology, with the main objective to evaluate groundwater inflow to surface water bodies and vice versa (recharge to groundwater from surface water). The methodology proposed by Carrera et al. (2004) assumes the existence of a known number of end-member waters, and looks for the mixing ratios that will better explain the composition of a set of mixed water samples. The major difference with the NMF/PMF technique is that additional information regarding the composition of the end-members is also given as postulate to the problem. It is assumed constant, but with a potentially large error which would mostly include conceptual modeling errors (e.g. the composition of the each real end member water may not be constant in reality, and the number of end members may also be different from what is assumed in the mixing model). Although Carrera et al. (2004) do not refer to the PMF/NMF literature, nor to applications in the field of atmospheric chemistry, the proposed additional postulates is a physically meaningful way to construct a quadratic function of *D* to ensure the unicity of the solution.

1.3 Scope of this paper

The main aim of this paper is to develop a methodology for identifying and quantifying end-members and their contribution to water chemistry from a purely chemical point of view. We named the method IQEA (Identification and Quantification of End members and their Apportionment)

A series of reasonable hypothesis drive to an objective function that was applied to the high-frequency chemical data set of the Oracle-Orgeval Observatory. The new method developed here, without any

preliminary assumption on the composition of the potential end-members, allows us analyzing the temporal variability of the end-member and their relationship to the different flow regimes

2 Material and method

2.1 Study site

The Avenelles catchment is located 70 km east of Paris, France. It is a sub-catchment of the Orgeval/ORACLE experimental catchment, managed by Irstea (https://gisoracle.irstea.fr/), covering an area of 46 km^2. The Orgeval/ORACLE observatory is one of the most instrumented and documented site in France, with more than 50 years of hydro-chemical record (Tallec et al., 2013) available from the ORACLE website (https://bdoh.irstea.fr/ORACLE/). The Avenelles catchment is homogeneous in terms of topography (see Figure 45a). The climate is temperate and oceanic, with annual average temperature of 11 ± 1 °C and a mean annual rainfall of 674 ± 31 mm. The average measured streamflow at the Avenelles outlet is about 0.2 m^3/s (1962 - 2017), with minimum flows in the summer (~0.1 m^3/s) and floods up to 10 m3/s in the spring (Tallec et al., 2013). The Avenelles catchment is covered with a 10-m silt layer, under which two tertiary aquifer formations are separated by a discontinuous grey clay layer (Figure 45b) (Garnier et al., 2014) : the shallower aquifer of the Brie limestone Oligocene formation, and the deeper Champigny limestone Eocene aquifer. The Avenelles streamflow is fed mostly by the Brie aquifer and by the drainage of the plateaus (Mouhri et al., 2012), with water chemistry dominated by $[Ca]^{2+}$, $[Mg]^{2+}$,$[SO_4]^{2-}$, et $[NO_3]^-$ ions (Floury et al., 2017). The land use (Figure 45c) is agricultural with few villages. The crops are essentially cereals (wheat, maize, barley), with conventional (i.e. intensive) farming practices, mainly based on mineral nitrogen fertilization (Garnier et al., 2016; Garnier et al., 2014).

The chemical concentrations are measured continuously at the Avenelles outlet, through a new concept of *in situ* laboratory (see Figure 45a), the "River Lab"(Floury et al., 2017). . The River Lab pumps and filters river water to feed a set of physical-chemical probes and ion chromatography (IC) instruments. The concentrations of the main dissolved species ($[Mg]^{2+}$, $[K]^+$, $[Ca]^{2+}$, $[Na]^+$, $[Sr]^{2+}$, $[F]^-$, $[SO_4]^{2-}$, $[NO_3]^-$, $[Cl]^-$ and $[PO_4]^{3-}$) are measured every 30 minutes (in average) and electrical conductivity (EC) each minute.

Streamflow is measured continuously by a gauging station at the Avenelles outlet (near River Lab see Figure 45a) and stored in French national database Banque Hydro (http://hydro.eaufrance.fr/). For our study we used half-hourly (30 min) flow data, as for the chemical measurements time step. The main data set (flow rates and chemical concentrations) covers the period between June 2015 and July 2017, i.e. 17300 measurements over 26 months (see

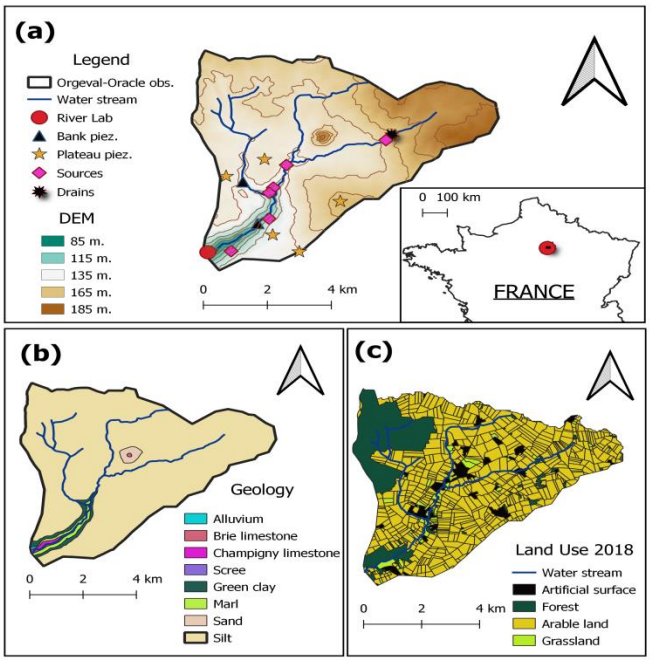


Figure 45 : a) Catchment localization; b) Geology and c) Land use

Table 22 : High frequency measurements of chemical concentrations (average, minimum and maximum values) from the River Lab at the Avenelles station (June 2015 - July 2017).

item	Unit	Avenelles Catchment		
		Mean	Min	Max
chloride	$mg.L^{-1}$	30	4	40
sulfate	$mg\text{-}S.L^{-1}$	19	2	32
magnesium	$mg.L^{-1}$	9	3	15
sodium	$mg.L^{-1}$	13	2	17
nitrates	$mg\text{-}N.L^{-1}$	12	3	18
calcium	$mg.L^{-1}$	124	37	202
EC	$\mu S.cm^{-1}$	704	267	1015

Last, we could also use field concentrations measurements from: (i) bank piezometers, (ii) agricultural drains that collect the superficial aquifer, (iii) deep piezometers (so-called plateau piezometers) which are all within the Brie aquifer, (iv) springs fed by the Brie aquifer**.** These measurements were made from January 2016 to April 2018, in the framework of the PIREN Seine program (see Mouchel et al., 2016; Mouhri et al., 2013). In Table 23 we summarize the field measurements dataset.

Table 23 : Chemical concentrations (median and standard deviation) measured in hydrological compartments in the framework of the PIREN Seine program (January 2016-April 2018) and at the Avenelles outlet by the River Lab (June 2015 - July 2017), from the ORACLE-Orgeval Observatory. Rainfall concentrations were taken from Floury et al. (2018)

		stream (River Lab)	banks	drains	plateau	springs	Rainfall
	unit	median ± sd					
Chloride	mg L^{-1}	30 ± 5	33 ± 12	32 ± 7	35 ± 8	34 ± 5	1.60
Sulfate	S-mg L^{-1}	19 ± 4	11 ± 26	10 ± 2	16 ± 50	18 ± 8	0.38
Magnesium	mg L^{-1}	9 ± 1	7 ± 6	6 ± 1	6 ± 2	8 ± 1	0.19
Sodium	mg L^{-1}	13 ± 2	12 ± 3	15 ± 3	18 ± 5	12 ± 4	0.75
Nitrate	N-mg L^{-1}	12 ± 1	4 ± 5	14 ± 3	13 ± 4	15 ± 3	0.75
Calcium	mg L^{-1}	124 ± 13	133 ± 31	67 ± 7	129 ± 60	145 ± 11	1.56
EC	µS m^{-1}	704 ± 76	711 ± 152	465 ± 77	742 ± 204	794 ± 62	NA

2.2 Data set processing

To test the methodology, we have chosen 2 ions the calcium and nitrate among the 5 ions and EC measured by River-Lab (see Table 22),. These two ions are the most chemically representative of our catchment (Floury et al., 2017), and they also have a marked seasonality variation with respect to the other ions (Floury et al., 2018; Tunqui Neira et al., 2019), which represents in our opinion a supplementary advantage to test the efficiency of our model.

Considering the large size of the dataset (~17500 samples), it is computationally complicated to work with all data simultaneously. For this reason that firstly, the data set has been subdivided into monthly samples. However, even with this partition, the quantity of concentration data continues to be important (700 concentration samples on average). Therefore, in order to improve the calculation time, we proposed the following approach: for each month we take 10 random samples, each containing 100 pairs of measurements of calcium and nitrate concentration (see Figure 46). In addition to improve the calculation time, this approach allows us to evaluate the evolution of potential source composition on a monthly scale and asses the robustness of the method by replicate computations. Finally, in order to avoid possible bias caused by the different orders of magnitude between the two ions studied, (the ratio of average calcium concentration over average nitrate concentration is 5), both concentrations where standardized, i.e. the measured concentrations were divided by their median:

$$C_s = \frac{C_j}{\widehat{C_j}}$$

C_s : Standardized concentration (-)

C_j : Stream water concentration (mg.L^{-1})

$\widehat{C_j}$: Median of the stream water concentration (mg.L^{-1})

For calcium the median is 123 mg.L^{-1} and for nitrates is 12.5 mg.L^{-1}

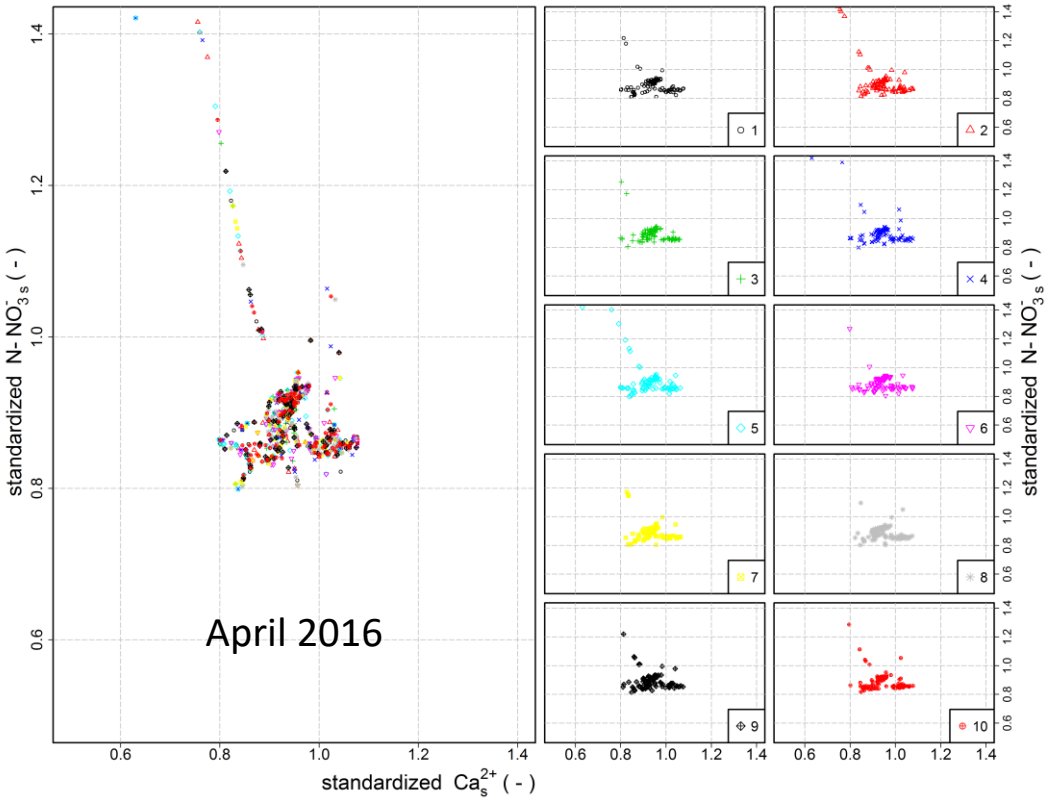


Figure 46 : On the left the cloud of the standardized concentration measured in April 2016 at the River-Lab station. On the right, the disaggregation of the dot cloud into 10 different random samples, each containing 100 pairs of concentrations of the calcium and nitrate ions.

2.3 Methodology

2.3.1 Optimization function

To compute the parameters of the IQEA model, namely the composition of the end members and the apportionment in each of the samples, a series of starting points and hypothesis must be set. To simplify the problem we decide to work in 2 dimensions, since working in one dimension reduces to apportioning into two end-members. The one-dimensional situation was already discussed in a previous publications regarding the Avenelles catchment and introducing baseflow separation techniques to better define the problem (Chapter 1 and 3 of part II of this thesis). In 2 dimensions, at least 3 end-members should be searched, since searching 2 end-members only would implicitly mean that the problem can be reduced to a one dimensional problem by some transformation of the

variables that could be obtained by PCA for instance. Introducing more than 3 end-members introduces a high level of uncertainty in the apportionment of each of the mixed samples since an infinite number of equally correct apportionments would be possible for each of the samples. Accordingly, 3 end-members exactly were used to solve the problem. Numerous researchers also proceeded to end-member characterization. The choice of 3 end-members for hydrological end-member apportionment in small catchments was also made in the past for geochemical and hydrological reasons (e.g. Genereux et al., 1993; Miller et al., 2017; Probst, 1985) to separate three main end-members, namely groundwater, soil water and runoff (or precipitation), while our basic argument in this paper is to focus on the most simple but not trivial problem. For easier identification of the end-members and comparison to field data (Mouchel et al., 2016) for a further interpretation of the results and avoid negativity it was decided to work with real chemical components with no previous PCA that could identify two linear combination of components as optimal factors to proceed with the analysis.

To avoid the non-unicity problem which is characteristic of PMF problems (see Figure 44) an additional criterion was introduced to limit the size of the set of end-members. This criterion can be represented by the variance of the *k* = 3 end members:

$$\frac{1}{k}.\sum_{k,j}\left(D_{kj} - G_j\right)^2$$

Where G_j is the center of gravity of the end member composition.

Now, to search for a end-member set (D_{kj}) and their respective apportionment (δ_{kj}), we minimize a function (F) composed of two terms: the residual variance of the chemical concentrations of mixed samples after apportionment to the end-members and the variance of the set of end-member composition. The ratio of two weights affected to each of the terms in the sum must be defined to fully characterize the objective function.

$$F = A'.\left(\frac{1}{n}\right).\sum_{i,j}\left(C_{ij} - \sum_{k}\left(D_{ik}.\delta_{kj}\right)\right)^2 + B'.\left(\frac{1}{k}\right).\sum_{k,j}\left(D_{kj} - G_j\right)^2 \qquad \textbf{Eq. (46)}$$

Where the first term $\left(\frac{1}{n}\right)\sum_{i,j}\left(C_{ij} - \sum\left(D_{ik}.\delta_{kj}\right)\right)^2$ is the unexplained variance of the C_{ij}, with C_{ij} the measured concentration and *n* the number of measured chemical concentrations. The ratio A'/B' as to be chosen to fully define the objective function. The optimization problem also includes constraints:

$$\sum_{k}\delta_{kj} = 1; 0 \leq \delta_{kj} \leq 1; D_{ik} > 0$$

For easier writing, we define

$A = \frac{A'}{n}$ and

$$B = \frac{B'}{k}$$

We do not *a priori* information on the best values for constants *A* and *B*. To evaluate the quality of the estimation of the optimization of *F* (Eq. (46)), we used the ratio of the residual variance to initial variance of the data set (VR):

$$VR = \frac{\sum_{i.j}\left(C_{ij} - \sum_k \left(D_{ik}.\delta_{kj}\right)\right)^2}{\sum_{i.j}\left(C_{ij} - G'_j\right)^2} \quad \textbf{Eq. (47)}$$

Where G'_j is the center of gravity of the chemical dataset C_{ij}.

VR is equivalent to the term 1-R^2 in a regression. In this paper, we have chosen an acceptable threshold for the VR ratio of 0.05 as roughly estimated of the chemical measurements errors.

VR is dependent on the A/B ratio, whose a-priori values are unknown. For this reason the method was applied to 100 different A/B ratios ranging from 0 to 10 to each of the monthly random samples. Then, with the 100 results obtained of VR for each month from the optimization of Eq. (47), we established a power-law relationship between the ratio A/B and VR. From this relationship we obtain the value of A and B for a VR of 0.05. With these optimal values, we solve one again the optimization problem (Eq**.** (46**)**) to calculate the set of end members (D_{kj}) and their respective apportionment (δ_{kj}) for each month.

To solve the Eq. (46) we used the "solnp" function of the RSOLNP package (version 1.16) on R (https://cran.r-project.org/web/packages/Rsolnp/Rsolnp.pdf). The "solnp" algorithm is based on the non-linear optimization of the parameters using the augmented Lagrangian method (Ye, 1988). The main benefit of this algorithm in our work is that it allows optimizing at the same time 2 variables with their respective equality and inequality constraints (D_{kj} and δ_{kj}).

2.3.2 Sensitivity analysis of end members

The optimization of each of the 10 samples with VR=0.05 produces a set of 3 unordered sources, which must first be grouped before the reproducibility of the source characterization. A cluster analysis was performed or this purpose. . Then, an analysis of variance compares the variance in each cluster of estimated source to the total variance. The following determination coefficient (R^2) was used:

$$R^2 = 1 - \frac{\sum_k \sum_i d^2(D_{ik}, G_k)}{\sum_{g=1} d^2\left(D_{ig}, G_g\right)} \quad \textbf{Eq. (48)}$$

Where the term $\sum_k \sum_i d^2(D_{ik}, G_k)$ represents the sum of variances of each of the end members in the cluster with respect to the center of gravity of the cluster (G_k), and the term $\sum_{g=1} d^2\left(D_{ig}, G_g\right)$ represents the variance between the whole set of end members with respect to the center of gravity of the whole set of end members (G_g).

To further enhance the robustness of the cluster analysis, we have also calculated the R^2 of the set of end members calculated for a set of a variable VR ($0.02 < VR < 0.10$).

Last, we have compared (graphically) the resulting end members with VR =0.05 of each month with the field data (mean ± sd) shown in the Table 23.

3 Results and discussion

3.1 Relation between the criterion VR and A/B ratio

For the ten different samples of each month (22 months), the relationship between VR and A/B ratio show a well-fitted power-law relationship (e.g. in Figure 47 for the month of April 2016)

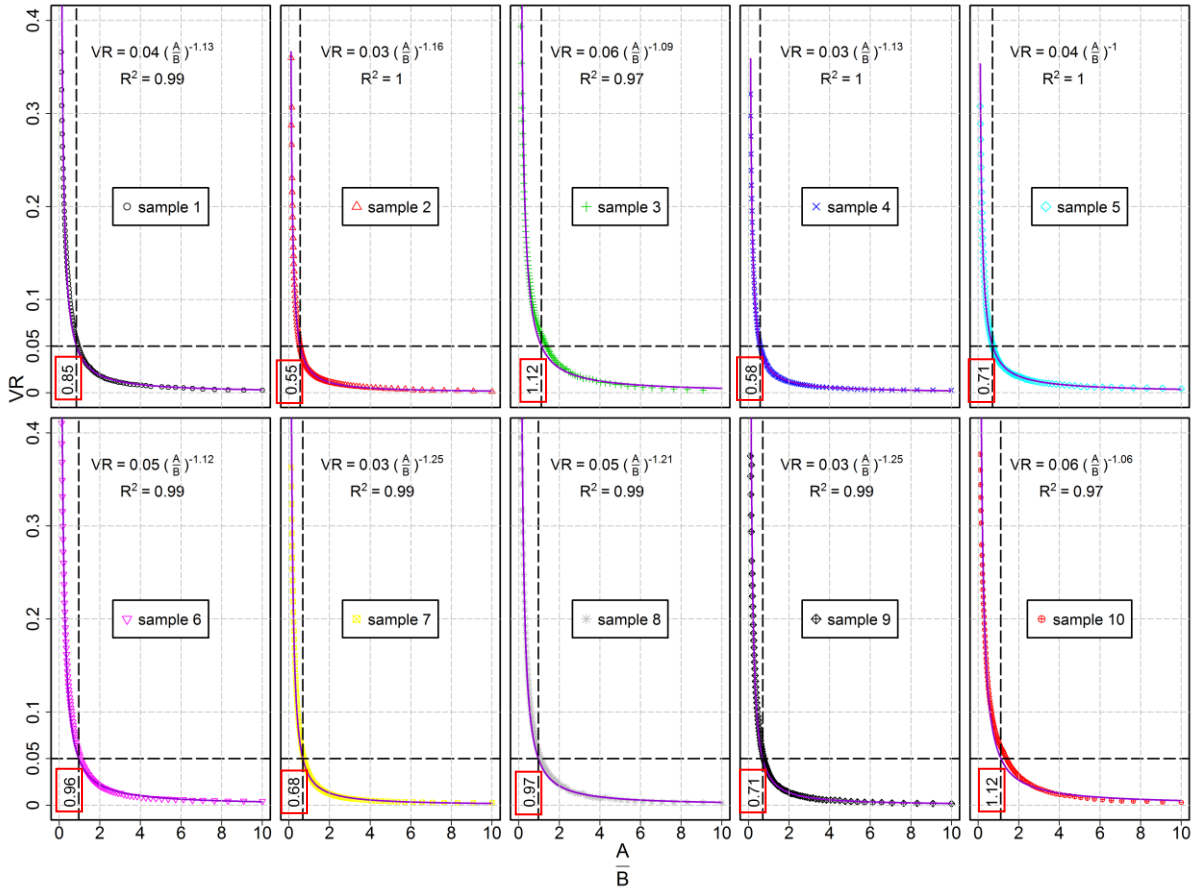


Figure 47: Relationship between the VR index and A/B ratio and for the 10 different samples of the month of April 2016. The numbers inside the red squares are the values obtained from the ratio A/B for a VR =0.05.

From these well-fitted power-law relationships, an ideal values for A and B (VR =0.05, red squares in the Figure 47) was computed. The optimization made with different values of A and B for a variable VR were also used later to analyze the robustness of the IQEA model. Figure 47 also illustrates that the different samples of the month (and also observed in the other months of the study) have different values of A and B (i.e. A/B ratio).

3.2 Average monthly concentrations of the potential end-members and their respective apportionment

For each month, we calculated the average value of each set of potential end-members (i.e. blue, red and green) shown in Figure 48. We have also calculated the average monthly apportionment given by each monthly potential end-member. The results are presented in. Table 24

Table 24 : Average monthly concentrations of the potential end-members and their respective apportionment calculated using the IQEA model

Month	Average ± sd end-member concentration for each month						Average end-member apportionment for each month (%)			Commentary/ predominant end-member
	Calcium (mgL^{-1})			Nitrate ($mgNL^{-1}$)						
	D_{blue}	D_{red}	D_{green}	D_{blue}	D_{red}	D_{green}	D_{blue}	D_{red}	D_{green}	
Jun-15	122 ± 1	117 ± 1	125 ± 0	12.8 ± 0.1	11.6 ± 0.1	12.5 ± 0.2	42.8	18.1	39.1	Blue
Jul-15	127 ± 2	111 ± 2	126 ± 2	12.8 ± 0.1	11.1 ± 0.4	11.7 ± 0.4	71.0	11.0	18.0	Blue
Aug-15	128 ± 0	103 ± 2	117 ± 2	12.6 ± 0.4	9.8 ± 0.1	11.1 ± 0.2	79.5	15.8	4.7	Blue
Sep-15	131 ± 1	96 ± 4	115 ± 5	13.4 ± 0	10.1 ± 0.3	11.6 ± 0.4	52.7	16.4	30.8	Blue
Oct-15	125 ± 0	99 ± 1	118 ± 4	12.8 ± 0.1	9.4 ± 0.2	11 ± 0.2	59.8	17.1	23.0	Blue
Nov-15	123 ± 4	90 ± 3	131 ± 0	12.7 ± 0.4	13.5 ± 0.3	10.3 ± 0.1	21.0	29.7	49.3	Green
Dec-15	103 ± 2	90 ± 3	119 ± 1	12.8 ± 0.1	10.5 ± 0.2	11.8 ± 0.1	37.6	11.2	51.2	Green
Jan-16	105 ± 2	99 ± 9	123 ± 1	13.1 ± 0.1	10.7 ± 0.7	13.1 ± 0.1	41.6	18.8	39.6	Blue
Feb-16	103 ± 1	79 ± 3	117 ± 1	12.0 ± 0.1	8.4 ± 0.3	10.9 ± 0.1	47.0	15.8	37.2	Blue
Mar-16	80 ± 3	95 ± 12	128 ± 1	12.2 ± 1.5	9 ± 0.3	10.9 ± 0.1	7.7	21.0	71.3	Green
Apr-16	95 ± 3	104 ± 2	128 ± 1	15.9 ± 0.4	10.3 ± 0.1	10.5 ± 0	10.8	35.5	53.7	Green
May-16	101 ± 20	66 ± 2	118 ± 11	14 ± 1.9	10.5 ± 0.6	10.9 ± 0.7	6.2	9.0	84.8	Green
Jun-16	105 ± 2	91 ± 3	158 ± 2	13.3 ± 0.2	9.1 ± 0.3	10.9 ± 0.1	19.9	24.5	55.6	Green
Jul-16	---	---	---	---	---	---	---	---	---	Insufficient data
Aug-16	124 ± 1	124 ± 0	129 ± 0	13.3 ± 0.1	12.6 ± 0	13.4 ± 0.1	32.4	27.2	40.5	Green
Sep-16	133 ± 0	120 ± 2	130 ± 0	12.8 ± 0	12.3 ± 0.2	13.8 ± 0.1	36.0	13.6	50.4	Green
Dec-16	---	---	---	---	---	---	---	---	---	Insufficient data
Mar-17	---	---	---	---	---	---	---	---	---	Insufficient data
Apr-17	111 ± 1	114 ± 0	119 ± 1	12 ± 0.2	9.8 ± 0.1	12 ± 0.1	40.1	13.2	46.7	Green
May-17	113 ± 4	110 ± 1	131 ± 1	15.3 ± 0.7	11.1 ± 0.1	11.3 ± 0.1	7.8	31.5	60.7	Green
Jun-17	150 ± 0	117 ± 4	130 ± 8	12 ± 0.1	11.1 ± 0.3	9.8 ± 0.2	48.2	33.9	17.9	Blue
Jul-17	147 ± 1	129 ± 1	127 ± 1	13.4 ± 0.1	13.4 ± 0.1	9.9 ± 0.2	42.5	15.4	42.1	Blue

The monthly average potential end-members for July, December, and March 2017 have not been calculated, since, for these months only a few sample points are available, thus the potential end-members calculated using our IQEA model are not representative.

Knowing that the main hydrological contributor of our catchment is groundwater (Floury et al., 2018; Mouchel et al., 2016; Tallec et al., 2013) seeing the alternation in the predominance of sources each month (Table 24) we can state that the green and blue potential end-members are representative of groundwater.

3.3 Sensitivity analysis of end members

Figure 48 illustrates the cluster analysis performed for the different set of k=3 end-members computed using the IQEA model for VR =0.05 from the 10 different samples of each month. In total for each month we have 30 end-members. The cluster analysis we have been able to group them (blue, red and green) according to their nearby position. The R^2 obtained from the cluster analysis for each is quite satisfactory, as almost all of them are proxy to 1.

To check the robustness of the cluster analysis, we also calculate the R^2 of the different end-members (k=3) for a variable VR (0.02<VR<0.10, see Figure 51 in Appendix 5.1). Although a decrease in the values of R^2, these remain high ($R^2 > 0.70$) thus we can affirm that the imposition of 3 potential sources for our 2 ions is well defined.

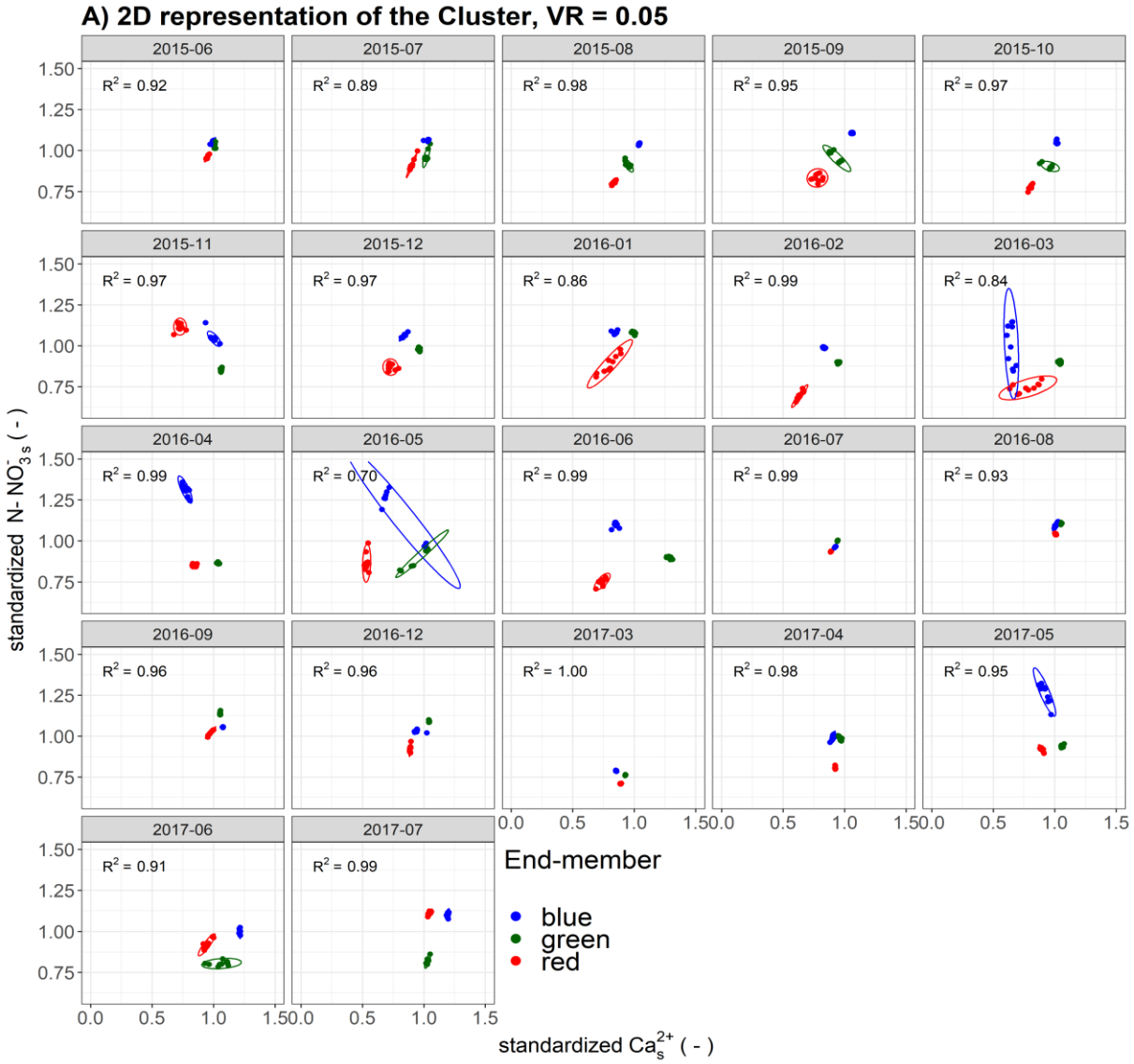


Figure 48 : Cluster analysis (ellipses show in the graph) and R^2 of the different end-member computed with the IQEA model for VR = 0.05.

3.4 Potential end-members to identify observed evolution in a synthetic manner

An apportionment could be computed for each of the downstream samples. It results in three fractions representing the estimated composition of the sample. Figure 48 shows the evolution of the estimated end-members along the two years studied period. The decomposition in potential end-members and the exploration of the time series of the dominant end-member computed for each sample enables a clear description of the major features of the coupled hydrology and chemistry in the catchment. Figure 49 shows the contribution of each of the potential end-members at every moment, as well as the dominant end-member. The time-series of dominant end-member enables to very clearly identify successive well defined periods, with a single highly dominant (or unique) end-member, which appear relevant from both the chemical and hydrological points of view, except in some specific situations

where the end-members are not precisely defined and may overlap (e.g. August 2016). The potential of this representation of the coupled hydrological and chemical behavior of the catchment using potential end-members is illustrated in the following paragraphs.

Between June and October 2015, our IQEA model indicates that the potential end-member that contributes most in the mixing of concentrations of the Avenelles stream the blue one; Then in the month of November there is a change and predominant potential end-member is the green one. This due to the fact that between October and November of every year it begins the water recharge of the Brie aquifer (Tallec et al., 2013).

During the first 5 months period (June to October 2015), quite similar compositions of end-members have been identified. It also appears that the three potential end-members are almost aligned which suggests that it might be delicate to properly define the contribution of each of them for a given sample. In June and July in particular the three potential end-members are positioned in a very narrow area. This situation is very much connected to the hydrology of the period with a low base flow (0.1 m^3s^{-1} or less) are rare higher discharge peaks (0.2 m^3s^{-1} in August and 0.5 m^3s^{-1} in September). Figure 49 also shows the dominant end-member for every water sample. It highly remarkable that the more diluted end-member, for calcium and nitrate shown as red on graphs, is dominant in all samples for most peak discharge conditions. This demonstrates a very strong relation between the major end-member as computed by the developed procedure and the hydrological features while both data sets (water quality and discharge) although they were treated independently. During this 5 months period, the intermediate end-member is rarely dominant, this mainly happen during post-peak discharge periods, in September, which was also the period with the highest peak discharge.

November 2015 is a transition period with two peak discharge periods with the same dominant end-member. However, unlike the previous period, the potential end-member which dominates both peak discharges is diluted regarding calcium but not regarding nitrate, on average, with complex concentration patterns during floods (see Figure 52 in Appendix-2). On the opposite, the major end-member during the low flow period has lower nitrate concentration than what was observed during the previous dry period (June to October). The following months (December 2015 to February 2016) turns back to a situation where all discharge peak appear more diluted.

Later on, the period March to June 2016 is highly different, with a quite rare event; the late-May flood had a 20 years return period in Paris (and also in the Orgeval catchment) and is very unusual in late spring. During this period, the peak flow correspond to a mixed domination of two potential end-members, always characterized by low calcium concentrations, but with more strongly fluctuating nitrate concentrations, which explains while two potential end-members are necessary to visualize these peak flow periods. Most of them reached 5 m^3s^{-1}, which was not observed during the previous

periods. The low water is characterized by higher calcium concentrations and average nitrate concentrations.

Following this period with an intense hydrological activity, the following period (July 2016 to April 2017) is a lower discharge period, with only one peak flow at 2 m^3s^{-1}, but with unfortunately no chemical data. During this last period, the situation is similar to the first 5 months period with low discharge and dilution process (diluted end-member) during the few recorded small peak discharge events. The rest of the time, during low flow periods, the concentrations are distributed between both, not much different, remaining end-members. The situation form July to December 2016 shows similar end-members to the previous (June to October 2016) low flow period.

Except for the months of January and February, according to our IQEA model, all the months of 2016 have the green end-member as the main contributor of the stream mixing concentrations. We believe that there has not been a change of predominant potential end-member this year because the 20 years return period flood (~12 m^3s^{-1}) produced in the catchment between late May and early June 2016. The enormous volume produced by this phenomenon was easily stored in the aquifer (paradox of the small catchment, see Kirchner, 2003), then this volume of water took the chemical characteristics of the green end-member (main contributor at that time) not seeing a change of end-member predominance until June 2017 (returns again the blue end-member to be the dominant, see Table 24).

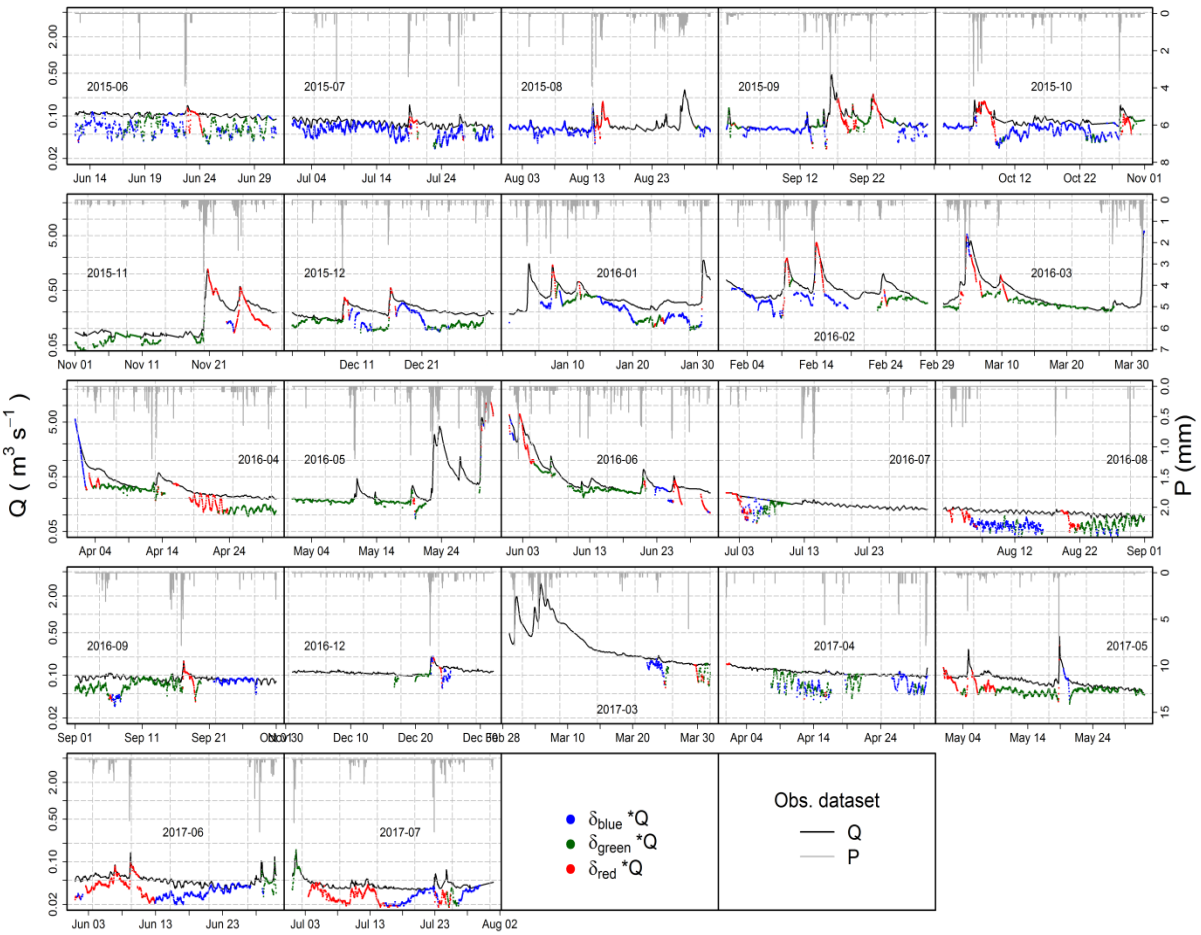


Figure 49 : Predominant end-member in the water sample (i.e. δ*Q)

3.5 Potential end-members versus pre-identified possible end-members

The potential end-members can interestingly be compared with possible end-members (ground water -GW-, spring water -SW-, tile drain water -DW-, riparian water -RiW- and rain water -RaW-). These data have been added to Figure 50, with their standard deviations (sd). Most of these end-members are rather well characterized but the highest ratio sd/mean is observed for groundwater of the Brie aquifer collected by piezometers on the plateau for calcium and sulfate. This is mainly due to one of the piezometers only were a highly variable gypsum dissolution process was observed, it mainly occurs during high water periods. Excluding data from this piezometer, the average concentration only slightly decreases (126 mg.L^{-1}, or 1.02 standardized calcium, see Figure 50) with a much lower standard deviation (22 mg.L^{-1}, or 0.19 standardized. For sulfate, the effect is much more significant since the average concentration decreases to 14 mg-S.L^{-1} with a standard deviation of 4 mg-S.L^{-1}. Gypsum lenses area common features in tertiary deposits in the Paris basin(Mouchel et al., 2016). Regarding Figure 50, excluding this piezometer leads to a better definition of the Brie ground water end-member. On the whole, the pre-identified end-members define 4 highly distinct poles: a pole with high calcium and

nitrate concentrations (ground water and springs end-members), a tile-drain end-member with high nitrate concentrations and lower calcium concentrations, a riparian pole with high calcium concentration and lower nitrate concentrations and a rainwater pole with almost zero concentrations of calcium and nitrate. It must be noted that rainwater could only be sampled in rain gages, there is no easily identified rainwater pool than could be sampled as a physical reservoir in the catchment. A theoretical composition of the rain water reservoir could be computed by taking into account evaporation and transpiration (about 2/3 of precipitation on average in this region, see Floury et al., 2018; Mouchel et al., 2016), however the budget of evaporation processes are highly variable along the year and during rain events or wet/dry periods and the composition of the rain water pool should be considered as highly variable. Under the simple evaporation hypothesis, the composition would move along a line sketched on Figure 50 (dotted brown line). However, this simple hypothesis is not much realistic from an hydrological point of view, highly evaporated rainwater must have spent a large amount of time inside the soil, and would consequently reach the Avenelles stream and tributaries as part of another end-member (drain water, ground and end-member water or riparian water) where the composition also reflects fertilizer inputs, biogeochemical processes in soils and water-rock interactions (Billy et al., 2013; Floury et al., 2018; Garnier et al., 2016; Garnier et al., 2014). Then the possible rain water end-member should be restricted to non-evaporated or slightly evaporated water.

Potential end-members are only barely outside of the polygon formed by the GW, SW, DW, RiW and RaW, suggesting that during most studied monthly periods the downstream concentration can be explained by a mixture of the identified end-members. Unexpected situations occurred in April and May 2016, with vary high nitrate concentrations and limited calcium concentrations that suggest and important contribution of DW with an above average nitrate concentration. The distribution of nitrate concentration in DW is quite high in April and the exceptional flood of May-June 2016 probably because a delayed contribution of tile drains to the river discharge in May during this specific year. Different unexpected situations occurred in June 2016 and 2017 with slightly too high calcium concentrations, but with non-exceptional hydrological situations. The variability of observed GW concentration should be retained as the only plausible explanation.

Another important feature regarding the position of potential end-members with regards to identified end-members is that most of them are inside the polygon that excludes riparian water. This would suggest that the contribution of RiW is unnecessary to explain river water composition except in some specific situations, the most notable being the period May to July 2017, a period with the lowest observed discharges reaching 0.04 m^3s^{-1} (or 1.840 $L.s^{-1}.km^2$), suggesting that riparian water would generally be very minor contribution except during very dry periods. Finally, the position of the most diluted potential end-member (rain) would never contribute to more than 30% of the total flow.

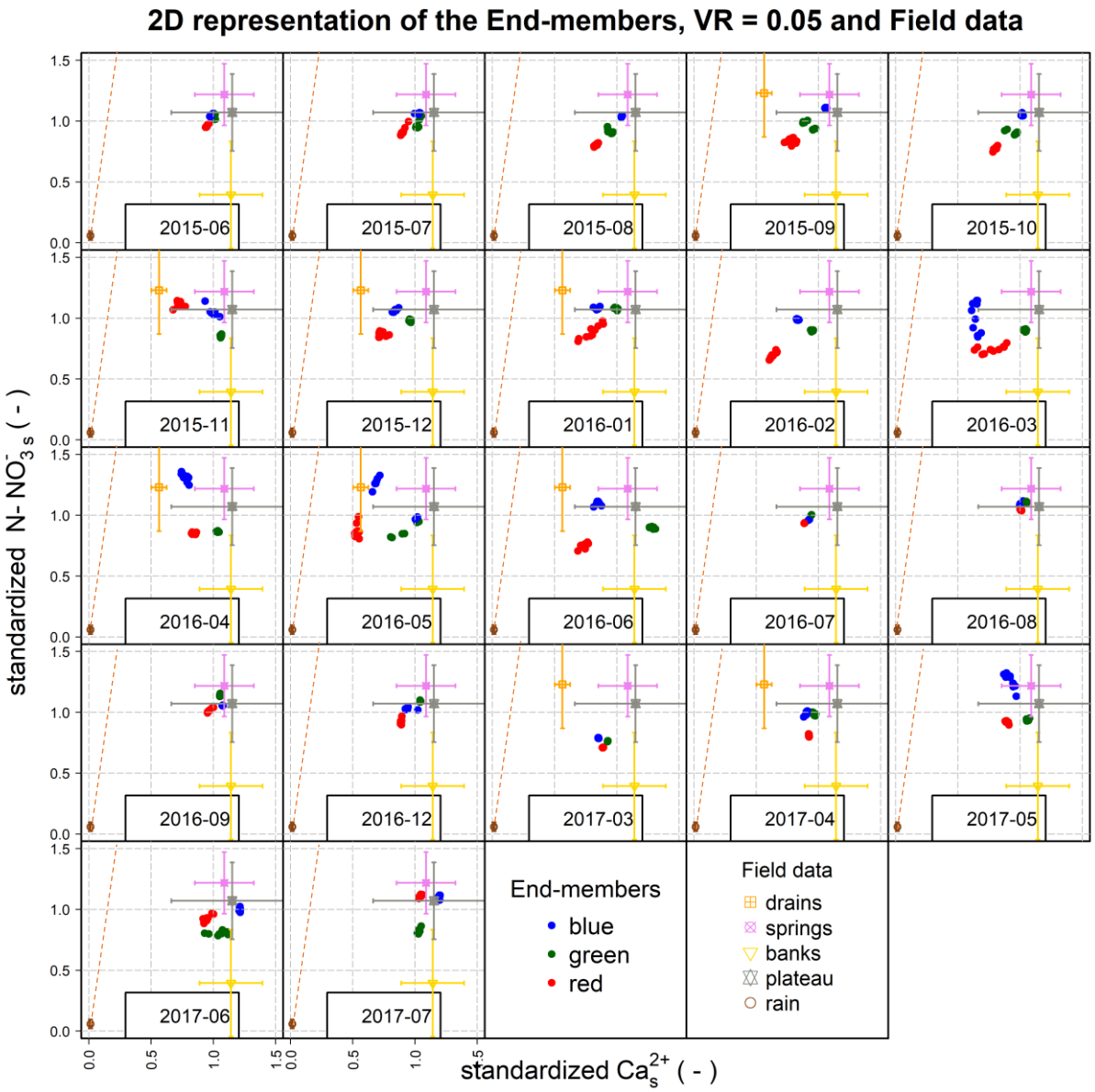


Figure 50 : Comparison between the end-members calculated using the IQEA model (VR =0.05) and the field data: ground water -GW-("plateau"), spring water -SW-("springs"), tile drain water -DW-("drains"), riparian water -RiW-("banks") and rain water -RaW-("rain").

4 Conclusions

In this article we developed a methodology (the IQEA model) that only uses downstream chemical data and a pre-defined fixed number of end-member (k =3) we can identify and quantify this end members (and their apportionment) that allow the formation of the chemical concentrations of the stream. We have been able to verify through random sampling in the data set and a further analysis of variance of the obtained end-members compositions, that the procedure we developed was robust regarding the composition of sources. In rare situations (e.g. march 2016), some possible overlap of potential end members could not be excluded, which also demonstrates a complex hydrological features. .

These calculated end-members are slightly different than the pre-identified possible end-members obtained from field data in the catchment, but can most of the time be computed as secondary mixtures of the pre-identified end-members. Thus they can be considered as secondary potential end-members that describe the way primary sources are combined during a given period in a manner that cannot be deciphered from downstream data (primary sources are never visible). One of the major achievements of this methodology proposed is that we can determine at a monthly scale the main potential sources that feed the chemical concentrations of the stream.

There must be a relation between the number of chemical components (nC) and the number of potential end-members (nD): n$D \leq$ nC+1. The additional 1 comes from an additional constraint that states that the sum of discharges from all end-members must be the total discharge. Solutions may be obtained n$D >$ nC+1; our experience was that these solutions are generally unstable, namely different random samples lead to different and overlapping solutions. In the example we treated for search of simplification nD was exactly equal to nC+1 (2 chemical components and 3 sources). The problem in and NMF (non negative matrix factorization), these problems have no unique solutions, an additional constraint is needed. We chose to reduce the variance of the source composition. (Carrera et al., 2004) proposed to add hypothesis regarding the position of the end-members, which is another possibility requiring additional information. More work is needed to improve the optimization algorithm, as we were practically limited to about 100 data. Several NMF algorithms have been proposed to solve NMF problems, the faster are based on alternate direction minimization but their convergence has been debated and depends on the additional constraint. More work is required is this direction.

5 Appendix

5.1 Cluster analysis for evolution of end-members using a variable VR (0.02<VR<0.10)

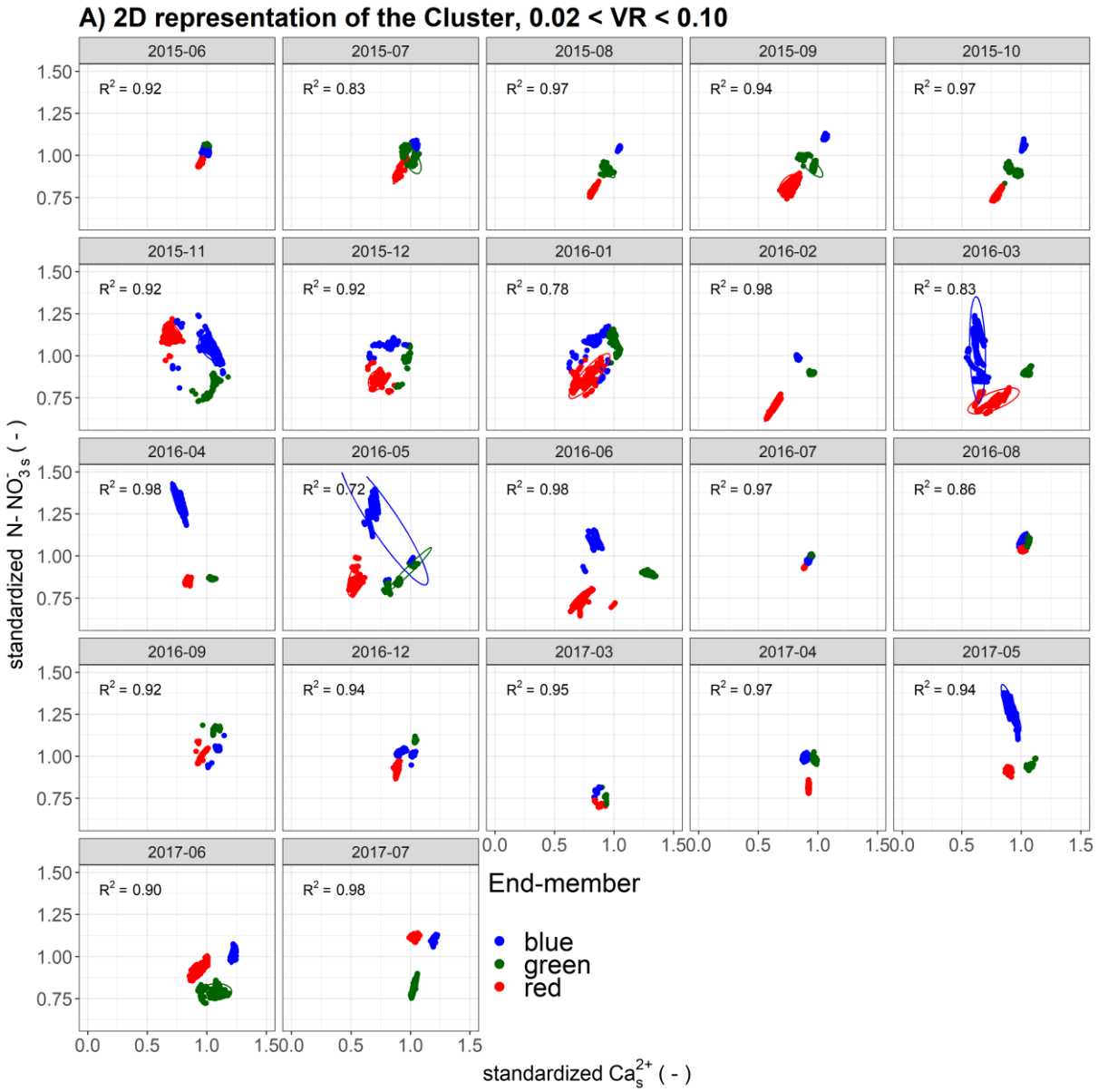


Figure 51 : Cluster analysis (ellipses show in the graph) and R^2 of the evolution of different end-members computed with the IQEA model for a variable.

5.2 Appendix-2: Summary for the month of November 2015 of: a) the initial position of the concentration with respect to the triangle of end-members. b) Position of the concentrations after their projection in the triangle. c) End-members apportionment. d) End-members apportionment multiplied by flow.

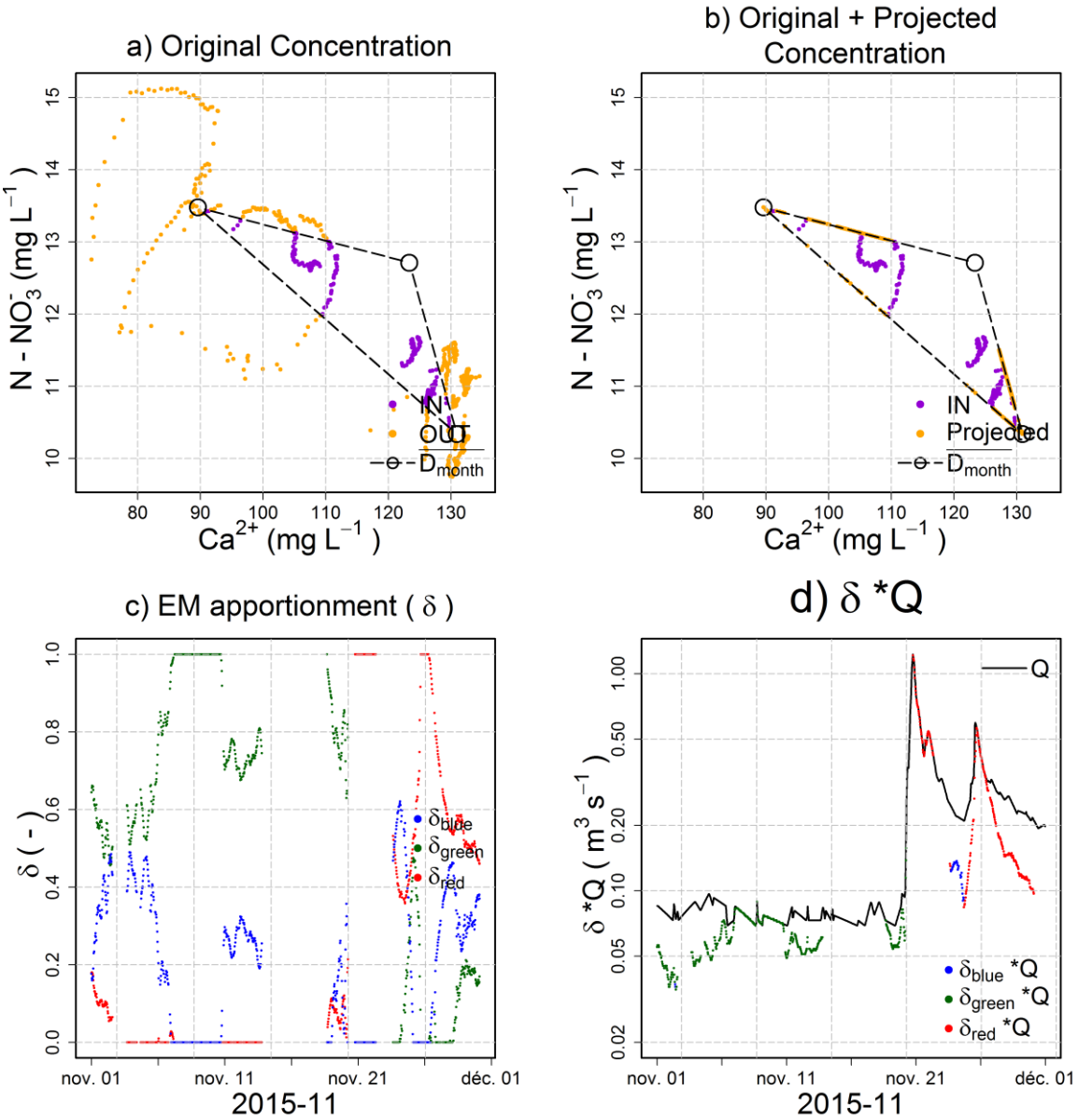


Figure 52 : Summary for the month of November 2015 of: a) the initial position of the concentration with respect to the triangle of end-members; b) Position of the concentrations after their projection in the triangle; c) End-members apportionment; d) End-members apportionment multiplied by flow.

6 References

Alberto, W.D., María del Pilar, D.a., María Valeria, A., Fabiana, P.S., Cecilia, H.A., María de los Ángeles, B., 2001. Pattern Recognition Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality. A

Case Study:: Suquía River Basin (Córdoba–Argentina). Water Research, 35(12): 2881-2894. DOI:https://doi.org/10.1016/S0043-1354(00)00592-3

Ameli, A.A., Beven, K., Erlandsson, M., Creed, I.F., McDonnell, J.J., Bishop, K., 2017. Primary weathering rates, water transit times, and concentration-discharge relations: A theoretical analysis for the critical zone. Water Resources Research, 53(1): 942-960. DOI:10.1002/2016wr019448

Anttila, P., Paatero, P., Tapper, U., Järvinen, O., 1995. Source identification of bulk wet deposition in Finland by positive matrix factorization. Atmospheric Environment, 29(14): 1705-1718. DOI:https://doi.org/10.1016/1352-2310(94)00367-T

Bansah, S., Ali, G., 2017. Evaluating the Effects of Tracer Choice and End-Member Definitions on Hydrograph Separation Results Across Nested, Seasonally Cold Watersheds. Water Resources Research, 53(11): 8851-8871.

Barakat, A., El Baghdadi, M., Rais, J., Aghezzaf, B., Slassi, M., 2016. Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. International Soil and Water Conservation Research, 4(4): 284-292. DOI:https://doi.org/10.1016/j.iswcr.2016.11.002

Barco, J., Hogue, T.S., Curto, V., Rademacher, L., 2008. Linking hydrology and stream geochemistry in urban fringe watersheds. Journal of Hydrology, 360(1): 31-47. DOI:https://doi.org/10.1016/j.jhydrol.2008.07.011

Barthold, F.K., Tyralla, C., Schneider, K., Vaché, K.B., Frede, H.-G., Breuer, L., 2011. How many tracers do we need for end member mixing analysis (EMMA)? A sensitivity analysis. Water Resources Research, 47(8). DOI:10.1029/2011wr010604

Berry, M.W., Browne, M., Langville, A.N., Pauca, V.P., Plemmons, R.J., 2007. Algorithms and applications for approximate nonnegative matrix factorization. Computational Statistics & Data Analysis, 52(1): 155-173. DOI:https://doi.org/10.1016/j.csda.2006.11.006

Billy, C., Birgand, F., Ansart, P., Peschard, J., Sebilo, M., Tournebize, J., 2013. Factors controlling nitrate concentrations in surface waters of an artificially drained agricultural watershed. Landscape Ecology, 28(4): 665-684. DOI:10.1007/s10980-013-9872-2

Blake, S., Henry, T., Murray, J., Flood, R., Muller, M.R., Jones, A.G., Rath, V., 2016. Compositional multivariate statistical analysis of thermal groundwater provenance: A hydrogeochemical case study from Ireland. Applied Geochemistry, 75: 171-188. DOI:https://doi.org/10.1016/j.apgeochem.2016.05.008

Botter, M., Burlando, P., Fatichi, S., 2019. Anthropogenic and catchment characteristic signatures in the water quality of Swiss rivers: a quantitative assessment. Hydrol. Earth Syst. Sci., 23(4): 1885-1904. DOI:10.5194/hess-23-1885-2019

Bowes, M.J., Smith, J.T., Neal, C., 2009. The value of high-resolution nutrient monitoring: A case study of the River Frome, Dorset, UK. Journal of Hydrology, 378(1–2): 82-96. DOI:http://dx.doi.org/10.1016/j.jhydrol.2009.09.015

Buttle, J.M., 1994. Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins. Progress in Physical Geography: Earth and Environment, 18(1): 16-41.

Capozzi, S.L., Rodenburg, L.A., Krumins, V., Fennell, D.E., Mack, E.E., 2018. Using positive matrix factorization to investigate microbial dehalogenation of chlorinated benzenes in groundwater at a historically contaminated site. Chemosphere, 211: 515-523. DOI:https://doi.org/10.1016/j.chemosphere.2018.07.180

Carrera, J., Vázquez-Suñé, E., Castillo, O., Sánchez-Vila, X., 2004. A methodology to compute mixing ratios with uncertain end-members. Water Resources Research, 40(12). DOI:10.1029/2003wr002263

Chanat, J.G., Rice, K.C., Hornberger, G.M., 2002. Consistency of patterns in concentration-discharge plots. Water Resources Research, 38(8): 22-1-22-10. DOI:10.1029/2001wr000971

Christophersen, N., Hooper, R.P., 1992. Multivariate analysis of stream water chemical data: The use of principal components analysis for the end-member mixing problem. Water Resources Research, 28(1): 99-107. DOI:10.1029/91WR02518

Christophersen, N., Neal, C., Hooper, R.P., Vogt, R.D., Andersen, S., 1990. Modelling streamwater chemistry as a mixture of soilwater end-members—a step towards second-generation acidification models. Journal of Hydrology, 116(1-4): 307-320.

Comero, S., Locoro, G., Free, G., Vaccaro, S., De Capitani, L., Gawlik, B.M., 2011. Characterisation of Alpine lake sediments using multivariate statistical techniques. Chemometrics and Intelligent Laboratory Systems, 107(1): 24-30. DOI:https://doi.org/10.1016/j.chemolab.2011.01.002

Devic, G., Djordjevic, D., Sakan, S., 2014. Natural and anthropogenic factors affecting the groundwater quality in Serbia. Science of The Total Environment, 468-469: 933-942. DOI:https://doi.org/10.1016/j.scitotenv.2013.09.011

Duncan, J.M., Band, L.E., Groffman, P.M., 2017. Variable nitrate concentration–discharge relationships in a forested watershed. Hydrological Processes, 31(9): 1817-1824. DOI:10.1002/hyp.11136

Dwivedi, R., Meixner, T., McIntosh, J.C., Ferré, P.A.T., Eastoe, C.J., Niu, G.-Y., Minor, R.L., Barron-Gafford, G.A., Chorover, J., 2019. Hydrologic functioning of the deep critical zone and contributions to streamflow in a high-elevation catchment: Testing of multiple conceptual models. Hydrological Processes, 33(4): 476-494. DOI:10.1002/hyp.13363

Evans, C., Davies, T.D., 1998. Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. Water Resources Research, 34(1): 129-137. DOI:10.1029/97wr01881

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., Roubaty, J.L., 2017. The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data. Hydrol. Earth Syst. Sci., 21(12): 6153-6165.

Floury, P., Gaillardet, J., Tallec, G., Ansart, P., Bouchez, J., Louvat, P., Gorge, C., 2018. Chemical weathering and CO2 consumption rate in a multilayered-aquifer dominated watershed under intensive farming: The Orgeval Critical Zone Observatory, France. Hydrological Processes: 1-19.

Frisbee, M.D., Phillips, F.M., Campbell, A.R., Liu, F., Sanchez, S.A., 2011. Streamflow generation in a large, alpine watershed in the southern Rocky Mountains of Colorado: Is streamflow generation simply the aggregation of hillslope runoff responses? Water Resources Research, 47(6).

Frisbee, M.D., Phillips, F.M., White, A.F., Campbell, A.R., Liu, F., 2013. Effect of source integration on the geochemical fluxes from springs. Applied Geochemistry, 28: 32-54. DOI:https://doi.org/10.1016/j.apgeochem.2012.08.028

Gaillardet, J., Dupre, B., Louvat, P., Allegre, C.J., 1999. Global silicate weathering and CO2 consumption rates deduced from the chemistry of large rivers. Chemical Geology, 159(1-4): 3-30. DOI:10.1016/s0009-2541(99)00031-5

Garnier, J., Anglade, J., Benoit, M., Billen, G., Puech, T., Ramarson, A., Passy, P., Silvestre, M., Lassaletta, L., Trommenschlager, J.M., Schott, C., Tallec, G., 2016. Reconnecting crop and cattle farming to reduce nitrogen losses to river water of an intensive agricultural catchment (Seine basin, France): past, present and future. Environmental Science & Policy, 63: 76-90.

Garnier, J., Billen, G., Vilain, G., Benoit, M., Passy, P., Tallec, G., Tournebize, J., Anglade, J., Billy, C., Mercier, B., Ansart, P., Azougui, A., Sebilo, M., Kao, C., 2014. Curative vs. preventive management of nitrogen transfers in rural areas: Lessons from the case of the Orgeval watershed (Seine River basin, France). Journal of Environmental Management, 144: 125-134.

Gaujoux, R., Seoighe, C., 2010. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics, 11(1): 367. DOI:10.1186/1471-2105-11-367

Genereux, D., 1998. Quantifying uncertainty in tracer-based hydrograph separations. Water Resources Research, 34(4): 915-919.

Genereux, D.P., Hemond, H.F., Mulholland, P.J., 1993. Use of radon-222 and calcium as tracers in a three-end-member mixing model for streamflow generation on the West Fork of Walker Branch Watershed. Journal of Hydrology, 142(1): 167-211. DOI:https://doi.org/10.1016/0022-1694(93)90010-7

Haji Gholizadeh, M., Melesse, A.M., Reddi, L., 2016. Water quality assessment and apportionment of pollution sources using APCS-MLR and PMF receptor modeling techniques in three major rivers of South Florida. Science of The Total Environment, 566-567: 1552-1567. DOI:https://doi.org/10.1016/j.scitotenv.2016.06.046

Hooper, R.P., 2003. Diagnostic tools for mixing models of stream water chemistry. Water Resources Research, 39(3).

Hopke, P.K., Paatero, P., 1994. Extreme value estimation applied to aerosol size distributions and related environmental problems. Journal of Research-National Institute of Standards And Technology, 99: 361-361.

Hrachowitz, M., Benettin, P., van Breukelen, B.M., Fovet, O., Howden, N.J.K., Ruiz, L., van der Velde, Y., Wade, A.J., 2016. Transit times—the link between hydrology and water quality at the catchment scale. Wiley Interdisciplinary Reviews: Water, 3(5): 629-657. DOI:10.1002/wat2.1155

Hubert, P., Martin, E., Meybeck, M., Oliver, P., Siwertz, E., 1969. Aspects hydrologique, géochimique et sédimentologique de la crue exceptionelle de la Dranse du Chablais du 22 sept. 1968. Arch. Sci. Soc. Phys. (Genève), 22(3): 581-603.

Kirchner, J.W., 2003. A double paradox in catchment hydrology and geochemistry. Hydrological Processes, 17(4): 871-874.

Kirchner, J.W., 2019. Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: theory and benchmark tests. Hydrology & Earth System Sciences, 23(1).

Klaus, J., McDonnell, J., 2013. Hydrograph separation using stable isotopes: Review and evaluation. Journal of Hydrology, 505: 47-64.

Koh, D.-C., Chae, G.-T., Ryu, J.-S., Lee, S.-G., Ko, K.-S., 2016. Occurrence and mobility of major and trace elements in groundwater from pristine volcanic aquifers in Jeju Island, Korea. Applied Geochemistry, 65: 87-102. DOI:https://doi.org/10.1016/j.apgeochem.2015.11.004

Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755): 788-791. DOI:10.1038/44565

Lenz, A., Sawyer, C.N., 1944. Estimation of stream-flow from alkalinity-determinations. Eos, Transactions American Geophysical Union, 25(6): 1005-1011. DOI:10.1029/TR025i006p01005

Li, T., Sun, G., Yang, C., Liang, K., Ma, S., Huang, L., Luo, W., 2019. Source apportionment and source-to-sink transport of major and trace elements in coastal sediments: Combining positive matrix factorization and sediment trend analysis. Science of The Total Environment, 651: 344-356. DOI:https://doi.org/10.1016/j.scitotenv.2018.09.198

Liu, F., Bales, R.C., Conklin, M.H., Conrad, M.E., 2008. Streamflow generation from snowmelt in semi-arid, seasonally snow-covered, forested catchments, Valles Caldera, New Mexico. Water Resources Research, 44(12).

Liu, F., Williams, M.W., Caine, N., 2004. Source waters and flow paths in an alpine catchment, Colorado Front Range, United States. Water Resources Research, 40(9). DOI:10.1029/2004wr003076

Lott, D.A., Stewart, M.T., 2016. Base flow separation: A comparison of analytical and mass balance methods. Journal of Hydrology, 535: 525-533.

McDonnell, J.J., Stewart, M.K., Owens, I.F., 1991. Effect of Catchment-Scale Subsurface Mixing on Stream Isotopic Response. Water Resources Research, 27(12): 3065-3073. DOI:10.1029/91wr02025

McGuire, K.J., McDonnell, J.J., 2006. A review and evaluation of catchment transit time modeling. Journal of Hydrology, 330(3): 543-563. DOI:https://doi.org/10.1016/j.jhydrol.2006.04.020

Miller, M.P., Tesoriero, A.J., Hood, K., Terziotti, S., Wolock, D.M., 2017. Estimating Discharge and Nonpoint Source Nitrate Loading to Streams From Three End-Member Pathways Using High-Frequency Water Quality Data. Water Resources Research.

Mouchel, J.-M., Rocha, S., Rivière, A., Tallec, G., 2016. Caractérisation de la géochimie des interfaces nappe-rivière du bassin des Avenelles. Tech. rep. PIREN Seine, France, 27 pp.

Mouhri, A., Flipo, N., Rejiba, F., de Fouquet, C., Tallec, G., Bodet, L., Durand, V., Jost, A., Guérin, R., Ansart, P., 2012. Stratégie d'échantillonnage des échanges nappe-rivière du bassin agricole de l'Orgeval. Tech. rep. PIREN Seine, France, 27 pp.

Mouhri, A., Flipo, N., Vitale, Q., Bodet, L., Tallec, G., Ansart, P., Rejiba, F., 2013. Influence du contexte hydrogéologique sur la connectivité nappe - rivière. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 89-98.

Muñoz-Villers, L.E., McDonnell, J.J., 2012. Runoff generation in a steep, tropical montane cloud forest catchment on permeable volcanic substrate. Water Resources Research, 48(9).

Neal, C., Kirchner, J.W., 2000. Sodium and chloride levels in rainfall, mist, streamwater and groundwater at the Plynlimon catchments, mid-Wales: inferences on hydrological and chemical controls. Hydrology and Earth System Sciences, 4(2): 295-310.

Neal, C., Reynolds, B., Rowland, P., Norris, D., Kirchner, J.W., Neal, M., Sleep, D., Lawlor, A., Woods, C., Thacker, S., Guyatt, H., Vincent, C., Hockenhull, K., Wickham, H., Harman, S., Armstrong, L., 2012. High-frequency water quality time series in precipitation and streamflow: From fragmentary signals to scientific challenge. Science of the Total Environment, 434: 3-12.

Norris, G., Duvall, R., Brown, S., Bai, S., 2014. Epa positive matrix factorization (pmf) 5.0 fundamentals and user guide prepared for the us environmental protection agency office of research and development, washington, dc. Washington, DC.

Paatero, P., 1999. The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear Problems, Including the n-Way Parallel Factor Analysis Model. Journal of Computational and Graphical Statistics, 8(4): 854-888. DOI:10.1080/10618600.1999.10474853

Paatero, P., Tapper, U., 1994. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. Environmetrics, 5(2): 111-126. DOI:doi:10.1002/env.3170050203

Pearce, C.R., Parkinson, I.J., Gaillardet, J., Chetelat, B., Burton, K.W., 2015. Characterising the stable (δ88/86Sr) and radiogenic (87Sr/86Sr) isotopic composition of strontium in rainwater. Chemical Geology, 409: 54-60. DOI:https://doi.org/10.1016/j.chemgeo.2015.05.010

Pelizardi, F., Bea, S.A., Carrera, J., Vives, L., 2017. Identifying geochemical processes using End Member Mixing Analysis to decouple chemical components for mixing ratio calculations. Journal of Hydrology, 550: 144-156. DOI:https://doi.org/10.1016/j.jhydrol.2017.04.010

Phung, D., Huang, C., Rutherford, S., Dwirahmadi, F., Chu, C., Wang, X., Nguyen, M., Nguyen, N.H., Do, C.M., Nguyen, T.H., Dinh, T.A.D., 2015. Temporal and spatial assessment of river surface water quality using multivariate statistical techniques: a study in Can Tho City, a Mekong Delta area, Vietnam. Environmental Monitoring and Assessment, 187(5): 229. DOI:10.1007/s10661-015-4474-x

Pinder, G.F., Jones, J.F., 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. Water Resources Research, 5(2): 438-445.

Polissar, A.V., Hopke, P.K., Paatero, P., Malm, W.C., Sisler, J.F., 1998. Atmospheric aerosol over Alaska: 2. Elemental composition and sources. Journal of Geophysical Research: Atmospheres, 103(D15): 19045-19057. DOI:doi:10.1029/98JD01212

Probst, J.L., 1985. Nitrogen and Phosphorus exportation in the Garonne basin (France). Journal of Hydrology, 76(3-4): 281-305. DOI:10.1016/0022-1694(85)90138-6

Reff, A., Eberly, S.I., Bhave, P.V., 2007. Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods. Journal of the Air & Waste Management Association, 57(2): 146-154. DOI:10.1080/10473289.2007.10465319

Rose, L.A., Karwan, D.L., Godsey, S.E., 2018. Concentration–discharge relationships describe solute and sediment mobilization, reaction, and transport at event and longer timescales. Hydrological Processes, 32(18): 2829-2844. DOI:10.1002/hyp.13235

Singh, K.P., Malik, A., Mohan, D., Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water Research, 38(18): 3980-3992. DOI:https://doi.org/10.1016/j.watres.2004.06.011

Singh, K.P., Malik, A., Sinha, S., 2005. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. Analytica Chimica Acta, 538(1): 355-374. DOI:https://doi.org/10.1016/j.aca.2005.02.006

Stewart, M., Cimino, J., Ross, M., 2007. Calibration of base flow separation methods with streamflow conductivity. Groundwater, 45(1): 17-27.

Tallec, G., Ansart, P., Guérin, A., Derlet, N., Pourette, N., Guenne, A., Delaigue, O., Boudhraa, H., Loumagne, C., 2013. Introduction. In: Loumagne, C., Tallec, G. (Eds.), L'observation long terme en environnement, exemple du bassin versant de l'Orgeval QUAE, Versailles, pp. 11-33.

Tubau, I., Vàzquez-Suñé, E., Jurado, A., Carrera, J., 2014. Using EMMA and MIX analysis to assess mixing ratios and to identify hydrochemical reactions in groundwater. Science of The Total Environment, 470-471: 1120-1131. DOI:https://doi.org/10.1016/j.scitotenv.2013.10.121

Tunqui Neira, J.M., Andréassian, V., Tallec, G., Mouchel, J.M., 2019. A two-sided affine power scaling relationship to represent the concentration–discharge relationship. Hydrol. Earth Syst. Sci. Discuss., 2019: 1-15. DOI:10.5194/hess-2019-550

Vázquez-Suñé, E., Carrera, J., Tubau, I., Sánchez-Vila, X., Soler, A., 2010. An approach to identify urban groundwater recharge. Hydrol. Earth Syst. Sci., 14(10): 2085-2097. DOI:10.5194/hess-14-2085-2010

Vega, M., Pardo, R., Barrado, E., Debán, L., 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Research, 32(12): 3581-3592. DOI:https://doi.org/10.1016/S0043-1354(98)00138-9

von Freyberg, J., Studer, B., Kirchner, J.W., 2017. A lab in the field: high-frequency analysis of water quality and stable isotopes in stream water and precipitation. Hydrology and Earth System Sciences, 21: 1721-1739.

Ye, Y., 1988. Interior algorithms for linear, quadratic, and linearly constrained convex programming, Stanford University, 120 pp.

Zanotti, C., Rotiroti, M., Fumagalli, L., Stefania, G.A., Canonaco, F., Stefenelli, G., Prévôt, A.S.H., Leoni, B., Bonomi, T., 2019. Groundwater and surface water quality characterization through positive matrix factorization combined with GIS approach. Water Research, 159: 122-134. DOI:https://doi.org/10.1016/j.watres.2019.04.058

Zhang, R., Li, Q., Chow, T.L., Li, S., Danielescu, S., 2013. Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method. Hydrological Processes, 27(18): 2659-2665.

Part III

Conclusions and Perspectives

Avant-propos

High-frequency measurements of concentrations in stream become increasingly used. However, they continue to be analyzed under methodologies developed for low and medium frequencies.

In this thesis, we have tried to adapt and develop different approaches of the concentrations-discharge (C-Q) relationships, better suited to high-frequency dataset.

Thus, we worked with different models, devised by different disciplines, all trying to help us to develop a parsimonious model that could simulate adequately chemical concentrations and flow components. Because no concept is more valuable than another we have tried to refine, and for some, to combine them. This Third part presents the conclusions and the perspectives of these works.

1 Conclusions

Today, we have exhaustive information of the variations of the concentrations and flows in the river. Nevertheless, only analysis of this knowledge will give us the keys to its understanding. Everything lies therefore, in the first place, in the method of analysis of these data. First, we will review the different adaptations and developments realized during the thesis on each methodology. However, it is clear that the high-frequency cannot solve the old problem of lack of data for all rivers around the world. Research must therefore find a solution to extract a maximum of information from this high-frequency data and integrate them into the models. Therefore, in the second part of this chapter, we try to evaluate, with the achievements of the thesis, the limits and perspectives of these works in the viewpoint of a C-Q relationships model the most parsimonious.

1.1 Main achievements of the thesis

1.1.1 The affine power scaling relationship

Across the years (50 years at least) it has been attempted to establish simple relations between the flow and the solute concentration of the stream to explain the different hydro-chemical processes occurred in the catchment. **The most popular and used approach is the power-law relationship.** This relationship has shown its effectiveness (especially with the long-term dataset) in different studies cited in the thesis. Although objective, this relationship is also “incomplete”: since it has been performed with low and medium frequency data, which produces a valuable loss of information. Do high-frequency measurements, recording in detail all the range of variations, allow a best fit to the power-law relationship? Works of the thesis show that, even with high-frequency, in the case of our study catchment, the high-frequency does not solve the non-linearity of the C-Q relationships sometimes encountered. However, we have shown that a three-parameter relationship, that we named two-sided affine power scaling relationship (2S-APS), had a better performance than the power-law relationship for all the solutes tested (sodium, sulfate, chloride, EC). With the suitable calibration process proposed, the 2S-APS is an alternative to the classic power-law relationship.

1.1.2 Calibration of the Hydrograph separation

The C-Q relationship represented through a mixing model is widely used to hydrograph separation. The hydrograph separation is usually seen as a mixture of two components (base flow and quick flow). Because it is the most intensely studied procedure in the field of hydrology, an endless number of methods have been developed. However, they are based on disputed concepts that are very difficult

to corroborate in physical reality. To overcome this lack of physical sense, based on the concept of mass balance, a chemistry calibration has been developed. Does the use of high-frequency data, increase this relation to the physical processes? Is the choice of the ion, used for the calibration, important? Does the detail of high-frequency data corroborate a two-component mixing representation of the river?

Calibration of hydrograph separation methods, using high-frequency data, only confirmed the suggestiveness of classic hydrograph separation methods. Nevertheless, despite this subjectivity, our work shows that these methods make it possible to simulate correctly the flows and the concentrations observed in the river, for chlorides and EC only. For the other ions tested (sulfates, magnesium, calcium and nitrate), the results obtained were not enough conclusive. The representation of a two-component mixing model therefore remains objective with representative ions of rainy events as chlorides and sodium or with the representative ionic charge as EC.

1.1.3 Combining of affine relationship and mixing model

Our works showed that the affine power scaling and mixing representations of the C-Q relationships are both objectives. Each represents some processes, as the dilution and concentration process in the affine power scaling relationship and a part of the functioning of the catchment, as the flow components in mixing models. However, studies of the C-Q relationships (see first Part of the thesis), have shown that all chemical tracers do not response to the mass balance concept; they are subject to chemo-dynamic processes not taken into account in these two representations. Can we establish a model that combines these two approaches and has a better performance than both models applied separately?

Although they use a similar concept (C = f (Q)), they have always been studied separately and in our recent knowledge no one has tried to establish a common link between the two approaches.

During the thesis**,** we developed a new conceptual representation of the C-Q relationships that combines the 2S-APS relationship and the mixing model. At inter-annual scale, the proposed combining model gives better estimate of the water chemistry than when the models are used separately. This is true for all tested ions, even for the nitrates that present much more complicated patterns. In addition, the analysis of different flood events (C-Q hysteresis) shows that the short flood events are not well simulated by the combining model. However, after removing the effects due to a short hydrodynamic process, we obtained good simulations. This shows the advantage of coupling a time dynamic hydrological model with a C-Q relationship for each of the flow components. It makes it possible to simulate time dependent relations that do not only rely on $Q(t)$ to simulate this dependence.

1.1.4 Identification and quantification of potential end-members

High-frequency measurements allow us to have an accurate and detailed record of the hydro-chemical interactions of the catchment. Therefore, it is not absurd to consider that with this information we could identify and quantify the chemical sources that produce these chemical data. Can the sources, through high-frequency chemical data taken downstream, be identified and quantified?
We attempted to answer this question with the development of a purely chemical model (named IQEA) using three end-members (without the use of discharge) at the monthly scale. One of the major achievements of the proposed method is that we can determine at a monthly scale the main potential sources that feed the chemical concentrations of the stream. Our work shows that, in the Avenelles catchment, three potential end-members could explain the chemical concentrations of the calcium and nitrates ions in stream water. These potential sources evolve each month, depending on wet or dry seasons and on exceptional hydro-meteorological events. During the dry season, the Avenelles catchment mobilizes the same sources from one month to another. During the wet season, from the first rains, the potential sources evolve from one month to another, with specific sources during exceptional events. Finally, the analysis of the chemical high-frequency data alone, also informs about the flow mixture and especially allows visualizing the different mixtures over time.

1.2 The develop of a parsimonious model

The coupled hydrological and chemical model we developed relies on two simple assumptions: (i) the hydrology can be disaggregated in two hydrological components which were computed by recursive digital filters and (ii) for each of the components a C-Q relation can be defined. In this work, we chose the 2S-APS model to relate the concentration and the total discharge. The general idea behind this attempt was to construct a model that simulates discharges as well as concentrations. Some of the limits of the model have been emphasized. One of them is the inability of the model to simulate short time lags that have important consequences on the highly utilized C-Q hysteresis curves. Another limit was the apparent equivocal relation between concentration and discharge although introducing two hydrological components has improved the simulation of concentrations. Comparing observed and simulated data suggested that introducing some seasonal variability of the C-Q relation would probably greatly improve the simulations; seasonality could also impact on the hydrological disaggregation by differentiating the type of flow occurring during different seasons. It was also shown that one single time constant could hardly simulate flow separation for shorter and longer flood periods, a double filter may be a solution, first computing a very quick flow with a small time constant and running the filter a second time with a longer time constant to disaggregate the obtained base flow into two additional components.

Clearly many lines of improvement exist. On the whole C(t) would not only be related to Q(t) or the Q(t) of disaggregated hydrological components (their number could be increased) but to additional variables : season, time lag which could vary as a function of discharge or estimated humidity of soils, on rain intensity add additional hydrological variables, or on external variables such as fertilizer inputs.

The source analysis starts from a very different hypothesis that states that there exist a finite number of sources, with constant concentration, that combine or mix to produce the water composition that is observed at every moment. However, the evaluation of potential source composition can be made for sequential time spans thus enabling to compute series of source definition that can vary with time. We proceeded this way at a monthly scale. Basically, source definition based on downstream composition data assumes that the variability due to the evolving distribution of discharge is stronger than the evolution of potential source composition. The method attributes all variability to the evolving distribution of discharges and computes an optimal potential source composition. The source analysis simulates only mixing processes but no reactivity of chemicals. Then all processes involving some reactivity would be included in the evolution of the sources.

More and more high-frequency database are becoming available, in France through the OZCAR structure, and elsewhere in the world, based on specific sensors or surrogate parameters or deploying on-site labs. These new data sets can be a basis for a coordinated effort to examine the pros and cons of each kind of model. Once validated, and if parsimonious enough, the models could certainly be tested on a larger number of stations with high-frequency hydrological data but less frequent water quality data.

2 Perspectives

The continuous measurements of chemical concentrations have given us the necessary tools to understand in a more precise way the hydrochemical functioning of the catchment. The high-frequency revives the interest that has never entirely failed for 70 years. With future works, our representation of the catchment functioning and its modelling, in a parsimonious or purely physical way, could be taken further. At our modest scale, this thesis offers a few perspectives to improve the methods developed, to test others ions, at others timescales or others catchments.

Improve validation of the developed methods. During the course of the thesis, we have adapted and developed all the methodologies, using only the high-frequency measurements of the Oracle-Orgeval observatory. Future works should test the methods on others catchments and others high-frequency dataset to analyze the efficiency and effectiveness of these methods. To begin with, electro conductivity (EC) could be used since it is measured continuously in many catchment A "split sample test", which verifies the modelling performance, and applied to the 2S-APS relationship (Chapter 1, Part II), could also be extended to the others methods of the thesis as the combining model or the end-member model (Chapter 4, Part II).
During the thesis, we also noticed that, with more time, even if our range of values tested were the widest possible, the calibration of the different parameters (i.e., n, a, b of the 2S-APS method; the τ value of the hydrograph separation method; etc.) could be further improved.
Except for the combining model, the methods developed have been tested with a small number of ions. It would be important to test each method with each ion measured with the *River Lab* of the Oracle-Orgeval observatory and other basins (a lot of work in perspective).

Test the methods under different time scales. Throughout this thesis, we worked with different time scales. Whereas the 2S-APS relationship (Chapter 1, Part II) and the combining model (Chapter 3, Part II) were developed at inter-annual time scale, the hydrograph separation methods (Chapter 2, Part II) have been calibrated at a daily time scale. Last, the identification of chemical end-members (Chapter 4, Part II) was conducted at a monthly time scale. All these time scales met specific needs or constraints, depending on the purposes. Research will need to be conducted on time scale issues. Indeed, some methods may be more or less effective depending on the time scale chosen, from the sub-hour to the inter-annual. However, it should be noted that the amount of high-frequency data increases the computation time.

Apart from the measurements frequency (low or high frequency), which would have to be compared for each method, the application of these methods over a long time (i.e. more than ten years) would be extremely enriching, at least on the ions of importance.

Improve the representation of the catchment. The methods developed during the thesis, as the combining model, allows taking into account more processes, and especially the chemo-dynamic processes, incorporating an assumption of the variability of concentrations across time, as well as the mixing flow components. However, a better representation of the mixing flow, as three components instead of two, would improve the model. In particular, this would make it possible to integrate different transit times, and better reproduce the hysteresis, even if this representation remains conceptual and to some extent arbitrary. Similarly, the end-member model allowed us to visualize the potential chemical end-members and their temporal variations over a time series of flow. There must be a relation between the number of chemical components and the number of potential sources. Because we supposed that the sum of discharges from all sources must be the total discharge, we chose to reduce the variance of the source composition. However, several algorithms have been proposed to solve multivariate problems and more work is required is this direction.

The continuation of this work can only clarify our knowledge and improve the modeling of the hydro-chemical functioning of the catchment.

As can be seen in Chapters 3 and 4 of Part II of this thesis, we had difficulties in understanding the hydro-chemical behavior of two ions: calcium and nitrates. In the case of the latter, its variable behavior in the Oracle-Orgeval observatory has been well studied and documented in many scientific articles, but not calcium (if we compare it with nitrates). We assume that this happens because calcium is the predominant geological element of the basin, thus it is assumed that its hydrochemical behavior is stable. Only since the installation of *River Lab* we have noticed its variable behavior, therefore it is important to try to develop a methodology to understand it, because to be an element "omnipresent" in the catchment, calcium could give us a valuable information about the hydro-chemical processes of the catchment.