Continental water cycle : climatic and non-climatic drivers of river discharge and evolution of water resources

Julie Collignan

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2023IPPAX138

Thèse de doctorat

Continental water cycle: climatic and non-climatic drivers of river discharge and evolution of water resources

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à École polytechnique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat: Mécanique des fluides et des solides, acoustique

Thèse présentée et soutenue à Palaiseau, le 04/12/2023, par

JULIE COLLIGNAN

Composition du Jury :

Catherine Ottlé Directrice de recherche, LSCE	Présidente
Maria-Helena Ramos Directrice de recherche, INRAE (Unité HYCAR)	Rapporteure
Joan Cuxart Rodamilans Professeur émérite, Universitat de les Illes Balears	Rapporteur
Stefan Kollet Professeur des universités, Universität Bonn, Forschungszentrum Jülich	Examinateur
Pere Quintana Segui Directeur de recherche, Observatori de l'Ebre (Universitat Ramon Llull - CSIC)	Examinateur
Jan Polcher Directeur de recherche, Laboratoire de Météorologie Dynamique	Directeur de thèse
Sophie Bastin Chargée de recherche, LATMOS	Co-directrice de thèse

Résumé en français

L'eau douce est une ressource essentielle tant pour les écosystèmes que pour les sociétés humaines. Détecter et quantifier les changements du débit des rivière et comprendre leurs causes est essentiel pour gérer et planifier l'utilisation des ressources en eau, anticiper les risques d'innondation ou les pénuries d'eau. Aussi, prévoir l'évolution de la ressource en eau est un défi majeur dans un contexte de changement climatique et de rivières hautement anthropisées.

Analyser, reproduire et projeter l'évolution des débits de rivière

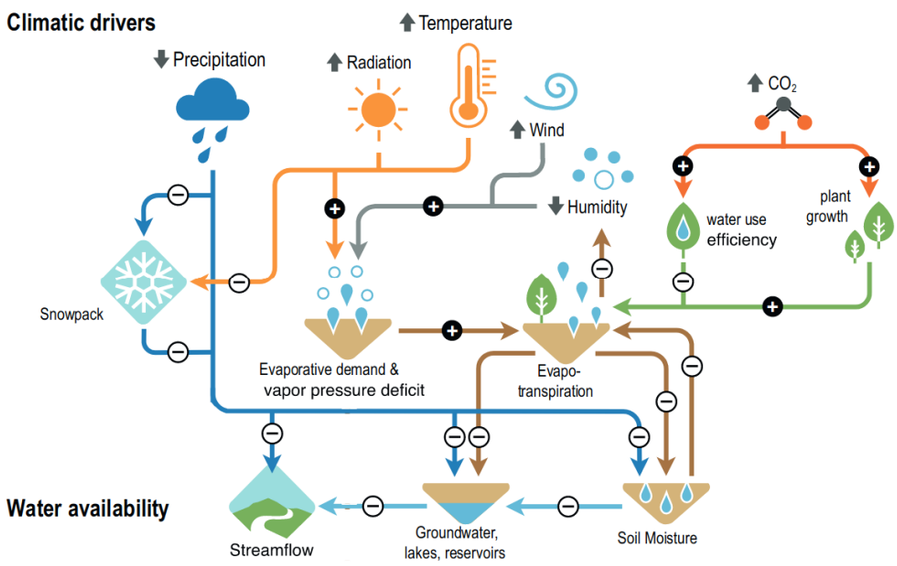


Figure 1 – Facteurs climatiques jouant un rôle dans la répartition de l'eau au cours de son cycle et entre les différents réservoirs continentaux (rivières, lacs, nappes phréatiques, sol...). Les signes plus et moins dénotent de la direction du changement associé à l'effet des vecteurs clilmatiques sur les différentes composantes du cycle. Source : Douville et al. (2021).

L'eau voyage en un cycle clos à l'échelle du globe, circulant entre plusieurs états et plusieurs réservoirs. Un miriade de facteurs entrent en jeux et influent sur ces changements d'états et sur les échanges entre les différents réservoirs (Fig. 1).

A l'échelle d'un bassin versant, il est possible de simplifier le cycle de l'eau à une répartition de l'eau entrante dans le système, sous forme de précipitations, entre l'évapotranspiration, l'eau stockée à l'échelle du bassin et le ruissellement qui s'intègre en débit à la sortie du bassin versant. L'équilibre de cette répartition repose sur de nombreux facteurs, tant climatiques (precipitations, demande évaporative...) que dépendant de l'état de la surface (couverture végétale, type de sol, pente, ...). Au cours du dernier siècle, les études ont montré que le climat a changé, avec au niveau de l'Europe, une augmentation des précipitations dans le Nord et l'Ouest, et une tendance à la baisse en moyenne autour de la Méditerrannée.

Sur l'ensemble de l'Europe, la demande évaporative a également augmenté avec notamment l'augmentation des températures. Par ailleurs, les charactéristiques, l'état et l'usage des sols a également changé, avec une évolution du couvert végétal, la fonte des glaciers, de l'anthropisation des sols en lien avec les techniques et l'expansion des surface irriguées, qui influent sur l'équilibre entre infiltration, évapotranspiration et ruissellement.

Le débit est le résultat intégré sur le bassin versant des effets combinés de tous ces facteurs influant l'équilibre entre évapotranspiration et ruissellement. En plus d'être directement lié à des enjeux environnementaux et sociétaux majeurs, comme les risques d'inondation ou encore l'assèchement saisonnier des cours d'eau, l'étude des débits permet également de considérer l'ensemble des changements du cycle de l'eau à travers une unique variable intégrative. Cependant cela complique également la compréhension et l'attribution des causes principales associées à ces changements. Différent types de modèles ont donc été construits pour comprendre, reproduire et prévoir les évolutions du cycle de l'eau et des débits.

Le premier type de modèle est celui des modèles parcimonieux. Ces modèles reposent souvent sur un nombre de variables climatiques réduits, pour exprimer des relations relativement simples entre ces variables et le débit. Des paramètres plus ou moins empiriques permettent d'ajuster les relations à l'échelle de la période et de la zone géographique étudiées. Cette paramétrisation empirique permet d'avoir de bonnes performances pour reproduire les débits et leur évolution à court terme. En revanche, due au manque de sens physique associé à ces paramètres, cela ne permet pas d'extrapoler ces modèles sur d'autres zones ou sur des projection à long terme, puisque l'on ne connait pas les processus, climatiques ou non, intégrés dans ces paramètres, et donc si ces paramètres peuvent être vraiment considérés comme indépendant du temps pour un système évolutif, notammant dans un contexte de changement climatique.

Un autre type de modèles correspond aux modèles reposant sur l'association de processus physiques définis pour essayer de reconstituer la dynamique du système terrestre et l'équilibre entre les différents processus intervenant dans la génération des débits. Ces modèles permettent quant à eux de comprendre les processus impliqués principaux et sont les seuls qui permettent une réelle attribution des causes de l'évolution des débits et donc des projections pertinentes selon des scénarios d'évolution climatique. Les performances de ces modèles ont été validés, notamment pour reproduire les processus atmosphériques et de surface naturels à grande échelle. Cependant, ces modèles ne sont pas assez complexes pour inclure toute la complexité du sytème terrestre. Ainsi les performances de ces modèles pour reproduire les débits et leur évolution effective sont moins bonnes, particulièrement dans les zones où les processus non représentés ont un fort impact. C'est notamment le cas de tous les processus associés aux activités humaines de gestion et d'usage de l'eau, qui sont encore manquant ou très grossièrement définis dans la plupart des modèles. Ces modèles évoluent donc vers des schémas de plus en plus intégratifs, avec une compréhension de plus en plus fine des processus, vers la reconstruction du sytème Terre intégrant différentes disciplines (sciences atmosphériques, agronomiques, hydrologiques, géologiques, sociales...).

Méthode pour isoler et décomposer les effets du changement climatique sur l'évolution des débits

Nous proposons une méthode innovante pour détecter et quantifier les changements dans le débit des rivières, climatiques et non climatiques (Fig. 2).

Le modèle de surface (LSM) ORCHIDEE est utilisé pour estimer la réponse "naturelle"

de la surface continentale aux fluctuations climatiques. Ce modèle repoduit les processus naturels et les interactions entre les variables atmosphériques et la réponse de la surface. Le cadre conceptuel de Budyko (équation de Fu) est ensuite utilisé, comme représentation plus simple de l'équilibre eau/energie, ajustée à l'échelle des bassins versants. Ce cadre permet de décomposer l'évolution du débit en une réponse directe aux fluctuations climatiques moyennes, et une réponse indirecte, due aux changements de l'efficacité évaporative du bassin versant (Fig. 3).

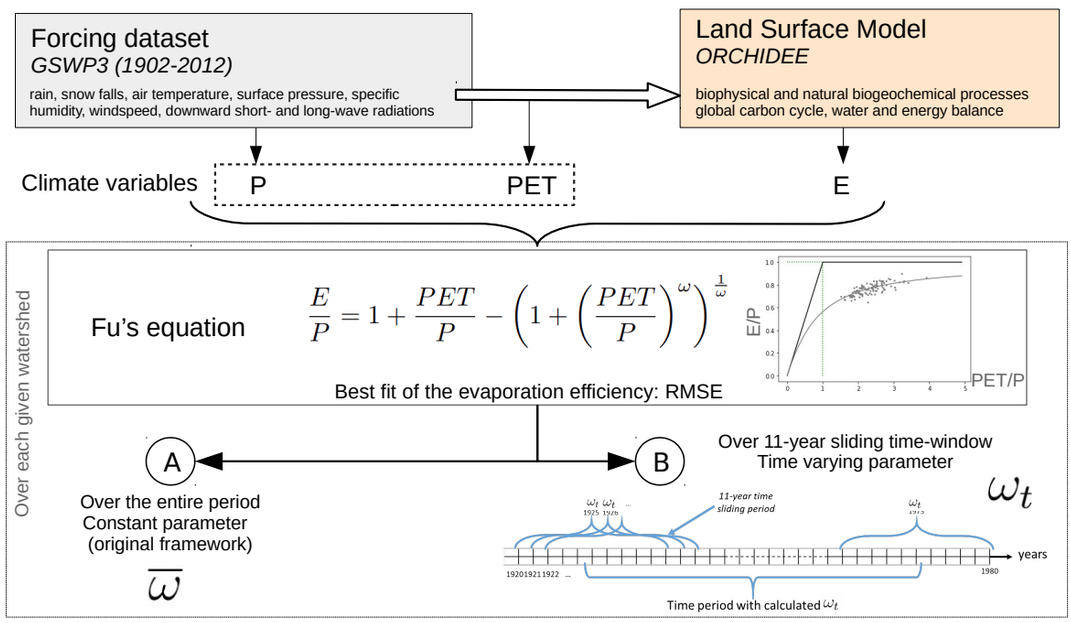
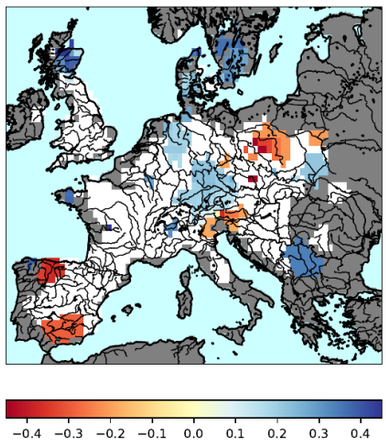


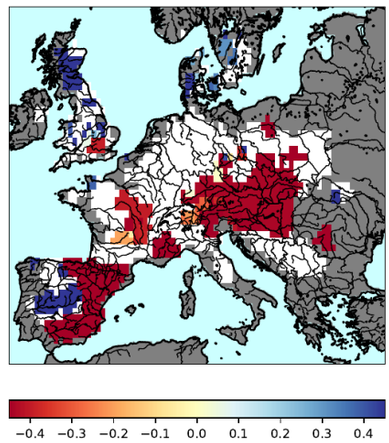
Figure 2 – Schéma de la méthode : le modèle de surface ORCHIDEE est lancé avec un forçage atmosphérique pour calculer la demande évaporative et l'évaporation effective à l'échelle des bassins versants. Il représente la "réalité climatique", dans laquelle les bassins versants ont des charactéristiques constantes. Ensuite les moyennes annuelles de précipitations P, évapotranspiration potentielle PET et évapotranspiration effective E sont utilisées pour ajuster le cadre conceptuel de Budyko (équation de Fu). Nous faisons deux ajustements : A- un ajustement considérant une efficacité évaporative constante sur toute la période. Cela nous donne l'effet sur les débits des changements de climat moyen. B- un ajustement avec une efficacité évaporative variable. Cela nous donne en plus l'effet des changements d'efficacité évaporative du bassin versant.

Dans un premier temps, nous appliquons ce cadre conceptuel aux sorties du modèle complexe ORCHIDEE. Dans ce modèle, les bassins versants sont représentés avec des caractéristiques surfaciques constantes et sans processus liés notamment aux activités humaines. Aussi, dans ce système, les seules sources de changement sont liées aux variables climatiques utilisées en entrée du modèle. A travers le cadre conceptuel de Budyko, nous pouvons séparer l'effet des variables climatiques (precipitations P et évapotranspiration potentielle PET) moyennes annuelles, de l'effet de leur distribution intra-annuelle sur la moyenne annuelle débit de rivière.

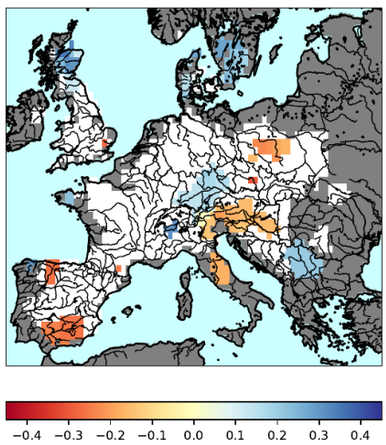
Les résultats obtenus en Europe montrent qu'au cours du dernier siècle, l'évolution des débits dus aux processus climatiques est dominée par la tendance sur les précipitations moyennes (P). Le deuxième principal facteur climatique est l'augmentation de l'évapotranspiration potentielle (PET) dans la majeure partie de l'Europe, sauf en Méditerranée. L'eau y est plus limitante, ce qui réduit fortement l'effet des changements de PET sur les débits, surpassé par l'effet de la répartition intra-annuelle de P.



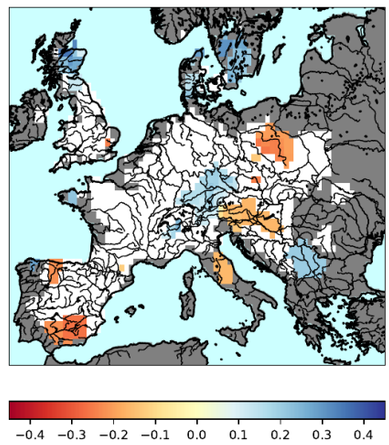
(a) Changements relatifs des débits dans le système de référence "climatique".



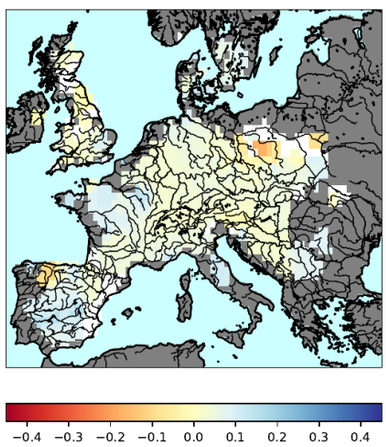
(b) Changements relatifs du débit dans le système réel.



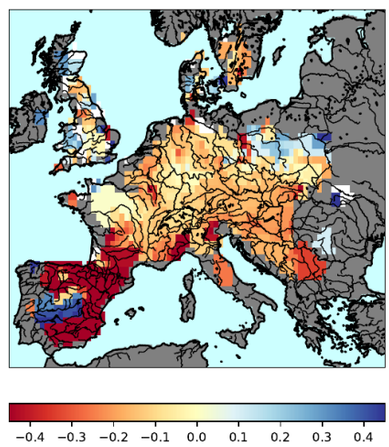
(c) Changements relatifs partiels des débits liés aux effets du climat moyen dans le système de référence "climatique".



(d) Changements relatifs partiels des débits liés aux effets du climat moyen dans le système réel.



(e) Changements relatifs partiels des débits liés aux effets des changements d'efficacité évaporative dans le système de référence "climatique".



(f) Changements relatifs partiels des débits liés aux effets des changements d'efficacité évaporative dans le système réel.

Figure 3 – Tendances significatives sur les débits de rivière, sur la période 1901-2012 (% de changement par an sur le siècle). La colonne de gauche représente les résultats pour le système de référence "climatique" (à partir des sorties d'ORCHIDEE), celle de gauche les résultats pour le système réel (reposant sur les observations de débit). Une tendance positive (vers le bleu) correspont à une augmentation significative du débit tandis qu'une tendance négative (vers le rouge) correspont à une diminution. Les échelles ont été forcées pour correspondre pour chacune de ces figures et faciliter les comparaisons, mais cela peut cacher des extrêmes en dehors des limites représentes, comme au niveau de la péninsule ibérique dans la colonne de droite.

Quantifier les effets des facteurs non climatiques sur les débits de rivière

Appliquer cette même méthodologie mais cette fois à des débits observés permet, en comparant aux résultats précédents qui incluent uniquement les effets de la variabilité climatique, de mettre en évidence les zones où la réponse "naturelle" des bassins versants à la variabilité climatique est insuffisante pour expliquer les changements enregistrés.

Nous montrons qu'en Europe l'évolution des débits au cours du dernier siècle a été significative, dominée par des facteurs non pris en compte dans le système "naturel", particulièrement dans le sud de l'Europe, au niveau de la péninsule ibérique (Fig. 3). Cependant, les changements attribuables aux variables climatiques demeurent importants et, sur des sous-périodes de dix ans, l'effet des facteurs non-climatiques devient faible, couvert par la variabilité climatique élevée. Il faut donc tenir compte à la fois des facteurs climatiques et non climatiques pour projeter les tendances futures des débits.

Ainsi, en combinant ORCHIDEE et le cadre conceptuel de Budyko, le premier nous permet d'identifier la variabilité climatique incluse dans les paramètres d'ajustement du second, tandis que le second permet de quantifier l'effet relatif des facteurs non-climatiques, non considérés par les processus reproduits dans ORCHIDEE.

Cette méthode ne permet cependant pas d'identifier plus finement les facteurs inclus dans "les facteurs non-climatiques manquant dans ORCHIDEE". Dans un premier temps, nous pouvons émettre des hypothèses. Par exemple, nous trouvons une bonne corrélation entre les variations de l'efficacité évaporative des bassins versants en Espagne et l'évolution de l'eau stockée dans les barrages, ce qui en fait un bon indicateur de l'effet des activités humaines sur l'hydrologie des bassins versants de cette région. L'ajout de l'effet des changements de couverture végétale et d'utilisation des sols, tels que mis en oeuvre actuellement dans le LSM utilisé, a un effet non significatif sur le comportement hydrologique des bassins versants à l'échelle étudiée dans la présente étude. Plus généralement, de nombreux processus liés aux facteurs humains influent sur l'efficacité évaporative apparente d'un bassin hydrographique, souvent avec des rétroactions complexes et corrélées avec le climat. Afin de valider ces hypothèses, un approfondissement des recherches est donc nécessaire pour attribuer les tendances non climatiques détectées. Les futurs développements des LSM permettront de mieux intégrer les facteurs anthropiques tels que l'irrigation et les réservoirs. Une fois que les LSM pourront reproduire les débits réels et leurs changements, ils seront en mesure de décomposer et d'attribuer les changements non climatiques détectés.

Vers une modélisation à l'échelle régionale : besoin de forçages atmosphériques à haute résolution

La plupart des activités humaines qui influent sur le cycle de l'eau se déroulent à petite échelle, celle des réservoirs ou des périmètres d'irrigation, et les forçages atmosphériques actuels limitent la résolution d'exécution des LSM. La première étape vers l'intégration de ces processus dans les LSM a une échelle pertinente consiste donc à construire un forçage atmosphérique à plus haute résolution, pour tester les performances des LSM aux échelles auxquelles les activités humaines modifient le cycle hydrologique.

Ce défi est abordé dans la dernière partie de nos travaux. Nous nous concentrons ici sur une zone géographique particulière, entourant les Pyrénées. Cette zone est contrastée, riche en données d'observations et déjà fortement impactée par les effets du changement

climatique et par une forte gestion de la resources en eau, avec de nombreux barrages et une forte demande irrigative.

Nous cherchons ici à décomposer spatiallement et temporellement un jeu de données atmosphériques reconstruit à partir d'un réseau d'observations , couvrant la zone, à une résolution approximative de 1000 km^2 en valeurs journalières (données SAFRAN, Fig. 4). Nous sommes également ici confrontés à la rareté des observations dans les régions orographiques, réduisant la qualité de la reproduction des variables atmosphériques comme les précipitations dans ces zones, selon les modèles de désagrégation classiques. Pour y pallier, nous utilisons ici les résultats de modèles climatiques (ici RegIPSL). Ces modèles sont constitués d'un modèle atmosphérique et d'un modèle de surface couplés, qui ont tourné ici sur la région du sud-ouest de l'Europe à haute résolution, bornés par une réanalyse de données d'observations atmosphériques sur la période 2000-2009 (Fig. 4). Cela permet de reconstituer des champs atmosphériques cohérent spatialement et en altitude à l'échelle kilométrique. Par contre, ces données sont biaisées du fait des biais et de la variabilité interne aux modèles. Nous développons donc ici une méthode qui combine les deux types de données : SAFRAN nous sert de référence pour les valeurs journalières à l'échelle du jeux de données et les sorties du modèle haute-résolution sont réajustées pour correspondre à cette référence. Cela permet donc de conserver une désagrégation spatiale et sub-journalière cohérente physiquement. Selon différentes hypothèses, nous pouvons également tester la différence apportée selon si nous privilégions la distibution en altitude des précipitations du modèle ou de SAFRAN, pour tester les biais d'observation dans les zones en hautes altitudes.

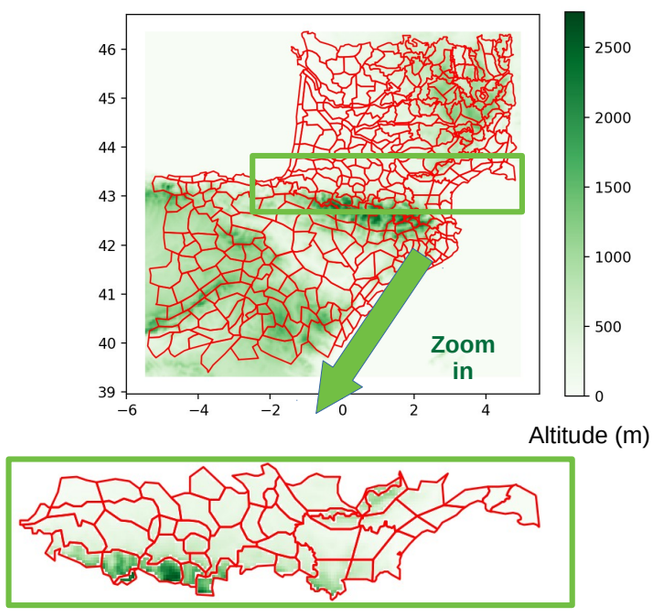


Figure 4 – Comparaison de l'échelle des différents jeux de données accessibles : SAFRAN qui correspond aux polygones rouges, et les sorties du modèles haute-résolution, ici la grille, colorée selon l'altitude associée à chaque point.

Les premiers résultats sont encourageant, avec notamment une distribution des précipitations en altitude en accord avec les hypothèses posées. Cependant d'autres tests de sensibilité et d'utilisation de ce jeu de données haute-résolution avec un modèle de surface sont nécessaires, pour valider et utiliser cette méthode dans un objectif de développer et améliorer les modèles de surface à haute résolution.

Conclusions et perspectives

Les progrès dans la représentation des processus physiques de surface intervenant dans la répartition et les dynamiques de la ressource en eau sont prometteurs. Les prochaines étapes de développement des modèles à haute-résolution vont permettre d'inclure des processus plus précis, à des échelles plus pertinentes, notamment pour les usagers et les décideurs. Cependant il ne faut pas perdre de vue les objectifs et les limites apportés par une trop grande complexité. Les cadres de modélisation plus simples et plus empiriques restent des outils de diagnostique à ne pas négliger.

Acknowledgments

I have many people to thank for these three years and for helping me achieve this work. First of all, I would like to thank my supervisors, Jan and Sophie. Jan, thank you for being always available, despite your heavy schedule. Thank you for introducing me to the topic, the problematics and for the always insightful discussions that we had, which would help me find new perspectives, new ideas and a regain of motivation! Sophie, thank you for the point of view and comments you brought to our monthly meetings, for your advice and literature recommendations and for your thorough review of all my writing! It was a great pleasure working with you two, I learned a lot, thank you for giving me that opportunity.

I would like to thank the members of my thesis comity, Fabienne and Antoine, for their support, along with the member of my jury for accepting to review my work and for their insightful questions.

I would also like to thank all the people who allowed me to broaden my experience and discover new work places and new ways to look at my work. Thank you to Fabien, Dimitri, Paul, Claire, Julio for letting me lend a hand in a field campaign and experience how it goes. Thank you Pere and Anaïs for your nice welcome in your Observatori in Roquetes. Thank you Anna and Luis for giving me the opportunity to discover your lab, Buenos Aires and a part of Argentina. Thank you Philippe and especially Riwal for helping me develop my teaching skills and getting to know more in interesting topics, thank you Riwal for your patience in answering all of my questions!

More generally, I would like to thank all the people, colleagues, who made these years very enjoyable. Thank you to all the youg researchers, PhDs and PostDocs paid to be nerds for the nice moments we spent together, the multi-cultural discussions, at lunch and around the Wednesday's cakes! Thank you to Samouro, Qiqi and Joan, for sharing an office with me and allowing me to talk to myself and to open the window once in a while. Thank you Imma, Patricia, Camille and teh others for the very nice moments at *La fête de la science*.

Finally I would like to thank my family and friends, for their support, their presence, for tolerating me when I wasn't realizing how tired I was (thank you Quentin!) and for coming to cheer me at the end!

Contents

CHAPTER 1

Introduction

Freshwater availability is a key resource for both ecological systems and human societies. Detecting and quantifying changes in streamflow series and understanding their drivers is essential, to organize water resources planning, and anticipate excess or scarcity in water availability (Coch & Mediero, 2016; Rodell et al., 2018), in a context of climate change and highly managed systems.

This manuscript addresses this issue, the current state of knowledge and the methods to better understand streamflow trends, and more generally the changes in the water cycle, due to both climate change and direct human influences on water management and land surface changes.

Contents

1.1 The water cycle

Water is a necessity for life on Earth. It circulates in a natural global cycle through the climate system, in different forms (solid, liquid, gaseous) and through different reservoirs (ocean, land, atmosphere, cryosphere). It is a closed cycle at global scale, which primarly involves precipitation and evaporation (Fig. 1.1). Water falling as precipitation was once evaporated from oceans and, to a lesser extend, from continental surfaces and transported through the atmosphere. As it is a cycle, evapotranspiration transports moisture back to the atmosphere and it eventually precipitates again (Riedel & Weber, 2020). More closely, a myriad of factors drive the equilibrium and exchanges between the different forms and reservoirs of water, playing a role at different spatial and temporal scales (Fig. 1.2). Terrestrial freshwater represents less than 2% of all water with only 4% of freshwater considered easily accessible and available for essential ecosystem functioning and human society's water resource needs (Douville et al., 2021).

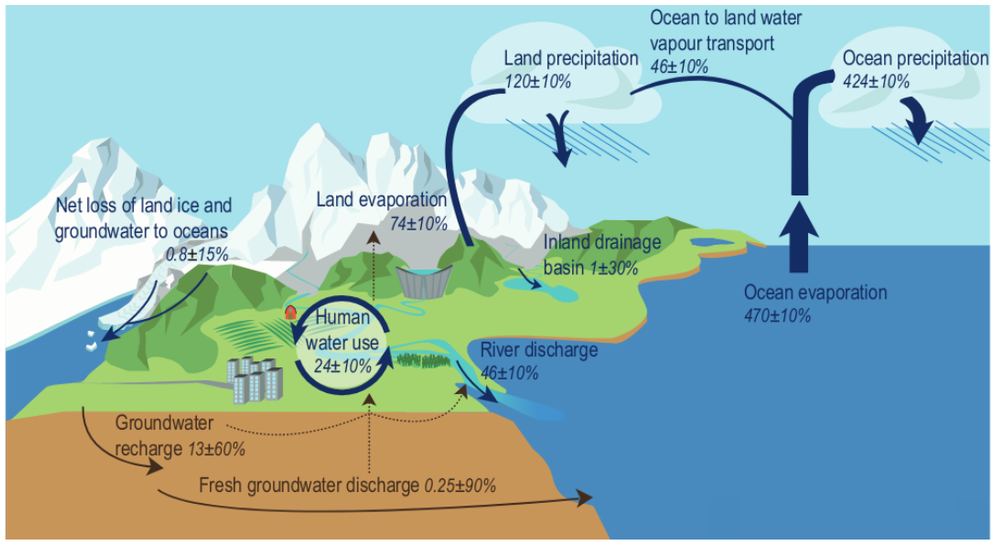


Figure 1.1 – Water fluxes for the present-day water cycle (in thousands of km^3 per year). From Douville et al. (2021).

If the water cycle is closed at global scale, with a conservation of water quantity along the cycle, it is also the case at more regional scale, such as the scale of a watershed. A watershed (or catchment) consist of all the land surface area channeling water towards the same given outflow point, generally to a stream or a river. For such system the water budget is balanced and we can summarize it at the scale of a river catchment with the following equation 1.1 over a given period, with incoming and outgoing water from a region in balance with the change in water storage (Müller et al., 2021):

$$P - E = Q + \Delta S \tag{1.1}$$

with P the integrated amount of precipitation, E the integrated amount of evapotranspiration, Q the discharge at the outlet of the catchment and ΔS the change in the amount of water stored over the catchment (Fig. 1.3). The amount of water is conserved and understanding water availability and variability in the cycle relies on understanding the partitioning of water between the different terms (Abatzoglou & Ficklin, 2017), the different reservoirs and the different forms, at different temporal scales.

The water cycle is closely linked to the energy cycle through energy budgets and thermodynamics processes at the surface and in the atmosphere, determining water states and exchange fluxes (Fig. 1.2). Incoming radiations warm up the air and the surface, playing

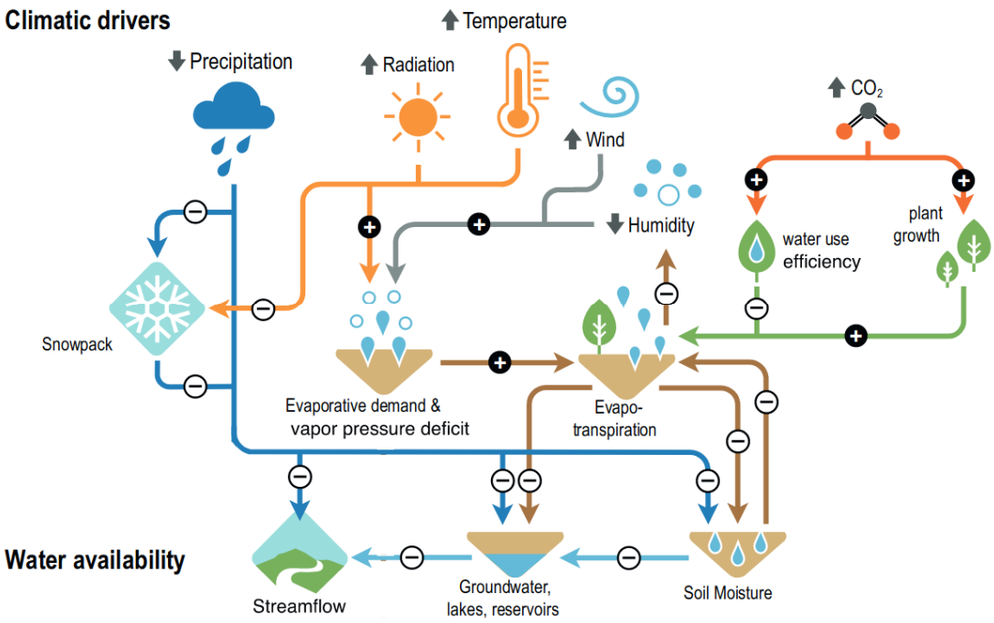


Figure 1.2 – Climatic drivers playing a role in water availability. Plus and minus signs denote the direction of change that drivers have on factors such as snowpack, evapotranspiration, soil moisture, and water storage. From Douville et al. (2021).

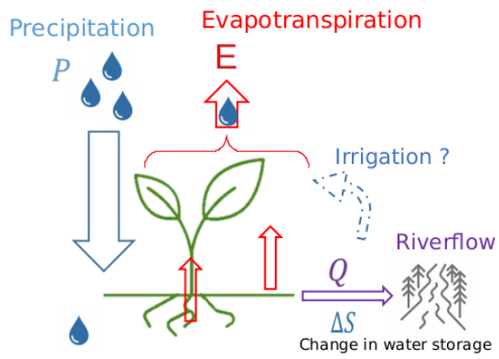


Figure 1.3 – Scheme of the main fluxes P, E, Q and ΔS over a catchment.

a role in defining air, surface temperature and heat fluxes. Clausius-Clapeyron equation defines the relationship between air temperature and how much water can be contained in vapor form in the atmosphere (water holding capacity), therefore defining atmospheric specific humidity. Atmospheric vapor content and vapor deficit (relative humidity) then drives condensation and evaporation processes, defining latent heat fluxes which are critical for driving atmospheric circulation at different spatial and temporal scales. In turn atmospheric circulation and turbulences drives horizontal moisture transport and therefore vapor-pressure deficit, relative humidity spatial distribution and regional $P-E$ patterns over land (Douville et al., 2021; Riedel & Weber, 2020; Sherwood & Fu, 2014). The dynamics of the water cycle is also affected by the different dynamics at different time scales. Synchronization between water availability and energetic atmospheric demand determines the partitioning between evaporation and runoff. For instance, the physical form of precipitation (rain vs snow) will change the timing of evaporation and runoff, depending on the residency time of water in the snow-cover surface-storage reservoirs.

This partitioning of water between the different terms is also dependent on land surface characteristics such as soil hydraulic properties and water holding capacity, slope, vegetation rooting depth, vegetation type which changes plants' water uptake and transpiration (water use efficiency) and thus soil moisture and evapotranspiration (Fig. 1.2) (Douville et al., 2021; Riedel & Weber, 2020). This links the water cycle to the carbon cycle as well, through

changes in vegetation dynamics and other related land surface characteristics.

All these dependencies result in very complex processes, dependent on the atmosphere, the soil, the vegetation, and which can be difficult to measure, such as evapotranspiration E (Quintana-Seguí et al., 2020). Different types of models have been developed to reproduce partially or fully the continental water cycle, with different levels of complexity, including different processes, at different temporal and spatial resolutions. These models allow different evaluations of how the different processes weight relatively to each other at different temporal scale. More details on the variety of existing models are given in chapter 2.

Despite the complexity of the processes in play, there is a necessity to understand and characterize the local and regional changes, to better address water management issues, which are critical for the environment and for societies.

1.2 Changes in the water cycle due to climate change

The continental water cycle is dependent on many climate variables, mainly through precipitation and through variables such as radiation, turbulences and water holding capacity, driving atmospheric circulation and vapor-pressure deficit. Therefore climate variability and long-term change impacts the water cycle and the partitioning of water between the different components of equation 1.1. In the following, we take an interest on how these variables have changed and are expected to continue changing, with a focus over Europe.

1.2.1 Observed changes in climate records over the past century in Europe

1.2.1.1 Changes in precipitation (P): average, seasonality, extremes

Different datasets exist to reconstitute past precipitation time series, based on extensive networks of weather observation stations and sometimes satellite observations over recent years. One widely used of such datasets is the Climate Research Unit (CRU) time series (TS) which is derived from the interpolation of monthly climate anomalies and covers the period 1901-2018 on a 0.5°latitude by 0.5°longitude grid (Harris et al., 2020). These datasets are regularly improved and updated. They allow to study how climate variables and mostly precipitation have been evolving over the past century, and to test the performance of climate models in reproducing these changes.

Using the CRU-TS dataset, Christidis & Stott (2022) study the pattern of change in P over Europe, over the different seasons (Fig. 1.4). They find that seasons gets dryer over the Mediterranean basin and wetter over the rest of the continent, except in summer where drying trends are more widespread. These trends are generally stronger in winter than in summer, showing that not only the average annual of P is changing but also its average distribution across seasons during the year. These results are in line with those of other studies covering the same area and a similar period as reported in the small review of P changes over Europe by Riedel & Weber (2020). Zveryaev (2004), using an old version of the CRU-TS dataset covering 1901-1978, completed with the Climate Prediction Center Merged Analysis of Precipitation (CMAP) over 1979-2001 study the seasonality variability of P over Europe. They similarly find in winter a decrease in precipitation over Italy and the Mediterranean region and an evident increase in other parts of Europe. In summer the precipitation decrease extends to the British Isles and most of western, central and eastern Europe with the increase in P localized to Scandinavia and most of European Russia. They

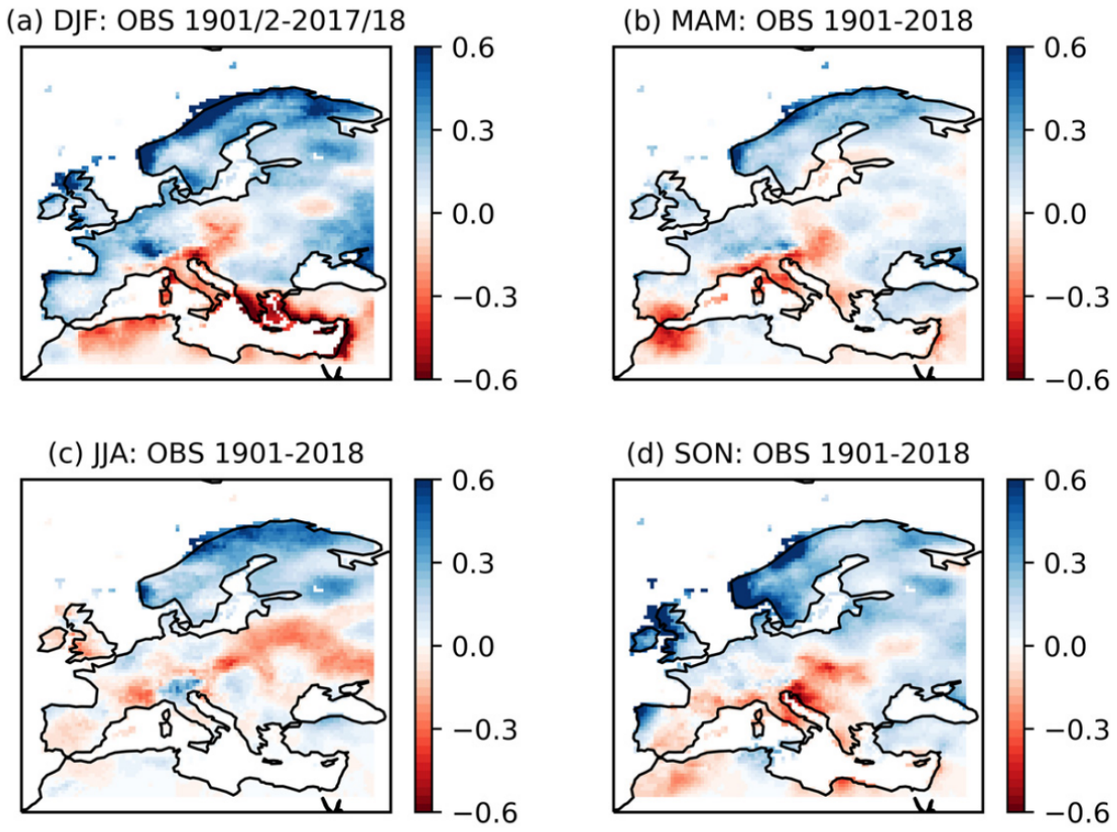


Figure 1.4 – Seasonal precipitation trends (mm yr^{-1}) calculated with CRU TS4 data over 1901–2018. Each panel corresponds to a different season, marked in the title by the first letters of the months within the season. From Christidis & Stott (2022).

also point out that year-to-year precipitation variability is smaller in summer than in winter over western Europe. In a review specific to the Mediterranean area for the period 1950-2002, García-Ruiz et al. (2011) sum up findings towards a general decrease in P, with greatest decrease in summer and spring, specifically over the Iberian Peninsula. Looking more precisely to changes in the intra-annual distribution of P is difficult to assess, and only a few indices exist to measure the inter-annual changes in the distribution of climate variables. For example, García-Barrón et al. (2013) define indices to assess the evolution of the intra-annual cycle of P over time throughout the Iberian peninsula. At the end of the century, they identify a shift of the main rainfall periods towards autumn, specifically over the Atlantic basins, and an increase in the inter-annual variability of the intra-annual cycle, especially over the Mediterranean basin.

Annual and seasonal trends in P are completed by changes in extreme P events and associated dry spell/ drought and water excess events. For instance, there is an increase in summer rainfall variability which broadens its distribution and therefore increases the likelihood of dry and wet extremes, even though the average P is decreasing over that season (Christidis & Stott, 2022). Over the Mediterranean area, most studies report an increase in precipitation intensity and dry spell/ drought events (Douville et al., 2021; García-Ruiz et al., 2011). More generally over Europe, there is a consensus that precipitation have intensified (Douville et al., 2021), with more frequent and longer intense rainfall events (Riedel & Weber, 2020). Focusing on the French Mediterranean area for the period 1961-2015, Ribes et al. (2019) find a substantial change in annual maximum of observed daily rainfall starting in the 1990's, with a significant increase in the number of events, specifically above high thresholds ($>$ 200mm). Overall their results are consistent with a broadening of P distribution, with an average drying and decrease in intensity of moderate and strong events but a significant increase in volume and extend of high threshold events.

There is also a change in the liquid/solid fraction of P and in the snowpack extent over the year, which overlaps with the changes in the total amount of rainfall over the Alps and the Pyrenees. Not only the precipitation, specifically in winter, are increasing over mountainous areas but the share of P falling as snow is decreasing (Rottler et al., 2020; García-Ruiz et al., 2011; Douville et al., 2021).

Generally, average trends per decade for all P characteristics are less significant due to the high inter-annual variability of precipitation P (Douville et al., 2021).

1.2.1.2 Changes in potential evapotranspiration (PET)

As previously explained, evapotranspiration (E) is flux dependent both on the water and energy cycle in very complex processes linked to the atmosphere, the soil and the vegetation. It is very difficult to measure (Quintana-Seguí et al., 2020), especially at large scale due to an added difficulty of high spatial heterogeneities which complicates extrapolations from targeted measurements. Therefore to estimate E most datasets and studies use modeling, sometimes complemented with satellite data, such as for the Global Land Data Assimilation System (GLDAS) (Fang et al., 2008). Modeling E most often relies on the concept of potential evapotranspiration (PET) to study the changes in climate variables driving E. PET represents the potential for E in a system not limited by water availability (Sherwood & Fu, 2014; Milly & Dunne, 2016). It corresponds to the atmospheric demand for water, limited by available energy and aerodynamic resistance (Barella-Ortiz et al., 2013; Yang et al., 2018). The hypothesis behind such a variable is that it can be defined relatively independently from the varying state of the surface, only dependent on atmospheric characteristics. However, the energy cycle and balance is also linked to surface processes, therefore such an hypothesis can only be an approximation. Still PET is a useful tool to estimate a change in atmospheric water demand. Several methods exist to calculate PET, dependent on the processes selected to define it. One common method is the Penman–Monteith equation, either over an open-water surface or an idealized reference crop (Milly & Dunne, 2016). More complex climate and land surface models can solve a more complex system of equation, including for instance changes in stomatal conductance, on shorter time-steps (Milly & Dunne, 2016; Barella-Ortiz et al., 2013). Depending on the method used and the variables it relies on, the magnitude of PET and PET changes can differ (Milly & Dunne, 2016). Still, there is a consensus that over Europe the atmospheric evaporative demand has increased (Riedel & Weber, 2020; Vicente-Serrano et al., 2019; Douville et al., 2021). This is probably due to the increased warming of the atmosphere and land surface, increasing atmospheric water holding capacity and increasing turbulence dynamics. However, the main drivers behind this evolution (changes in radiation, wind, temperature...) are still under debate (Riedel & Weber, 2020; Vicente-Serrano et al., 2014). Over the Iberian Peninsula, between 1961-2011, Vicente-Serrano et al. (2014) relates changes in PET and E measured on sites to a decrease in relative humidity, driven by an increased atmospheric water holding capacity because of higher temperature and by a decrease in moisture supply, specifically in summer. Other studies negatively correlate changes in reference PET to sunshine duration and cloudiness anomalies (Kitsara et al., 2013). These are correlations and no true attribution to actual causes, which are complicated by the inter-related processes involved and the many existing expression of PET.

To sum up, we observe, over the past century in Europe, an increase in PET and significant changes in P, rather decreasing in the South and increasing in the North, and with a contrast between seasons. It now raises the question of how it will continue to evolve in the

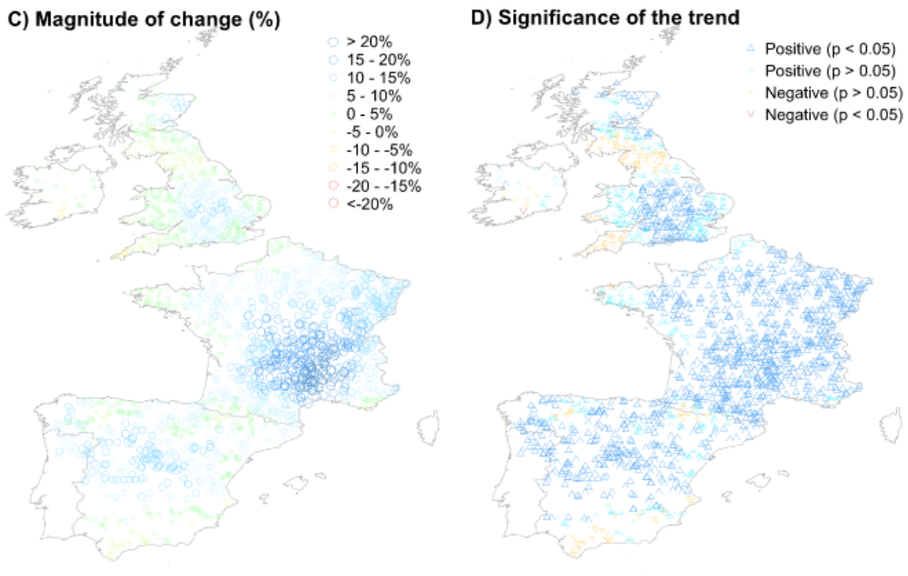


Figure 1.5 – Spatial distribution of the magnitude of change (1961–2012) in annual atmospheric evaporative demand (AED) (here equivalent to PET, monthly series of AED were obtained using the FAO56-Penman Monteith equation) and their corresponding significance (at p-value < 0.05). Each point represents an individual gauging station at which the annual AED is integrated throughout the entire drainage basin. From Vicente-Serrano et al. (2019).

future.

1.2.2 Modeling and projection of future changes in climate variables

Complex models have been built and are continuously improved. Several projects gather their outputs and projected changes in climate variables under different projection scenarios, such as the World Climate Research Programme (WCRP) Coupled Model Intercomparison Project (CMIP). CMIP experiments (last phase CMIP6) consist of an ensemble of state-of-the-art climate models performing common experiments and historical simulations (1850-near present), with their outputs standardized and made accessible to the scientific community (Eyring et al., 2016).

Many studies, using ensemble outputs from General Circulation Models (GCMs) mainly from CMIP5 and CMIP6 experiments, analyze the projected changes in climate variables over Europe (Douville et al., 2021; Christidis & Stott, 2022; Knutson & Zeng, 2018; García-Ruiz et al., 2011; Riedel & Weber, 2020; Schneider et al., 2013; Dai, 2016; Dezsi et al., 2018). Most often, the average over the model ensemble is used to smooth out the internal variability of each model and alleviate the noise effect (Christidis & Stott, 2022; Dai, 2016). All studies project an increase in the tendencies already observed over Europe. Models project consistent tendencies toward increasing temperatures and extreme hot events all over Europe (Fig. 1.6) (Douville et al., 2021; García-Ruiz et al., 2011; Schneider et al., 2013; Dezsi et al., 2018) along with projected increase in PET (Dai, 2016; Riedel & Weber, 2020; Arnell, 1999). Since P is a more chaotic variable, projections of future evolution of temperature (error ± 1 °) are more reliable than for precipitation (error ± 25 %) (García-Ruiz et al., 2011). Still projections under different scenarios, from 2050 to 2100, project a continuation of the drying over the Mediterranean area and the rest of Europe getting wetter (Christidis & Stott, 2022; Knutson & Zeng, 2018; Arnell, 1999; Riedel & Weber, 2020; Dai, 2016; Dezsi et al., 2018). Seasonal contrasts are expected to increase with a higher increase in P in winter (Fig. 1.6) (Christidis & Stott, 2022; García-Ruiz et al., 2011), leading to

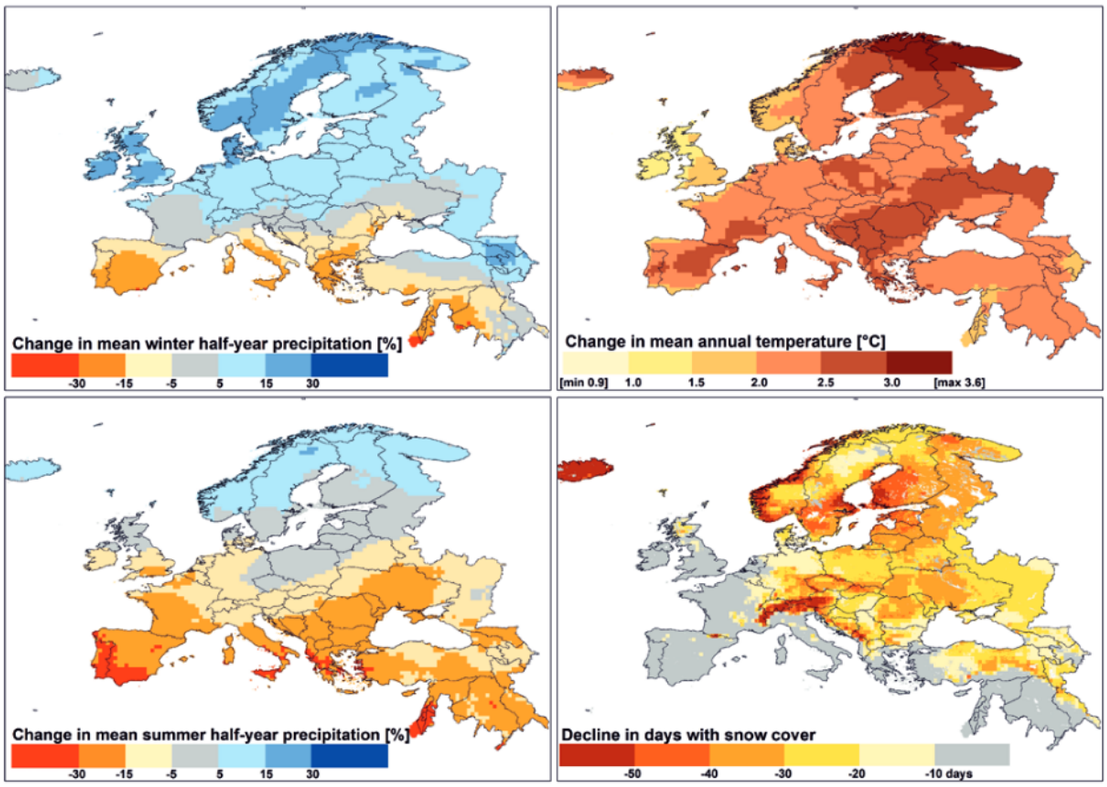


Figure 1.6 – Climatic changes in the 2050s featuring changes from the baseline (1971-2000) in mean precipitation of winter and summer half-year (left), mean annual temperature and snow cover duration (right). The maps represent the ensemble mean of the climate projections with three state-of-the-art GCMs (ECHAM5/MPI-OM, model from the Max-Planck Institute for Meteorology, Germany,; IPSL-CM4 model from the Institute Pierre Simon Laplace, France; CNRM-CM3 model from Centre National de Recherches Meteorologiques, France) under SRES A2 emission scenario. From Schneider et al. (2013).
These results are not from the most recent simulations but still represent well the expected pattern of changes in these variables, since the main differences with recent results are the amplitude of the changes.

increased contrasts between wettest and driest months (Fig. 1.6) (Douville et al., 2021). Amplified seasonality is expected to be more marked over the Mediterranean area where the atmospheric evaporative demand is high (Douville et al., 2021). Not only marked seasonality are expected to increase but also more generally the variability of P, increasing the likelihood of extreme events. Seasons with extremely high precipitation anomalies are expected more frequently over most of Europe, specifically in winter, less likely over the Mediterranean area (Christidis & Stott, 2022). Extreme dry events increase in likelihood and duration over most European areas (Christidis & Stott, 2022), especially in autumn and summer and over the Mediterranean area (García-Ruiz et al., 2011). Snowfall and the duration of snowpack are expected to reduce (Fig. 1.6) (García-Ruiz et al., 2011) (Schneider et al., 2013).

All these projections simulate an intensification of the water cycle and significant changes in P and climate variables related to E, leading to expected significant changes in water availability (Christidis & Stott, 2022; García-Ruiz et al., 2011; Dezsi et al., 2018).

1.3 Changes in land surface characteristics

The dynamics of partitioning of water between the different reservoirs and between runoff and evapotranspiration does not only depend on atmospheric and climate characteristics but also on many characteristics of the land surface, from soil and vegetation characteristics to water management activities lead by human societies. Furthermore, along with climate

change, these characteristics may evolve and have inter-related feedbacks with atmospheric variables, complicating the understanding on how they will impact the different components of the water cycle.

1.3.1 Vegetation cover and land use changes

Vegetation organizes itself according to energy and water availability (Fig. 1.7). For instance high water availability leads to higher vegetation growth, such as in valleys and riparian wetlands. There is also a difference in vegetation organization between sunny and shady slopes, due to differences in exposure and balance between water and energy availability (Fan et al., 2019). In turn, vegetation also plays a role in the organization of catchment, retaining soil and water through their rooting system, limiting erosion over slope hills and enhancing evapotranspiration through transpiration. Natural landscapes features co-evolve with other natural drivers of catchments' hydrological response, such as climate, with empirical evidence of interaction and feedbacks (Troch et al., 2015). Vegetation impacts several hydrological processes such as interception, infiltration and the partitioning of water between runoff, groundwater recharge and evapotranspiration (Douville et al., 2021; Riedel & Weber, 2020; García-Ruiz et al., 2011).

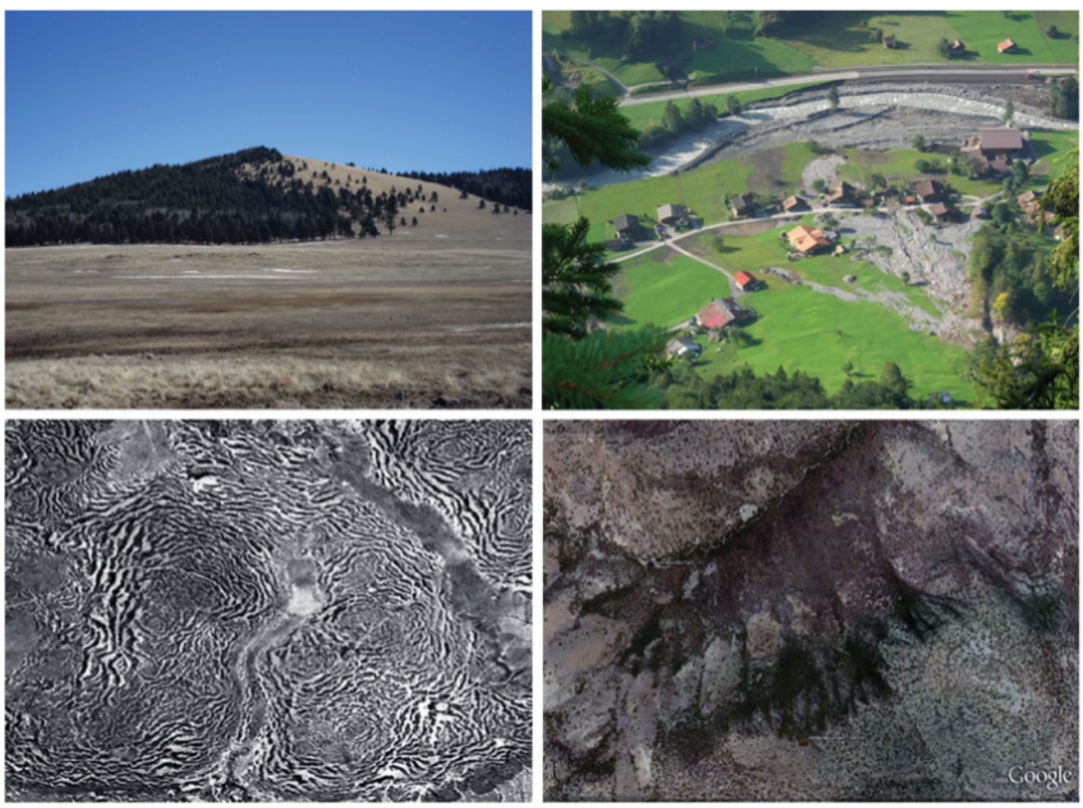


Figure 1.7 – Spatial patterns that reflect a coevolving landscape. From top left clockwise: Hill in the White Mountains of Arizona where coniferous trees grow on the north face and grasses on the south face; mudflow induced by landslides destroy road and property, and flood alters course of a river; riparian vegetation grows along washes in semiarid environments; aerial view of a tiger bush plateau in Niger reflecting surface runoff processes. From Troch et al. (2015).

Changes in vegetation and more generally land use and the state of the surface are therefore expected to highly alter the continental water cycle. Land cover and land use (LCLU) is affected by the economic development of countries and regions. In Europe and over the past century, rural exodus led to farmland abandonment and to a concentration

of human population and activities in urban centers, along coastal and more productive lands (valley floors). Along with reforestation policies, it led to an expansion of shrubs and forest areas, specifically in mountainous and hilly areas (García-Ruiz et al., 2011; Riedel & Weber, 2020). With climate change, vegetation activity has also changed and is expected to continue changing in the future. A dryer climate and a general increase in CO_2 atmospheric concentration impacts stomatal closure and reduces plants transpiration which is estimated to count for 55-67% of evapotranspiration (Riedel & Weber, 2020). Over the past few decades, the length of the growing-season has increased in Central Europe due to earlier start of flowering (Riedel & Weber, 2020). A vegetation more active in spring and less active in the end of summer will change the yearly balance of evapotranspiration. Species abundance and diversity are also already evolving, following changes in climate conditions (García-Ruiz et al., 2011; Riedel & Weber, 2020).

1.3.2 Irrigation, human water management and other human influences

In Europe, along with farmland abandonment, the agriculture practices have intensified over the maintained crop lands, with an increasing use of irrigation, specifically in Mediterranean regions and over the Iberian Peninsula (largest irrigated area in the European union) during the second half of the 20th century (García-Ruiz et al., 2011; Hurtt et al., 2020). Since the 2000's, the extend of irrigated land has stabilized in developed countries with expected evolution rather around irrigation techniques and practices towards more water efficient systems (Angelakιs et al., 2020). Irrigation is the largest anthropogenic global freshwater use. It has an impact on temperature in the crop canopy, crop heat stress and more generally changes the water and energy balances by increasing vegetation activity and enhancing evapotranspiration. Water used for irrigation either comes from uptake from rivers or from groundwater mining. Therefore it is associated to hydraulic infrastructures making water available and regulating and transporting water resources, such as dams. In the case of groundwater abstraction, it concerns mostly Spain, the south of France and Italy in Europe (Wada et al., 2012), with a considered unsustainable mining in Spain, where the aquifers' recharge does not balance the abstraction to cover the water demand (Holtz & Pahl-Wostl, 2012; Llamas et al., 2015; Wada et al., 2012; Custodio et al., 2016; Esteban & Albiac, 2012). Irrigation can have a positive effect on groundwater recharge through leakage and increased infiltration but the total effect on groundwater storage is negative when irrigation water is sourced from groundwater reservoirs (Riedel & Weber, 2020).

In Europe, 40% of withdrawals are used for agriculture (Sordo-Ward et al., 2019). Other activities use freshwater and play a role in water withdrawals and changes in ecosystem dynamics such a domestic consumption and hydro-power in the case of dams. Rivers have also been artificially modified (channelization, embanking, straightening, widening, deepening) with further impacts on flow and flow velocity (Schneider et al., 2013). As another impact of human activities, soil sealing has been continuously increasing in the past decades with the extension of urban areas. It has a tendency to increase runoff and reduce infiltration (Douville et al., 2021), even if high uncertainties remain on understanding recharge dynamics under urban areas due to the large spatial heterogeneities in these areas (Douville et al., 2021).

More generally, land use changes and human activities affect the water cycle by changing the partition of water between runoff, evapotranspiration and groundwater recharge (Fig. 1.8) and confound the identification of climate-driven changes (Stahl et al., 2010).

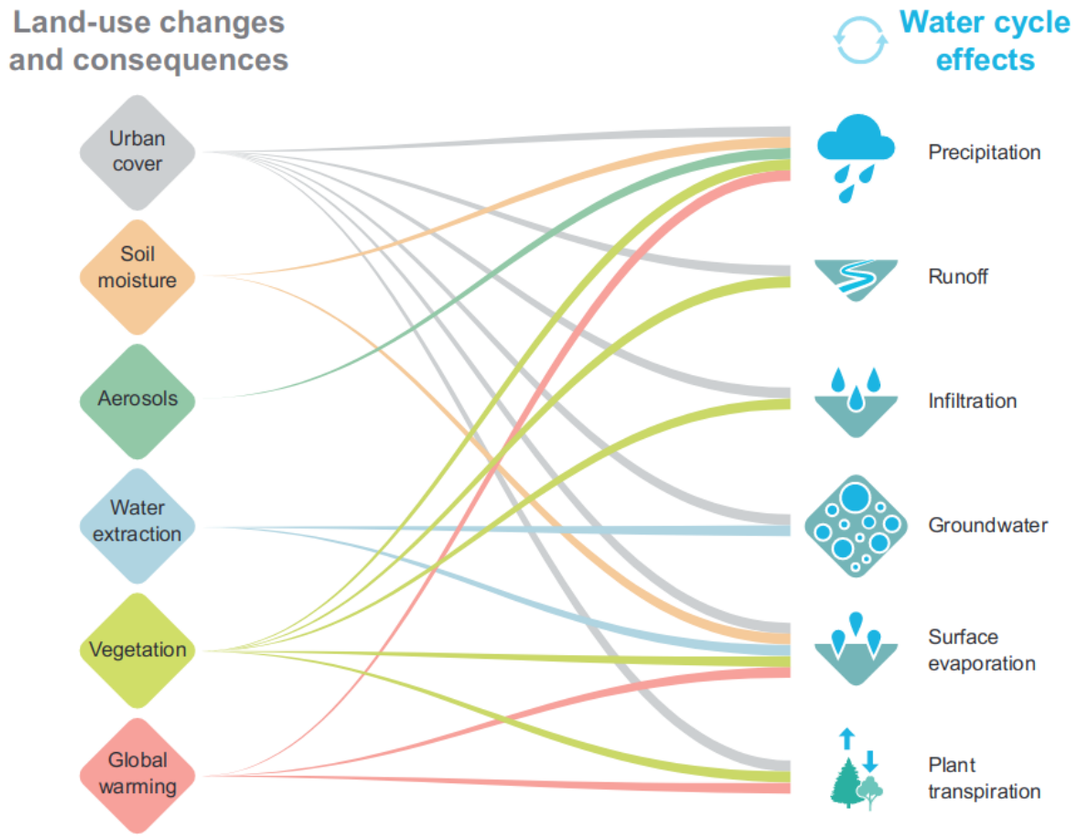


Figure 1.8 – Land-use changes and their consequences on the water cycle. As all the components or the water cycle are tightly connected, changes in one aspect of the cycle affects almost all the cycle. From Douville et al. (2021).

Projecting the future of these practices and activities depend on many factors. It depends on future management strategies which will adapt also to future climate conditions, to changes in policies and demand, in irrigation and reservoirs performances due to both improved technologies and climate dynamics... (García-Ruiz et al., 2011). It highly complicates the projections of changes in the water cycle.

1.4 Problematic and objectives

Both changes in climate and in land surface characteristics impacts the water cycle and its balance with the energy and the carbon cycle. Furthermore all these processes are interrelated and concurrent in space and time. Different drivers intervene on the water cycle on different components, at different spatial and temporal scales.

Of the main components considered in the water balance (equ. 1.1), evapotranspiration E and water storage ΔS are the most difficult to measure and are most often estimated through remote sensing and modeling (Riedel & Weber, 2020; Simons et al., 2020; Pan et al., 2019; Quintana-Seguí et al., 2020). To get a more comprehensive view of the changes in the water cycle and more easily compare models outputs to observations, most studies focus on the discharge Q. Q is also directly related to key environmental and societal challenges such as the rising flood risks, the seasonal dry up of river beds impacting local biodiversity or the changes in freshwater flows to the ocean, playing a role in ocean circulation and salinity along with ocean nutrient supply (Wang & Polcher, 2019).

Q is the temporally lagged, spatially integrated resultant of runoff and surface water fluxes over a catchment (Milly et al., 2005; Rottler et al., 2020). This is a strength to detect regional trends, by freeing ourselves from point measurements spatial and temporal

heterogeneities (Rottler et al., 2020). It allows to evaluate the resulting effect of the complex interplay between atmospheric drivers and multiple terrestrial processes at regional scale (Gudmundsson et al., 2017b). However that latter strength is also the main limit, since it complicates the interpretation of trends (Rottler et al., 2020; Milly et al., 2005; Stahl et al., 2010). The identification of climate driven changes is complicated by the confounding effects of direct anthropogenic disturbances and changes in watershed properties (Rottler et al., 2020; Milly et al., 2005; Stahl et al., 2010). This leads to the main question we wanted to address:

How to separate the effect of the climate and anthropogenic drivers in streamflow trends?

Understanding the relative role of such factors also plays a determining role for constructing adaptation scenarios, with a better comprehension of the weight of human activities on streamflow changes.

Most existing methodologies addressing this issue rely on modeling with their inherent strengths and limits. We want to find a new and innovative way to use these methods to better assess the relative magnitude of these different drivers on discharge changes with an objective of better understanding and improving how streamflow is projected.

Since more and more models start to include direct anthropogenic water management, which intervene at local scale, we also address one current limit, the necessity to have atmospheric forcings at higher resolution to use models at such scales and test the newly added processes. It is a step towards attributing the changes, which will complement detection methods. It is part of the LIAISE project (Boone, 2019), which aims at improving the understanding the human impact on energy and water cycles at regional scale (more details in chapter 5).

1.5 Introducing the thesis structure

This thesis manuscript is structured as follows:

- Chapter 2 first reviews the literature of streamflow studies, assessing observed streamflow changes, and different types of models, with their respective strength and limits. This review shows the difficulty linked to assessing trends as best as possible and interpreting and attributing these trends to possible climatic and anthropogenic drivers. Some type of models aim representing the current trends as best as possible to the detriment of understanding all the underlying processes. Other are physical-based and aim at understanding the different drivers and interactions but are a lot more complex to use and are limited by the processes they represent and therefore often lack accuracy.

- Chapter 3 presents the methods we developed to isolate and quantify the effect of climate change on discharge over the past century, benefiting from a parsimonious model simpler framework of interpretation and from the quality in process representation of a complex state-of-the-art land surface model. Our method allows to separate the effect on discharge trends, of precipitation inter-annual and intra-annual variability.

- Chapter 4 details a second application of our method, to isolate and quantify the effect of non-climatic drivers over Europe, compared to the effects of climatic drivers addressed in the previous chapter. The use of the parsimonious model and observations

allows to quantify the effect of missing processes in the complex land surface model. Hypotheses to identify the main non-climatic drivers are detailed, but attribution is not possible as long as these drivers are not included in physical-based models.

- Chapter 5 develops an essential step towards adding anthropic disturbances to physical-based models: the need to work at regional scale and therefore of an adequate high-resolution climate forcing. We use the results from a high-resolution run of a land surface model coupled to an atmospheric model, to get an adequate high-spatial and sub-daily disaggregation of climate variables. They are then bias-corrected at daily scale using a broader-resolution observation-based dataset.

- Finally, I conclude in chapter 6 with a summary of the main results obtained during this thesis work and with my understanding of the perspectives, next steps and challenges, as I understand them at the end of this thesis.

CHAPTER 2

Review of literature: Analyzing, reproducing and projecting discharge changes

As introduced, discharge is the resulting integration of surface flows over a catchment. It is therefore influenced by all phenomena and activities changing the balance between all water flows. We have seen that both climate variables and land surface characteristics have been concurrently evolving and are expected to continue to change in the following years. Since streamflow is a more integrated variable, it is not easy to understand and predict its resulting evolution. This chapter summarizes the current state of knowledge on observed streamflow changes and the methods developed to understand the processes behind these changes and predict their future evolutions.

Contents

2.1 Observed streamflow over Europe

2.1.1 Changes in streamflow over the past century

2.1.1.1 Observation network of gauging stations

Rivers are the main and easiest access to fresh water for human societies (Dai, 2016). Monitoring streamflow is therefore highly developed in most countries and relatively easy compared to the monitoring of other water fluxes such as evapotranspiration. Gauging stations exist all around the world, with many institutions gathering their data at different scale (region, country, world). For instance, the Global Data Runoff Center (GRDC) was established in 1988 in the German Federal Institute of Hydrology (BfG) in Koblenz, Germany, with the support of the World Meteorological Organization (WMO) to gather and archive global streamflow data. They have access to this day to over 10,000 stations from 159 countries (Federal Institute of Hydrology (BfG), 2023). However, not all area of the world are equally well represented: the spatial coverage, length and period covered by the time series vary across different regions (Fig. 2.1, 2.2). Europe, Americas, Australia, Russia, South Africa are well covered with long time period covered up to the last decade. Their is a clear lack of data in South and Eastern Asia and in recent years in the rest of Africa. There is overall a decline in discharge data availability in recent years (Fig. 2.2) due to the reduction in gauging stations with the development of satellite data and due to the increasing reluctance of some countries such as in Asia or Africa to share streamflow data, considered as strategic resources. A sparse coverage can lead to biased inferences of observation-based studies over large spatial domains wherever gauges are not a representative sample (Do et al., 2020). Moreover, not all time series have the same quality and many contain gaps of variable length (Dai, 2016). Despite their limitations, such databases exist and can be completed with other databases available at national scale in some part of the world. New developments can help to improve the measurement deficits in the future, such as the Surface Water and Ocean Topography (SWOT) satellite mission, launched in December 2022, which should be able to detect surface water elevation, slope and water mask for many surface water bodies with a goal for good performances for 50 m to 100 m wide rivers and water bodies (Biancamaria et al., 2016), and therefore enable to improve discharge estimation over areas with no access to in-situ observations.

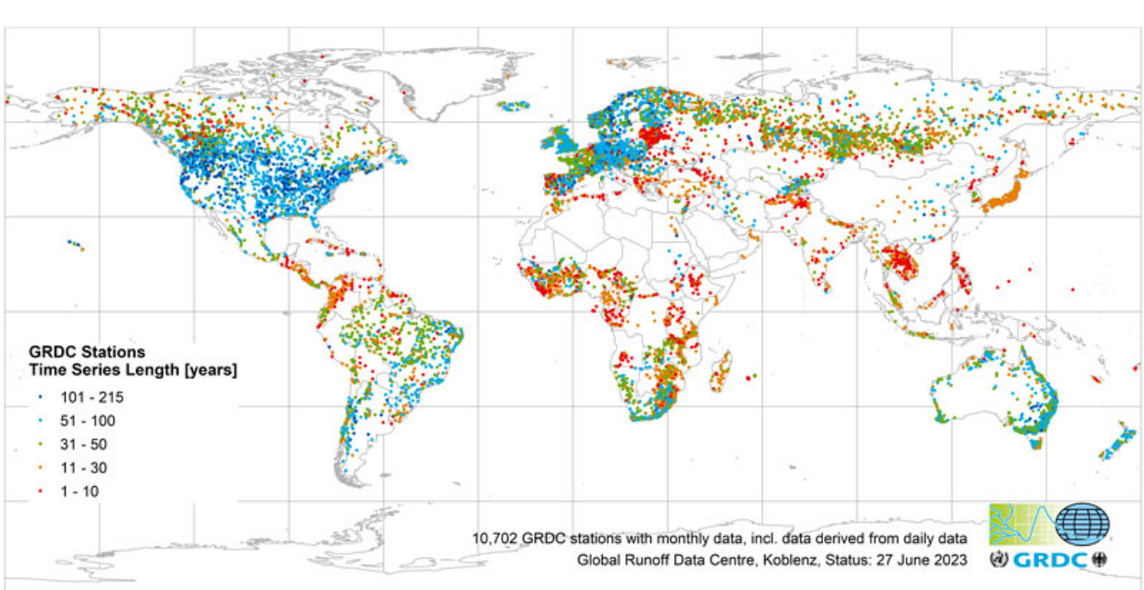


Figure 2.1 – Length of the time series for the 10 702 stations with monthly data, GRDC, from (Federal Institute of Hydrology (BfG), 2023), (27/06/2023)

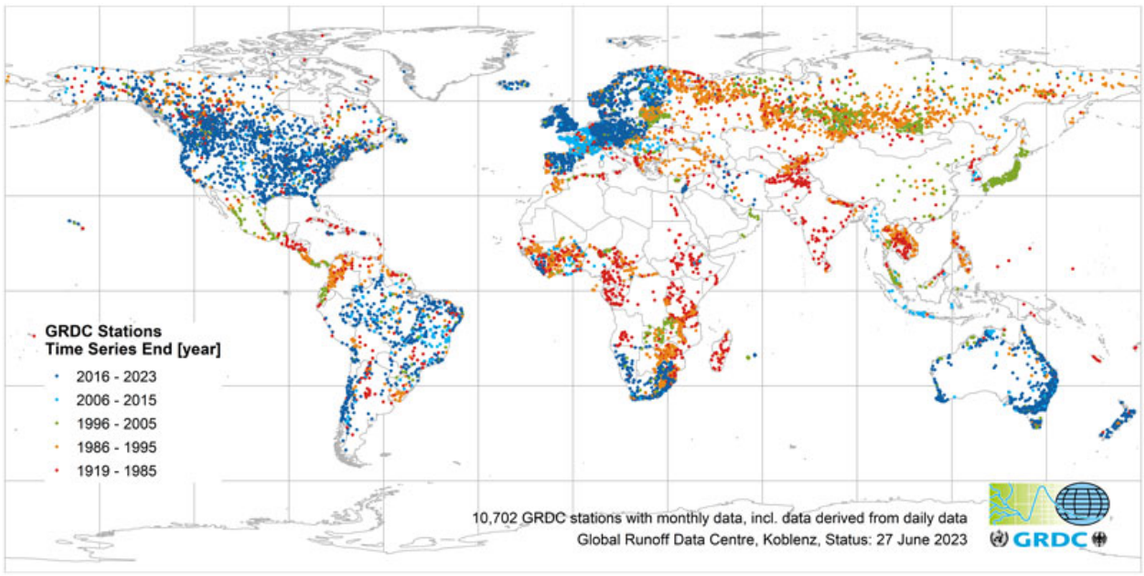


Figure 2.2 – End date of the time series for the 10 702 stations with monthly data, GRDC, from (Federal Institute of Hydrology (BfG), 2023), (27/06/2023)

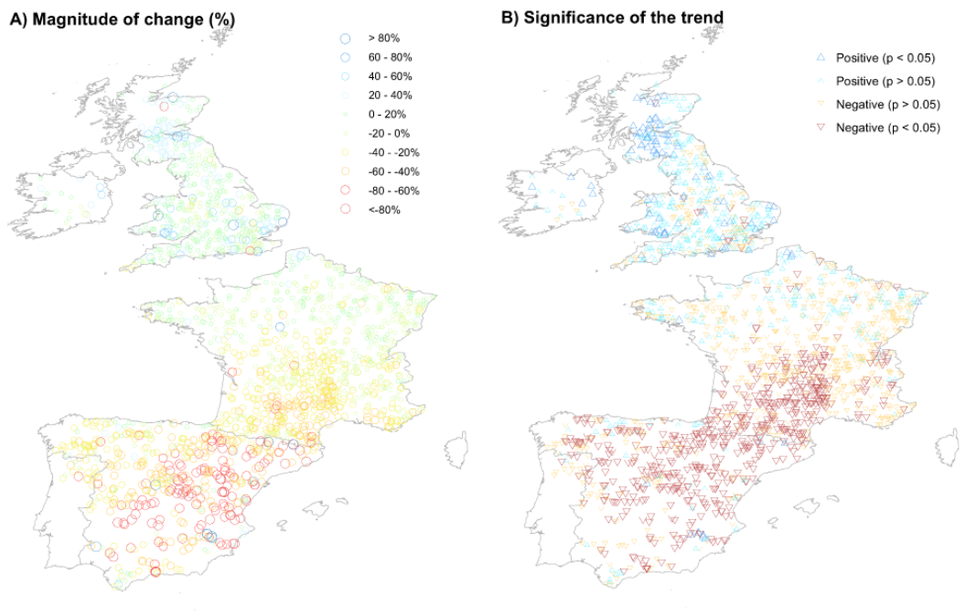
2.1.1.2 Changes in average annual discharge

Such databases allows for global scale studies to estimate the global continental freshwater discharge (Dai et al., 2009; Labat, 2010; Milliman et al., 2008). Such studies consider the largest rivers to estimate the general tendencies in continental freshwater discharge. However, depending on the number of rivers considered, the period covered and the method to fill observational gaps, discrepancies in results arise (Alkama et al., 2013). Furthermore, usually large main rivers may not be representative of unmonitored areas which are an important part of the freshwater streamflow and may have very different behaviors (Wang & Polcher, 2019; Hannerz & Destouni, 2006).

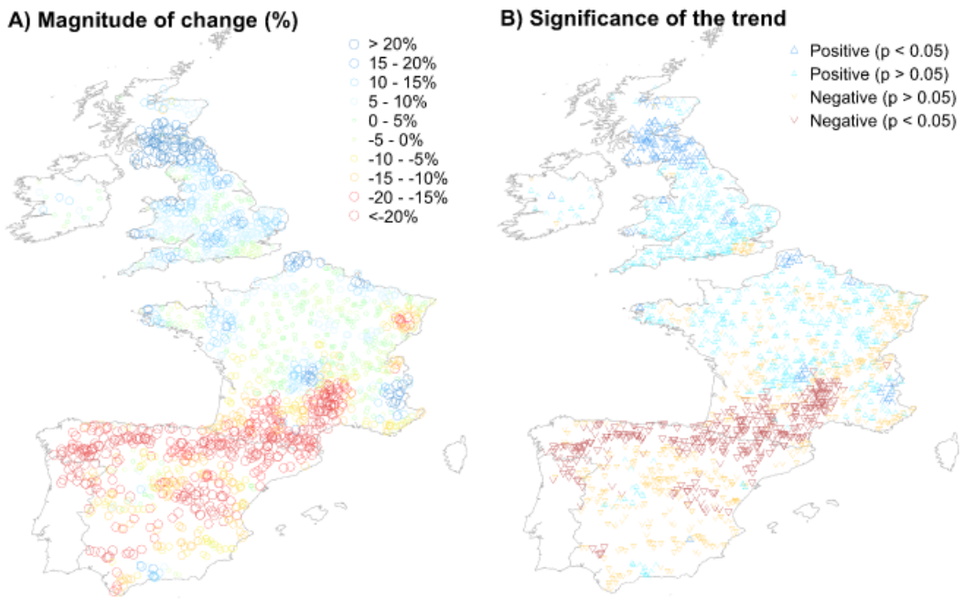
Observational datasets are also interesting to look at spatial patterns in discharge and trends or for more regional studies. Looking at more than 200 largest rivers worldwide from 1949 to 2012, Dai et al. (2009) find non significant trends for a majority of rivers and varying results depending on the time period studied. Streamflow variability and longterm changes are highly correlated to precipitation integrated over the catchment (Dai et al., 2009; Dai, 2016; Alkama et al., 2013; Milly et al., 2005) and therefore to atmospheric processes driving interannual variability of land precipitation such as El Niño-Southern Oscillation (ENSO) (Dai, 2016; Milly et al., 2005). This explains the small significance of trends for most large rivers due to the high resulting year-to-year variability (Dai et al., 2009) and to the integration effect.

If overall trends are mostly non significant, they are noteworthy regionally and for some individual rivers, highlighting that streamflow trends are more a regional/ basin-scale issue (Alkama et al., 2011). Such scale better account for regional specificities and differences between large and small rivers, upstream and downstream of catchments. It is also at that scale that the effect of human activities and land use change are most influential (Alkama et al., 2011).

Focusing on Europe and the Mediterranean area, several studies find average drying trends in Spain, southern France, more generally in the Mediterranean area (Gudmundsson et al., 2017b; Coch & Mediero, 2016; García-Ruiz et al., 2011; Vicente-Serrano et al., 2019; Stahl et al., 2010) and in central/eastern Europe (Stahl et al., 2010) and average wetting weaker trends in northern Europe (Gudmundsson et al., 2017b; Vicente-Serrano et al., 2019; Rottler et al., 2020). Similarly to the results at global scale, these trends are highly correlated



(a) Trends in annual streamflow (1961 to 2012)



(b) Trends in annual precipitation (1961 to 2012)

Figure 2.3 – Trends in annual streamflow and precipitation from 1961 to 2012. (A) Spatial distribution of the magnitude of change and (B) their statistical significance (at $p < 0.05$). Each circle represents one gauge station. From Vicente-Serrano et al. (2019).

and consistent to changes in precipitation (Fig. 2.3) (Gudmundsson et al., 2017b; Arnell, 1999; Zanardo et al., 2012; García-Ruiz et al., 2011; Vicente-Serrano et al., 2019; Stahl et al., 2010) and to the atmospheric circulation variability (Bouwer et al., 2008) and the North Atlantic Oscillation (NAO) (Gudmundsson et al., 2017b; Arnell, 1999).

2.1.1.3 High flows, low flows and river regimes

Regional studies do not focus only on annual average discharge. Indeed seasonal trends can be different from annual trends. For instance, Stahl et al. (2010) find the trends in discharge over the end of the century to be disconnected between summer flows and winter flows for an ensemble of small near-natural catchments in Europe, with a dominance of positive trends in winter and negative trends in summer. Focusing on the North of the Alps, Rottler et al. (2020) find changes in seasonality with an increase of streamflow in winter (and spring) and lower decrease or even an increase in summer (and autumns).

Other important aspects of streamflow studied are low and high flows (Tuel et al., 2022; Bouwer et al., 2008), associated to drought (Douville et al., 2021; Vicente-Serrano et al., 2014) and flood events (Douville et al., 2021; Rottler et al., 2020; Milly et al., 2002). They can show different tendencies than average trends due to their dependency on different phenomena, often more localized in time and space. Tuel et al. (2022), for catchments in Switzerland, study the link between extreme precipitation events and peak discharge events. They show that the discharge response does not only respond to the magnitude of the precipitation events but also to the clustering of these events. Clustered events generally result in larger, longer discharge response. It could be because of the saturation of soil moisture, as isolated events allow for more time for the soil to dry up, and therefore the next precipitation event generates less runoff than when events are clustered. However, they also show that this effect of clustering is attenuated for large catchments and snow-dominated catchments. Therefore the amplitude of these phenomena also depends on scale and on the type of flow regime followed by the river.

Current European hydrological regimes are mainly rain-fed and snow-dominated regimes (Arnell, 1999; Rottler et al., 2020). In the first types of regime, the intra-annual discharge variability is highly correlated to precipitation events, with usually a peak discharge in winter (Bouwer et al., 2008). The second type of regime concerns mostly continental regions where the peak in runoff follows snow melt. In these regimes, peak flows are less directly correlated to precipitation events. There is a delay due to the water storage in snow packs in winter, with discharge peaks following snow melt in summer. Reality is also never as simple and many catchments fall in between these classifications. For instance in their classification of river regimes, Poschlod et al. (2020) identify 6 different regimes: (glacio-)nival, nival (transition), nivo-pluvial and three different pluvial classes. Moreover, not only these regime determine different responses to climate drivers, themselves are expected to change with climate (Poschlod et al., 2020), mostly toward more rain-fed classes with the decrease of snowfall and snow cover. The increase in temperature results in less snow accumulating in winter in headwaters and to more rainfalls in winter, while snow melt also occurs more rapidly, earlier in the year. Both effects results in changes in river regimes, towards greater runoff in winter and lower earlier high flows in spring (García-Ruiz et al., 2011).

Blöschl et al. (2019b) study the trends in highest peak discharge from 1960 to 2010 over Europe (Fig. 2.4). They show that peak discharge trends follow similar patterns than the trends in average discharge: an increase of flood discharge in northwestern Europe, a decrease in southern Europe and eastern Europe. They distinguish three zones where these trends follow patterns of different possible climate drivers. In northern Europe, flood discharge increase matches the increase in winter and autumn rainfall. In southern Europe, the decrease in flood discharge matches the decrease in precipitation and increase in evaporation. In eastern Europe, the decrease matches rather the changes in snow melt dynamics (earlier snow melt and decreasing events). Similarly, and over the same period, Berghuijs et al. (2019) observe that winter annual floods in large parts of western Europe and the Mediterranean rather correspond to peaks in precipitation events and soil moisture excess. In northeastern Europe and around the Alps, the regimes are different, with respective annual flood events rather in spring and summer. In these regions, it does not match the peaks in soil moisture excess (also rather in winter) or the peak in rainfall events (rather in summer). It matches the annual maxima of snow melt and rain on snow. These results show that the dominating regime in western and southern Europe is rain-fed rivers, while in north-eastern Europe and around the Alps most rivers follow snow-dominated regimes, and that both types of catchments respond differently to climate change.

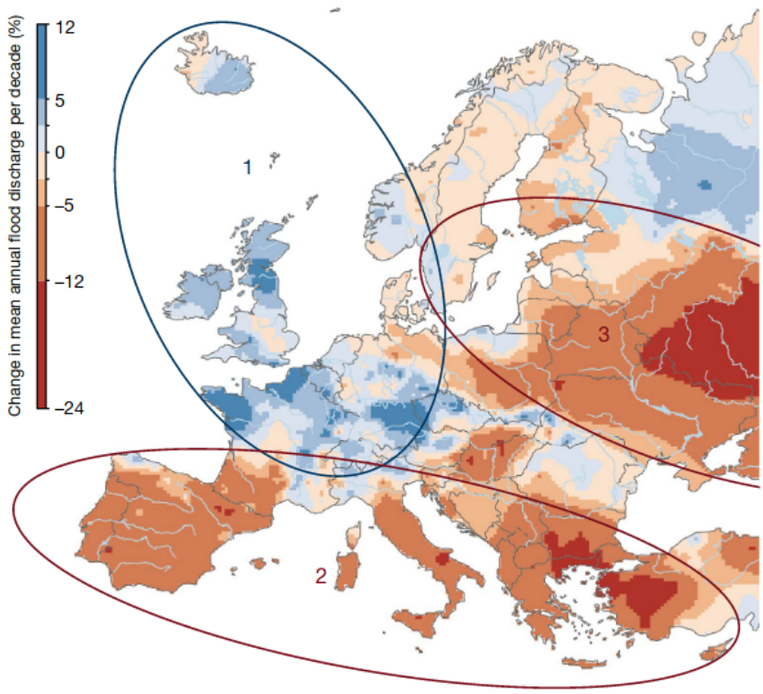


Figure 2.4 – Observed regional trends of river flood discharges in Europe (1960–2010). Blue indicates increasing flood discharges and red denotes decreasing flood discharges (in per cent change of the mean annual flood discharge per decade). Numbers 1–3 indicate regions with distinct drivers. 1, Northwestern Europe: increasing rainfall and soil moisture. 2, Southern Europe: decreasing rainfall and increasing evaporation. 3, Eastern Europe: decreasing and earlier snowmelt. From Blöschl et al. (2019b).

Most studies looking at the evolution of drought events in Europe focus on the most sensitive regions: the Mediterranean area. For instance, Coch & Mediero (2016) show a trend toward a decrease in low flows in Spain between 1949 and 2009, with similar patterns observed in the South of France, Greece and Turkey. Looking at the same region between 1961 and 2011, Vicente-Serrano et al. (2014) find an increase in drought severity and extend area, associated to a severe decline in streamflow, related to both a decrease in precipitation and to an increase in evaporative demand.

To sum up, this section shows that there is a general consensus in observed regional streamflow tendencies in Europe: drying trends in the Mediterranean and eastern area, wetting trends in north and western Europe. Trends in extremes events (floods and drought) follow more or less similar patterns than average flows but not always consistently, with for instance increasing high flows in drying areas, due to the heightened variability of river flows. An added difficulty when studying streamflow trends are the changes in seasonality and sub-annual dynamics. Most studies making assumptions describing potential processes underlying streamflow changes are localized. Indeed it is easier to make assumptions in smaller-scale specific systems, but it does not easily allow generalizations. Here we only give a descriptive overview of the climatic drivers assumed to play a major role in streamflow changes: mainly precipitation, snow dynamics and associated changes in soil moisture. In the following part, we will give an overview of how to better assess the role of climate and surface processes in streamflow dynamics, with the associated difficulties raised.

2.1.2 Separate climate vs human drivers on observations

2.1.2.1 Catchments complexity and potential intricated drivers

Streamflow characteristics and trends are dependent on the area, the climate characteristics, the climate regime, the scale of the catchment (Tuel et al., 2022). Discharge is the integrated result of non linear phenomena relating climate characteristics and landscape features. As introduced in chapter 1, the landscape organization determines how a catchment filters the climate into an hydrologic response. Troch et al. (2015) introduce the concept of hydrological age, linked to the concept of a co-evolution between landscape features and climate, which may depend also on how active the drivers are. Fluxes are driven by spatial characteristics which in turn can be modified by said fluxes. For instance, erosion increases slope which increases runoff which increases erosion. And the magnitude of that feedback chain will depend on how active the runoff flow and how steep the slope were in the first place. Vegetation also organizes itself according to water and energy availability and in turn changes evapotranspiration and runoff dynamics (Fan et al., 2019). All terrestrial water reservoirs are connected and trends in global freshwater availability are dependent on the interaction between all of them: groundwater, soil moisture, surface waters, snow and ice (Rodell et al., 2018). They have different dynamics in space and time, often explaining that trends in average flows and trends in high and low flows are not always similar, depending again on the catchment characteristics, the river regimes and the period studied. To understand changes in discharge, it is therefore important to take that complexity into account.

On top of these intricated non linear relationships, as we have seen in chapter 1, climate is changing and humans intervene to modify the functioning of catchments. Here we sum up the main climate and then human related drivers, which supposedly impact streamflow. If, as previously described, precipitation seems to be a key driver of streamflow changes, with similar patterns observed in discharge and precipitation changes, it is not sufficient to explain all of streamflow variability and trends. Vicente-Serrano et al. (2014) show that the decrease in streamflow over Spain is higher than the decrease in precipitation. They associate that difference in magnitude to the concurrent increase in evaporative demand. Still considering climatic potential drivers, we already discussed the potential effect of changes in timing and quantity of snow melt, towards increased runoff, earlier in the year and dryer summer months. Associated to that effect, the changes in glaciers dynamics is expected to impact streamflow in short and long-term. Glaciers contribute to river discharge, with a typical glacier runoff showing a seasonality with a minimum in snow-accumulation season and a maximum in melt season, which can compensate for possible low flow and drought events in lowlands (Huss & Hock, 2018). Schaner et al. (2012) estimate that glaciers contribute from 5 to 50% to streamflow in some catchments in the Alps, with a peak contribution in summer month when seasonal snowmelt from the non-glacierized part of a river basin is low, glacier melt is high, and other non-glacier sources of runoff are low. However, the acceleration of glaciers mass loss (Vincent et al., 2017) results in an expected short term increase in downstream streamflow (peak water) and an expected decrease in the long term due to glaciers depletion, especially a decrease in seasonal runoff maxima, also expected to shift towards earlier in the season (Huss & Hock, 2018). Peaks and changes in regime depend on the size of the glacier and its share to streamflow contribution (Douville et al., 2021; Huss & Hock, 2018). In regions with smaller glaciers as in Europe, the peak water years are expected to have passed or to occur within the next decade (Huss & Hock, 2018). These changes in regimes are still difficult to identify and quantify, remaining amongst the important questions to solve for hydrologists (Blöschl et al., 2019a).

Moreover, water systems are highly managed and under influence of human activities, masking potential effects of climate change on runoff and streamflow (Rottler et al., 2020; Ficklin et al., 2018). Slater et al. (2015) relate trends in floods both to trends in streamflow and to trend in river channel capacities. Rottler et al. (2020) show that the changes in seasonality in discharge coincide both to changes in climatic drivers such as precipitation and snow cover and to the constructions of reservoirs and anthropogenic alteration of river networks. There is a consensus that dams reservoirs and management change the seasonality of river flow (García-Ruiz et al., 2011; Rottler et al., 2020; Ficklin et al., 2018), as they are designed for such purposes. They retains water in winter to recharge and results in more uniform flows throughout the year (Ficklin et al., 2018), with more reduced peak and low flows due to their controlled releases and to maintained ecological flows. In the Pyrenees, they are associated to decrease in discharge due to divergence of water towards irrigation channels (García-Ruiz et al., 2011). More generally, flow alterations and reservoir managements, withdrawals for irrigation and changes in land cover are expected to affect both the average discharge and low/high flows. For instance over the Iberian Peninsula, García-Ruiz et al. (2011) review several studies making the assumption that the increase in the forest cover explains the decorrelation between streamflow and climate, with specific studies looking at unregulated rivers in the Pyrenees finding the expansion of forested areas to be responsible from up to 25% of the decrease in streamflow over the end of the twentieth century (Beguería et al., 2003).

These assumptions on climatic and non-climatic factors and processes driving streamflow changes are tested with different methods, described in the following parts.

2.1.2.2 Identifying the effect of potential drivers using observations

Observation-based studies to identify a signal and its potential drivers all rely on similar methods. These methodologies have been developed at first to look at the anthropogenic signal in climate change. The first step is to look at spatial association between different variables and identify long-term trends in time series usually dominated by short-term dynamics. If the trends and association are stronger than random variability, it identifies a possible cause of change (Hegerl & Zwiers, 2011). Similarly for streamflow, most observation-based studies compare spatial and temporal patterns of streamflow to those of possible drivers.

Several studies show high spatial correlation between streamflow trends and precipitation trends (Dai, 2016; García-Ruiz et al., 2011; Blöschl et al., 2019b; Bouwer et al., 2008; Vicente-Serrano et al., 2019). Correlation with temperature patterns are harder to assess (García-Ruiz et al., 2011), probably due to a more indirect link between streamflow changes and temperature changes. Other study directly look at the graphical patterns to identify more complex characteristics. Over the Po River in Italy, Montanari (2012) graphically studies average, peak and low flows, intra-annual variations of seasonality. He identifies an increasing drought risk in summer even though the traditional methods for trends detection show no statistically significant trends due to the high variability in discharge. Other studies look for breakpoints in streamflow series to separate a baseline and change period and associate breakpoints to potential drivers (Adeyeri et al., 2020; Ahn & Merwade, 2014; Fenta et al., 2017).

All these methods face the same limitations. Patterns and correlation can be hidden by confounding effects and internal climate variability (Hegerl & Zwiers, 2011; Ficklin et al., 2018). Possible long-term natural fluctuations may not be effectively captured by the length of the historical records (Montanari, 2012), a possible delayed response to a driver would not be identified (Hegerl & Zwiers, 2011). Not all drivers influence streamflow at the same

spatial and temporal scale. The result therefore may depend on the scale studied. More generally, these methods do not allow to attribute trends and patterns to specific physical processes, only to identify correlations, coincidences and similarities (Montanari, 2012).

As an attempt to attribute changes, several studies separate catchments in categories based on known characteristics. To separate the effect of climatic drivers from direct anthropic drivers, several studies categorize natural or near-natural catchments vs human-modified catchments (Ficklin et al., 2018; Coch & Mediero, 2016; Stahl et al., 2010; Beguería et al., 2003). Such a method allows to isolate the effects of climate change (Stahl et al., 2010) or other remaining drivers in near-natural catchments, such as reforestation (García-Ruiz et al., 2011; Beguería et al., 2003).

However these methods face other limitations. There is no global database of "naturalized" streamflow available (Alkama et al., 2011). Several countries have recently established so-called "reference" or "benchmark" networks but access to such a database regularly updated in Europe is complicated by the many jurisdictions responsible for data gathering and their willingness to share data nationally as well as internationally (Stahl et al., 2010). Therefore different studies use different database and definition of "near-natural" catchments. More generally, to minimize the impact of human disturbances, "natural" catchments are typically small and sparse since larger watersheds are more likely to have water infrastructure and management (Vicente-Serrano et al., 2019; Stahl et al., 2010; Ficklin et al., 2018). It highly reduces the surface covered by the analysis, especially over Europe due to its high population density and long history of water infrastructure development (Fig. 2.5) (Stahl et al., 2010).

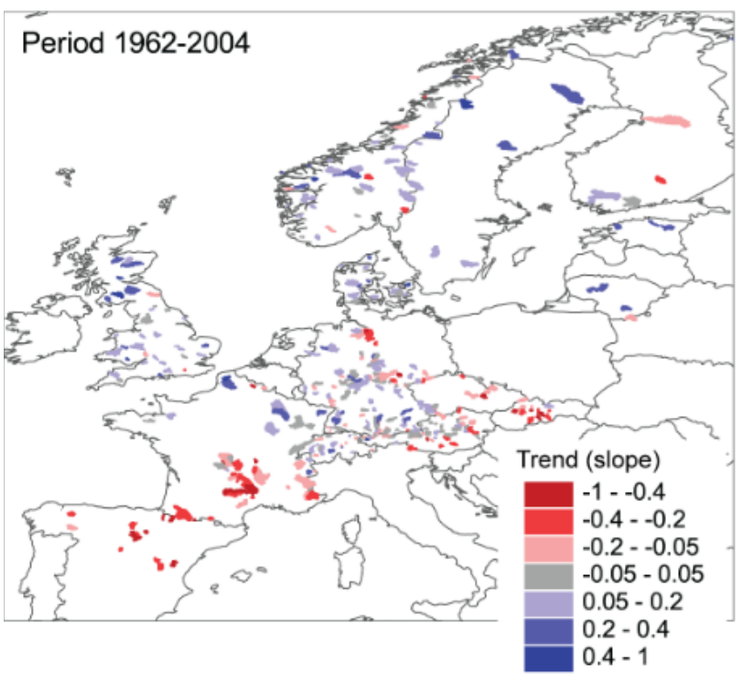


Figure 2.5 – Trends in annual streamflow for the period 1962-2004 (trends are given in standard deviations per year), for a data set of near-natural streamflow records of 441 small catchments. From Stahl et al. (2010).

The lack of representativity of such "near-natural" catchments to show the overall effect of climate change may not only be due to the small spatial cover of such defined catchments. By definition, "near-natural" catchments are often small catchments, upstream of river channels, and focusing on them to represent the effect of climate change may not represent well differences other than human intervention, such as differences in landscape characteristics, slope, altitudinal climate, and overall differences in climate reactivity between mountainous

and lowland areas. Such studies also do not allow to consider large catchments and integrated effects. Those result in relatively smaller trends and a dampen signals of climate change for large watersheds (Ficklin et al., 2018; Dai, 2016; Milly et al., 2005).

2.2 Reproduce, analyze and project streamflow

Methods only observation-based do not allow to attribute discharge changes to specific drivers. Furthermore, they only allow to study past climate and not to effectively project trends: a current trend may not continue in the future (Dai et al., 2009), future variability and trends may not have an equivalent in historical records (Hegerl & Zwiers, 2011). To better understand the drivers of changes and how discharge will evolve in the future, we need to use methods based on models. Different approaches to simulate the continental cycle exist. These range from easy-to-calibrate simple water balance models to the more complex land surface models (LSM) (Quintana-Seguí et al., 2020).

2.2.1 Calibrated parsimonious hydrological models

These models rely on usually few variables, depending on their level of complexity. They aim at reproducing well but as simply as possible the relationship between discharge and climate variables. The first step is to identify pertinent climatic and environmental drivers, then the model is adjusted over the area of study through free parameters to best fit observed data.

The simplest of all models matching that description are linear regression models. They go a step further than comparing correlation patterns to explore how much of discharge variability the selected variables can explain. Comparing model outputs to observations (Vicente-Serrano et al., 2019) and indices of goodness-of-fit such as r-squared allow to assess how much of the discharge variation is explained by the selected variables. The linear coefficient weights the relative sensitivity of discharge to each terms (Gardner, 2009). Many studies use such methods to relate changes in discharge to changes in climate variables such as precipitation P and potential evapotranspiration PET (Ficklin et al., 2018; Vicente-Serrano et al., 2019; Blöschl et al., 2019b; García-Ruiz et al., 2011), or indices of atmospheric circulation variability (Bouwer et al., 2008). Other studies includes variables for watersheds characteristics, groundwater storage and human activities (dams storage, population density, agriculture...) (FitzHugh & Vogel, 2011; Rodell et al., 2018). However if linear models can help for a first analysis, they lack physical justification to quantify the temporal evolution of hydro-climatic variables (Rottler et al., 2020). The relationships between discharge and climate drivers are not linear. Some studies use machine learning methodologies to create non linear models to better fit discharge trends. However these models are "black box" models, the relationships are not physically justified either.

More complex non linear models are built to include some physical boundaries. For instance the Budyko framework is widely used in hydrology to study long-term equilibrium in the partitioning of precipitation (P) between evaporation and runoff over catchments (Oldekop, 1911; Budyko, 1974; Andréassian et al., 2016b; Andréassian & Sari, 2019). It accounts for two physical boundaries: the atmospheric water demand ($E < PET$) and the atmospheric water supply ($E < P$) limit (Greve et al., 2015). It was first established over multiple catchments, to represent their average behavior over a long period of time (several years). It has been adjusted at the catchment level by introducing a parameter standing for catchment specific evaporation efficiency. This model can only be used for long-term studies, at least a year or more (more details in chapter 3). Another example of lumped models is

the model GR4J (Perrin et al., 2003), which was developed to reproduce streamflow at the catchment level for a daily time step. It is soil moisture accounting model, relying on four adjusted parameters. Perrin et al. (2003) find that a complexity of three to five parameters in the model is sufficient to obtain satisfactory performances at a daily time-step and settled on four. Plenty other lumped models exists, often resulting of a continuous development process, and from a compromises between flexibility and efficiency. The goal of such models is to reproduce as best as possible streamflow characteristics over the area, period and time step (daily, monthly, annual,...) of interest. The more parameters, the more flexibility but too many parameters can lead to over-fitting and to a loss in robustness and performance. The number of parameters often relates to the number mathematical functions and physical boundaries accounted for. Their is also a limitation linked to the data available for the fit. There is a limited number of pertinent parameters which can be fitted on a given time series, to represent at best its main characteristics and without over-fitting.

Their performance to reproduce historical discharge has been widely tested against observations (Jiang et al., 2015; Andréassian et al., 2016a; Perrin et al., 2003) and the scale at which they perform is depending on how they were designed. To stay with the same examples, the Budyko framework performs well to reproduce a long-term equilibrium (minimum annual time scale), for which we can neglect the variation in the water storage over catchments. GR4J was designed to take into account water store and be used at the daily time step, therefore allowing to reproduce finer intra-annual variability. More generally, mass balance models perform at different temporal and spatial scale depending on which water reservoirs are considered, on their degree of freedom and on how they were adjusted.

Lumped mass balance models are very functional for direct applications for water management due to their efficiency and relative simplicity of use, the choice of the model depending on the goal, the scale and the area of the study. They are fitted over a specific area and time period and they are used to extrapolate streamflow to unmonitored areas with similar characteristics or to predict streamflow over the same area in short/medium/long-term. The accuracy to reproduce streamflow at least in short-term makes them good candidates for studies focusing on water quality, often dependent on concentration estimations and therefore on the accuracy of modeled quantitative discharge (Perrin et al., 2003).

These models have also been used to try to separate the effects of climate from the effect of human activities on discharge through different approaches. To weight the relative impact of climate change, many studies estimate an elasticity to describe the sensitivity of streamflow to specific climate variables included in the model (Andréassian et al., 2016a). The effect of human-related drivers is most often considered to correspond to the residual variability when comparing the model outputs to observations (Vicente-Serrano et al., 2019; Fenta et al., 2017; Ficklin et al., 2018; Roderick & Farquhar, 2011; Jiang et al., 2015; Zhao et al., 2018). The remaining variability or the residuals are associated to the empirical parameters of the models. Therefore relating residuals or remaining variability to the effect of non-climatic drivers is based on the assumption that these parameters are independent from climate variability and change. However, since these empirical parameters have no clear physical meaning, they may not be independent of climate and there is no attribution possible of what is included in the residuals.

To better assess the climate part within the residuals, similarly to observation-based methods, studies compare the outputs of the model fitted over catchments with different characteristics or over different periods (Wang et al., 2020; Jiang et al., 2015; Zhang et al., 2023; Andréassian et al., 2016a; Gardner, 2009). Studies compare a model fitted over "near-natural" vs "human-modified" catchments (Ficklin et al., 2018; Wang et al., 2020; Palmer

et al., 2008). One fit represents the reference without human activities (either the "natural" catchment or the "pre-change" period) and the other the effect with ("human-modified" or "post-change"). Then for both fits, the elasticity to climate variables can be compared to study how human activities change the way streamflow reacts to climate variability (Ficklin et al., 2018). The reference identifies the residual variability due to climate and when comparing both fits, the difference in residuals and remaining variability is attributed to human-related drivers. However, these methods face the same limitations as observation-based methods, link to the definition of the reference "natural" catchments. Others identify a reference period vs a post-changes period (Ahn & Merwade, 2014; Zheng et al., 2018; Luo et al., 2020; Zhao et al., 2018). The climate sensitivity defined by the model over the reference period allows to project a reference for the post-change period without human intervention. They rely on the assumption that the adjusted parameters are independent of climate over time and that the changes in the adjustment relate all to human intervention between the pre- and post-change period. Therefore these methods attribute the difference between the fit over the post-change period and the modeled reference period to human activities. More details on the elasticity and decomposition method applied to the Budyko framework are given in chapter 4.

Using that same assumption that the adjusted parameters of the lumped models and the resulting climate elasticities remain constant under climate change, these models are run with climate projections, to project future streamflow (Teng et al., 2012; Kirby et al., 2013; Zhang et al., 2023). However, we will see that this assumption is questionable. Studies have shown that models perform best over long periods with time varying parameters (Zheng et al. (2018) for GR4J, more example for the Budyko framework in chapter 3). With the preceding assumption, we relate that result to human intervention. However even free of all human intervention, since the adjusted parameters lack a physical definition, we don't know how the adjustment reacts to a change in climate variables and to the related adaptation of the catchment. Therefore the heightened performance of time-varying parameters models could be due to an influence of climate variability and change in their calibration. Similarly, elasticities of streamflow to different climate variables (such as P and PET) does not truly independently separate its sensitivity to those variables, since their changes can be correlated (Andréassian et al., 2016a). If climate change induces an evolution of that dynamic and correlation, then the relative elasticities would change. More generally, the majority of hydrological models fit and calibration are not entirely independent of climate conditions (Nicolle et al., 2021). Calibration is therefore not transferable, with a potentially huge impact in non-stationary conditions (Coron et al., 2014). All the more in systems where stationarity-based design are contestable due to both climate change and human disturbances in river basins (Milly et al., 2008). These models are valid over specific space and time scale (over which they were developed). They do not account for potential changes to climate-runoff relationship (Teng et al., 2012). They may not be flexible enough to deal with changes not observed in historical records (Andréassian et al., 2016a) such as changes in extreme (which lie outside of the range during the period of calibration) or abrupt changes in climate regimes (Douville et al., 2021; Milly et al., 2008).

As a first step toward a better understanding of streamflow sensitivity to human-related drivers, some studies try to use direct data on human intervention. For instance, Vicente-Serrano et al. (2019) correlate trends in residuals to area covered by irrigated lands, together with areas of increased satellite vegetation activity. Rodell et al. (2018) relate fresh water availability linear trends at global scale to Landsat imagery and published reports of human activities including agriculture, mining, reservoir operations and inter-basin water transfers.

Those are correlation and no attribution. Other studies try to include some of these potential drivers in the lumped models. This allows to attribute them an elasticity and study the sensitivity of streamflow to these drivers. FitzHugh & Vogel (2011) integrate dam storage and population in their regional regression models of annual 1-day maximum flow. Han et al. (2011) add irrigation as a complementary source of water in the Budyko framework. Kirby et al. (2013) work with the Simplified Monthly Hydrology and Irrigation Water Use Model which is a calibrated mass balance model, including dams, irrigation demands and diversions, calibrated on monthly streamflow. They focus on scenarios of decreased diversions as policy makers recommended with rainfall-runoff parameters held constants to study the performance of those to adapt to climate change. Similarly, Palmer et al. (2008) use the WaterGAP integrated water-resource model, calibrated over annual runoff historical data. It has sub-models computing water withdrawal and consumption for rivers classified as dam impacted. They run scenarios based on climate projections with assumptions on population growth, and technologies development; they aim at identifying catchments likely to require action (with insufficient natural flow projected). Again these methods are limited since the assumption to held parameters constant under climate change can be contested, therefore the effect of climate change could be underestimated and the evaluation of the adaptation strategies not adequate.

Parsimonious models are simple of use and accurate to represent streamflow over the period and scale over which they were designed and adjusted. They are very useful to predict streamflow in the short/medium term, or over unmonitored areas for water management purposes. However their use for projecting long-term streamflow evolution under climate change or for attributing changes to specific drivers is questionable due to the lack of clear physical definition behind the adjusted parameters, which may not be independent from climate.

2.2.2 Physical based hydrological models and land surface schemes

Another type of models is physical-based distributed models, such as Land Surface Models (LSM). These models are structured around a set of mathematical functions representing different physical, chemical and biological processes connected with different cycles (surface energy, water, carbon) (Zhao & Li, 2015). They are more demanding in development and computational power (Stephens et al., 2023). There are different generations of such models, following the improvements in mechanisms structures and numerical techniques. In their review, Zhao & Li (2015) distinguish three generations of models. The first consists of the "bucket models" (Zhao & Li, 2015; Stephens et al., 2023; Nazemi & Wheater, 2015a) with spatially constant soil properties and based on a simple energy balance to predict evaporation and runoff when soil moisture exceeds saturation (Fig. 2.6). The second generation models start to include physical-based land surface processes, to illustrate interactions in the soil-vegetation-atmosphere system. The first two generations focus on the energy and water exchanges between the land surface and the lower atmosphere. The third generation includes the carbon cycle and the biological controls to evapotranspiration such as photosynthesis and stomatal resistance (Fig. 2.6).

In the end, LSMs were initially designed to provide realistic land boundary conditions to global climate models (GCMs) (Tafasca et al., 2020) and explicitly simulate water and energy exchanges at the interface of the soil with vegetation and the interaction with the atmosphere (Quintana-Seguí et al., 2020). They are run either coupled with an atmospheric model (coupled mode) or with an atmospheric forcing dataset (off-line mode) to represent the latter.

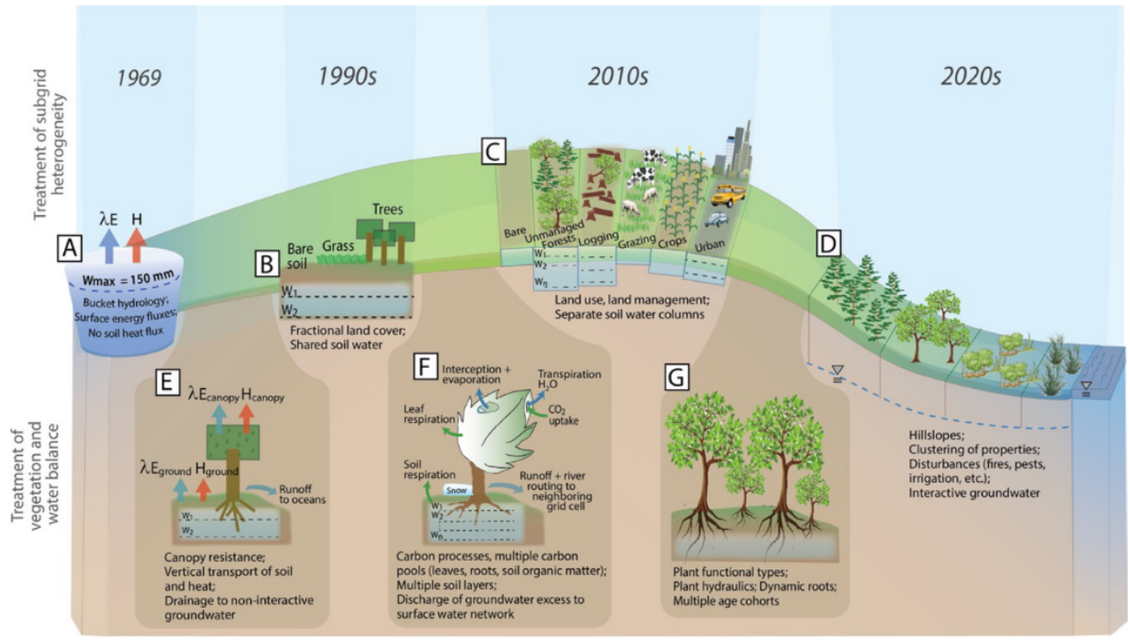


Figure 2.6 – The evolution of land model formulations, beginning with the Manabe bucket model in 1969 (label A), gradually improving the treatment of water, heat, and vegetation, while also including increasingly complex and heterogeneous representations of vegetation and soil processes both above and below the land surface. Dates are approximate. Blue arrows:λE = evaporative flux (where λ = latent heat of vaporization of water, E = evaporation rate). Red arrows: H = sensible heat flux. Green arrows: carbon fluxes. From Stephens et al. (2023).

They reproduce land surface dynamic information from sub-hourly to yearly scales, from the diurnal cycle of the biogeochemical fluxes, pools, vegetation/soil temperature, water content, to yearly variation of vegetation dynamics, and land use/cover change (Zhao & Li, 2015). Other physical-based models are large-scale hydrological models, socalled global hydrologic models (GHMs), which focus more on representing larger scale hydrological responses and the water cycle but have simpler structure to represent the other connected cycles (energy, carbon) (Nazemi & Wheater, 2015b).

The parameters of the physical-based models include information about land surface characteristics necessary for the model initialization and run (Zhao & Li, 2015) (Fig. 2.7). Indeed these models rely also on a parametrization but not as empirical as for the calibrated models: the parameters are associated to known physical and biological processes. Often these processes and the associated parameters are tested in controlled environments which may not match processes in real conditions. Therefore an uncertainty remains around the values associated to these parameters in real conditions. Further improvements in data acquisition and real condition measurements should improve the way these parameters are estimated (Stephens et al., 2023).

Due to their processes-based structure, these models allow for long-term simulations and can be used for two major purposes. First LSMs were designed to be run in coupled-mode with atmospheric models, to represent the interaction and feedbacks between the atmosphere and the land surface, under different climate scenarios. Coupled with an atmospheric model to run different greenhouse gas concentration scenarios, these physical-based models allow to project future land surface reactions to climatic changes. Secondly, physical-based models allow to isolate the effect of a specific process by fixing alternatively one factor or several factors (Zanardo et al., 2012; Alkama et al., 2010; Do et al., 2020). For instance, they are used to attribute the effects of "anthropogenic climate change" over the past century in fingerprint methodologies (Douville et al., 2021; Christidis & Stott, 2022; Knutson & Zeng,

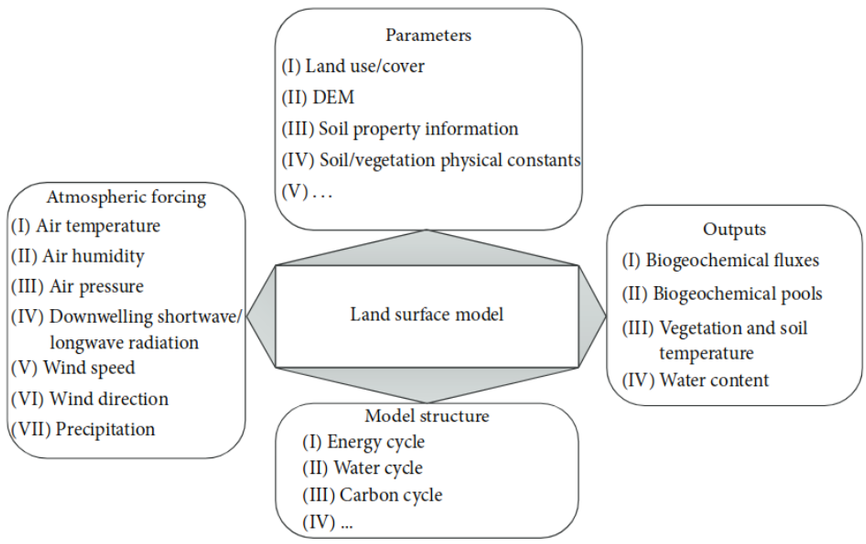


Figure 2.7 – General component of a typical land surface model. From Zhao & Li (2015).

2018; Hegerl & Zwiers, 2011). These methods compare the outputs of the LSM, coupled with atmospheric models run without and with human induced CO2 emissions. The first run gives a "baseline" and allows to detect trends in the second run standing out from the background climate noise (Douville et al., 2021). When comparing the different runs with observations, it allows to look for spatial and temporal patterns matching (or not) the observations and therefore attributing the source of change (Hegerl & Zwiers, 2011). These are the only methods which can effectively attribute a source of change to a specific process (Douville et al., 2021). Run with a forcing covering past climate, they also allow to recreate a coherent image of the land surface system in the past, including variables that are difficult or impossible to observe (Quintana-Seguí et al., 2020; Simons et al., 2020), or covering areas with no or scarce observations (Do et al., 2020; Wang et al., 2018).

To validate these models, they are run in off-line mode with forcings datasets (reanalysis or bias-corrected outputs of GCMs), to assess their realism, their sensitivity to climate variables, and to remove biases and improve simulations (Zhao & Li, 2015; Decharme et al., 2019). They are evaluated against satellite estimates (which also rely on modeling to convert the signal into the desired variable) and in-situ observations for the variables which can be measured (Decharme et al., 2019). High uncertainties remain for these models to reproduce a realistic land surface system. Their strengths relate to their limits. There are uncertainties behind the parameters in the parameterized processes (photosynthesis, soil diffusivity...), the forcing (Hegerl & Zwiers, 2011). The processes included in each models are not exhaustive and the structural uncertainty of model design, due to implementation and formulation differences, leads to higher uncertainties than uncertainties due to the forcing used (Quintana-Seguí et al., 2020; Sordo-Ward et al., 2019). The different models often better match in patterns than in amplitudes of the detected trends (Hegerl & Zwiers, 2011; Do et al., 2020). Several studies analyze the average result of a model ensemble to predict future changes (Milly et al., 2005; Nohara et al., 2006; Gudmundsson et al., 2017b; Alkama et al., 2013; Do et al., 2020), as it was shown that such average is better at reproducing historical systems (Nohara et al., 2006; Alkama et al., 2011). Several projects aim at running different LSMs under similar protocols to compare their outputs. It can be off-line runs forced with reanalysis datasets covering historical periods as for the Global Soil Wetness Project (GSWP) (Dirmeyer et al., 1999, 2003; Oki et al., 2013). Other projects allow to test both historical periods and projections, running LSMs in coupled-mode as in the previously introduced CMIP projects (chapter 1) (Eyring et al., 2016), or off-line but forced with bias-

corrected outputs from coupled runs as in the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) (ISIMIP, 2012; Frieler et al., 2017), covering both historical period and different projection scenarios. In the case of coupled runs, there is an added uncertainty due to the atmospheric modeling, while in off-line mode, that uncertainty is reduced but the feedbacks between the land surface and the atmosphere are not represented. In all cases, working from model ensemble averages can also lead to a bias and to an attenuation of the variability and of the extremes (Do et al., 2020). More generally, these models aim first at reproducing at best the processes they include and not at being as close to the real system as possible. Therefore an inadequate projection should not be interpreted as a failure but rather as diagnostics revealing the strength of the controls that may have been left out (Zanardo et al., 2012).

In the case of streamflow, studies working with model ensemble under climate change scenarios agree on the direction of streamflow changes at the global scale, with the amplitude of change dependent on the model ensemble, the scenario, the projection period. For coupled land–atmospheric simulations, the majority of LSMs in CMIP5, under high carbon emission scenario RCP 8.5, simulate an increase in runoff over South Asia, northern Europe, northern Asia and North America, and a decrease over southern Europe by 2100 (Alkama et al., 2013) in agreement with the results of Milly et al. (2005) (12 model ensemble, projection for the period 2041–60, based on the Special Report on Emissions Scenarios A1B (SRESA1B) scenario scenario), Nohara et al. (2006) (19 coupled atmosphere–ocean general circulation models based on the same SRESA1B scenario, which projects a CO2 concentration of 720 ppmv by the year 2100). Other studies work from LSM or hydrological models run in off-line mode, forced with projected, bias-corrected climate scenarios (outputs of the coupled runs) (Schneider et al., 2013; Do et al., 2020; Arnell, 1999; Sordo-Ward et al., 2019; ISIMIP, 2012). Such studies can include any GHM or other hydrological model, since there is no need to compute the feedback interactions with the atmosphere (Nazemi & Wheater, 2015a). Do et al. (2020) look at changes in streamflow extrema consistent across a six-model ensemble for the scenarios RCP2.6 and 6.0 and find that significant increasing trends are projected consistently over southern and south-eastern Asia, eastern Africa, and Siberia, while high agreement of decreasing trends is found over southern Australia, north-eastern Europe, the Mediterranean and north-western North America. With a similar methodology, with the physical-based hydrological model PCRGLOBWB run with the output of five climate models under different emissions scenarios following ISIMIP protocol, Sordo-Ward et al. (2019) project severe negative changes in the Iberian Peninsula, from the Black Sea in the South almost to the Baltic Sea in the North, and predominantly positive changes in western to central Europe and in northern Europe. Fig. 2.8 shows their results for the emission scenario RCP 4.5.

At large scale (global scale), when run with a forcing over the end of the past century, models which do not include any processes related to water management (irrigation, dams regulation, water withdrawals) are generally good at reproducing the spatial patterns of observed trends in streamflow annual averages (Do et al., 2020; Milly et al., 2005; Nohara et al., 2006; Dai et al., 2009; Arnell, 1999). It shows that the external forcing has a significant effect on discharge annual mean trends, with a pattern of change qualitatively similar to the one of precipitation changes (Milly et al., 2005; Arnell, 1999), with a non negligible influence of changes in evapotranspiration in projections for some regions (Arnell, 1999).

However streamflow is an integrated variable and its variability is subjected to long-term changes in atmospheric forcings, climate variability (at different time scales, intra-

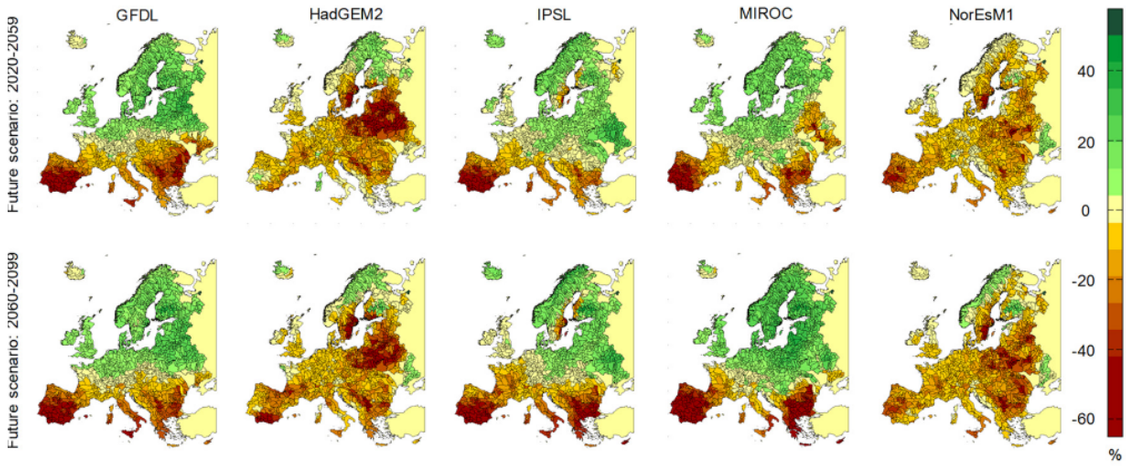


Figure 2.8 – Changes (percentage) of mean annual runoff in future scenarios (2020–2059 and 2060–2099) compared with the reference scenario (1960–1999), according to different climate models and for the emissions scenario RCP4.5. Red shading represents a decrease of the mean annual runoff and green shading an increase. Model PCRGLOBWB, ISIMIP. From Sordo-Ward et al. (2019).

annual, inter-annual, inter-decadal), as well as human activities across the drainage basin and other changes in basins' behavior. Therefore the capacity of such models to represent physical features of hydrological regime may not be sufficient to perform in simulating the characteristics of trends (Do et al., 2020). Nohara et al. (2006) find that their river model performs well except in area more sensitive to human water management and irrigation, such as arid areas. Models underestimate the amplitude of observed changes in streamflow (Gudmundsson et al., 2017b), with observed trends less spatially coherent than for modeled trends (Milly et al., 2005). The major differences are mainly at local scale, with still a strong global scale relation (Milly et al., 2005; Do et al., 2020). It could be due to the impact of human activities at catchment level which could be significant (Gudmundsson et al., 2017b) for given catchment and over given periods. For instant over snow-fed rivers at a continental time scale, the effect of direct management of river flow is estimated to be comparable in magnitude to climate change effect (Douville et al., 2021). More generally, models are designed and have a proven ability to capture the natural water cycle but more progress still needs to be made to take into account human processes (Wang et al., 2018). Other processes are also still missing or misrepresented in models, such as groundwater processes (Douville et al., 2021), often absent or crude, disconnected from other processes in most LSMs and GHMs (Nazemi & Wheater, 2015a), or glaciers and permafrost melting (Alkama et al., 2013). This explains why parsimonious hydrological models offer better results in streamflow simulation and are so far best suited to detect actual changes in basin behavior as opposed to physically-based models, which are limited by the processes they include (Perrin et al., 2003; Quintana-Seguí et al., 2020). Zhang et al. (2023) find different elasticities for streamflow when based on observations or when based on the output of CMIP6 models ensemble. It shows that these models fail to reproduce all current influences on streamflow.

To palliate to the effect of these missing processes, many studies combine the physical based models to empirical models or add calibration/bias correction steps. For instance, Kour et al. (2016) present a study where streamflow outputs of the model ensemble are bias corrected to fit the observation in the reference period and then the bias correction (considering the same algorithm and parametrization) is used to bias correct streamflow projections. Similarly, Schneider et al. (2013) use bias-corrected climate data from different General Circulation Models (GCMs) to force an hydrological model for which the river

discharge was calibrated over the reference period (1971-2000). That calibrated model is run with climate projections of these same GCMs for the 2050s to get streamflow projections. In their study, Sordo-Ward et al. (2019) use a similar bias correction for the average runoff of their hydrological model in the control and the projected series. Zhang et al. (2023) calculate an elasticity for streamflow sensitivity to climate drivers over the modeled reference historical period and use that elasticity to project future streamflow changes under climate change scenarios. However these methods takes us back to the limits associated to calibrated models for projection, with the issue that the bias correction or the fitted parameters and associated elasticities may not be adapted to future situations due to their lack of a physical meaning. These methods can be used without such limitation when the goal of the study is not to project but for instance to estimate streamflow over un-gauged rivers. Wang et al. (2018) incorporate a river discharge assimilation process in the LSM ORCHIDEE which allows to compensate for model systematic errors or missing processes similarly to a bias-correction process and therefore to better extrapolate streamflow over ungauged rivers to estimate total current freshwater flows.

Physical-based models are continuously improving, with on-going studies aiming at better representing the currently missing processes. Continuous upgrades are made to improve soil schemes, snowpacks, river routing systems and flood schemes such as in the Land Surface Models ISBA (Decharme et al., 2019) and ORCHIDEE (Schrapffer et al., 2020). The goal for the hydrological cycle and the simulation of the closed global water budget is to get runoff and streamflow correct for the right reasons (Fan et al., 2019) and therefore to get the full coupled human-natural system for conservation purposes (Nazemi & Wheater, 2015a). New anthropogenic processes are included to study how man-made irrigation and/or dams alter the river flow and continental evapotranspiration. Including human-related processes allows to compare "naturalised runs" (where human water management is not taken into account) and "human impact runs" (which includes water management inputs) (Do et al., 2020). In online mode, it also allows to better understand the feedbacks with atmospheric variables (Nazemi & Wheater, 2015b). The challenges are both in modeling the water demand and the corresponding water supply and allocation (Nazemi & Wheater, 2015a).

Some models are already developed to include water demand and link it to socio-economic drivers to built projections scenarios. To give a few examples, Wada et al. (2012) couple the process-based hydrological model PCR-GLOBWB to a model estimating the water demand, simulating irrigation needs based on crops and country-specific efficiencies and other water needs estimated from population and socio-economic drivers. These studies aim at looking at catchment over which water management will be a challenge in the future. Similarly, Sordo-Ward et al. (2019) combine the same process-based hydrological model PCR-GLOBWB to the model "Water Availability and Adaptation Policy Assessment (WAAPA)" considering reservoirs and demand for water withdrawals. The model PCR-GLOBWB is used to simulate naturalized streamflow while WAAPA components are reservoirs, inflows and demands which are linked to nodes of the river network. Withdrawals are spatially distributed using the map of areas equipped for irrigation (Siebert et al., 2015) and population density. Future withdrawals are estimated using projections of population and gross domestic products. In the Model S-HYPE over Sweden, Arheimer et al. (2017) includes the modeling of human regulation (artificial lakes, dams) for hydro-power in a process-oriented integrated catchment model, for snow-dominated river systems. They compare the output to the natural flow modeled when removing all regulations and extrapolate to other snow-dominated areas of the world, assuming similar methods of water management. They find that in snow-fed rivers, hydropower regulation impacts flow regime by dampening high flows and redistributing it

to other times of the year. Downstream of large reservoirs that regulation can affect flow regimes much more than climate change. Other studies include dams (Hanasaki et al., 2006) and water demand schemes and allocation paths in surface river routing systems such as in ORCHIDEE LSM (Zhou et al., 2021; Baratgin et al., 2023). Voisin et al. (2017) use a regional integrated assessment (IA) model to spatially distribute allocation of sectoral water demands to surface and groundwater systems in a regional earth system model.

As irrigation is the main consumptive use of freshwater, it has been the first human activity represented in LSMs. It is also a good example to illustrate the current progress and difficulties faced by modelers to include human-related processes, with a goal to respect water conservation. More and more process-based models integrate irrigation schemes to get a temporal and spatial information and understand the interaction and feedbacks between irrigation and the Earth system. Irrigation remains underrepresented or simplistic. The first step is to estimate irrigation demand (Nazemi & Wheater, 2015a). Most models estimate irrigation water requirements using either a root-zone soil moisture deficit approach or a crop-specific potential evapotranspiration approach (McDermid et al., 2023). Both approaches estimate irrigation requirements as the additional water needed to fill the deficit of water, deficit of moisture or deficit in crops needs. These methods estimate an optimized irrigation and do not account for the actual amount of water applied to the field with possible losses in runoff and irrigation system. Other models work without an external irrigation prescription. The second step is to allocate water requirements to water supply (Nazemi & Wheater, 2015b). This requires an implementation of river routing systems with pertinent reservoirs and more performing groundwater systems (Nazemi & Wheater, 2015b). As a first "naive" approach, models can fulfill irrigation requirements with available surface water or implicit groundwater sources (McDermid et al., 2023). With such methods however, the water budget may not be closed over the system. Other models are improving their land surface and river schemes to include irrigation withdrawals associated to specific reservoirs. This has currently mostly been developed for GHMs, which focus on the water cycle and lack processes related to the energy and carbon cycle. These models can also only run in off-line mode and do not allow to analyze the feedbacks with the atmosphere (Nazemi & Wheater, 2015b). In LSMs, it is slowly introduced, mostly at large scale, allowing for a first overview of the feedback effects of irrigation, when water scarcity is less localized and allocations of withdrawals to specific water reservoirs are less sensitive (Guimberteau et al., 2012). To improve, models need to represent spatio-temporally detailed water infrastructure and management, including reservoir operation, inter-basin transfer and managed aquifer recharge, along with the consideration of different irrigation techniques (McDermid et al., 2023). Another challenge would be to implement an explicit representation of human decision-making and to create associated "policy" scenarios for future projections (Nazemi & Wheater, 2015a).

Different barriers still limit such improvements. First there is a barrier on data availability surrounding these issues (Do et al., 2020), forcing to find a surrogate for actual use of water and introducing high uncertainties (Nazemi & Wheater, 2015a). For irrigation for instance, most models currently work with, as input data, irrigated area in best cases or area equipped for irrigation in most. These are gathered from institutional database from agricultural census which may not always be reliable and induce a large source of uncertainty. Furthermore the irrigated areas do not account for water withdrawals which also depend on the irrigation techniques, the timing of irrigation, the volume of water used. An area equipped for irrigation may not be irrigated or depending on the crop or the season. Remote sensing techniques can help improve irrigation maps (Brocca et al., 2018) but many

challenges and uncertainty remains surrounding these data. Data on use and operation of water resources are also difficult to gather at regional and global scale (Nazemi & Wheater, 2015a). Furthermore, most of LSM have been conceived at global or large scale, and these new anthropogenic processes need to be implemented at relevant regional and local scale (Voisin et al., 2017; Nazemi & Wheater, 2015a). The models therefore need to be tested and adapted at such scales (Do et al., 2020; McDermid et al., 2023; Guimberteau et al., 2012; Douville et al., 2021; Simons et al., 2020). The main limits to that latter issue is currently the lack of high resolution climate forcing datasets and computational power (Nazemi & Wheater, 2015b; Zhou et al., 2021).

In the end, process-based models are a useful tool to understand processes, to attribute specific effects to specific drivers and to study links and feedbacks between different processes. They are improving and growing more and more complex. However there is also the need to find a balance between the search for accuracy and realism in the included processes and the model complexity, to keep the model useful and understandable (McDermid et al., 2023).

KEY POINTS TO REMEMBER

- Observations shows significant trends in streamflow over Europe, along with a growing variability and more extremes high and low flows events.
- However, the complexity of intricated climate and land surface drivers make these trends more variable and less coherent in time and space than trends in climate variables such as P. Changes in land surface characteristics and human activities need to be taken into account.
- There is a need to use models to understand the main processes involved in streamflow generation and identify the main drivers of discharge changes.
- Parsimonious calibrated models allow to accurately reproduce streamflow and trends but are of limited use to understand the underlying processes and their evolution and for long-term projections. They only allow for an exploration of correlation between streamflow changes and potential drivers.
- Physical-based models are the only way to effectively attribute streamflow trends to specific physical and biological processes. However, they are limited by the processes they include and their increasing complexity. Mainly, the effects of human activities lack in most models or are still coarsely designed. The current major limitations for improvement are the lack of data on human-related processes and simulation resolution.

By design, all models have strengths and limitations, which need to be known and understood to improve them and allow to choose the best model according to the goal of the study. In the case of studying streamflow, it is important to keep in mind that it is a flux dependent on a variety of complex drivers, at the intersection of several scientific fields including hydrology, climatology, biology and sociology when we start including human decision making.

CHAPTER 3

Isolate and decompose the effect of climate change on discharge records

This chapter was the subject of an article accepted for publication in the journal *Water Resources Research* (Collignan et al., 2023a). It is mainly composed of part of the article, with some added details in the method part and in the data analysis. The full article is in Appendix.

Contents

3.1 Introduction

As discussed in chapter 1, water is a key resource for the whole of society and understanding the hydrological cycle and how it evolves due to a changing climate has been and will continue to be a significant challenge.

Over the past century, several studies have shown the impact of climate change on climate variables in Europe (see chapter 1). To recap, over the past century, average annual precipitation increased over most of Europe except the Mediterranean area where they tend to decrease (Douville et al., 2021; Knutson & Zeng, 2018; Christidis & Stott, 2022). Trends per decade are less significant due to the high inter-annual variability of precipitation P (Douville et al., 2021). Trends in potential evapotranspiration PET are linked to an increase in the energy available at the surface, which is highly correlated to rising temperatures (Douville et al., 2021; Vicente-Serrano et al., 2014), and have shown a significant increase in PET over the end of the century, especially over the Mediterranean area (Vicente-Serrano et al., 2019; Vicente-Serrano et al., 2014; Kitsara et al., 2013).

For precipitation, studies have shown that not only is the annual average of P changing, but there is an increasing seasonality with contrasted trends between summer and winter, depending on the area (Zveryaev, 2004; Christidis & Stott, 2022). Moreover, over the past few decades, extreme precipitation events have significantly intensified (Ribes et al., 2019). The intra-annual variations of climatic variables are more difficult to assess, and only a few indices exist to measure the inter-annual changes in the distribution of climate variables. For example, García-Barrón et al. (2018) defined indices to assess the evolution of the intra-annual cycle of P over time throughout the Iberian peninsula. At the end of the century, they identify a shift of the main rainfall periods towards autumn, especially over the Atlantic basins, and an increase in the inter-annual variability of the intra-annual cycle, especially over the Mediterranean basins. Therefore, it is important to investigate the effects of changes in the annual averages of climate variables along with the effect of changes in seasonality and intra-annual distribution of these variables. The distribution of P within the year and its coupling or decoupling from the atmospheric demand PET will influence water partitioning between evapotranspiration and discharge on the annual scale.

Transformations in different climate variables governing the water cycle alter the equilibrium in the water balance over the different watersheds, thus impacting the discharge of rivers (see chapter 1 and 2). As detailed in chapter 2, worldwide discharge trends are and will continue to be significantly impacted by changes in climatic factors. Over Europe, statistically significant trends in discharge are observed in historical records (positive in the northern region and negative in the south and east). These trends are spatially coherent with precipitation changes (Stahl et al., 2010; Vicente-Serrano et al., 2019) and less coherent with PET changes patterns (Yang et al., 2018).

We have seen that P intra-annual distribution is also likely to change and impact the partitioning of water between evapotranspiration and discharge. This effect is primarily considered in the literature through the study of seasonality and annual extremes of P and PET in order to examine their impact on floods (Douville et al., 2021; Rottler et al., 2020; Milly et al., 2002), drought events, (Douville et al., 2021; Vicente-Serrano et al., 2014) and, more generally, on discharge peaks (Tuel et al., 2022; Bouwer et al., 2008). Seasonal flows are separately analyzed in several studies, showing that trends in summer and winter flows are disconnected for an ensemble of small near-natural catchments in Europe (Stahl et al., 2010) and that increasing autumn and winter rainfall led to increased floods in northwestern

Europe (Blöschl et al., 2019b).

However, rivers are also highly managed, and human activities are an important driver of change in how watersheds function. As previously explained, a significant difficulty in analyzing the effect of climate on historical discharge changes is decomposing the effects of the different drivers of change and isolating them from each other to better understand their relative importance (chapter 2).

To focus on the effect of climate change, several studies have concentrated on catchments that are considered as near-natural or unimpaired in order to investigate the effects of climatic changes on discharge (Stahl et al., 2010; Yang et al., 2018). However, this highly limits the areas studied (Vicente-Serrano et al., 2019), especially in Europe, where the high population density and long history of water management limit these studies to small catchments (Stahl et al., 2010).

As described in chapter 2, another approach is to use models to separate the factors involved in discharge changes. Parsimonious models relying on a few variables and adjusted parameters are favored for their simplicity of use and interpretation. However, they are empirical: they rely on adjusted parameters over the area and the time period studied and lack a clear physical meaning.

Other methods are physical-based hydrological and Land Surface Models (LSM). They have grown increasingly complex and are able to reproduce the behavior of watersheds and to model "natural flow" regimes Decharme et al. (2019); Wang et al. (2018); Gudmundsson et al. (2017b); Schneider et al. (2013). They do not always accurately represent a whole real hydrological system depending on which processes are included in them but allow a meaningful assessment of hydrologic aridity Yang et al. (2018). However, due to their complexity, it is more difficult to decompose the effects of individual climate factors and to interpret their outputs than with other simpler models.

In light of this, we propose a tool that combines the simplicity of the more empirical model with the heightened performance and complexity of the physical-based model to better understand the phenomena encapsulated behind the adjusted parameters.

We use here the well-known and widely used empirical Budyko framework (Mianabadi et al., 2020). It is predicated upon utilizing the annual mean of water and energy balances at the watershed scale (Tian et al., 2018), taking into account the water and energy limitations of the physical system. It was initially conceived over multiple catchments. Parametric equations were developed to introduce an empirical parameter adjusting the framework to the specific evaporation efficiency of each catchment over an equilibrium period (Zhang et al., 2004; Yang et al., 2008). However, equilibrium disruptions, due to climate change or any other direct human activities and vegetation change, highlight limitations to the model. Moreover, as detailed in chapter 2, most disruptive features are concurrent and interrelated with climate change. The parameter introduced has no evident physical meaning and is just a well-adjusted proxy to E/P over a specific catchment and period. There is no straightforward method to attribute changes in the adjusted parameters to specific climatic or non-climatic features (Berghuijs et al., 2020; Reaver et al., 2022), as for any parsimonious model with calibrated parameters.

To focus on the effects of climate change, the present study applies the Budyko framework to the outputs of a state-of-the-art LSM. The latter represents the constant physical behavior

of watersheds. The only source of change in the dynamics of the modeled watersheds is the evolving climate variables introduced. Using LSM outputs also allows for adjusting the near-surface atmospheric variables to more adequately decompose the effects of the different elements of climate variability and change.

This chapter is organized as follows: the first section 3.2 describes the Budyko framework, with its underlying hypotheses and limitations, the chosen LSM, the reference forcing describing the climate over the past century and our methodology to combine the parsimonious model and the physical-based model. Here, we create synthetic forcings to test if our methodology yields an optimal analysis of the effects of different aspects of climate change. We also explain the use of the time-moving window to examine temporal trends in the different climatic effects. In the next section 3.3, we present the results of the effect of different elements of climate change across Europe (changes in annual averages against changes in the intra-annual distribution of climate variables) on discharge trends over the past century. Section 3.4 provides a comprehensive analysis of the advantages of our findings, while also highlighting the present constraints and areas for further investigation. Finally, in section 3.5, we summarize our conclusions.

3.2 Our "climatic" reference method

3.2.1 The Budyko framework

3.2.1.1 General presentation

Over watersheds considered as closed systems, the water balance equation (3.1) applies when explaining the equilibrium between the variables of the hydrological cycle: the river discharge (Q), the evapotranspiration (E), the precipitation (P) and the change in the water storage over the watershed between two-time steps (ΔS).

$$P - \Delta S = Q + E \tag{3.1}$$

Long-term, ΔS can be negligible. Ideally, this hypothesis should be applied over a long enough period that the system's equilibrium is reached (Zhang et al., 2008). It also supposes no external disturbances impact the water budget, such as groundwater mining or water transfers to or from other basins.

The Budyko framework, which is frequently used in hydrological research to study the partitioning of P into E and Q, draws from this long-term equilibrium of water balance over a catchment coupled with the energy balance. It postulates that the partition of the annual water budget between runoff and evapotranspiration over catchments, represented by the evapotranspiration E, is a function of the relative water supply (rainfall P) and the atmospheric water demand (potential evapotranspiration PET) (Tian et al., 2018; Xing et al., 2018; Yang et al., 2007). The latter depends on both available energy and aerodynamic resistance (Barella-Ortiz et al., 2013). Therefore, this framework considers the system's water and energy limitations, which cannot evaporate more than the atmospheric demand allows and more water than the catchment receives from the water source (P). In short, it defines the "Budyko space" (Berghuijs et al., 2020; Reaver et al., 2022).

This framework relies on a closed water budget in time and space, neglecting ΔS. Therefore, it must be applied over a closed watershed and fitted on a long-term equilibrium. To be freed from seasonal water storage variations, we use a time series of a yearly resolution (hydrological year) in this study. For the region considered, the hydrological year starts in

September, at the end of the dry season, when the reservoirs are supposedly at their lowest. It minimizes the differences in ΔS from year to year. Later on, unless specified otherwise, the variables P, E, and Q represent the annual averages over the hydrological year. We then apply the framework over minimum periods of 11 years, considered a long enough period for ΔS to be negligible over most catchments, dependent on the area Han et al. (2020). We tested this hypothesis with the outputs of the LSM, and we found that ΔS is about a hundred times smaller than Q when 11-year sub-periods are considered (infobox 1).

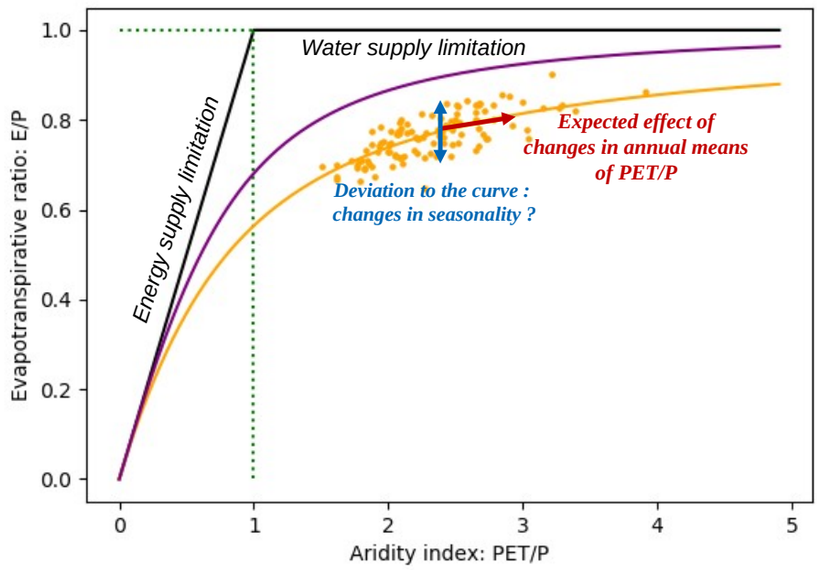


Figure 3.1 – Budyko framework: relationship between evapotranspirative ratio (E/P) and aridity index (PET/P) (Fu's equation). E, PET, P are annual averages. ω associated with the purple curve is larger than ω associated with the orange curve and translates into a higher evaporation efficiency above the watershed. For a given watershed with constant characteristics, there is still a dispersion around the curve of the dots for a given year due to intra-annual variations of the climate cycle (orange dots). The curve and its associated ω represent the average behavior of the watershed. The framework includes trends in annual climate variables by a displacement along the curve (red arrow). However, it doesn't include trends that could impact the way water is partitioned over the catchment such as long-lasting trends in the intra-annual distribution of P and PET (blue arrows).

INFO BOX 1

About the hypothesis $\Delta S \approx 0$

If ΔS is not negligible, it induces an artificial change in the evaporation efficiency when we fit the parameter with the wrong hypothesis. Because we are using in this chapter the output of a land surface model (LSM), we can test that hypothesis.

We define $\Delta S = P - E - Q$ with the LSM outputs and compare the magnitude and variance of ΔS to Q over the period. First, we can assess that there is no significant trend in ΔS over the period for any of the basins modeled in our analysis. We then considered the annual ratio $\Delta S/Q$ to look at the relative importance of that term for all our basins and study the magnitude of the effect of storage change on Q. For annual values, that term average is between -0,35 and 0,097 for all basins considered but with 91% of the stations with an average $\Delta S/Q$ between -0.05 and 0.05 and a standard deviation between 0.015 and 0.83. For the majority of basins, the annual ΔS represents less than 5% of Q (Fig. 3.2, boxplot 1). This shows that even at the annual

time scale, the effect of ΔS is rather small for most basins. It should indeed better be accounted for over some basins where ΔS can have the same order of magnitude as Q over given years but these basins are all very small basins and don't weigh a lot in our final results. Furthermore, when we consider the same ratio over 11 years with the moving time window we use in our study, the standard deviation of that same ratio drops to between 0,0015 and 0,12 for all basins and the average is divided by 10, between -0,011 and 0,017. ΔS drops to less than 1% of Q on average for all basins (Fig. 3.2, boxplot 2).
This validates our hypothesis that over a decade we can consider ΔS to be negligible since its a hundred times smaller than Q.

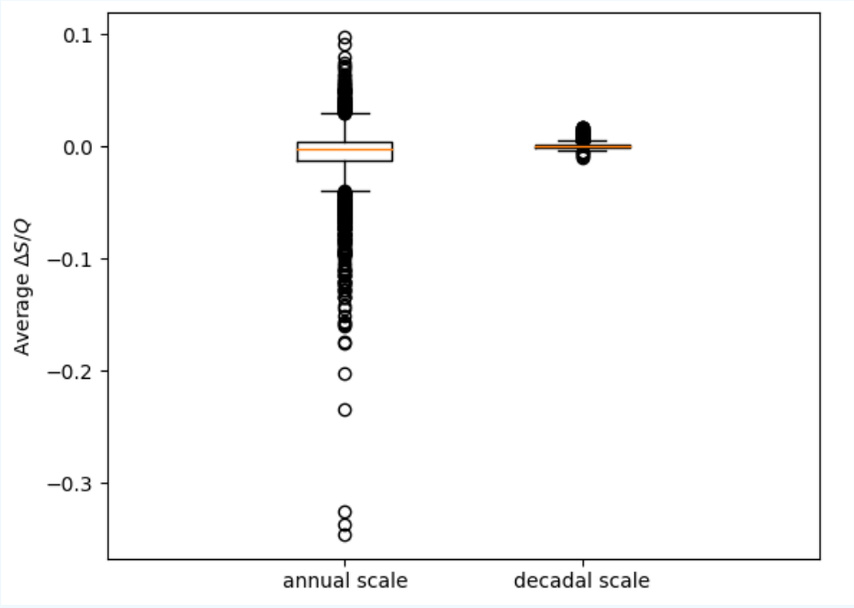


Figure 3.2 – Distribution of the relative magnitude of modeled $\Delta S/Q$ over the catchments studied, on average over the full period. The first box represent the distribution of the average of the annual ratio of $\Delta S/Q$ while the second box shows the distribution of the average of the ratio $\Delta S/Q$ over a decade.

3.2.1.2 One parameter equation

The original Budyko framework was empirically constructed over a set of catchments to define a curve followed, on average, by catchments in the Budyko space. Different analytical approximations to this hypothesis (Budyko curves) have been developed, expressing the evapotranspiration rate (E/P) as a function of the aridity index (PET/P) over a catchment (Fig. 3.1).

More specifically, the framework was extended to analyze individual catchments over a stable period. Parametric equations were developed which introduced an empirical parameter representing the specific position of the catchment within the Budyko space (Yang et al., 2008).

Two of the most widely used are the Fu equation (3.2) (Zhang et al., 2004; Ning et al., 2019; Simons et al., 2020; Zhang et al., 2008; Zheng et al., 2018) and the Mezentsev–Choudhury–Yang equation (3.3) (Yang et al., 2008; Luo et al., 2020; Roderick & Farquhar, 2011; Wang et al., 2020; Xing et al., 2018; Xiong et al., 2020). These can be found under different names in the literature such as the Tixeront-Fu equation for (3.2) or Turc-Mezentsev for (3.3) (Andréassian & Sari, 2019).

$$\frac{E}{P} = 1 + \frac{PET}{P} - \left(1 + \left(\frac{PET}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} \tag{3.2}$$

$$E = \frac{P * PET}{(P^n + PET^n)^{\frac{1}{n}}} \tag{3.3}$$

The two parameters derived from equation (3.2) and (3.3) are linearly correlated, implying that both equations are almost equivalent (Yang et al., 2008; Andréassian & Sari, 2019; Du et al., 2016; Roderick & Farquhar, 2011). We examine the sensitivity of the results to the parametric equation used. A limitation exists when fitting Fu's equation for watersheds with a particularly high dryness index, such as in arid climates. In these areas, the estimated ω uncertainty will increase as the values used to fit the curve are all too close to the plateau (Fig. 3.1) and not scattered enough to fit the curve correctly. We fit the parameter n with Choudhury's equation (3.3) and a set of E/PET and P/PET. This method uses the ratio of P to PET and gives us a plateau in humid areas as opposed to the previous fit of PET/P. We obtain very similar results for the methodology with either equation used (not shown). We conclude that this issue does not strongly impact our study area and that we could use either equation. For the rest of the study, we use results obtained with Fu's equation (3.2).

E measurements are not available over large spatial and temporal scales. Therefore, most studies work from the analysis of Q, which can be calculated from the water balance equation (3.1) where ΔS has been neglected. With Fu's equation (3.2) used to express E in (3.1), it yields (3.4):

$$Q = P * \left(1 + \left(\frac{PET}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} - PET = f(P, PET, \omega) \tag{3.4}$$

3.2.1.3 Discussion of the evaporation efficiency parameter

The evaporation efficiency parameter is empirical; it is obtained by fitting data from a specific catchment during a period of assumed equilibrium state. It determines the position of the catchment in the Budyko space.

The specificity of the parameter relates to all factors impacting the evaporation efficiency of the watershed other than changes in the average aridity index (Zhang et al., 2004; Padrón et al., 2017; Donohue et al., 2012). The most common hypothesis is that it reflects the various hydrological characteristics of the watershed, such as topography, vegetation coverage, and soil properties, which play a part in the annual partitioning of P into E and Q over the catchment (Gudmundsson et al., 2017a; Reaver et al., 2022). Some are considered time-invariant (soil type, topography, etc.), while others are possibly affected by long-lasting changes. These can occur in the hydrological properties of the surface water system, most likely due to direct anthropogenic activities such as river management, irrigation, and land cover changes. It leads to the "catchment trajectory conjecture" (Reaver et al., 2022), which suggests that the watersheds would follow an average Budyko-curve (Fig. 3.1, red arrow) if it were not for changes in hydrological properties independent of changes in the average aridity index.

Several studies attempted to analyze the evolution of watershed behavior between two equilibrium states, a period of reference and a period of post-changes (Jiang et al., 2015; Luo et al., 2020; Wang et al., 2020; Zhao et al., 2018; Zheng et al., 2018) and then fit the parameter independently over each period. Two distinct curves (Fig. 3.1) were acquired using distinct evaporation efficiency parameters to characterize the pre- and post-change equilibrium states. As a first hypothesis, they then considered that deviation from the initial

curve (period of reference) is only due to changes in the land surface, such as the effect of anthropogenic activities and land cover variations. Assuming ω to be climate invariant, the changes due to climate are considered in the framework only through the modifications of the average P/PET (Fig. 3.1, red arrow). It follows the hypothesis that watersheds follow their Budyko curve if the catchment's surface characteristics remain constant.

However, studies have shown that not all catchments under climate change exhibit this behavior. There is a climate dependence of the deviation to the initial curve. Reaver et al. (2022) showed that reference catchments with the long-term stability of land use did not always follow their Budyko-curve. With the previous hypothesis, this deviation could be misinterpreted as a change in the land-surface characteristics. Padrón et al. (2017) found that the variability in the parameter is highly correlated to climate features such as snow fraction precipitation and the storm arrival rate. Over their extensive global database, the correlation to vegetation indices and direct anthropogenic influence factors is only secondary. Jaramillo et al. (2022) used CMIP6 multi-models ensemble to fit Budyko curves over several basins for the period 1901-1950 and to calculate ET/P PET/P for 2051-2100. They compared the results of the ensemble to those obtained with the hypothesis that catchments should follow their initial Budyko curve. Most basins will not follow the curve under climate change, showing a climate dependence of the deviation from the initial curve.

To circumvent the limitation due to the hypothesis of ω being climate invariant, several studies have tried to locate an expression of the evaporation efficiency parameter as a function of pertinent factors. It would allow us to express the evolution of ω over time and decompose the effects of climate and human activities through the different factors chosen. If valid, it would also allow transposing the expression to unmonitored catchments where ω cannot be directly fitted or to future catchment conditions. Different methods, such as step-wise regressions and neural networks, were used to identify pertinent factors across a set of catchments. Such methods require enough information on the chosen factors; strong hypotheses stand behind the expression. Some factors were selected in some studies and rejected in others and not all studies tested the same factors. It shows a high dependence of the final expression on the choice of factors tested on the area of study. Another strong hypothesis is that such a relationship defined over spatial differences is applicable to explain temporal differences (Luo et al., 2020; Abatzoglou & Ficklin, 2017). Other studies (Jiang et al., 2015; Zhao et al., 2018) looked at time-varying human activities and climate change factors to construct expressions, using a time-moving window to fit the evolution of the catchment parameter over a basin. This approach faces another limitation due to the availability of information on the different factors' time evolution. Ning et al. (2019) used a mixed technique, applying their fit across 30 basins at different time scales using moving time windows and found that the impact of vegetation cover and climate seasonality on the evaporation efficiency parameter was stronger over longer time steps, showing that the weight of different factors varies with the time scale. Ultimately, the pertinent factors highly differ among studies, regions, and climate types (Padrón et al., 2017). More details on these methods and their limits will be given in chapter 4).

Moreover, recent studies question the hypothesis underlying these studies and "the catchment trajectory conjecture" (Berghuijs et al., 2020; Reaver et al., 2020, 2022). The study demonstrates that the parameter exhibits a lack of independence from climate but also depends on the biophysical characteristics of the catchment directly due to the dependence of E, PET, or P on those features (Reaver et al., 2020). The highly non-linear relationships between all the features involved in the evaporation efficiency of the catchment and the average P and PET contradict the hypothesis that the parameter ω can be expressed as a simple

function of independent parameters. It also explains why previous studies were so different from one another. The catchment parameter is, therefore, a mathematical tool to represent the evaporation efficiency of a catchment over a given period and has no physical meaning in itself (Abatzoglou & Ficklin, 2017; Reaver et al., 2022). It is not transferable through time and space. It only positions the catchment within the Budyko space (Reaver et al., 2022). It can still be used to study the position of the catchment in the Budyko space and how its evaporation efficiency changes. Studying the deviation to the curve may provide insight into how factors besides aridity affect the water balance (Berghuijs et al., 2020).

3.2.2 The Land Surface Model (LSM) ORCHIDEE: a "natural reference" simulation

To isolate the climate change effect from other factors that could affect watersheds, we work with the outputs of a Land Surface Model (LSM). The model constructs watersheds with constant hydrological properties and represents an idealized watershed without any direct changes from human activities and other non-climatic disturbances. Therefore, the only source of long-term change would be due to a difference in response to an evolving climate.

This study uses the LSM Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) from the Institut Pierre Simon Laplace (IPSL). It includes biophysical and biogeochemical processes to simulate the global carbon cycle and quantify terrestrial water and energy balance. It runs coupled to an atmospheric model or in offline mode. In the latter case, an independent dataset forces the atmospheric conditions. Here, we use the model in stand-alone conditions, with the forcing dataset GSWP3 covering 1901 to 2013 (Hyungjun, 2017) with a resolution of 0.5°for all climate variables.

The hydrological network of the ORCHIDEE LSM is constructed from the hydrological elevation model HydroSHEDS (Lehner et al., 2008), which covers the area studied with the resolution of 30 arc seconds (approximately 1 kilometer at the equator). The hydrological information is then upscaled to the resolution of the atmospheric grid, the hydrological coherence being preserved by the construction of Hydrological Transfer Units (HTU) at the sub-grid level (Polcher et al., 2022). From a database of gauging stations, upstream basins are reconstituted on the hydrological elevation model grid and then projected on the atmospheric grid during the process. We have access to 2134 stations over the area studied for which the LSM calculates a discharge and for which we have the reconstituted upstream basin (Fig. 3.3).

The LSM ORCHIDEE, more specifically the SECHIBA (Schématisation des EChanges Hydriques à l'Interface Biosphère-Atmosphère) module, uses the USEB (Unstressed Surface Energy Balance) method to model Potential Evapotranspiration PET (details in (Barella-Ortiz et al., 2013)). This method relies on the turbulent diffusion equation to calculate the potential soil evaporation PET_{soil}, obtained from the air density, the aerodynamic resistance, and the humidity gradient. The USEB method estimates the virtual surface temperature from an unstressed surface-energy balance, computing a new energy balance considering a saturated surface (Barella-Ortiz et al., 2013). Potential transpiration is driven by PET_{soil} but limited by vegetation resistance, calculated in LSM ORCHIDEE and based on Plant Functioning Types (PFTs) maps and LAI calculations (Guimberteau et al., 2012). Then, the total potential evapotranspiration PET is calculated by summing the potential evaporation and the potential transpiration. PET is reduced to the actual evapotranspiration E by a "moisture availability function" (Barella-Ortiz et al., 2013).

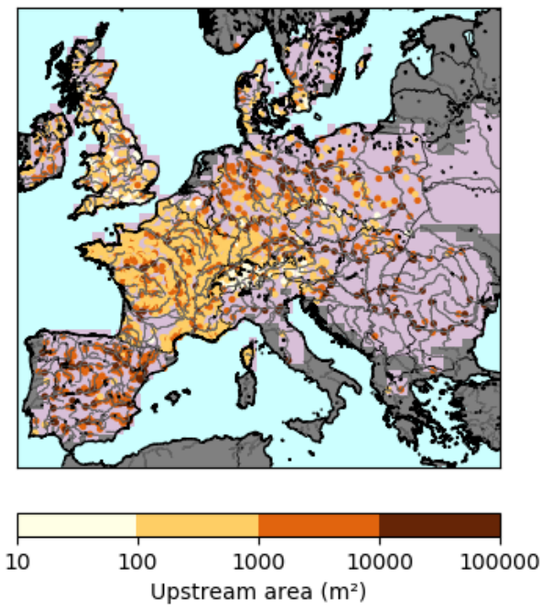


Figure 3.3 – Area studied. The dark grey background shows the area outside the scope of the study. The purple background shows the area covered by the sum of all catchments upstream of the available gauging stations. The points position the outlet of 2134 stations considered in the study, coloured by the size of their upstream catchment area.

Over the course of several years, the model has been tested and validated on many aspects of the land surface processes (hydrology, vegetation, and carbon cycle processes). This attests to its quality to reproduce the water and energy balance and also discharge over different areas over the globe (Tafasca et al., 2020; Nguyen-Quang et al., 2018; Wang et al., 2018; Guion et al., 2022; Polcher et al., 2022). Comparing the LSM outputs directly to observations for discharge is challenging, mainly due to the absence of certain processes in the models, including those resulting from direct human activities and the extensive water and river management (Wang et al., 2018), as it is the case in our area of interest. Based on previous literature, we can assume that the model proficiently emulates the mechanisms underlying actual evaporation, thereby effectively replicating the "natural" response of watersheds with persistent physical attributes to the past climate conditions prevalent in Europe. We study Q variations and not the absolute value of Q since we know that the output of the LSM does not represent the complete processes over real catchments and there could be systematic biases in forcing data. We focus here on the impact of the changes in atmospheric parameters on land surface responses with constant characteristics. The modeled watersheds react to the climate data input at each time step (30 min time step). Therefore the LSM output depends on both the evolving annual average and the evolving distribution over the year of the climate variables.

INFO BOX 2

Details on the basin's reconstruction

Here are more detail on how to reconstitute the watersheds over which we will then integrate climate data. This procedure was later on directly implemented in the construction of the routing system in ORCHIDEE.

We have :

- gauging stations' location (lat, lon) and reported upstream area (more details in

chapter 4).

- the HydroSHEDS (Hydrological Data and Maps Based on Shuttle Elevation Derivatives at Multiple Scales) developed by the Conservation Science Program of the World Wildlife Fund and based primarily on elevation data obtained during NASA's Shuttle Radar Topography Mission (SRTM) (Lehner et al. (2008)), with the 30 arc seconds grid (approximately 1 kilometer at the equator) resolution.

For each grid cell of the HydroSHEDS grid, we have the indication of where the river flow is directed (1=North, 2=North-East, ..., 7=West, 8=North-West). Placing the gauging station on the grid, looking from neighbouring cells to neighbouring cells, we regrouped all the cells constituting the catchment upstream of that point. Considering an uncertainty margin on the station positioning on the grid, we tested several grid points as starting point for the reconstitution, in a radius of 10 km around the station location and kept the best reconstituted catchment when comparing the observed and modeled upstream areas.

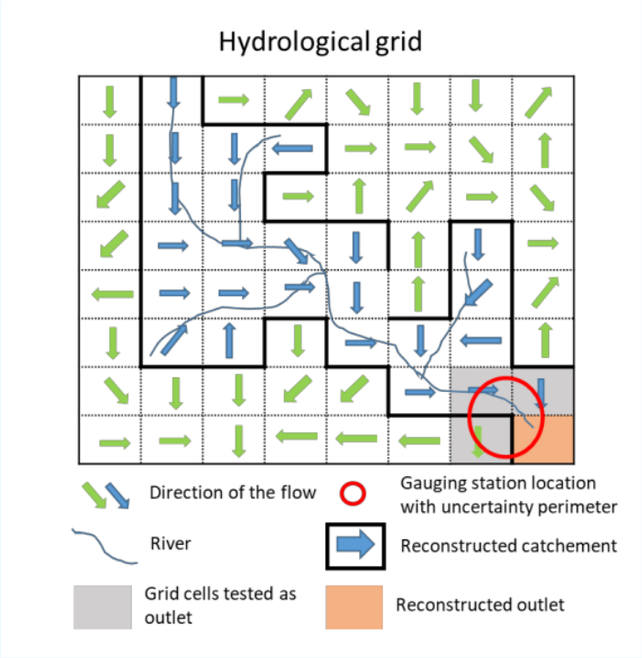


Figure 3.4 – Scheme presenting the reconstruction methodology of the upstream basin of a station on the HydroSHEDS grid. The reconstituted basin was considered as correct when there was less than 20% difference between the reconstituted area and the one provided on the metadata of the gauging station.

Still we were not able to correctly reconstruct catchments for all gauging stations. We considered the reconstruction was correct if we had less than 20% difference between the reported and the modeled upstream area.

For each station with a reconstituted watershed on the HydroSHEDS grid, we re-projected the basin on the lower resolution forcing data grid.

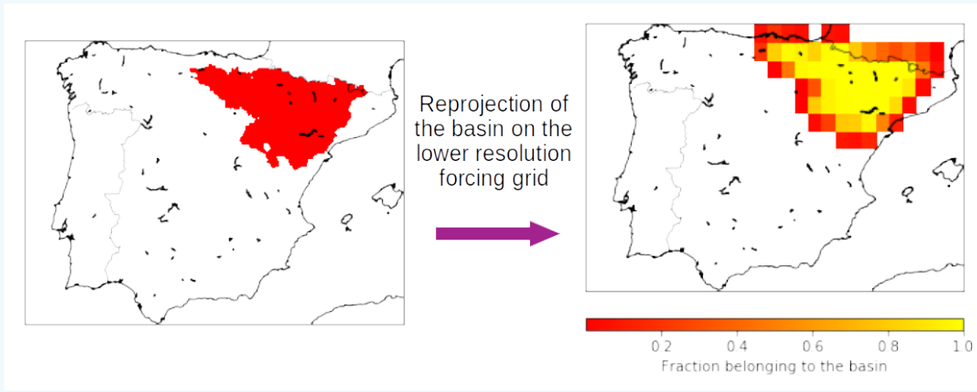


Figure 3.5 – Reprojection of a basin (Ebro basin, upstream of the gauging station Tortosa) fom the HydroSHED grid to the lower resolution grid of the GSWP3 dataset.

3.2.3 The forcing dataset GSWP3

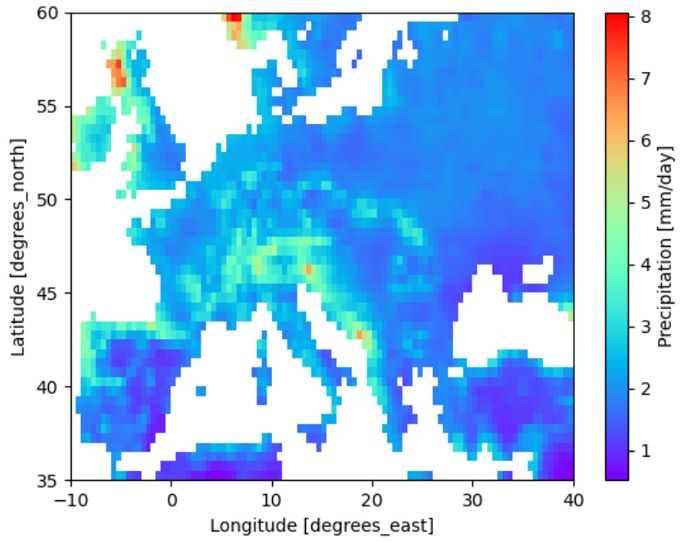
GSWP3 stands for : Global Soil Wetness Project Phase 3, Atmospheric Boundary Conditions, Institut of Industrial Science, University of Tokyo (Hyungjun (2017)). This dataset contains retrospective atmospheric boundary conditions for 9 climate variables for 1901-2010 at a 3-hourly resolution. It draws on observational data such as Global Precipitation Climatology Center (GPCC) (Rudolf et al., 2005) for precipitation, SRB for short/long wave radiations and Climate Research Unit (CRU) (Harris et al., 2020) for air temperature and daily temperature range and uses correction algorithms to produce the final dataset. The final grid has a geographic resolution of 0.5°x 0.5°.

We analyze this climate forcings in the following part, focusing on the climatic variables found to be most impactful for discharge changes: the average precipitation P, the average PET and the intra-annual distribution of P over the Mediterranean area, especially the Iberian Peninsula. We compare the results to those from the literature to verify the forcing as an adequate representation of the evolution of climate variables.

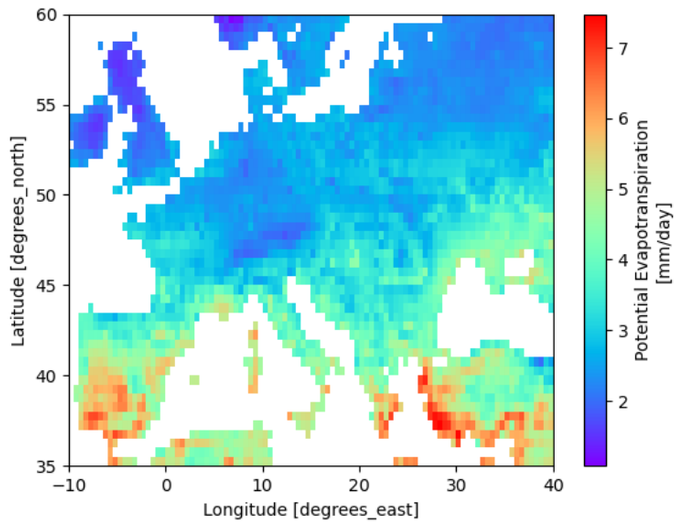
3.2.3.1 Analysis of annual average and trends of climate variables P and PET

Over the past century (1902-2010), the average P is higher in the north of Europe, especially over Norway and north of England, and over mountainous areas (Alps, Pyrenees) (Fig. 3.6a). When run with the LSM, the average PET grows lower as we go to the north of Europe and over mountainous areas (Fig. 3.6b), as expected if it follows the distribution of higher average temperature. Average evapotranspiration (E) fluxes modeled with the LSM have a smaller range of magnitude (Fig. 3.6c) than P and PET fluxes. The pattern in the North of Europe resemble PET patterns while it is closer to P patterns over the Mediterranean area. It is as expected, with the north being an energy limited area: PET is the limiting evaporative factor, while the south of Europe and especially the Mediterranean area is water limited, P being the limiting factor.

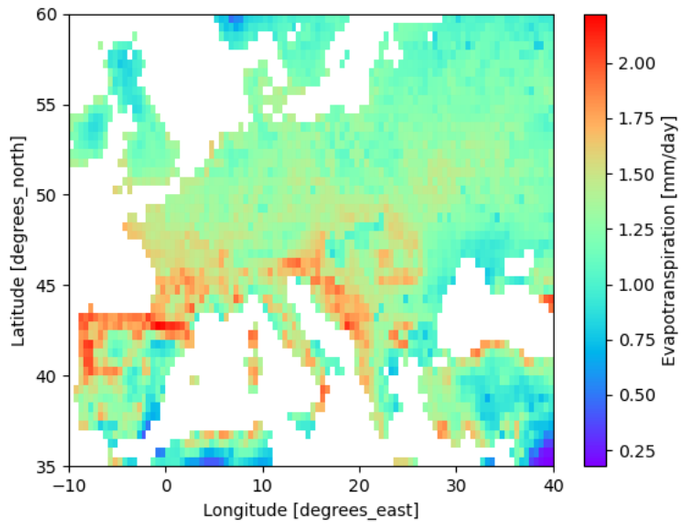
Looking now at the trends in the climate variables over the century. There are significant positive trends in P over Northern Europe, most of France and Germany, and negative trends over the Mediterranean area and Eastern Europe (Fig. 3.7b). Few basins have significative trends (Fig. 3.7a, 3.7b), due to the high inter-annual variability of P. Trends in PET are more significant due to a lesser inter-annual variability and show an increase over most of Europe, especially northern France and southern England, and a decrease in Eastern Europe,



(a) Average daily precipitation rate (mm/day)

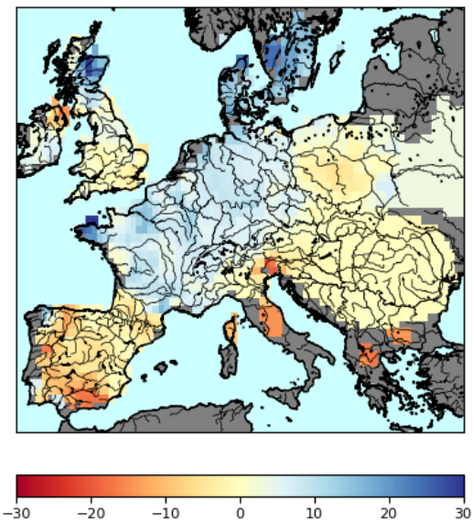


(b) Average daily potential evapotranspiration rate (mm/day)

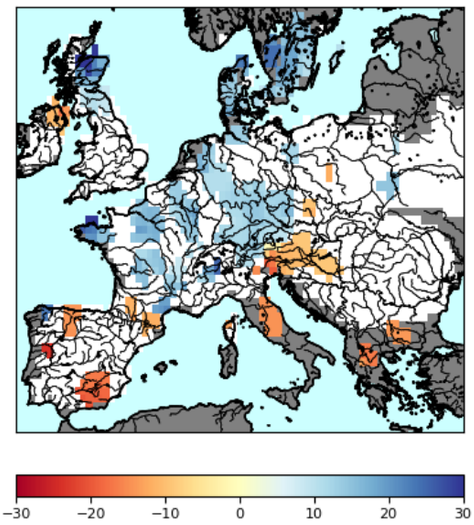


(c) Average daily evapotranspiration rate (mm/day)

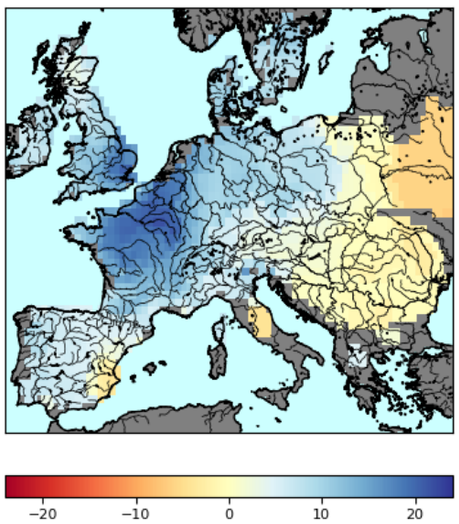
Figure 3.6 – Average daily flux rate (mm/day) with GSWP3 between 1901 and 2010. Evapotranspiration flux rate have a smaller rate than the two other fluxes, the scale of the colormap is shorter.



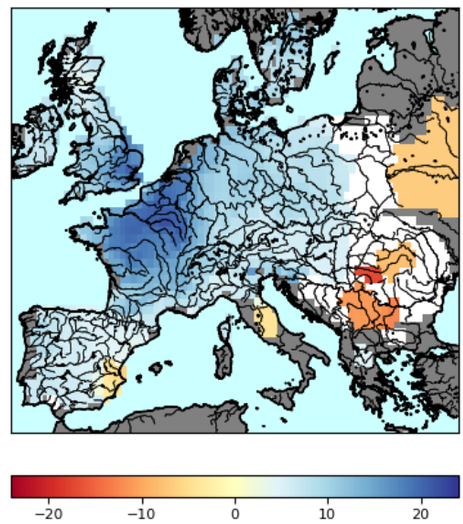
(a) Relative trend in daily precipitation rate over the century (%)



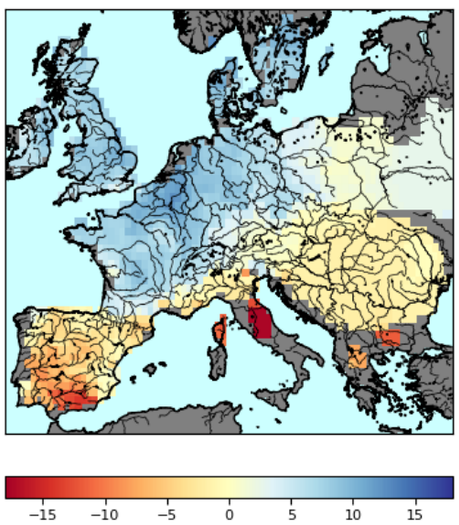
(b) Relative significant trend in daily precipitation rate over the century (%)



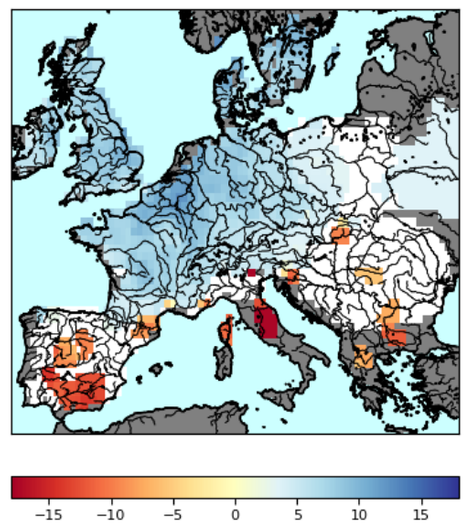
(c) Relative trend in daily potential evapotranspiration rate over the century (%)



(d) Relative significant trend in daily potential evapotranspiration rate over the century (%)



(e) Relative trend in daily evapotranspiration rate over the century (%)



(f) Relative significant trend in daily evapotranspiration rate over the century (%)

Figure 3.7 – Relative trends (%) in average water fuxes P, PET and E in ORCHIDEE run with GSWP3 between 1901 and 2010. Left column represent all trends, right column only the significant ones. The white areas are where the trends are not significant. All trends have been calculated at the catchment level and mapped over the entire catchment (see infobox 5).

Italy and Eastern Spain (Fig. 3.7d). Therefore P and PET both contribute to a similar trend in E: increase in the North and decrease in the South and the East of Europe (Fig. 3.6c). The significance of E trends is lower than for PET trends due to the effect of P variability.

The patterns and trends in average P and PET match rather well the patterns and trends found in other studies using other climatic datasets (García-Ruiz et al., 2011; Vicente-Serrano et al., 2019; Arnell, 1999). GSWP3 is therefore coherent with what is found in the literature. It appears that the variable P and PET seem to explain most of E pattern and trends modeled with the LSM but not completely and not linearly. It explains why parsimonious models mostly relying on these variables to explain E (and therefore Q) such as the Budyko framework are useful. However they are still missing some effects and partial variability, such as the effect on annual averages of intra-annual variability.

3.2.3.2 Analysis of intra-annual variations of P over the Iberian Peninsula

For that analysis, we focus on the Mediterranean area and more specifically on the Iberian Peninsula, where the effect of changes in intra-annual distribution of P appears to play a more important role on discharge trends.

Indicators

We use the indicators defined by García-Barrón et al. (2013), in order to evaluate the evolution of the intra-annual cycle over time. In their study, they analyze the output of a network of weather stations over the Iberian Peninsula. We want to assess whether GSWP3 produces similar results. These indicators are in days, for a given year n; the first one C_n represent a centered equilibrium day of the intra-annual distribution of P while the second D_n measures the dispersion around that centered day. More details in the infobox 3.

We can also calculate C_N and D_N, the value of the indicators on average over the entire period. Comparing C_n to C_N shows either an advance or a delay in the centralization of P. Comparing D_n to D_N shows either increase or a decrease in the concentration of P events around C_n.

Average intra-annual distribution of P

Figure 3.8 illustrate these indicators on average over the Iberian Peninsula, with two chosen stations representative of the North (Ebro river) and the South (Guadalquivir river) average intra-annual distribution regimes.

For the North-East of the area, there are higher values of C_N. It corresponds to more major P events occurring in late season (spring, summer). There is also a high D_N which can have two meaning: either there is rather a uniform distribution across the year of P events, either the distribution is bimodal. If we look at the average distribution across the year for the Ebro river, it corresponds to the second option. It matches the results of García-Barrón et al. (2013), who found a bimodal distribution along the Mediterranean coast, with a main peak in autumn for most of the coastal region but in spring for the Ebro river.

For the South-West, a lower C_N correspond to an average peak in winter, and a smaller D_N to a rather unimodal distribution.

It matches the results of García-Barrón et al. (2018), who found an asymmetric unimodal profile with maximum values at the beginning of winter (December), predominantly in the

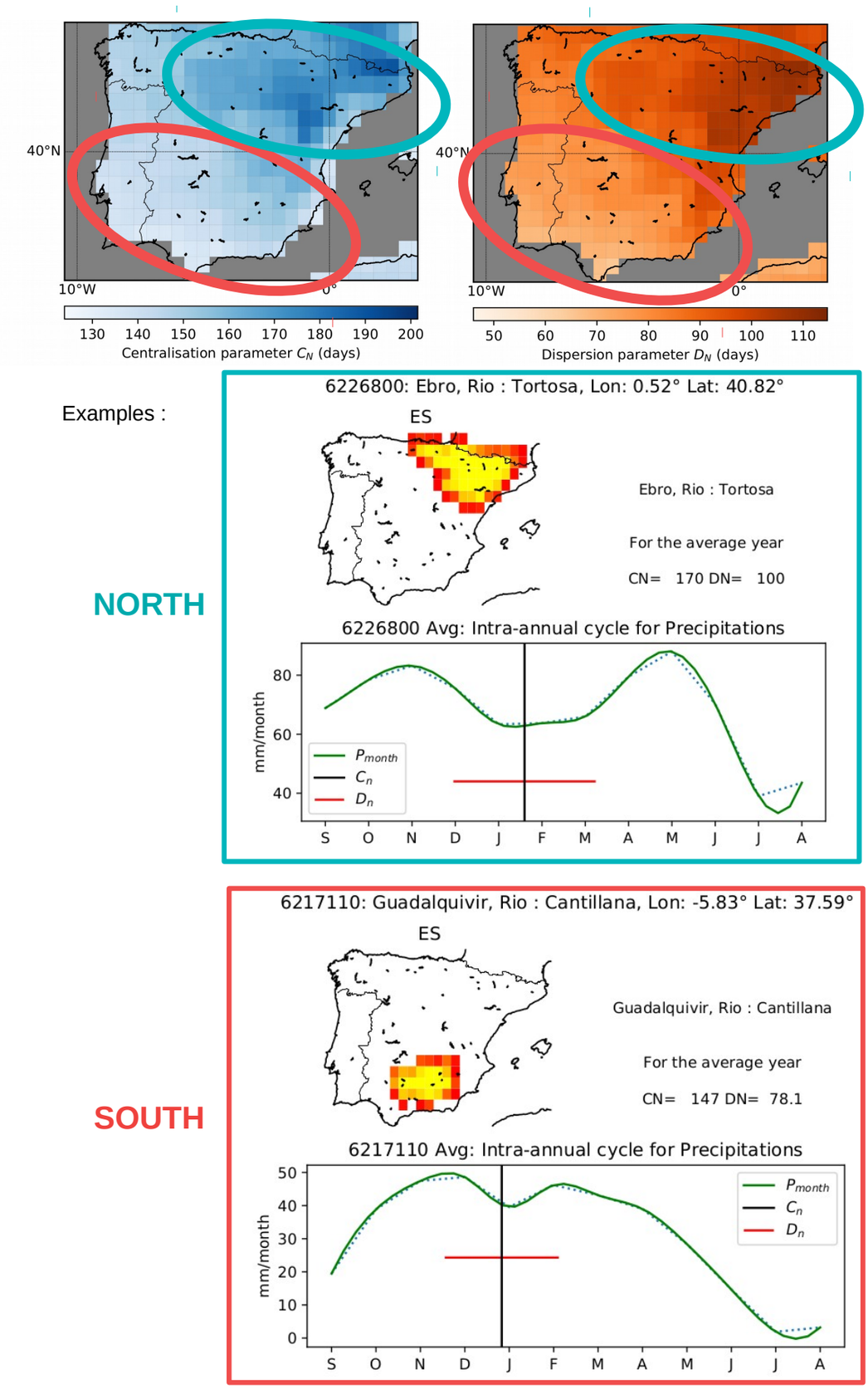


Figure 3.8 – Average characteristics of the intra-annual distribution of precipitation over the Iberian Peninsula (IP) and example of this average distribution for two representative catchments. The individual graphs show the average distribution of P (mm/month) over a year. The green dotted line represents the monthly distribution profile and the full green line the polynomially smoothed distribution of average monthly P. The black vertical line corresponds to the C_N value, and the red horizontal line to D_N, centered on C_N. The wider the line, the larger the D_N.
Lower values of the centralization parameter (C_N) correspond to more concentrated rain events in autumn and winter. Lower values of the dispersion parameter (D_N) correspond to a higher proportion of precipitation around the centralization date. In the south of the IP corresponds to a unimodal distribution (low D_N) with a peak in winter month (low C_N). Oppositely, in the north, it corresponds to an asymmetric bimodal distribution (high D_N) with a small peak in autumn and the larger peak in spring (higher C_N but tempered by the autumn peak).

INFO BOX 3

- Centralisation Parameter (C_n)

  This first indicator corresponds to a day in the year for which if the whole annual rainfall had occurred on that date, the first-order moments distributed throughout the year would be equivalent (García-Barrón et al., 2013).

$$C_n = \sum(x_i.p_i)/\sum p_i \tag{3.5}$$

  With p_i the amount of P for a given month i and x_i the order of the day respecting of the origin chosen (number of days corresponding to the middle of the month since the origin). Here the origin is September 1st.

  C_n represents the equilibrium day of the intra-annual distribution of P for the year n.

- Dispersion parameter (D_n)

  The second indicator measures the dispersion of precipitation around the centered parameter (García-Barrón et al., 2013).

$$D_n = \sqrt{\sum(d_i^2 p_i)/\sum p_i} \text{ with } d_i = |x_i - C_n| \tag{3.6}$$

  Allows to evaluate the dispersion compared to the central position. Low value: precipitation concentrated around C_n. High value: precipitation distributed throughout the year or peaks towards initial (autumn) or final (summer) months.

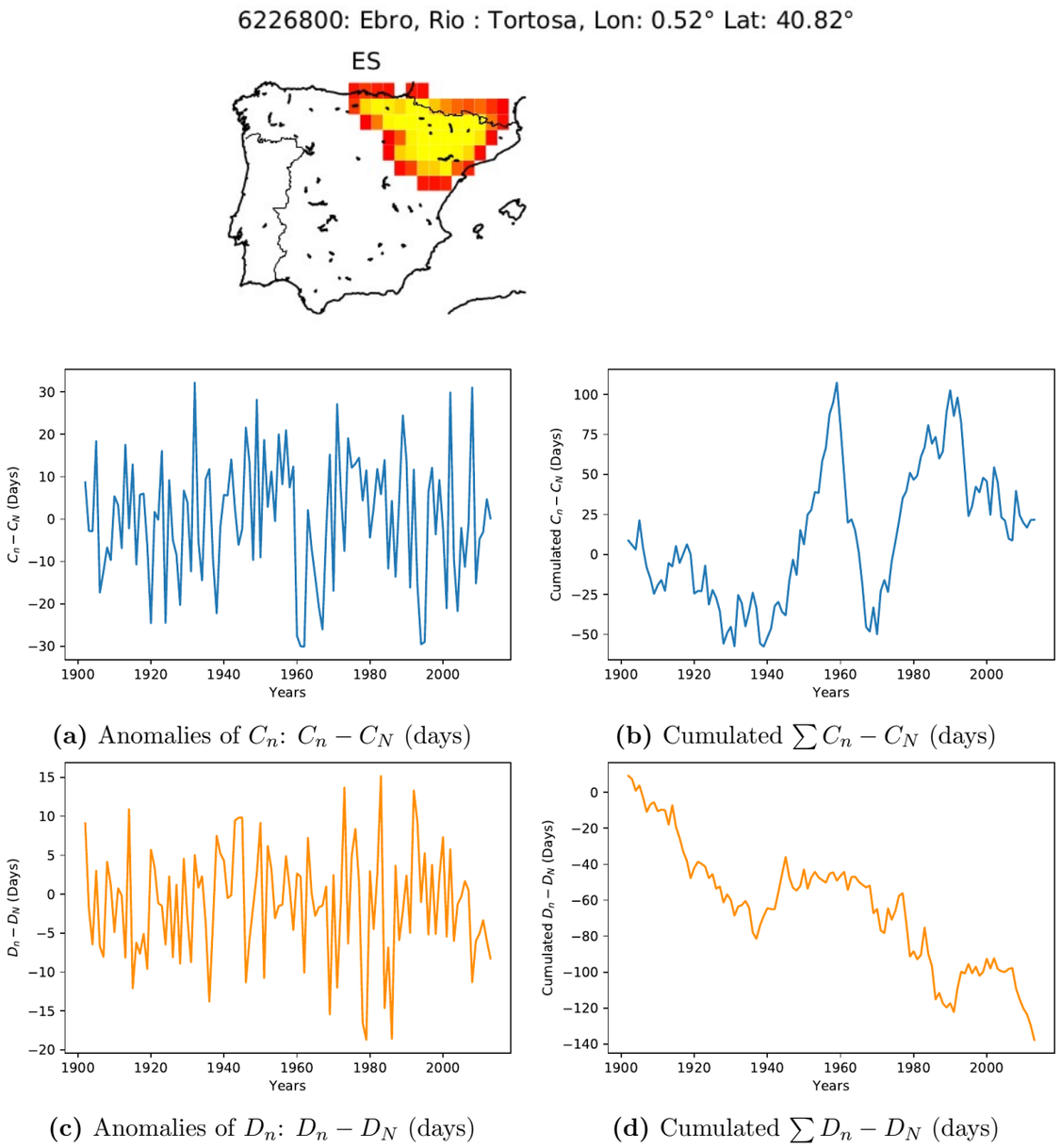
Atlantic region and the bimodal profile with maximum values in autumn and spring, predominantly in the Mediterranean region.

Trends in the intra-annual distribution of P

If we analyze the trends in C_n (Fig. 3.9a, 3.10a) and D_n (Fig. 3.9c, 3.10c) anomalies for the north (Fig. 3.9) and the south (Fig. 3.10), there is no clear signal or significant trend.

Since there is a high inter-annual variability of these parameter, in order to better see if over the period, there is a tendency toward an advance or a delay over the year, we look at the cumulated $C_n - C_N$ over the years. If the cumulated $C_n - C_N$ tends to increase, it means that the centralisation of P is delayed towards spring season. It seems to be what is happening in the south (Guadalquivir basin, Fig. 3.10b) from 1940 to 1980. If it decreases, it shows an increase of rainfall events in autumn. Here it matchs the trend from 1980 to 2010. Along with an increase in D_n anomalies (Fig. 3.10d), it can correspond to a shift in the climate distribution of P from an unimodal distribution with a peak in winter to a more bimodal distribution with a higher D_n, with a peak shifting from spring to autumn. Therefore the P intra-annual distribution regime of the south and West of the Iberian Peninsula seems to be changing to look more like the average Mediterranean regime of the north and East (see Fig. 3.8). These results again match with the ones from García-Barrón et al. (2013) who found a qualitative tendency of rainfall displacement towards autumn in the southwest of the Iberian Peninsula and relate that effect to an accentuation of the Mediterranean climate and to a decline of Atlantic influence in the area.

Trending periods are less clear for the North of Spain and the Ebro river (Fig. 3.9b), with



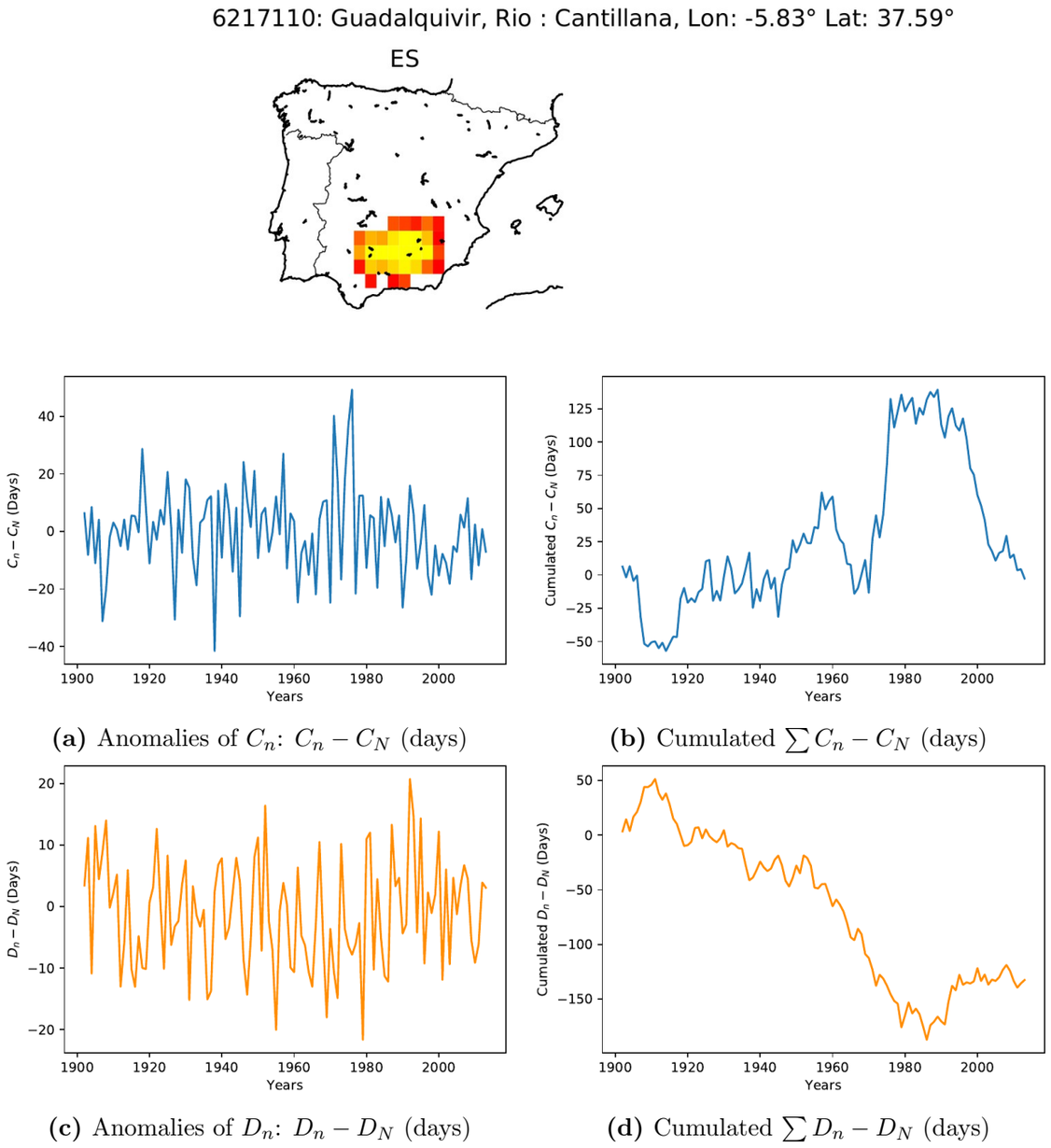
(a) Anomalies of C_n: $C_n - C_N$ (days)

(b) Cumulated $\sum C_n - C_N$ (days)

(c) Anomalies of D_n: $D_n - D_N$ (days)

(d) Cumulated $\sum D_n - D_N$ (days)

Figure 3.9 – Catchment Ebro, Tortosa: evolution over the century of the centralisation (C_n) and dispersion (D_n) parameters. The left column present the yearly anomalies of each parameter. The right column present the cumulated anomalies over the year.
For C_n, a positive anomaly corresponds to an increase of in the proportion of late precipitation (spring, summer) while a negative anomaly corresponds to an increase in early precipitation (autumn and winter). The cumulative graph allows to better look at long term trends towards negative or positive anomalies.
For D_n, a decreasing trend corresponds to an increased variability and an increasing trend to a decrease in the variability of precipitation distribution around the year.



(a) Anomalies of C_n: $C_n - C_N$ (days)

(b) Cumulated $\sum C_n - C_N$ (days)

(c) Anomalies of D_n: $D_n - D_N$ (days)

(d) Cumulated $\sum D_n - D_N$ (days)

Figure 3.10 – Catchment Guadalquivir, Cantillana: evolution over the century of the centralization (C_n) and dispersion (D_n) parameters. The left column present the yearly anomalies of each parameter. The right column present the cumulated anomalies over the year.
For C_n, a positive anomaly corresponds to an increase of in the proportion of late precipitation (spring, summer) while a negative anomaly corresponds to an increase in early precipitation (autumn and winter). The cumulative graph allows to better look at long term trends towards negative or positive anomalies.
For D_n, a decreasing trend corresponds to an increased variability and an increasing trend to a decrease in the variability of precipitation distribution around the year.

an increase in D_n anomalies (Fig. 3.9d). This could be that the bimodal regime becomes even more pronounced (increase in D_n), with the highest peak shifting regularly between autumn and spring (no clear trend in C_n but large anomalies).

More generally, García-Barrón et al. (2018) found a displacement towards spring starting around 1970 and a new displacement phase towards autumn for the past last decades. It is similar to what we found when analyzing the changes in GSWP3 P intra-annual distribution. The trends are however not statistically significant.

Since we found trends similar to what is reported in the literature for the climate factors tested, we can conclude on the coherence of the reanalysis used in this study to represent historical climate. This analysis also highlight that the analysis and the quantification of trends in intra-annual distribution of P is difficult and rarely significant. Despite this difficulty, the method we present in this chapter is able to identify and quantify their effect on discharge changes through the use of a LSM outputs and synthetic forcings.

3.2.4 Synthetic forcings to analyze the effect of variation of seasonality

In order to better understand the effect of inter and intra-annual climate variations on the Budyko framework and on discharge Q, we construct synthetic climate forcings, fixing one or the other (Fig. 3.11).

The calculation of PET includes many related climate variables and non-linear relationships, making it very difficult to anticipate how a change in a given climate variable may influence its behavior. It is, therefore, too complicated to create synthetic forcings for which we can modify climate variables to fix PET seasonality, for instance. Therefore, we only modify the precipitation P in the synthetic forcings to see how it impacts our results compared to the reference forcing.

The reference forcing is the GSWP3 dataset from September 1901 to September 2012 (3h time step). Then we constructed three forcings, which were modified over hydrological years (Tab. 3.1, Fig. 3.11):

- ***f2000***: A forcing where all 3h values of P are set to the values of the year 2000 (September 1999 to September 2000) for each year. Therefore, all components of P (average and intra-annual variations) are set constant.

- ***cstmean***: A forcing for which we keep the relative intra-annual distribution of P of each year, but where the average P of each year is set constant. The 3h values of P are scaled so the yearly hydrological year average is set to the one of the year 2000 (September 1999 to September 2000).

- ***cstintravar***: A forcing for which we keep the annual average of P for each year, but where the relative intra-annual distribution of P is set constant. The 3h values of P are set to the values of the year 2000 (September 1999 to September 2000) for each year and then scaled so the hydrological year average is set to the one of the corresponding year in the reference forcing.

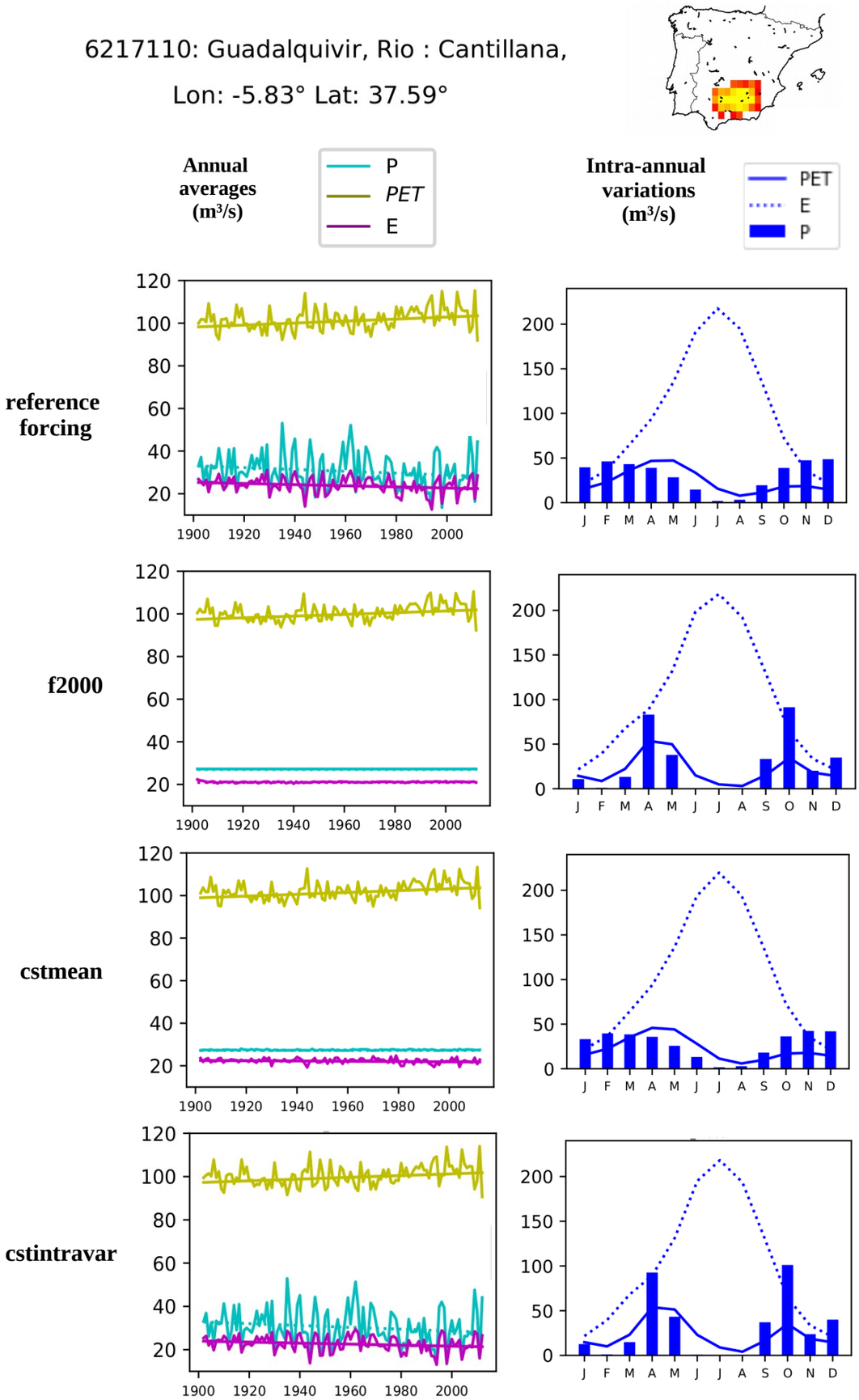


Figure 3.11 – Synthetic forcings over a given basin: the first row of graphs shows the inter-annual variability of annual averages of climate variables P, PET, and E modeled by the LSM, forced with the different synthetic forcings. E mostly relates to P. The second row shows the average seasonal distribution of P over the catchment for each forcings over the entire century. The average monthly distributions of P shown here are computed over the century. It however varies from one year to another for the reference forcing and the forcing *cstmean* which is not illustrated here.

Table 3.1 – Synthetic forcings

	forcing name	Average P	Intra-annual variation of P	Description*
1	*ref*	-	-	Reference forcing: GSWP3 (1901-2012)
2	*f2000*	fixed	fixed	P has been entirely fixed for each year, equal to the precipitation and the seasonality of the year 2000.
3	*cstmean*	fixed	-	Only the average value of P has been fixed for every year to the one of year 2000
4	*cstintravar*	-	fixed	Only the intra-annual variations of P have been fixed for every year to the one of year 2000

*For forcings 2 to 4, P has been modified compared to the reference: the average value of P over the year and/or the distribution of precipitation over the year have been fixed for each year to the value of the year 2000.

INFO BOX 4

Details on the creation of the synthetic forcings

- **Fixing P or its intra-annual variations to the one of a reference year**

We wanted to fix the average P and the intra-annual distribution of P over the years. To do so we had to pick a year as our standard (an hydrological year, starting in September of the previous year).
The choice of the year will influence the absolute values of the different climatic outputs of the model run with the synthetic forcings. However in this study, we are mostly interested in the changes than in the absolute values. Therefore it didn't matter much which year we choose. It could also have a small influence on the trends studied later but considered negligible (see 3.4).
Choosing a year at the end of the period would allow to contrast more with the beginning of the period, under the assumption that climate change had a significant effect, if we were interested in looking at absolute values. Also, GSWP3 is more likely to be accurate over more recent years, since it relies on observational datasets which improved over the years.

To fix the intra-annual distribution of all years, once we have our reference year, we calculate for each time-step (here a 3-hr time-step) of that reference year, the fraction of P fallen over this time step, out of the total P fallen that year. This fraction is used then to redistribute P over the year for each year. We choose the year 2000.
For each year time step i of year y_i:

$$P_i = f_i^{2000} * P_{yi}$$

with f_i^{2000} the fraction of P attributed to that same time-step i over the year 2000 and P_{yi} the total P fallen for the year yi.

$$\sum_{i=1}^{n} f_i^{2000} = 1 \text{ so } \sum_{i=1}^{n} P_i = P_{yi}$$

with n the number of time step over the year 2000.
We keep the annual average of the year P_{yi} but the intra-annual distribution is the one of the year of reference 2000.
This is only possible if for all time step i in each year, we can find the equivalent time-step in our reference year. **Therefore our reference year has to be a leap year**. In the case where yi is also a leap year, there will always be a matching time step in the reference year. In the case where yi is not a leap year, we will also always find a matching step for all time step. However the number of time step for that year will be lower than the one from our reference year $n - l < n$. The final sum $\sum_{i=1}^{n-l} f_i^{2000}$ won't be exactly equal to 1 since we are missing the steps of one day, so the final average of P for the year i in the synthetic dataset will be slightly lower than the initial P_{yi}. We consider that the effect of one day missing is however negligible and induces no trend since gap years are regularly distributed.

To fix the annual average of P for all years, we calculate the fraction f_i^{yi} of P over each time step for each year yi and then we rescale P_i to reach, on average over the year, the average of our reference year P_{2000}.

$$P_i = f_i^{y_i} * P_{2000}$$

Similarly, we keep the annual average of the year of reference P_{2000} but the intra-annual distribution is the one of the year yi.
Again due to the mismatch of the number of time steps between leap and non-leap years, in the final synthetic product, only leap years will truly have a final average P equal to P_{2000}. For all other years, the final average P should be slightly lower but constant for all non-leap years. Since leap years happen regularly, it doesn't introduce any trends.

- **Understood limits**

In the initial variables of the dataset GSWP3, we don't have P but we have rainfall $Rainf$ and snowfall $Snowf$ separately. Therefore we didn't exactly fix P but rather each $Rainf$ and $Snowf$ separately. In the end it effectively fixes P but the ratio $Rainf/Snowf$ may have changed.

Fixing PET is not possible because it is non-linearly dependent on too many different climate variables. Therefore we only fix $Rainf$ and $Snowf$ to study the effect of P. Also since we fix these variables independently from all the other climate variables, there will be inevitable inherent incoherences in synchronization between all variables over the year.

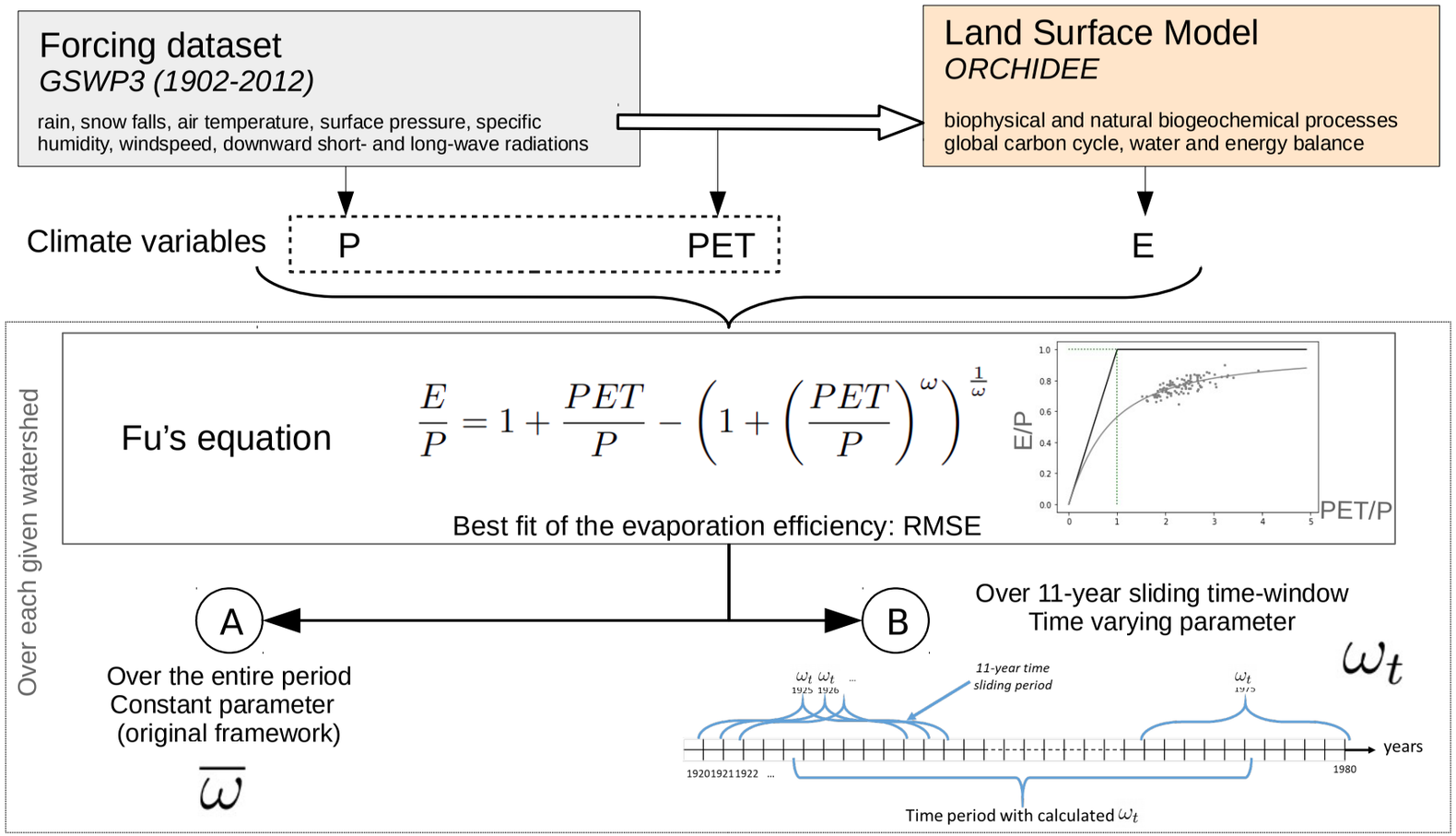


Figure 3.12 – Scheme of the method: The Land Surface Model (LSM) is obligated with the forcing dataset to calculate E. The LSM is considered to represent the "climatic reality" over a catchment without any changes in the watershed characteristics. We then average P, PET, and E and integrate them over each watershed to get annual averages for all catchments. Then we fit Fu's equation. A) The fit of the equation over the entire century results in the calculation of an empirical parameter $\overline{\omega}$ which represents the average catchment evaporation efficiency. B) To have an evolution of ω_t over time, the fit was then successively applied over an 11-year sliding time period.

3.2.5 Combining the Budyko framework to LSM outputs

3.2.5.1 Aim of our methodology

In this study, we apply the Budyko framework to the output of an LSM to explore the sensitivity of the empirical parameter ω to climate change and the resulting effect on discharge. Without expressing ω as a function, we can still consider that the changes in the fitted parameter (i.e., the deviation to the initial curve) relate to changes in the overall evaporation efficiency of the catchment. We use the changes in ω as a proxy for changes in the partitioning of P into Q and E other than direct changes in average PET/P. We want to capture the non-linear effects of climate change on the evaporation efficiency and analyze the main climatic factors involved. We use the changes in ω as a proxy for changes in the partitioning of P into Q and E other than direct changes in average PET/P; it focuses on the deviation from the initial curve and attempts to decompose its dependence on climate. In this case, any deviation to the curve is only due to climate effects. Since ω has no clear physical meaning, we don't analyze directly the changes in ω but rather how they impact the evolution of discharge.

The watersheds in the LSM have constant biophysical characteristics. The LSM then reproduces the interaction of the land surface with climate parameters. It is affected by climate change and no other source of change. Using an LSM, we can also change various climate parameters to better address how they weigh in the modeled changes. We develop a varying ω_t to capture part of the change in the evaporation efficiency of the watersheds due to climate. We compare its effects to the magnitude of change in discharge already captured with the traditional framework, which only considers changes in annual averages of PET/P.

3.2.5.2 Integrated data at the catchment level: P, E, PET

For consistency in the calculation of E and PET, we take both from the output of ORCHIDEE forced with GSWP3. The gridded outputs (PET, E) are at the resolution of the forcing dataset (0.5°). P is the sum of rainfall and snowfall in GSWP3. Then we consider the annual mean P, PET, and E over hydrological years, integrated over each catchment. The catchments' shape has been reproduced at a finer resolution and then projected on the 0.5°grid (infobox 2).

3.2.5.3 Fit of the Fu's equation of the Budyko framework

Fit over the entire period

The evaporation efficiency parameter of the Budyko curve is calculated over each catchment with a fit of the equation curve $E/P = f(PET/P)$ (equation 3.2), using the minimum root mean square error (RMSE) for a given set of annual averages of evapotranspiration E, precipitation P and potential evapotranspiration PET data (Jiang et al., 2015; Yang et al., 2007). We fit the parameter once with all points over the entire period covered by the climate dataset to obtain $\overline{\omega}$ representing the average behavior for each catchment (Fig. 3.12, A).

Introducing a varying evaporation efficiency ω: fit with a sliding time window

For a watershed with constant hydrological properties (which is the case when considering modeled watershed in ORCHIDEE), if we consider the "catchment trajectory conjecture",

INFO BOX 5

Details on mapping of the results

To map the results at the catchment level and have a better visual for the mapped results, all grid points of a basin are coloured with the metric computed over the catchment. For sub-basins included into larger ones, the metric is averaged to color the shared grid points.

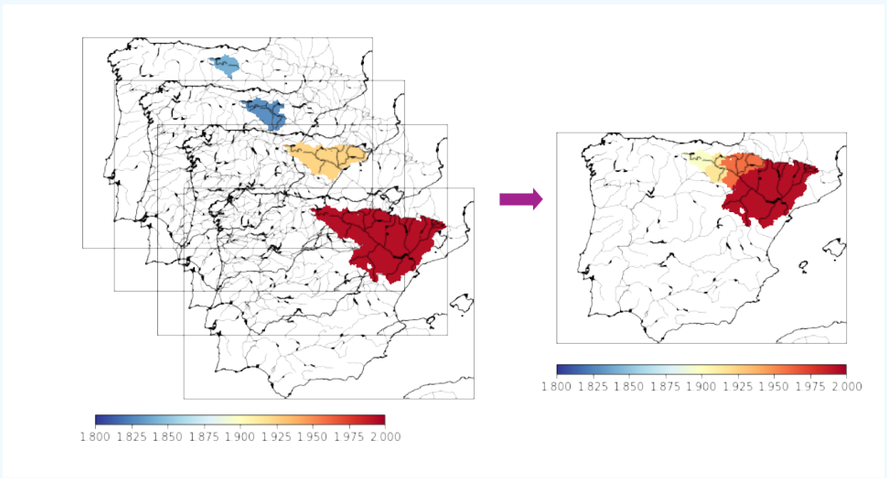


Figure 3.13 – Method applied to map the results, example over the Ebro river: we have one given value for each basin. It is applied to the entire upstream area. When sub-basins are contained in larger basins, the different values are averaged over the overlapping area.

ω is independent of climate, and the catchment follows its initial curve. However, ω varies for a given watershed because of climate.

For instance, a difference in storm depth over a catchment can change the capacity of the soil to store water, the response of vegetation, and change the dynamic of the water partition into runoff and evaporation even if the annual amount of precipitation stays constant Donohue et al. (2012). More generally, the intra-annual synchronization of P water available) and PET (energy demand) (or the annual covariance between PET and P) impacts the annual mean of E and Q for the same average climate Abatzoglou & Ficklin (2017); Li et al. (2022). With its simple framework, the Budyko model does not cover possible changes at intra-annual time scales. The average effect of this synchronization is included in the adjustment parameter ω, which is, therefore, not completely independent of climate. Therefore, long-term changes in seasonality should induce a climatic time dependence which is not accounted for in the framework with a constant ω. Therefore, considering a varying parameter should improve the Budyko model to reproduce E/P and its climatic evolution.

To obtain a varying parameter ω_t for each catchment, we carry out several fits over successive 11-year time-sliding sub-periods (Fig. 3.12, B).

We chose 11 years as the smallest time length to apply the Budyko framework relevantly, considering that each 11-year sub-period is stationary ($\Delta S = 0$, see infobox 1). This allows us to focus on long-term changes and to minimize the impact of year-to-year "transient" effects (e.g., soil storage and groundwater changes) (Yang et al., 2018). Tian et al. (2018) found that below a certain time length, the fit of the ω parameter was too unstable to be relevant.

3.2.5.4 Decomposing the impact of climate on discharge trends

The evaporation efficiency ω is a conceptual variable that provides little insight into the magnitude of discharge changes. Thus, we examine the impact of ω_t changes on the river mean annual discharge Q and compare these changes to the impact of annual averages of climate variables (P and PET) changes on Q over time. To simplify the discussion, we gather the annual averages of P and PET in a "climate" variable $C = (P, PET)$ (infobox 7).

Following our previous hypothesis (equation 3.4), Q can be estimated with the Budyko framework using C and ω: $Q = f(C, \omega_t)$. Q can be decomposed with first-order partial derivatives (equation 3.7), with the first term of the right-hand side representing the partial derivative due to climate variables C and the second term for the partial derivative due to changes in the evaporation efficiency ω_t. We then estimate the partial derivatives due to C and due to ω independently.

$$\frac{dQ}{dt} = \frac{\delta Q}{\delta C}\frac{dC}{dt} + \frac{\delta Q}{\delta \omega}\frac{d\omega}{dt} \qquad with \;\; C = (P, PET) \tag{3.7}$$

To independently estimate the partial derivative due to climate variables C, we must cancel the second term (equation 3.7, left side). To do so, we calculate the mean annual discharge $Q_c = f(C, \overline{\omega})$, with a constant value of ω. The trend of that discharge $\frac{dQ_c}{dt}$ matches the term with the partial derivative due to C in equation (3.7). To estimate the partial annual discharge trend due to ω_t, we need to eliminate the trends in annual averages of P and PET over the century to cancel the first term (equation 3.7, left side). We randomly draw P and PET pairings for each year. We do so several times and average the results for each year. It gives us a random climate without trends over the century. We then apply Fu's equation 3.2 with the resulting random annual averages of P and PET and the varying ω_t calculated with the forcing before the random drawing. It gives $Q_\omega = f(C_{rand}, \omega_t)$ for which the climate trends are only due to variations captured by the time-varying parameter ω_t. The trend $\frac{dQ_\omega}{dt}$ matches the term with the partial derivative due to ω in equation 3.7. In the end, we get:

$$\frac{dQ}{dt} = \frac{dQ_c}{dt} + \frac{dQ_\omega}{dt} \tag{3.8}$$

We calculate the trends of each term and their significance using the Mann-Kendall non-parametric test, associated with the Thiel-Sen slope estimator (infobox 6). It gives us time series and associated trends for each studied watershed. Fig. 3.14 shows an example of a watershed in southern Spain.

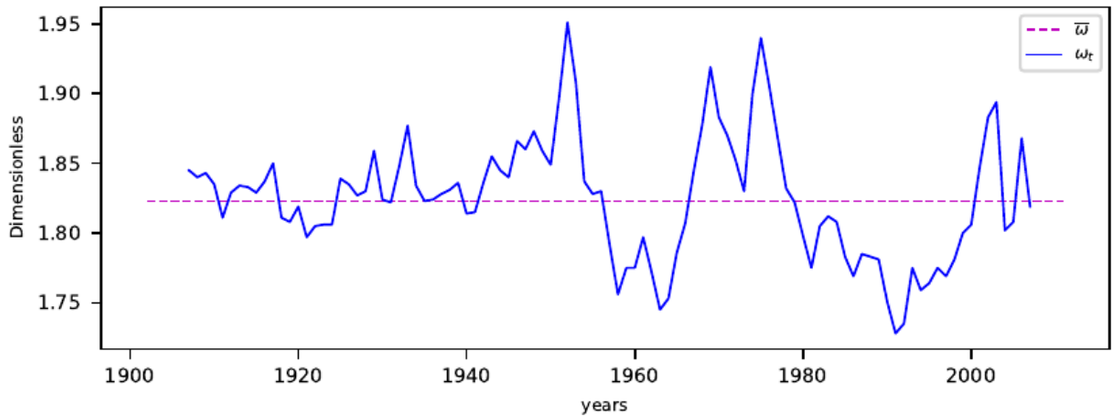
3.3 Results

3.3.1 Performance of Budyko with or without a variant ω parameter

We hypothesize that for watersheds with constant hydrological properties, the dispersion of annual points around the curve is due to intra-annual variations of climate. If these variations did not exist, catchments would follow their Budyko curve, and we could use it to model the discharge almost perfectly.

To test this hypothesis, we examine the performance of the Budyko curve with a constant parameter $\overline{\omega}$ to reproduce the discharge from the LSM for the reference forcing compared

(a) Evaporation efficiency $\overline{\omega}$ fitted over the entire time period (dashed purple line) and ω_t fitted successively over a sliding 11-year time-window (blue line) for the reference forcing.



(b) Annual mean discharge estimated with Budyko for the reference forcing: $Q = f(P, PET, \omega_t)$ (blue line), $Q_c = f(P, PET, \overline{\omega})$ (purple line), $Q_\omega = f(P_{rand}, PET_{rand}, \omega_t)$ (black line) with their associated trends. Unsignificant trends are dashed. Here all trends are significant.

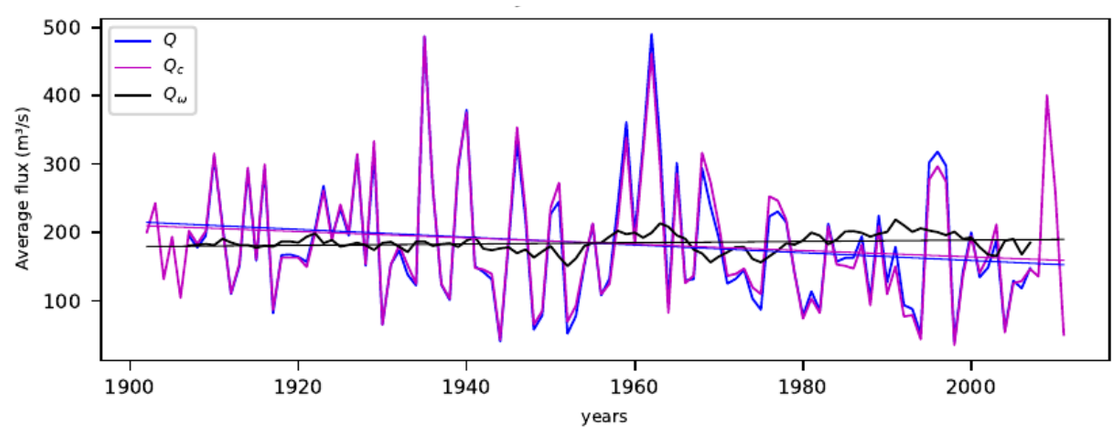


Figure 3.14 – Time series obtained through the full application of our methodology for a given basin in Spain. (a) and (b) are results for the reference forcing. (a) shows the varying ω_t resulting from the time-sliding window calculation (blue curve), compared to $\overline{\omega}$ calculated with one fit over the entire century (dashed purple line). (b) shows the decomposition of the discharge, comparing the full discharge to partial discharges and their respective trends. The full mean annual discharge Q is modeled Fu's equation with annual averages of P and PET from the reference forcing and ω_t. The first partial annual discharge Q_C is the one calculated with the constant parameter $\overline{\omega}$. It covers most Q variations for the given basins. The second partial annual discharge Q_ω covers some of the missing variations of Q and some of the missing trends due to deviations to the average curve. From that figure, we can conclude that most variations and trends of the discharge in this basin are explained by $C = (P, PET)$.

INFO BOX 6

Mann-Kendall (MK) non-parametric test and Thiel-Sen slope estimator

The MK test is a non-parametric test often used to analyze temporal trends in hydro-meteorological series (Xiong et al. (2020)). For n values identically temporally distributed of a given variable x, the given statistic test S and its corresponding variance are given by:

$$S = \sum_{i=1}^{n-1} \sum_{j=1}^{n} sgn(x_j - x_i) \text{ with } sgn(\theta) = \begin{cases} 1 & \text{if } \theta > 0 \\ 0 & \text{if } \theta = 0 \\ -1 & \text{if } \theta < 0 \end{cases}$$

$$Var(S) = n(n-1)(2n+5)/18$$

For the normal distribution, the standardized test Z is calculated by:

$$Z = \frac{1}{\sqrt{Var(S)}} * \begin{cases} (S-1) & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ (S+1) & \text{if } S < 0 \end{cases}$$

for a two-tail test:

$$p_{value} = 2 * (1 - P(X > |Z|)) \text{ where } X \sim \mathcal{N}(0,\, 1)$$

If $p_{value} < 0.05$, then the null hypothesis that there is no trend in the dataset is rejected for a significance level of 5%.
When the trend is considered significant, the slope β can be estimated with Thiel-Sen estimator:

$$\beta = median\left(\frac{x_i - x_j}{i - j}\right), \forall i, j,\, 1 \leq j < i \leq n$$

to the forcing *cstintravar*. For that latter forcing, we removed the intra-annual variations of P from one year to another, which should render the performance of the Budyko curve model close to perfect if the hypothesis is valid. We use the Nash-Sutcliffe coefficient (NSC) as a performance indicator (equation (3.9), Fig. 3.15). We consider a NSC> 0.5 to be satisfactory (Moriasi et al., 2007).

We obtain NSC values above 0.5 for 89.9% of all 2134 watersheds tested for the Budyko curve with a constant parameter (Q_c, calculated with a constant $\overline{\omega}$) applied with the reference forcing (boxplot on the left, Fig. 3.15). Therefore, the average curve model is rather effective in reproducing the annual discharge over watersheds with constant hydrological properties reacting to an evolving climate.

For the forcing *cstintravar*, NSC for Q_c increases to above 0.6 for all watersheds (boxplot on the right, Fig. 3.15). It confirms our hypothesis: the average Budyko curve model is even more effective if there are no intra-annual variations of P from one year to another. Therefore, most of the variability that is not captured by the average Budyko curve over the past century is due to the intra-annual variability of P and the covariance of P and PET.

When looking at NSC for the framework applied to the reference forcing with a varying

INFO BOX 7

Discussion around notations

1- About the parameter C

We gather the annual averages of P and PET in a "climate" variable $C = (P, PET)$. This allows to simplify the discussion around the variables P and PET. The notation around P and PET can be confusing, since they stand respectively for the annual average of precipitation and potential evapotranspiration most of the time, except when explained otherwise as when we refer to "the intra-annual variations of P". In that case, P stands for all precipitation characteristics.
In the case of the variable C, we use it to simplify the discussion as it always stands for the annual averages of P and PET, considered together.

INFO BOX 8

Nash-Sutcliffe coefficient (NSC)

We used the Nash-Sutcliffe coefficient (NSC) to attest the quality of streamflow modeled using the Budyko framework to adequately reproduce the LSM outputs.
NSC < 0 indicates that the use of the average value would be more accurate than the model tested (Ahn and Merwade 2014; Moriasi et al. 2007).

$$NSC = 1 - \frac{\sum_{i=0}^{years}(Ql_i - Qb_i)^2}{\sum_{i=0}^{years}(Ql_i - \overline{Ql})^2} \tag{3.9}$$

$$\begin{cases} \text{with } Ql = \text{reference discharge, here dicharge from the LSM} \\ \text{and } Qb = \text{discharge to test, here result from the fit with Fu's equation} \end{cases}$$

This indicator is considered as satisfactory for streamflow when NSC > 0 and good when NSC > 0.5.

parameter $Q(ref) = f(C(ref), \omega_t)$, we gain up to 0.26 points of NSC for the tested watershed and reach 94.1% of all watersheds with NSC>0.5 (boxplot on the center, Fig. 3.15). It does not reach the performance to reproduce Q_c with the forcing *cstintravar*. However, it enables to catch some of the deviation to the curve due to intra-annual trends of climate variables. We capture long-term trends following our choice of the 11-year time-moving window. It validates our hypothesis that introducing a varying evaporation efficiency ω_t improves the framework to better encompass climate variability and the effect of climatic trends on annual mean discharge, including the effect of climate change on the intra-annual distribution and covariance of climate variables (P and PET).

To sum up, for watersheds with constant hydrological properties under historical climate, most of the deviation to the average curve model (i.e., changes in the evaporation efficiency of catchments) is due to variations in the intra-annual distribution of climate variables (P and PET). Our varying parameter improves the framework by allowing us to capture the long-term trends of these variations. We will now analyze their effect on the annual mean discharge and compare them to the direct effect of trends in the annual average of climate variables.

INFO BOX 9

Complementary illustration of Q and ω variations at the catchment level, comparing synthetic forcings

To illustrate the part of variation of Q captured in C and in ω, we graph Q and ω at the catchment level for each forcing (Fig. 3.16).

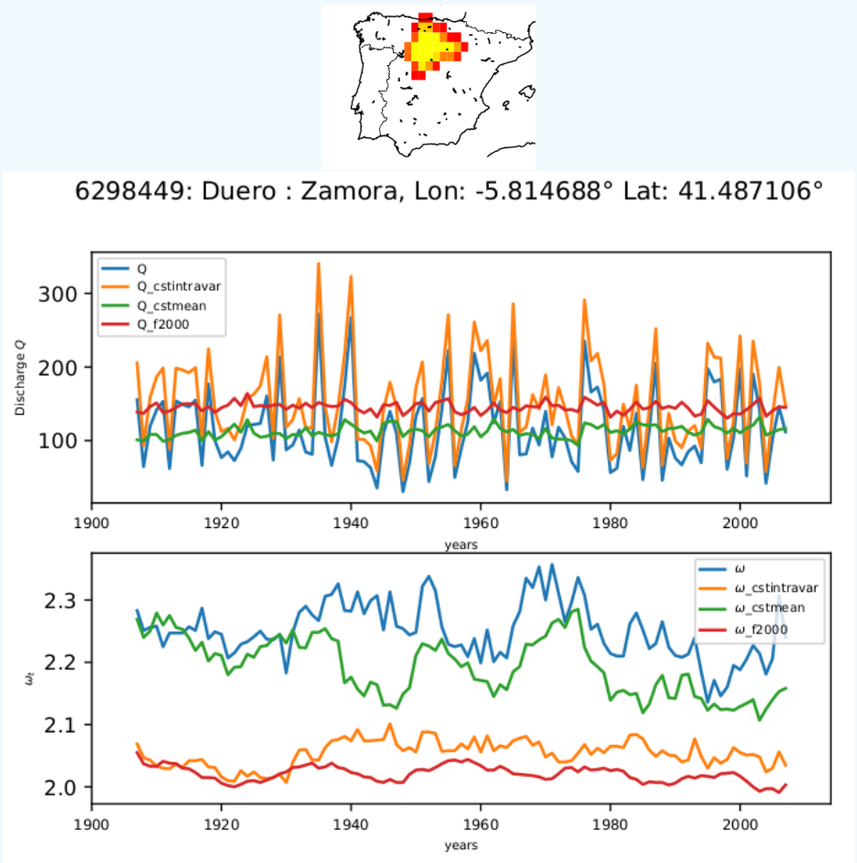


Figure 3.16 – Inter-annual variations of the annual mean discharge Q (top graph) and the associated ω parameter for each synthetic forcing over a catchment over the Duero river (basin coverage at the top).

We see that inter-annual the variations of Q are similar for the forcings reference and cstintravar (top graph, Fig. 3.16). The forcings f2000 and cstmean both have very little inter-annual variations, as they are designed as such, with no inter-annual variation of the annual average of P. It shows that most of inter-annual variations of Q are due to C (see section 3.3.2). However for the variations of ω_t (bottom graph, Fig. 3.16), the variation of each forcings are paired differently: reference/cstmean and f2000/cstintravar. For each pair of forcings, the annual average of C differ but intra-annual distribution of P are the same, with the same variations for the pair reference/cstmean (which results in higher variations in ω_t) and a fixed distribution for the pair f2000/cstintravar. Therefore it shows that while the main variability of Q is due to the inter-annual variability of annual average of P, the varying parameter ω_t captures well the effect of intra-annual distribution of P.

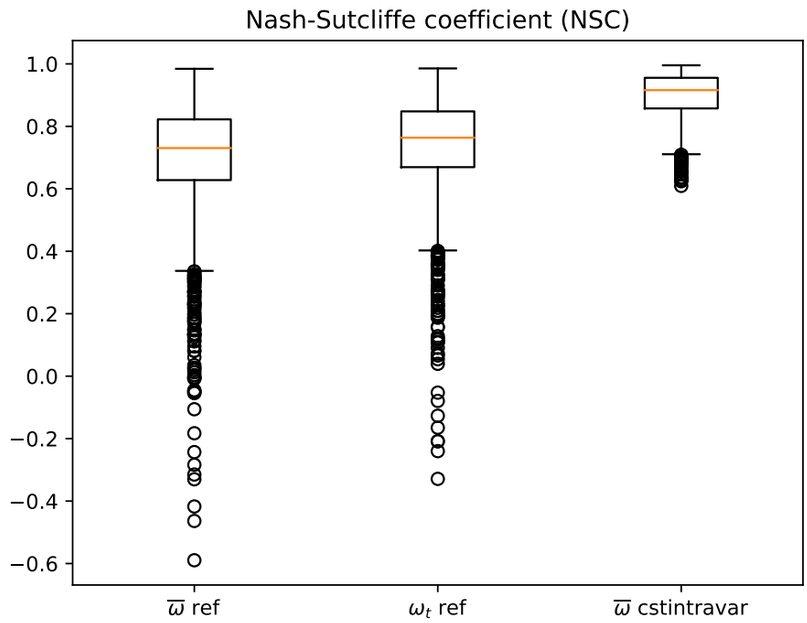


Figure 3.15 – Boxplot of Nash-Sutcliffe coefficient (NSC) for all watersheds: for the forcing of reference with the constant parameter $\overline{\omega}$, with the varying parameter ω_t and for the forcing *cstintravar* (where the seasonal distributions of P have been fixed over the entire time period) with a constant $\overline{\omega}$. It represents how well the Budyko model reproduces the discharge output from ORCHIDEE. A value above 0.5 is considered satisfactory. Very similar results are found when looking at R^2 from a linear regression.

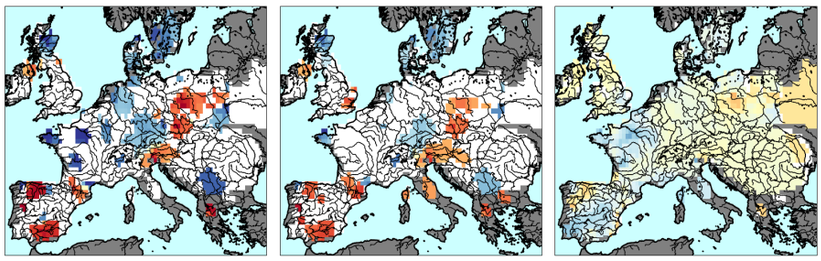
3.3.2 Comparing the effects of intra-annual variations of P on discharge Q to the effects of variations in annual averages of P in Europe

We consider our area of study, Western Europe (2134 watersheds modeled) (Fig. 3.17). To better illustrate our results, we also take two contrasted basins: one in Italy (Fig. 3.18) and another in England (Fig. 3.19).

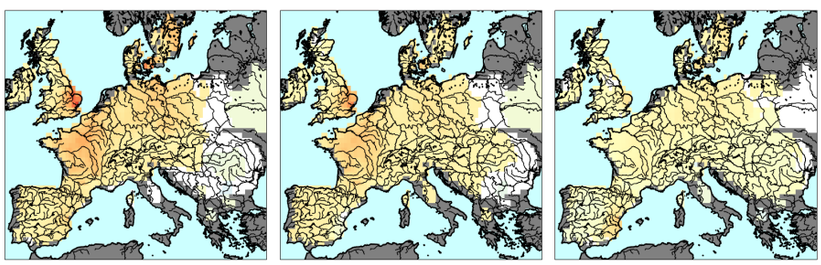
Figure 3.17a, 3.17b, 3.17c show the relative trends over each basin, respectively of Q, Q_c and Q_ω, for the reference forcing. There are significant decreases in the total discharge Q (Fig. 3.17a) (-0.3% to -0.4% per year over the past century) over scattered basins in Spain, the Pyrenees, Italy, Slovenia, Greece, and Eastern Europe. There are significant increases (Fig. 3.17a) (+0.2% to +0.4% per year over the past century) over sparse basins in France, Germany, Denmark, Sweden, Northern UK, and Serbia. These trends are primarily due to changes in the annual averages $C = (P, PET)$ since the average Budyko curve model Q_c captures most of the signal (Fig. 3.17b). The inter-annual variability of C is high, making the trends less than 95% significant over most basins for Q and Q_c. Both selected catchments better illustrate it (Fig. 3.19b and 3.18b), with the reference forcing (top left), the dominant effect in the variations of annual mean discharge Q (blue line) is due to the annual mean of climate variables C (purple line). Clearly, the blue and the purple curves have very similar high inter-annual variations and trends.

Changes in the mean annual climate, designated C here, are the dominant factors explaining the climatic trends in Q over the past century in Europe. The results obtained with the forcing *cstintravar* (bottom right for Fig. 3.19b and 3.18b and maps Fig. 3.17j to 3.17l) confirm it. It shows that without inter-annual changes in P distribution (in other words, with a maximum reduction of the inter-annual changes in the annual covariance of P and PET), the annual mean discharge Q obtained and the associated relative trends are very similar to the results obtained with the reference forcing. Therefore, the effects of changes

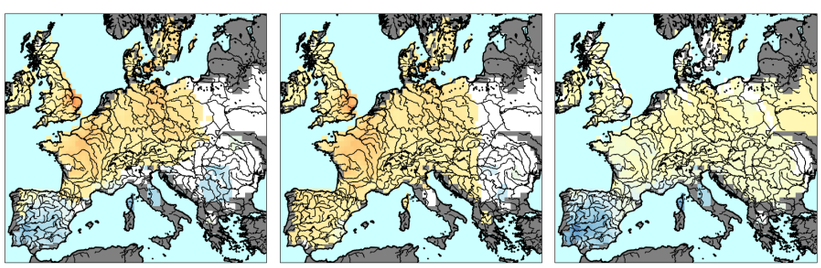
(a) Relative trends $\frac{dQ/\overline{Q}}{dt}$: ref **(b)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: ref **(c)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: ref



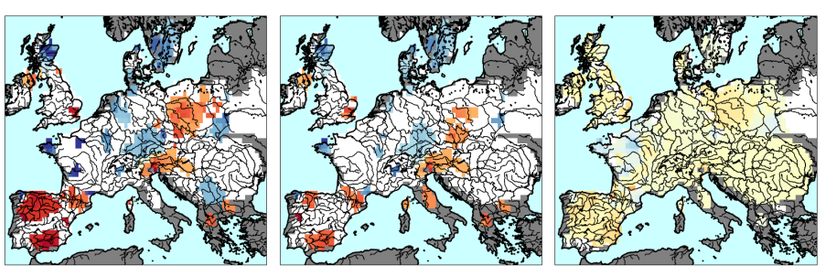
(d) Relative trends $\frac{dQ/\overline{Q}}{dt}$: f2000 **(e)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: f2000 **(f)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: f2000



(g) Relative trends $\frac{dQ/\overline{Q}}{dt}$: cstmean **(h)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: cstmean **(i)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: cstmean



(j) Relative trends $\frac{dQ/\overline{Q}}{dt}$: cstintravar **(k)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: cstintravar **(l)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: cstintravar



(m) Colorbar: significant relative trends (% of change per year over the century)

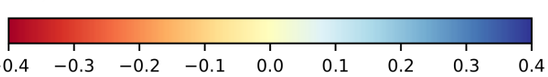
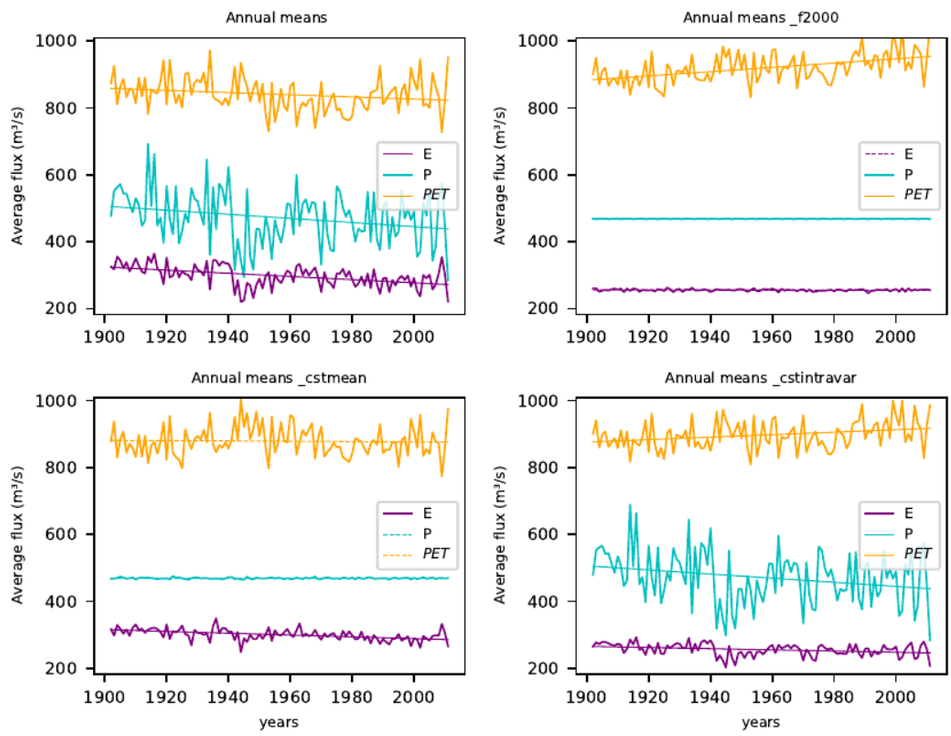


Figure 3.17 – Decomposition of significant relative annual mean discharge Q trends (% of change per year over the century) for all the tested forcings: the first line is the reference forcing. The first column is the total change in Q, the second is the partial change due to trends in the annual average of P and PET, and the last column is the partial change due to changes in the evaporation efficiency, mostly due to trends in the intra-annual distribution of P and PET. For the modified forcings: $f2000$ has the annual average and intra-annual distribution of P fixed for every year to their value for the year 2000. *cstmean* has only the annual average of P fixed. *cstintravar* has only the intra-annual distribution of P fixed. White areas don't have significant trends.

Tiber River:
Roma

(a) Annual average of climate variables P (light blue line), PET (yellow line) and E (purple line) modeled with the LSM for each academic forcing. Not shown here, the intra-annual distribution of P has been fixed for the forcings *f2000* and *cstintravar*.



(b) Annual mean discharge estimated with Budyko and their respective trends: full annual mean discharge modeled $Q = f(P, PET, \omega_t)$ (blue line), annual mean discharge $Q_c = f(C, \overline{\omega})$ (purple line) with only variations in C accounted for, and $Q_\omega = f(P_{rand}, PET_{rand}, \omega_t)$ (black line) with only the variations of evaporation efficiency, mostly due to the intra-annual covariance of P and PET, accounted for. Unsignificant trends are dashed. Here the results for each academic forcing are shown. We represented the normalized anomaly of discharge $((Q - Q_{mean})/Q_{mean})$ in order to better compare the plots to each other. The scale of the y-axis changes and is divided by 5 for the focings *f2000* and *cstmean*.

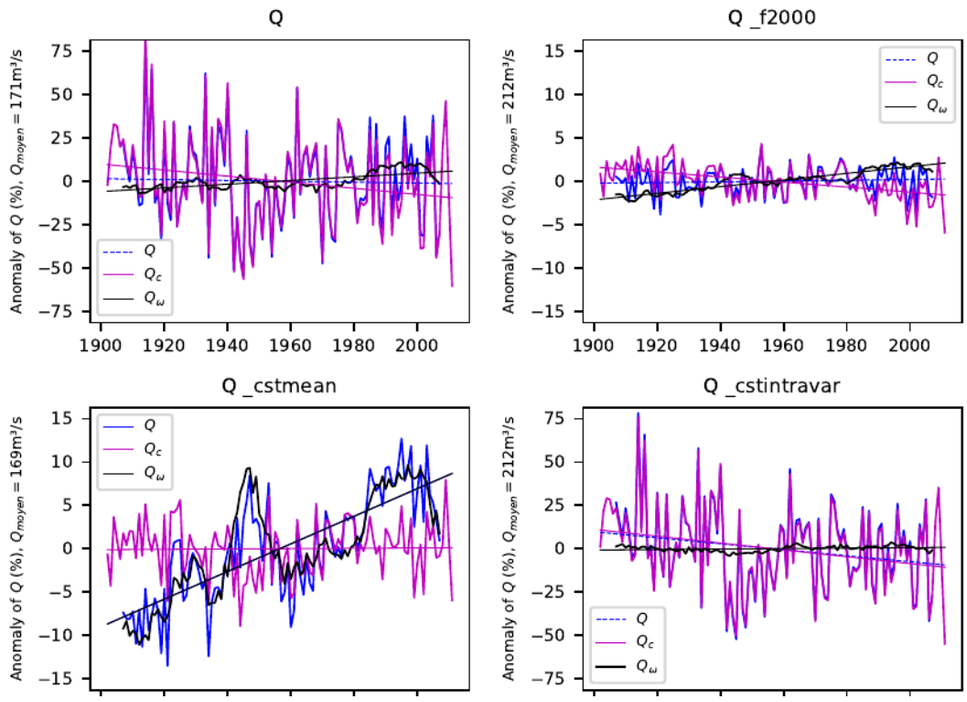
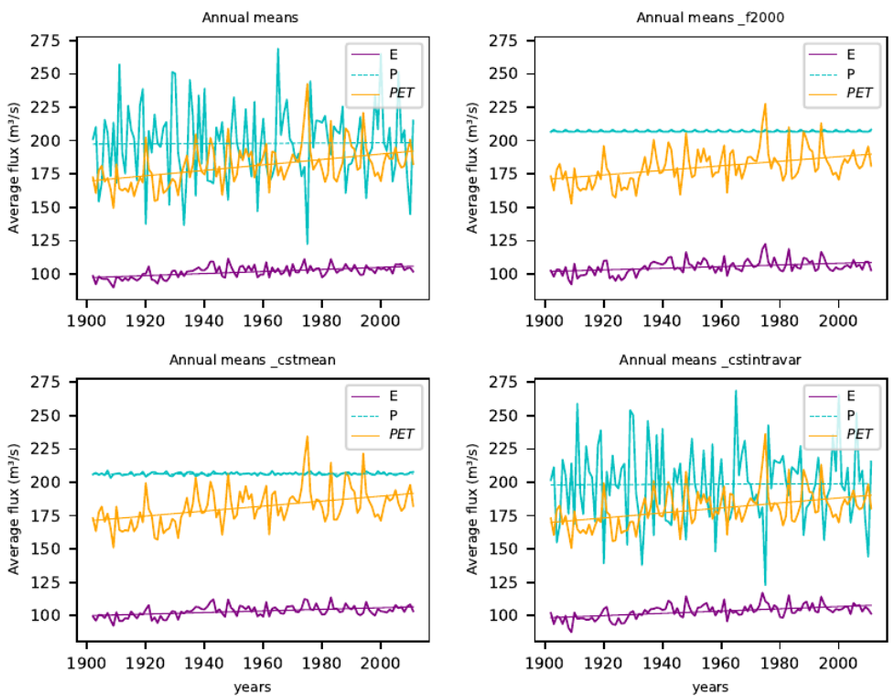


Figure 3.18 – Example 1: Time series obtained through the full application of our methodology for a given basin in Italy.

Trent River:
Colwick

(a) Annual average of climate variables P (light blue line), PET (yellow line) and E (purple line) modeled with the LSM for each academic forcing. Not shown here, the intra-annual distribution of P has been fixed for the forcings *f2000* and *cstintravar*.



(b) Annual mean discharge estimated with Budyko and their respective trends: full annual mean discharge modeled $Q = f(P, PET, \omega_t)$ (blue line), annual mean discharge $Q_c = f(C, \overline{\omega})$ (purple line) with only variations in C accounted for, and $Q_\omega = f(P_{rand}, PET_{rand}, \omega_t)$ (black line) with only the variations of evaporation efficiency, mostly due to the intra-annual covariance of P and PET, accounted for. Unsignificant trends are dashed. Here the results for each academic forcing are shown. We represented the normalized anomaly of discharge $((Q - Q_{mean})/Q_{mean})$ in order to better compare the plots to each other. The scale of the y-axis changes and is divided by 5 for the focings *f2000* and *cstmean*.

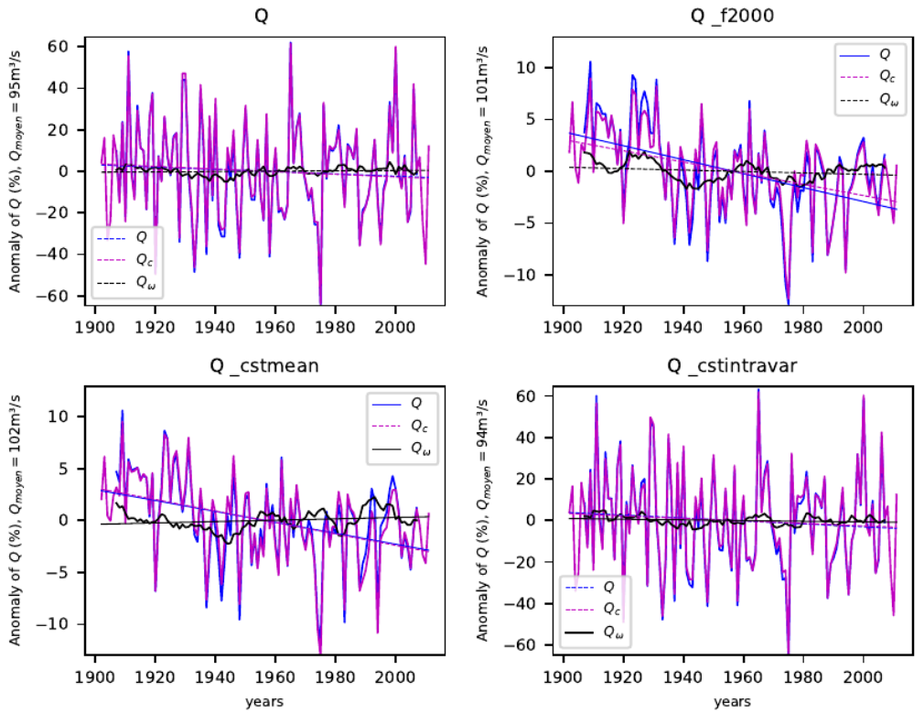


Figure 3.19 – Example 2: Time series obtained through the full application of our methodology for a given basin in England.

in the covariance of P and PET are minor compared to the effects of changes in the annual mean of climate variables C in most of Europe.

However, in some areas, the effects of the intra-annual distribution of P should be addressed. If we look at the Tiber river in Italy (Fig. 3.18b), the trend in Q_c (purple line) is significant for both the reference forcing and the forcing *cstintravar*. However, the total annual discharge Q (blue line) trend is only significant for the forcing *cstintravar*. For the reference forcing, the decreasing trend in the discharge due to C (Q_c) is counteracted by the increasing trend due to changes in the evaporation efficiency (Q_ω), making the final trend in discharge Q insignificant.

More generally, over Europe, when we erase the inter-annual variability of C, we capture the effect of trends in the intra-annual distribution of P and PET, through changes in the evaporation efficiency, in Q_ω (Fig. 3.17c). It tends to increase discharge, especially in south-western Spain, Italy, and the west of France (+0.1% per year over the century). It corresponds to the increasing trend of the black line in Fig. 3.18b, top left graphs for the Tiber river. It has an opposite trend towards decreasing discharge in eastern Europe and has a relatively neutral effect in the rest of the continent (Fig. 3.17c and, in the example of the English basin, Fig. 3.19b, top left graph, black line). It amplifies the trends due to changes in annual averages of C over certain watersheds such as the Duero basin (north-western Spain, decrease in discharge), western France, and northern Germany. Indeed, we note a significant increase in discharge over certain watersheds where the effect of changes in C alone was insignificant. In other areas, such as the Tiber river in Italy, or in southern UK, the intra-annual variability of P and PET counteracts the effect of C, making the relative total Q trends lose their significance due to opposite signals. We note the decreasing trend is due to the evolution of C while the effect of the change in the intra-annual distribution of the climate variables tends to increase the discharge.

In order to investigate the impacts of intra-annual variations of P on discharge, we analyse the results of the synthetic forcing *f2000* and *cstmean* (respectively top right and bottom left Fig. 3.19b and 3.18b and maps 3.17d to 3.17f and 3.17g to 3.17i).

For the synthetic forcing $f2000$ (Fig. 3.17d to 3.17f), P have been entirely set for each year to P of the year 2000. Therefore, this only yields the trends due to changes in PET, both for changes due to annual climate variables and changes in the evaporation efficiency of the catchment. For the synthetic forcing *cstmean*, only the annual mean of P has been set. In this case, the trends are due to PET and changes in the intra-annual distribution of P.

For the forcing *f2000*, the effect of PET is towards a decrease in discharge over all of Europe (less than -0.1% to -0.2% per year over the century) (Fig. 3.17d). For both the chosen examples, the effect of PET (top right graphs) tends to decrease discharge (purple line, Q_c when P has been fixed). It is consistent with the significant increase in PET (Fig. 3.19a and 3.18a, top right). The effect of intra-annual variations of PET on changes in the evaporation efficiency (Fig. 3.17f and black lines, top right graph Fig. 3.19b and 3.18b) has the same order of magnitude, if not a little smaller (less than -0.1% per year over the century), than the effect of inter-annual change of the annual average of PET (Fig. 3.17e or purple line top right graph Fig. 3.19b and 3.18b). It tends to amplify the latter's effect, especially over western France and southern UK. It has a slightly opposite effect towards increasing trends in Q (less than +0.08% per year over the century) over the east of Europe, west of Spain, and for the Tiber river. The effect of changes in the annual mean of PET, in this specific case, is canceled in the total discharge (blue line) by the effect of the changes in the intra-annual distribution of PET captured in Q_ω (black line) (Fig. 3.18b).

For the forcing *cstmean*, we now add the effect of changes in the covariance of P and PET due to changes in the intra-annual distribution of P. Depending on the area, there are two different responses. The two basins chosen in the example each correspond to one type of response. In the case of the basin in England (Trent river), the results obtained for the forcing *cstmean* (Fig. 3.19b, bottom left) are very similar to the results obtained for *f2000* (Fig.3.19b, top right). This means that the effect is due to changes in intra-annual synchronicity of P and PET has little impact compared to the effect of the annual mean of PET over that particular basin. It matches the results over northern Europe, especially over France, Germany, and southern UK, where the trends in Q (Fig. 3.17g) are mainly driven by changes in the annual mean of PET (Fig. 3.17h). However, over the Tiber river in Italy, the results obtained for the forcing *cstmean* (Fig. 3.18b, bottom left) shows that the changes in the total annual mean discharge Q (blue line) match the changes due to the evolution of ω_t (Q_ω, black line). In this latter case, the effect of the intra-annual variations of P is dominant compared to the effect of changes in PET. This matches the results over southern Europe (Spain, Italy) where for the forcing *cstmean*, the trends in Q (Fig. 3.17g) are driven mainly by changes in the evaporation efficiency (Fig. 3.17i). This increase in discharge diverges from the trends due to changes in C in the area (reference forcing and forcing *f2000*, purple lines).

The discharge trends for both forcings, namely *f2000* and *cstmean*, are statistically significant across multiple watersheds, independent of the high inter-annual variability observed in the annual mean of P.

Trends are significant for 1883 basins with the forcing *f2000* and 1756 for the forcing *cstmean* against only 352 basins with significant trends in Q out of 2134 for the reference forcing. However, the magnitude of these trends is also quite smaller. Comparing the discharge obtained with the reference forcing shows that the main factor driving Q is the annual mean of P since the discharge trends look entirely different when free of its variations.

To sum up, the results obtained with the synthetic forcings, the annual mean of P is the first driver of change in the annual discharge over all of Europe. However, its high inter-annual variability tends to hide the trends in most areas. The second most important climatic driver of discharge change depends on the area. Over southern Europe (Italy, Spain), where water is the limiting factor to evapotranspiration, the second most important climatic factor driving discharge changes is the intra-annual distribution of P. Over the rest of Europe, where water is less limiting, the second most important factor driving discharge changes is the increasing PET.

3.4 Discussion

Similarly to the results of several studies (Abatzoglou & Ficklin, 2017; Li et al., 2022; Xing et al., 2018; Jaramillo et al., 2022; Reaver et al., 2022; Padrón et al., 2017), we find that the average Budyko curve model with a constant evaporation efficiency $\overline{\omega}$ does not capture climate-related changes in the watershed behavior impacting its evaporation efficiency. Even with constant hydrological land surface characteristics, most catchments do not follow their average curve over the past century. The deviation to the curve can significantly affect Q's long-term trends over the past century if we free our analysis from the high inter-annual variability of P. It is in accordance with the results of Reaver et al. (2022) criticizing the "catchment trajectory conjecture".

Parameter ω has no direct physical meaning but is a proxy to represent the evaporation efficiency of catchments (Berghuijs et al., 2020; Reaver et al., 2022). However, since it cannot

be expressed as a function of clearly defined factors, it is difficult to attribute the changes in the evaporation efficiency to specific climatic features (Berghuijs et al., 2020). Yang et al. (2018) assume that further reductions in Q declining trends due to changes in catchment properties are likely associated with elevated atmospheric CO2 concentration or increased rainfall intensity. Other studies find correlations between changes in the evaporation efficiency of catchments and storm depth, the portion of precipitation such as snow (Donohue et al., 2012; Padrón et al., 2017; Xing et al., 2018). Using the outputs of an LSM, our studies allow to test a selection of hypotheses by adjusting climate parameters. We find that the climatic deviation to the average Budyko curve over the historical records is mainly due to variations in the intra-annual distribution of P.

We introduce a time-varying window to fit the parameter of Fu's equation in order to capture trends in the deviation to the average curve in the Budyko space. The choice of window size determines the size of accounted for trends. This functions as a frequency filter and only captures the effect of variations over periods the temporal scale of the window or larger. We must balance the length of our dataset and the appropriate length of the trends we choose to analyze. Since our aim is to investigate the effects of climate change, we do not need to capture the high inter-annual variability and can focus on decadal trends or longer (Yang et al., 2018). Furthermore, a shorter time window would not be adapted to the hypothesis of the Budyko framework which needs a long enough period to be fitted. So the window cannot be shorter (Tian et al., 2018). An exploration of an extended time window could be conducted to investigate the limited duration of time that captures the most significant impact on discharge. However, the longer the time window, the fewer points we will have to evaluate the trends.

In our methodology, we decompose the trends due to climatic changes in evaporation efficiency and the trends due to changes in average climate variables P and PET. One limitation in our decomposition method is that the variations in the evaporation efficiency captured in the deviation to the average curve are not entirely independent from the variations of average P and PET. Here, the relationship between P, PET, and evaporation efficiency is complex and relies on many interrelated factors (Reaver et al., 2020). We find in our study that the changes captured in the varying ω_t are mostly due to changes in the covariance of the intra-annual distribution of P and PET. However, the effect of the intra-annual distribution of climate variables on discharge is not completely independent from the annual mean of P and PET because of the difference in sensitivity of the system to a change in water availability. It can impact the magnitude of the identified trends. It is shown by the slight differences observed in Q_ω between respectively the reference forcing and the forcing *cstmean* (Fig. 3.17c and 3.17i) and between the forcing *f2000* and the forcing *cstintravar* (Fig. 3.17f and 3.17l). For each pair of forcings, the intra-annual distribution of P is the same, but the inter-annual mean of P differs. The difference in Q_ω for each pairing is due to a link between the annual mean and the intra-annual distribution of P. Therefore, the amplitude of the effect of the intra-annual distribution of P and PET quantified here may depend on the choice of the fixed average P (again, P from the year 2000 in this study). The observed differences were found to be comparatively insignificant in light of the identified trends, indicating that the fundamental findings regarding Europe would remain unchanged; therefore, we opted to disregard them. When studying specific basins, it could be interesting to choose specific pairings of intra-annual distributions/annual averages of P to construct synthetic forcings, to compare how specific associations combine.

Furthermore, we can't simply fix PET or its intra-annual variations in our synthetic forcings due to its non-linearity dependence on a number of climate variables. Therefore, we

are unable to decompose the effects of PET as easily as for the effects of P, which would be interesting to do, especially in the areas where P is less limiting, such as in western France or northern Europe.

Our methodology allows for the separation of the effect of primary and secondary climatic drivers on discharge trends. We look at the trends in P, and PET for the forcing GSWP3. Our results concur with those in the literature, validating that this forcing reasonably reproduces the climatic trends of the past century over Europe. The trends in PET are significantly (95% level) increasing over Europe. However, the trends in P are most often non-significant because of its high inter-annual variability, with a significant trend in the annual average of P for 413 catchments out of 2134 selected. The present study finds that the main driver of annual mean discharge Q (trends and inter-annual variability) is the annual mean of P. As expected with the increase in P over western Europe and the decrease in P observed in the Mediterranean area (Douville et al., 2021; Knutson & Zeng, 2018; Christidis & Stott, 2022), the trends in Q have followed the same direction. It concurs with the findings of Vicente-Serrano et al. (2019); Stahl et al. (2010) who find strong spatial consistency between streamflow changes and global rainfall changes.

Yang et al. (2008) show that Q is universally more sensitive to changes in P than to changes in PET, for a fixed land surface condition. Similarly, we find that over most of Europe, the second most important climatic factor on discharge changes is PET, which leads to a decrease in discharge due to the increasing evaporative demand by the atmosphere. Over the Iberian peninsula and the Mediterranean area, however, PET trends have a lesser impact. There, the water limit is the prevailing factor, having been attained by the end of spring and persisting throughout the entirety of summer. Therefore, a warmer summer does not have a strong impact. The evolution of intra-annual variations of P is the second most important factor impacting the changes in the annual discharge, with a higher effect on discharge than the increase of PET over the past century. The intra-annual covariance of P and PET impacts the annual behavior of the catchment and the annual balance between evapotranspiration and discharge since it changes the timing between water and energy available throughout the year. The evolution of the intra-annual cycle of P tends towards decreasing discharge in the Mediterranean area. It partially counteracts the effect of decreasing P and increasing PET on discharge. Therefore, the intra-annual distribution of P deserves more attention when studying the evolution of annual discharge. In most studies, it is only considered to look at changes in discharge peaks, floods, or droughts (Douville et al., 2021; Rottler et al., 2020; Milly et al., 2002; Douville et al., 2021; Vicente-Serrano et al., 2014; Tuel et al., 2022). We calculate the indices defined by García-Barrón et al. (2013) to evaluate the trends in the intra-annual cycle of P for the forcing GSWP3 (see 3.2.3). Similarly to the authors' findings, in Spain, we identify a shift over the end of the century towards a more bimodal distribution of precipitation throughout the year. However, the trends in the intra-annual cycle are mostly qualitative. The tendencies of the annual cycle to have an increasingly marked seasonality, concentrating rain events in fewer, but more extreme, events over the year, can explain the increasing runoff and relative discharge. Our methodology allows identifying these effects despite the only qualitative trends observed in the indices that measure the intra-annual distribution of P.

We apply our parametric model to LSM outputs to isolate the discharge variations due to changes in climate factors. This methodology relies on the capacity of the chosen LSM to reproduce the "natural" response of a catchment to climate, such as its behavior and response to changes in the intra-annual distribution of P. The amplitude of our results could depend on the choice of the LSM or the forcing data. We tested the use of other forcing datasets:

WFDEI (Weedon et al., 2014), which covers the period from 1979 to 2010, with the same resolution as GSWP3, and E2OFD (Beck et al., 2017), while also covering 1979 to 2010 but at a lower resolution. We also tested another model, SURFEX Quintana-Seguí et al. (2020), forced with SAFRAN (Quintana-Seguí et al., 2017), over the Ebro river (see chapter 4). This yielded similar results over the overlapping period with little differences in the trends' significance and amplitude. This indicates that the resolution of the forcing exerts a greater influence on the results compared to a specific forcing or model employed. This confirms the suitability of utilizing an LSM as a climatic reference in accordance with our methodology. In the future, when looking at specific basins, it would be interesting to use higher resolution forcings to obtain a more accurate picture of the effects of climate change on discharge. In this case, the diversity of behaviors exhibited among sub-basins within a given catchment could be elaborated upon by distinguishing the behavior of upstream sub-basins within mountainous regions from that of the downstream portion, which may display differential responses to climate change.

3.5 Conclusion

Our methodology combines a physical-based model to a parsimonious model. The first allows to identify the climatic changes in the empirical parameter of the second. The second allows for a simple decomposition of the relative changes in discharge. In this case, the Budyko framework and a one-parameter equation. The deviation from the average curve corresponds to a change in the evaporation efficiency of the catchment. A state-of-the-art LSM was used to simulate changes in the evaporation efficiency under the climate of the past century, independent from any other disruptive process. The successive fit of the parametric equation allows us to find the climatic dependence of the deviation to the average curve in the Budyko space over time. For a given catchment, we quantify its effect on annual mean discharge Q. Over the past century, the primary climatic source of deviation to the average curve is the change in the intra-annual distribution of P. We compare the impact of that deviation on changes in the average annual discharge compared to the change due to average climate variables P and PET. Over Europe for the past century, the main climatic driver of change in the average Q is the change in the average P. The second main driver of discharge change is PET over most of Europe except the Mediterranean area, where a change in the intra-annual distribution of P weighs more on Q changes than PET. Therefore, the effect of the intra-annual distribution of P should be addressed when studying the evolution of the average discharge and water availability under climate change, especially over the Mediterranean.

If we were to work from observations instead of model outputs, there would be other non-climate-related sources of variability, such as direct human activities or vegetation changes which would modify watershed behavior. Our next step is to apply the methodology to quantify these human-induced changes and compare their magnitude to those attributed to climate change in the present study's responses.

KEY POINTS TO REMEMBER

- Combining an LSM to a parsimonious model allows to assess the climate variability included in the calibrated parameter(s) of the latter.
- In the case of the one parameter Fu equation: the climate variability of the parameter is mainly due to the variability in the intra-annual distribution of P.
- The main climate driver of annual mean discharge over Europe is the annual average of P.
- The second most important climatic driver is dependent on the area: it is the annual average of PET over most of Europe; the intra-annual distribution of P over the Mediterranean area.

CHAPTER 4

Separate the effect of climate from the effect of non climatic drivers

This chapter is the subject of an article in progress. It is mainly composed of what will be part of the article, with some added details in the method part and complementary hypotheses and results. The methodology is the same as the one presented in chapter 3. This methodology was actually first conceived for the case presented in this chapter. We then thought that it should be applied to further study the effects of different climatic drivers and to test which part of climatic fluctuations were effectively covered, as it was explained in the chapter 3. It shows that we cover most of climatic fluctuations at intra- and inter-annual time scale. Here it will focus on comparing the effects of climate drivers to non-climatic drivers.

Contents

4.1 Introduction

As we described in introduction and in the chapter 2, changes in climate and catchment hydrology are concurrent. They both impact the partitioning of water between the different components of the water cycle, and therefore, discharge (Troch et al., 2015; Rottler et al., 2020; Ficklin et al., 2018; Fan et al., 2019).

To understand and interpret discharge trends, we need to be able to separate the effect of the different types of drivers. As previously described (chapter 2), some studies try to do so from discharge observation, by isolating catchments considered as "natural" and relatively untouched by anthropogenic water management (Stahl et al., 2010; Ficklin et al., 2018; Coch & Mediero, 2016; Beguería et al., 2003). However these studies are limited, especially in spatial extend due to the limited catchments corresponding to such a definition (Vicente-Serrano et al., 2019; Stahl et al., 2010). Therefore there is a need to use models to better understand and attribute the role of climate fluctuations and human activities, land surface and land use changes on discharge.

Physical-based models are very useful to attribute patterns of change and trends to specific processes (Douville et al., 2021; Zanardo et al., 2012; Alkama et al., 2010; Do et al., 2020). However, to this day, they fail to effectively include most of anthropogenic land surface changes and water management, even if progress are made in that direction (Wang et al., 2018; Nazemi & Wheater, 2015a). Parsimonious models on the other hands, are more suitable tools to represent and detect accurate trends in discharge (Perrin et al., 2003). However the physical interpretation of their output is more difficult, which limits their use for extrapolations and projections. Different methods exists relying on such models, to separate the effect of climate variable from the effect of anthropogenic activities. These methods mostly rely on two hypotheses. One common hypothesis is to identify a reference period or area where the effect of anthropogenic activities is considered negligible (Ahn & Merwade, 2014; Jiang et al., 2015; Wang et al., 2020; Fang et al., 2008). Similarly as for observation-based methods, this is limited by the choice of reference. Another current hypothesis is to calculate a sensitivity of discharge to specific factors, assuming that these sensitivities are independent from each other and can be extrapolated over time and space (Andréassian et al., 2016a; Roderick & Farquhar, 2011; Zhang et al., 2023). In each case, the effect of human activities is studied through the residuals changes, not attributable to climate factors (Vicente-Serrano et al., 2019; Jiang et al., 2015). Both hypotheses are disputable, mostly because of the empiricism of the calibrated parameter(s), with no well-defined physical meaning (Andréassian et al., 2016a; Milliman et al., 2008; Coron et al., 2014; Reaver et al., 2022), which can encompass effects of many interrelated factors not well understood and still include a sensitivity to climate variability.

With the method developed in chapter 3, we use the output of a physical-based land surface model as a "climatic" reference. It allows to get a reference free from anthropogenic activities and which covers the full area and period studied. Then using that reference, we attribute the climate effects encapsulated in the empirical parameter of a calibrated parsimonious model, in our case Fu's equation from the Budyko framework. Based on such a framework, we can then use the method to separate the effect of climate variability from the effects of non-climatic factors, without the limits of the previously presented methods. In this chapter, we use a comparison with the parsimonious model fitted on observations, to identify and quantify the changes in discharge not covered with the reference system, due to missing processes in the LSM.

The chapter is structured as follows: first it details more the existing methods based on

the Budyko framework and more generally parsimonious models, to separate the effects of climate factors from anthropogenic factors, with their associated limits. Then it explains how we apply our method to quantify the relative effects of climate change and other non-climatic factors, with the method, the data used and the validation tests. It is followed by the main results, comparing the changes in discharge in the "climatic" reference and the "actual" system. In secondary results, hypotheses are detailed on which non-climatic drivers are dominant over the different European regions, at different period of the past century. In the final section, we discuss the results and hypotheses raised with our methodology.

4.2 Existing method to separate the effect of climate change from the effect of anthropic drivers

In our method developed in chapter 3, we use a parsimonious model from the Budyko framework, the parametric equation of Fu. Here we focus on the existing methods using that same framework to separate the relative effect on discharge of climate change and anthropogenic factors, to illustrate the strengths and inherent limits of such methods.

4.2.1 Existing methodologies based on the Budyko framework

This framework was introduced with more details in the previous chapter 3. It relies on balancing the water and energy fluxes through only a few variables (precipitations P and potential evapotranspiration PET) to express the partitioning of water between evapotranspiration E and runoff. It needs to be applied to a closed system where the boundaries can be defined, such as a watersheds at an equilibrium/stationary state (the variation of water storage within the catchment is supposed to be very small) and the water balance equation can be simplified (Zhang et al., 2008) as such:

$$P - E = Q \tag{4.1}$$

Due to that stationarity hypothesis, we estimate an apparent E, which can differ from the effective E if that hypothesis is not justified.

Here we focus on parametric equations developed to describe that framework, reducing for each catchment its evaporation efficiency to a single specific parameter ω (equation 3.2). For the same climatic conditions P, PET, a catchment with a higher ω (Fig. 4.1 red curve) will evaporate more than another one with a smaller ω (Fig. 4.1 blue curve). In the original framework, this parameter is constant since the watershed is considered to be at an equilibrium state with constant characteristics.

Several studies have used the Budyko framework to try and separate the effects of climate change and the effect of human activities on the evolution of the discharge at the outlet of catchments. An added difficulty is that most often both types of changes happen over a similar time period and interact.

$$\Delta Q = \Delta Q_{climat} + \Delta Q_{human} \tag{4.2}$$

The two main methods used are the decomposition method (Jiang et al., 2015; Tian et al., 2018; Xiong et al., 2020; Zhang et al., 2019; Zhao et al., 2018) and the elasticity method (Jiang et al., 2015; Luo et al., 2020; Roderick et al., 2014; Roderick & Farquhar, 2011; Wang et al., 2020; Zhang et al., 2019; Zhao et al., 2018; Zheng et al., 2018). Most study divide the

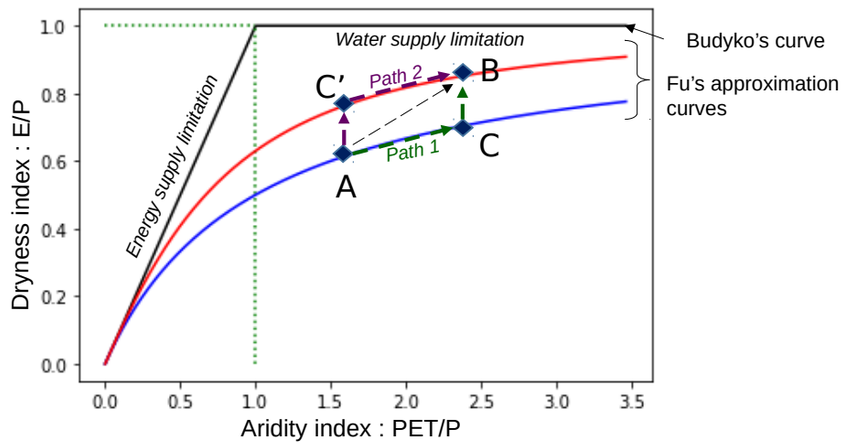


Figure 4.1 – Budyko framework: relationship between evapotranspirative ratio (E/P) and aridity index (PET/P) (Fu's equation). ν associated to the red curve is larger than ν associated to the blue curve and translate in a higher evaporation efficiency above the watershed. Points A, B, C and C' represents the steps used in the decomposition method for a changing path toward a drier climate and a more optimized evapotranspiration ratio.

period study in a pre- and post-change period, with the assumption that during each sub-period, the climatic conditions and the human activities are relatively stable (Jiang et al., 2015; Luo et al., 2020; Wang et al., 2020; Zhao et al., 2018; Zheng et al., 2018). Other use as a reference period, an area with close characteristics but considered as less impacted by human influences, such as Wang et al. (2020) who consider the source region of the studied catchment as a "natural" reference compared to downstream of the river.

4.2.1.1 The decomposition method

This method relies on the comparison between a pre-change (point A where $Q = f(P_1, PET_1, \omega_1)$) and a post-change time period (point B where $Q = f(P_2, PET_2, \omega_2)$) (Fig. 4.1). The equation is fitted twice, once over each period, giving two values of the watershed evaporation eficiency ω, pre- and post-changes.

In the original method, the change can be decomposed into two parts. The first part is a change due to a shift in climate variables, which is a change following the curve for a constant ω. On Fig. 4.1 it is either the part of the path 1 from point A to point C (where $Q = f(P_2, PET_2, \omega_1)$) or the equivalent on path 2, from point C' (where $Q = f(P_1, PET_1, \omega_2)$) to point B. The second part of the change is the part due to a change in the evaporation efficiency ω (vertical shift, either from point A to point C' or from point C to point B) (Tian et al., 2018; Zhao et al., 2018).

The hypothesis of this methodology is that the first part of the change is attributed to climate change while all changes in the catchment characteristics reflected by ω are attributed to human interference. This relies on the evaporation efficiency being climate invariant and it does not account for the modification of ω due to climate variability.

Different paths can be followed from point A to point B. Path 1 and path 2 (Fig. 4.1) are only two examples, C and C' being fictional steps. Some studies (Xiong et al., 2020; Zheng et al., 2018) used the complementary budyko model (Zhou et al., 2015) to introduce a weighting factor to account for the probabilities of all different paths (all included between path 1 and path 2, Fig. 4.1). Depending on the value of this weighting factor, it introduces uncertainty bounds framing the climate change term and the human induced change term, the maximum of one corresponding to the minimum of the other and vice versa.

4.2.1.2 The elasticity method

This method relies on the first order partial derivatives of Q (4.3). Inspired from the "elasticity" concept, an analytical expression is derived from it with sensitivity coefficients of the relative changes of Q to those of P, PET and ω (4.4) (Luo et al., 2020; Roderick & Farquhar, 2011).

$$dQ = \frac{\partial Q}{\partial P}dP + \frac{\partial Q}{\partial PET}dPET + \frac{\partial Q}{\partial \omega}d\omega \tag{4.3}$$

$$\frac{dQ}{Q} = \varepsilon_p \frac{dP}{P} + \varepsilon_{PET}\frac{dPET}{PET} + \varepsilon_\omega \frac{d\omega}{\omega} \tag{4.4}$$

The hypothesis is that the climate change contribution can be estimated from the sensitivity coefficient to P and PET.

Therefore the hypothesis here is that:

$$dQ_{climat} = \frac{\partial Q}{\partial P}dP + \frac{\partial Q}{\partial PET}dPET$$

With the original framework, the sensitivity to ω can't be directly assessed. It is associated to the human induced changes:

$$dQ_{human} = \frac{\partial Q}{\partial \omega}d\omega$$

and is deduced from observations by the residual term in equation (4.2) once ΔQ_{climat} has been calculated. For this method too, all changes in the catchment evaporation efficiency are attributed to human activities only.

4.2.1.3 Decomposing the variation of ω

Both methods in their standard definition assume that the changes in evaporation efficiency ω of the catchment are independent from climate (Zhao et al., 2018; Roderick & Farquhar, 2011). However as discussed in the previous chapter 3, this hypothesis is questionable. Longterm changes in climate should have an impact on evaporation efficiency ω. For instance changes in the seasonality of precipitation changes the synchronization between energy demand and water availability and can also alter soil and vegetation characteristics, therefore changing the evaporative dynamic of catchments (Xing et al., 2018; Ning et al., 2019; Li et al., 2022). In the previous chapter, we show that the parameter ω is influenced by the intra-annual distribution of precipitation.

Some studies (Jiang et al., 2015; Tian et al., 2018; Xing et al., 2018; Zhang et al., 2019) adapted the methodologies described above by using the expression of the evaporation efficiency ω as a function of different time varying factors: $\omega = f(factors_{climate}, factors_{human})$. Most studies construct their expression of ω by testing the relationship between ω and different factors for different watersheds. They use different methods, mostly step-wise regressions (Jiang et al., 2015; Tian et al., 2018; Li et al., 2022; Abatzoglou & Ficklin, 2017) but also neural networks (Simons et al., 2020), spatial multivariate adaptive regression splines (Xing et al., 2018) to identify the pertinent factors through a range of factors tested, available over the selected watersheds.

All these methods are however contested (see chapter 3). The evaporation efficiency parameter can't be so simplistically related to independent factors. It can not clearly relate to physical processes, and it is therefore not possible to construct a function that can be extrapolated in time (for predictions) or space (for unmonitored catchments) (Reaver et al.,

2022). These studies can only describe the particular cases where they were developed and show correlations which are highly depend on the area, the size of the catchments (Li et al., 2022), and the time scale studied (Ning et al., 2019).

In the end, across all studies, a wide variety of factors where tested and found to be correlated to ω accross catchments. Some factors are characteristic of the watersheds such as infiltration capacity (Tian et al., 2018; Li et al., 2022), relief (Ning et al., 2019). Other factors can change with the climate such as temperature (Jiang et al., 2015; Zhang et al., 2019), aridity index, drought index (Li et al., 2022; Xing et al., 2018), precipitation seasonality (Xing et al., 2018; Li et al., 2022; Ning et al., 2019). Finally factors representing the impact of human activities vary across studies from GDP per capita (Zhang et al., 2019) to ratio of irrigated area (Xing et al., 2018; Jiang et al., 2015).

Changes in the land use and vegetation cover (Xing et al., 2018; Xiong et al., 2020; Luo et al., 2020), vegetation fraction and routing depth (Gentine et al., 2012; Li et al., 2013; Ning et al., 2019), are shown to also impact ω and depending on the area and the study, are either attributed to climate change or to human activity (Tian et al., 2018; Xiong et al., 2020) or to both effects.

Over all, the correlations between changes in the evaporation efficiency and the different factors highly varies with time and space and doesn't allow to identify and relate to each other the different physical processes involved in a change of evaporation efficiency of catchments.

4.2.2 Generalization to all parsimonious models

The methodologies presented here can be apply with similar hypotheses to all kind of parsimonious models with few empirical parameters, as already introduced in chapter 2.

As long as the models are simple enough, their is possibility to express the elasticity or sensitivity of streamflow to the climate variables included in the model (Andréassian et al., 2016a; Vicente-Serrano et al., 2019; Gardner, 2009). The effect of human influences is considered in the residual variability of streamflow, dependent on the calibrated parameter(s). Most often, elasticities are considered to be constant over time, to quantify the effect of climate change as a function of the climate variables. This hypothesis is contestable, since the elasticity to different variables may not be independent from each other, and therefore not constant in time (Andréassian et al., 2016a).

Other studies use a period or area of reference to define the effect of climate variables on streamflow and compare the way the models fits over the human impacted system. The reference can be either the same area but in the pre-industrial era (Fenta et al., 2017; Ahn & Merwade, 2014), the "untouched" upstream of a catchment (Wang et al., 2020) or catchments in an area considered as similar as the studied catchment but with no strong human activities (Ficklin et al., 2018). In the end, the effect of human activities is associated to the difference in the fit of the models between both systems. Again the residuals associated to human interference are dependent on the calibrated parameter(s). These methods also allow to test how elasticity to climate variables reacts in time and space, associated to human influences.

All these methods rely on the assumption that the parameter(s) are independent of climate variability and change. If so, then the fit of the model over the natural reference or the climate elasticity defined with the model can be extrapolated to other time or area. However, they all have the same common limit: since the parameter(s) are empirical, they have no well-defined physical meaning associated to understood processes and that hypothesis may not be valid (Coron et al., 2014; Nicolle et al., 2021). Studies show that time-varying parameters are better at reproducing changes and can correlate these changes to specific factors, with other

parsimonious models such as GR4J (Zeng et al., 2019), but these are not an attribution, just a correlation of concurring changes.

4.3 Comparing an "actual" system to a "climatic" reference

In chapter 3, we use a different approach to study the climate-dependence of the empirical parameter ω in the Budyko framework. We use a time-sliding window to successively fit ω, defining a time series for each catchment. We apply it to a physical-based Land surface model (LSM) outputs to quantify the impact of climate change on the variations of ω without having to express ω in a closed form. It allows to relate changes in the evaporation efficiency ω to the physical processes represented in the LSM. As the LSM does not include human water management or land use changes, the resulting evolution in ω can be assumed to be only depend on climate.

In this chapter, we propose to use this technique to go further and identify the relative impact of non climatic factors compared to the overall effect of climate change on the catchment evaporation efficiency and on the associated discharge.

4.3.1 Method

The methodology introduced in chapter 3 allows to separate, for a given system: the part of discharge changes due to direct changes in mean annual variables P and PET (later on we group this variables in our notation under the variable $C = (P, PET)$, infobox 7); and the part due to changes in the evaporation efficiency ω of each catchments.

We compare two systems (Fig. 4.2):

- **A "climatic" system**: the LSM stands as our "climatic" reference. It allows to identify the climatic part of evaporation efficiency changes ω_c. It gives us for this system:
  - $\Delta Q_{climat}(C, \omega_c)$: overall trend in Q in the "climatic" system.
  - $\Delta Q_c(C, \overline{\omega_c})$: partial trend due to C changes
  - $\Delta Q_c(C_{rand}, \omega_c)$: partial trend due to climatic changes in evaporation efficiency ω_c
- **An "actual" system**: we fit the framework to historical records of discharge. We successively fit the framework to discharge observations, getting another time series of the evaporation efficiency parameter ω_a.
  - $\Delta Q_{actual}(C, \omega_a)$: overall trend in Q in the "actual" system.
  - $\Delta Q_a(C, \overline{\omega_a})$: partial trend due to C changes
  - $\Delta Q_a(C_{rand}, \omega_a)$: partial trend due to changes in evaporation efficiency ω_a. In that case, these changes are due to all changes in the watershed characteristics and climatic responses.

We illustrate the differences in both systems over a given catchment, with an illustration of Q_{climat}, Q_{actual}, ω_c and ω_a in section 4.5.1, Fig. 4.11.

For both systems, the effect of C changes on discharge changes should be the same. We consider that the LSM accurately reproduces dynamic changes but may be off to reproduce

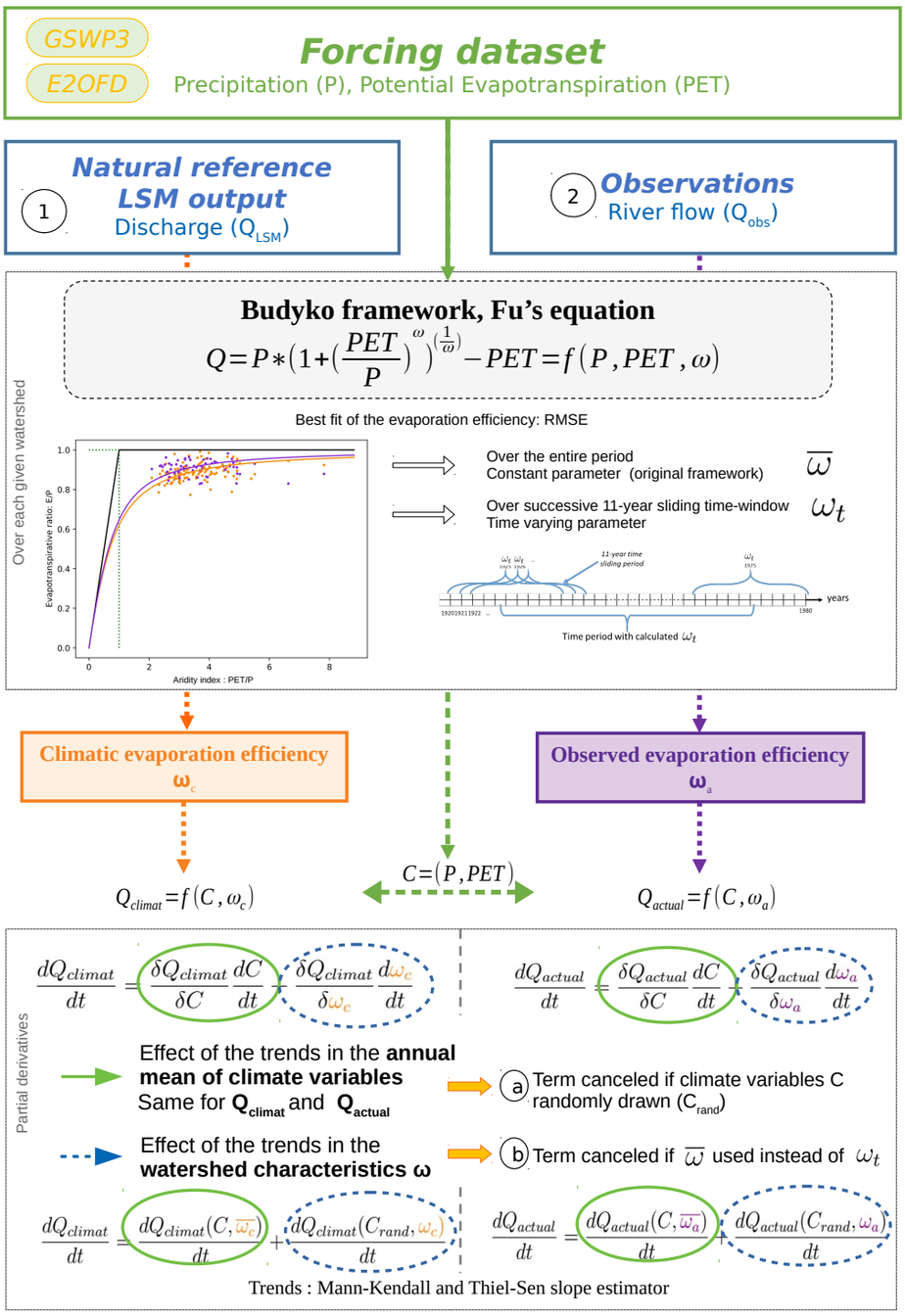


Figure 4.2 – Method to decompose the effect of climate drivers and non-climatic factors on the river discharge. The full methodology is applied twice: once to discharge output of a land surface model (Q_{LSM}) and once to discharge observations (Q_{obs}) over the same watersheds. The climate variables P and PET come from the same forcing dataset for both applications. The LSM reproduces the "climatic" behaviour of the watersheds and their response to changes in climate processes while the observations represent the behaviour with the effect of all factors impacting the watersheds' apparent evaporation efficiency. Fu's equation is fitted over both Q_{LSM} and Q_{obs}, once over the full period and once successively with an 11-year sliding window to get the evaporation efficiency. In the end we get the discharge modeled with the Budyko framework over the "climatic" reference watershed (Q_{climat}) and over the actual watershed (Q_{actual}). For each we then decompose the effect of annual mean of climate variables and the effect of changes in the evaporation efficiency by estimating the partial derivatives corresponding to both. The partial the response to changes in the evaporation efficiency only is estimated by randomly drawing annual mean of the climate variables (a) and the partial derivative due to changes in annual mean of climate variables is estimated by fixing the evaporation efficiency parameter over the full period (b).

INFO BOX 10

Discussion around notations

2- About the "climatic" and "actual" system

Our "climatic" system stands for the modeled reference system, which represents the natural functioning of catchments where the only source of change is due to the inputed near surface atmospheric variables. We hesitated with calling it the "natural" system but we wanted to avoid a possible confusion with the terminology used in observation-based methodologies, which refers to monitored real catchments, classified as not or very little impacted by anthropogenic activities.

We call "actual" system the one represented by our method based on observations. We did not call it the "observed" system since it is not a direct study of observed discharge (Q_{obs}) but it goes through a parametrization to get Q_{actual} and its associated partial trends.

absolute values of "climatic" discharge. Therefore we only compare trends in both systems. All trends are computed using the Mann-Kendall non-parametric test, associated to Thiel-Sen slope estimator (infobox 6). We consider significant trend in a 95% interval.

Table 4.1 – Description of the variables used

Variable	Description	Origin	Scale
P	Precipitations (annual averages)	From the climate dataset	Grid point/ Averaged over watershed
PET	Potential Evapotranspiration (annual averages)	From the Model ORCHIDEE, forced with the climate dataset	Grid point/ Averaged over watershed
Q_{LSM}	River discharge modeled (annual averages)	From the Model ORCHIDEE, forced with the climate dataset	One value per watershed
Q_{obs}	River discharge observations (annual averages)	From the river discharge observation database	One value per watershed
ω_c	Evaporation efficiency fitted over Q_{LSM}: "climatic" parameter	Calculated from equation (3.2) fitted over P, PET and Q_{LSM}	One value per watershed
ω_a	Evaporation efficiency fitted over Q_{obs}: "actual" parameter	Calculated from equation (3.2) fitted over P, PET and Q_{obs}	One value per watershed
Q_{climat}	River discharge modeled with the budyko framework fitted on Q_{LSM} (annual mean)	Calculated from equations (3.4) with P, PET and ω_c	One value per watershed
Q_{actual}	River discharge modeled with the budyko framework fitted on Q_{obs} (annual mean)	Calculated from equations (3.4) with P, PET and ω_a	One value per watershed

4.3.2 Data

4.3.2.1 Land Surface Model

The model used is still ORCHIDEE, described in the previous chapter 3.2.2.

4.3.2.2 Forcing datasets

We use three different datasets for climate reanalysis:

- GSWP3 (Hyungjun, 2017), which covers 1901 to 2010 at 0.5°x 0.5°resolution (more details in chapter 3).
- WFDEI-GPCC (Weedon et al., 2014). This dataset is another reanalysis covering 1979 to 2012 with a similar 0.5°x 0.5°grid, with either a daily or 3-hourly resolution. It was created using the same methodology as the WATCH Forcing Data (Weedon et al., 2011) but using ERA-Interim reanalysis data to replace ERA-40. It uses more surface observation and satellite data which improves the estimation of meteorological variables. With higher resolution, the initial ERA-Interim grid is also "closer" to the final 0.5°x 0.5°grid targeted, therefore limiting interpolation errors. Then a sequential elevation and a monthly bias correction are applied to the dataset, based on monthly gridded observations from CRU and GPCC for precipitations totals (Weedon et al., 2014).
- E2OFD (Beck et al., 2017) This dataset is another reanalysis covering 1979 to 2014. It is coupled with a global precipitation dataset: MSWEP, a 3-hourly 0.25°global gridded precipitation (1979–2015) (Beck et al., 2017). The final products covers from 1979 to 2014 with 0.25°spatial resolution, with either a daily or 3-hourly resolution.

These datasets are used as input in the LSM and to get tha variables needed in Fu's equation 3.2. P is taken directly while PET is calculate through the LSM based on these climatic data at the sub-annual scale (Barella-Ortiz et al., 2013). Testing the methodology with different independent climate datasets allows to verify the robustness of our results comparing the two systems and their sensitivity to the choice of climate forcing used.

We mainly focus on the GSWP3 dataset which runs over more than a hundred years to study the entire twentieth century. The other forcing datasets allow us to compare the results to those obtained with independent climate forcing data for the last part of the century. It tests the sensitivity of our results to the climate data used in input. E2OFD had a finer resolution which increased the variability of the results obtained using that forcing compared to the other two coarser climate datasets.

4.3.2.3 Watersheds and discharge observation datasets

The river discharge observations data used for the present study were the one collected by the Global Runoff Data Center (GRDC) from gauging stations all over Europe (GRDC, 2020). It was completed over Spain with data obtained from the Geoportal of Spain Ministerio (Ministerio para la Transición Ecológica y el Reto Demográfico, 2020) and over France with data from the database HYDRO (Ministere de l'ecologie, du developpement durable et de l'energie, 2021). The dataset therefore gathers data from 4475 stations in 25 countries over Europe (Austria, Bulgaria, Belgium, Switzerland, Czech Republic, Germany, Denmark,

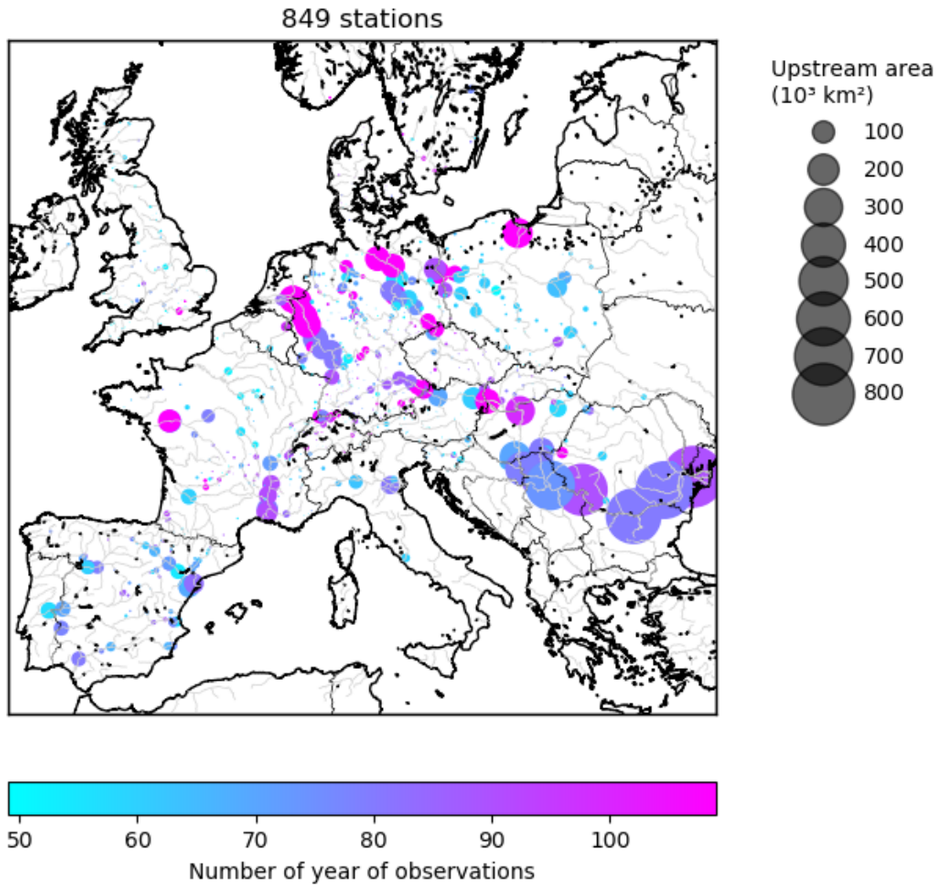


Figure 4.3 – Map of the dataset of observation stations over Europe. Each station has at least 50 years of river discharge observations over the time period 1900-2010 and a reconstructed basin. They are represented proportionally to the size of their upstream catchment area.

Spain, France, Great-Britain, Greece, Croatia, Hungary, Ireland, Italia, Netherlands, Norway, Poland, Portugal, Serbia, Romania, Sweden, Slovenia, Slovakia, Ukraine) with 1152 stations with at least 50 years of observations over the time period 1900-2010.

In the final analysis, we only keep the stations for which we were able to satisfyingly reproduce the upstream catchment in the hydrological routing of the LSM (Polcher et al., 2022; Nguyen-Quang et al., 2018), based on the dataset HydroSHEDS (Hydrological Data and Maps Based on Shuttle Elevation Derivatives at Multiple Scales) (Lehner et al., 2008) (infobox 2). In the final dataset, 3373 stations out of 4475 were placed on the grid, 849 out of 1152 stations with at least 50 years of observation data between 1900 and 2010.

We summed up P, PET over the watershed to get one average value per basin each year. To map the results and have a better visual for the results presented later on, all grid points of a basin are colored with the metric computed over the catchment. For sub-basins included into larger ones, the metric is averaged to color the shared grid points (see infobox 5).

ORCHIDEE outputs used in this and the previous chapter at the annual time step for each catchment are gathered in a file freely available on Zenodo.org (Collignan et al., 2023b), along with the list of the stations used and their main related metadata: their location and the size of the upstream area used to position the station on the grid.

4.3.3 Validation of the method

4.3.3.1 Validation of the Budyko framework to correctly reproduce river discharge

In this chapter, we want to compare the effect on annual mean discharge of climatic variations to the partial trends due to other sources, which are all playing a role in the observed trends (Q_{obs}). To better decompose the trends due to climate, we compare Q_{LSM} and Q_{obs} through the Budyko framework and Fu's equation. In chapter 3, we have shown that Fu's equation with a varying parameter can encapsulate the climatic tendencies in discharge, for modeled natural catchments with constant characteristics. In this case, Q_{LSM} from the LSM. We also need to validate that this framework is adequate to encapsulate most of the trends in Q_{obs}, which represent all conditions at once, the "actual" conditions. We therefore need to attest the quality of the budyko framework to reproduce both Q_{LSM} and Q_{obs} through their parametric representation Q_{climat} and Q_{actual}.

INFO BOX 11

Complements on performance indices

The Nash-Sutcliffe coefficient (NSC) is presented in infobox 8.

Percent Bias (PBIAS)

This index tests whether the model overestimates or underestimates the reference.

$$PBIAS = \frac{\sum_{i=0}^{years}(Ql_i - Qb_i)}{\sum_{i=0}^{years} Ql_i} * 100 \tag{4.5}$$

$$\begin{cases} \text{with } Ql = \text{reference discharge,} \\ \text{here dicharge from the LSM } Q_{LSM} \text{ or from observation } Q_{obs} \\ \text{and } Qb = \text{discharge to test, here result from the fit with Fu's equation} \end{cases}$$

A positive value indicates that the model tested (here the Budyko framework) underestimates the discharge compared to the reference (either Q_{LSM} or Q_{obs} in our case), and a negative values indicates that the model tends to overestimate the discharge.

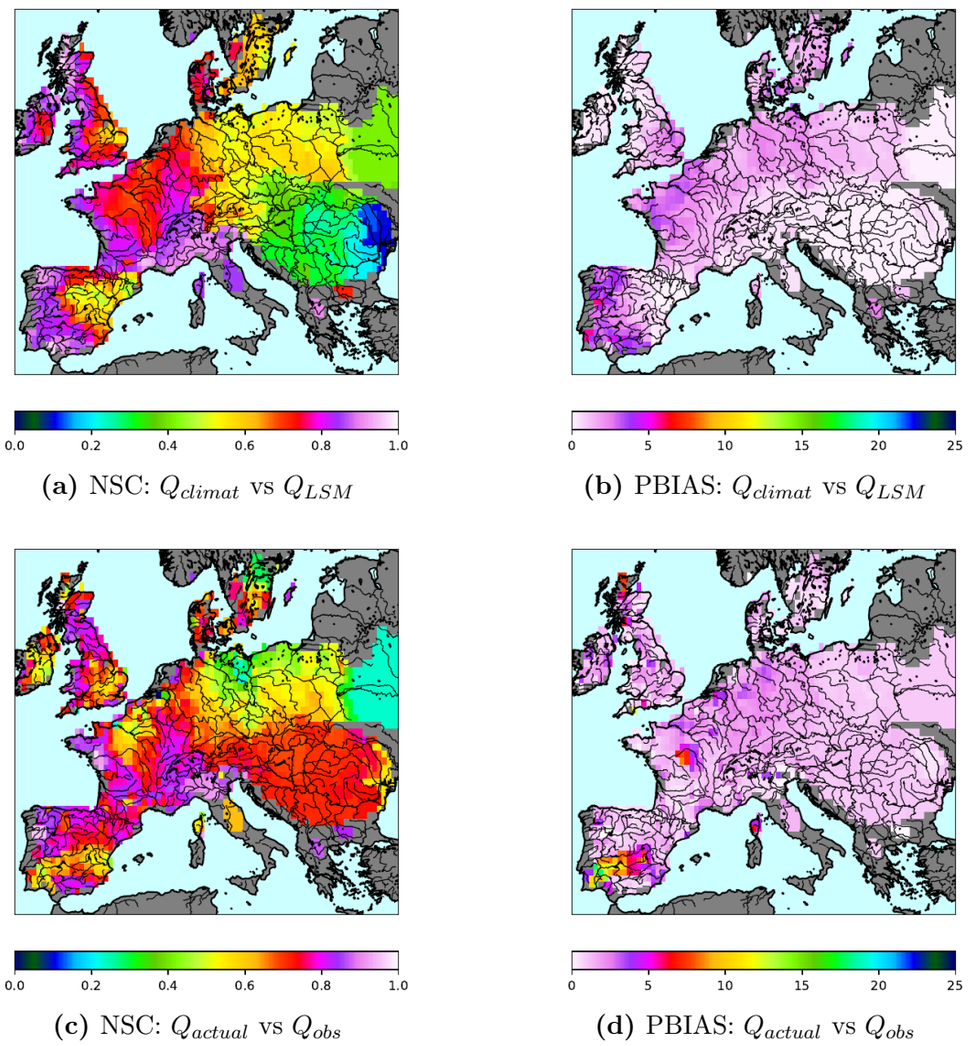
Performance ratings for discharge estimation

We need to compare the validity of the Budyko framework in river discharge estimation.

Table 4.2 – Significance of the test used (Nash-Sutcliffe coefficient (NSC) and Percent bias (PBIAS)) to estimate discharge for monthly time-step (Moriasi et al., 2007). For an annual time-step, the rating can be stricter.

Performance Rating	NSC	PBIAS
Very Good	$0.75 < \text{NSC} \leq 1$	$\text{PBIAS} < \pm 10$
Good	$0.65 < \text{NSC} \leq 0.75$	$\pm 10 \leq \text{PBIAS} < \pm 15$
Satisfactory	$0.50 < \text{NSC} \leq 0.65$	$\pm 15 \leq \text{PBIAS} < \pm 25$
Unsatisfactory	$\text{NSC} \leq 0.50$	$\text{PBIAS} \geq \pm 25$

We use the Nash-Sutcliffe coefficient (NSC) and the Percent bias (PBIAS) (Moriasi et al.,



(a) NSC: Q_{climat} vs Q_{LSM}

(b) PBIAS: Q_{climat} vs Q_{LSM}

(c) NSC: Q_{actual} vs Q_{obs}

(d) PBIAS: Q_{actual} vs Q_{obs}

Figure 4.4 – Using Nash-Sutcliffe coefficient (NSC) and absolute Percent bias (PBIAS) to compare river discharge modeled Q_{LSM} or observed Q_{obs} to river discharge Q_{climat} and Q_{actual} calculated with Fu's equation, to attest the quality of the Budyko framework. Colors from yellow to pink are considered as satisfactory (Tab. 4.2). The construction of the colormaps are detailed in infobox 5

2007) (Tab. 4.2) (Fig. 4.4) to test how good the Budyko fraemwork is at reproducing either the LSM or the observations. It is able to reproduce correctly the annual mean observed discharge over all European basins with a very good PBIAS (<10 % for all river basins) (Fig. 4.4d) and a good NSC > 0.5 for 569 stations out of 849, except for north-eastern Europe (Fig. 4.4c). This second test is more demanding and attests to the quality of Budyko framework to reproduce the inter-annual variations of the discharge. It is also efficient to reproduce the "climatic" river discharge from the model (Fig. 4.4a and 4.4b) with $NSC > 0.5$ and $PBIAS \leq 15\%$ apart for a few basins and still an under-performance for NSC over Eastern Europe. Therefore the Budyko framework is an adequate parametric representation of annual mean discharge.

Since the Budyko framework correctly reproduces the annual mean river discharge simulated with the model (Fig. 4.4a, 4.4b) and the annual mean river discharge from the observations (Fig. 4.4c, 4.4d), we can use Q_{climat} and Q_{actual} derived from this framework to compare the "climatic" behavior of the watershed and its "actual" behavior.

For the rest of the results, we filter the stations for which NSC < 0.5. We only keep the 569 stations for which the Budyko framework is efficient for both reproducing Q_{LSM} and Q_{obs}. Therefore the analysis when comparing Q_{climat} and Q_{total} will not be tinted by the ability of Budyko framework to effectively reproduce Q_{LSM} and Q_{obs} respectively.

4.3.3.2 Robustness of the method

- **Sensitivity to the climate forcing**

Table 4.3 – Comparison of the evaporation efficiencies time-series calculated with the different forcings for each system: ω_c for the "climatic" system and ω_a for the "actual" system.

Average variance over all catchments

	ω_c	ω_a
GSWP3	0.0023	0.039
WFDEI	0.0033	0.036
E2OFD	0.0110	0.031

Correlations:

% of stations with average correlation > 0.6		and median correlation between all catchments		
		ω_c		ω_a
E2OFD/GSWP3	38%	0.50	53%	0.65
WFDEI/GSWP3	73%	0.75	77%	0.99
E2OFD/WFDEI	64%	0.70	59%	0.73

Over the 1979-2010 period, when three different independent atmospheric datasets are available, we verified the sensitivity of the results to the input data. We applied the methodology using the different forcing datasets. We want to attest if the "climatic" trends captured with our method are similar for all forcings used.

Over such a short period, trends are mostly non significant and can't be appropriately statistically compared. However looking at their pattern, they are very similar for all forcings considered. Here we focus on the efficiency parameters ω_c and ω_a correlation and variance for each forcing, to analyze the impact of the forcing choice on how our method attributes variations of ω to "climatic" behaviors with our LSM of reference.

Comparing ω_c and ω_a obtained for the three forcings over the common period and for each system, we obtain very similar results when looking at the average variance over all basins for each evapotranspiration efficiencies time-series and the two-by-two correlations (Tab. 4.3).

The variances have a similar order of magnitude no matter the forcing used to calculate ω_c and ω_a with a larger ω_a by a factor of ten with all forcings. E2OFD had a finer resolution which increases the variability of the results relative to the other two coarser climate datasets. Given the limited number of observations, the forcing datasets are not fully independent. For instance GSWP3 and WFDEI use the same precipitation product to bias correct the re-analyses on which they are based. E2OFD and WFDEI use the same re-analysis but interpolated to different resolutions and corrected with two distinct observational precipitation estimates. Given that the results are closer for GSWP3 and WFDEI, we can hypothesize that the method is more sensitive to the precipitation data used than the other variables.

These results show globally that the method is robust since it is not very sensitive to the forcing used. The differences in variance between forcings are smaller than the differences between the variance of ω_a and ω_c for all forcings tested. The poorest correlation is between E2OFD and GSWP3 (the most different forcings from one another) and mostly for ω_c which has the smallest average variance. Therefore it will impact our results less when comparing trends. However the absolute values of ω are significantly different depending on the forcing used, comforting the idea that this method can only be used to assess and compare trends.

- **Sensitivity to the Land Surface Model**

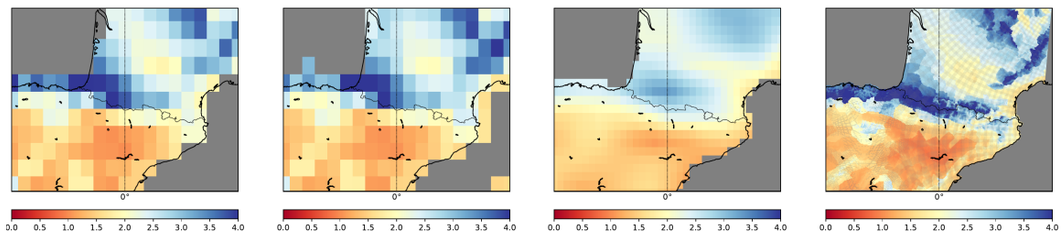
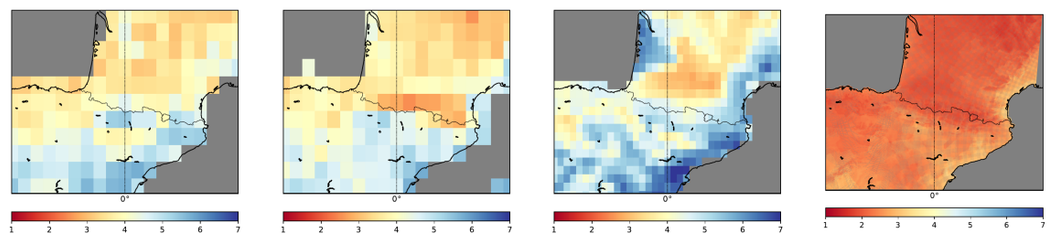
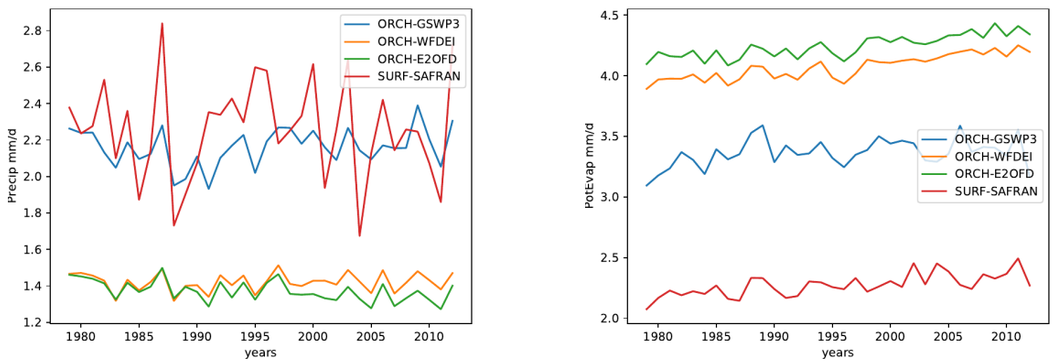
Variable	ORCHIDEE GSWP3	ORCHIDEE WFDEI-GPCC	ORCHIDEE E2OFD	SURFEX SAFRAN
Average P (mm/ day)				
Average PET (mm/ day)				

Figure 4.5 – Comparison of climate variables P and PET for the different forcings over the Ebro area, 1979-2013. The scale are different for all forcings. WFDEI-GPCC and E2OFD have similar order of magnitude. GSWP3 and SAFRAN tend to larger fluxes on average.



(a) Precipitation (mm/day) **(b)** Potential evapotranspiration (mm/day)

Figure 4.6 – Average annual climate variable (P and PET), integrated over the Ebro area, for the different Model/Forcing: ORCH/GSWP3, ORCH/WFDEI, ORCH/E2OFD, SURF/SAFRAN.

We tested different forcings, therefore the sensitivity of our method to the reference choice of climate variables. Here we go one step further to add a comparison when a different LSM is used. Different forcings imply differences in P and PET through the climate variables needed to calculate it. However in the case of a different LSM, it implies a difference in how the catchment behaviors are modeled and in PET due to how it was calculated in the different models. We had access to a different the LSM, SURFEX Quintana-Seguí et al. (2020), forced with SAFRAN (Quintana-Seguí et al., 2017) from 1979 to 2013, over the Ebro river.

Figure 4.5 shows the average climatic fluxes over the area and time period for all forcings and both models tested. Figure 4.6 shows the same variables but this time their inter-annual variation, integrated over the full area. For P, the average is closer for E2OFD and WFDEI than for GSWP3 and SAFRAN, slightly more humid, and with a higher variability with SAFRAN (Fig. 4.6a). All show similar patterns (Fig. 4.5), with slightly lower contrast over the Pyrenees for ORCHIDEE forced with E2OFD. PET is the result of a calculation and therefore is highly sensitive to the method used to calculate it. It explains the major difference between ORCHIDEE and SURFEX in the average PET over the area, with an average PET 2.5 lower with SURFEX/SAFRAN (Fig. 4.5 and 4.6b). This shows that the main difference in PET is due to how the models calculate it. However, when looking at the

spatial trends for that variable, the pattern (not shown) and magnitudes (average slopes in Fig. 4.6b) of these trends are similar.

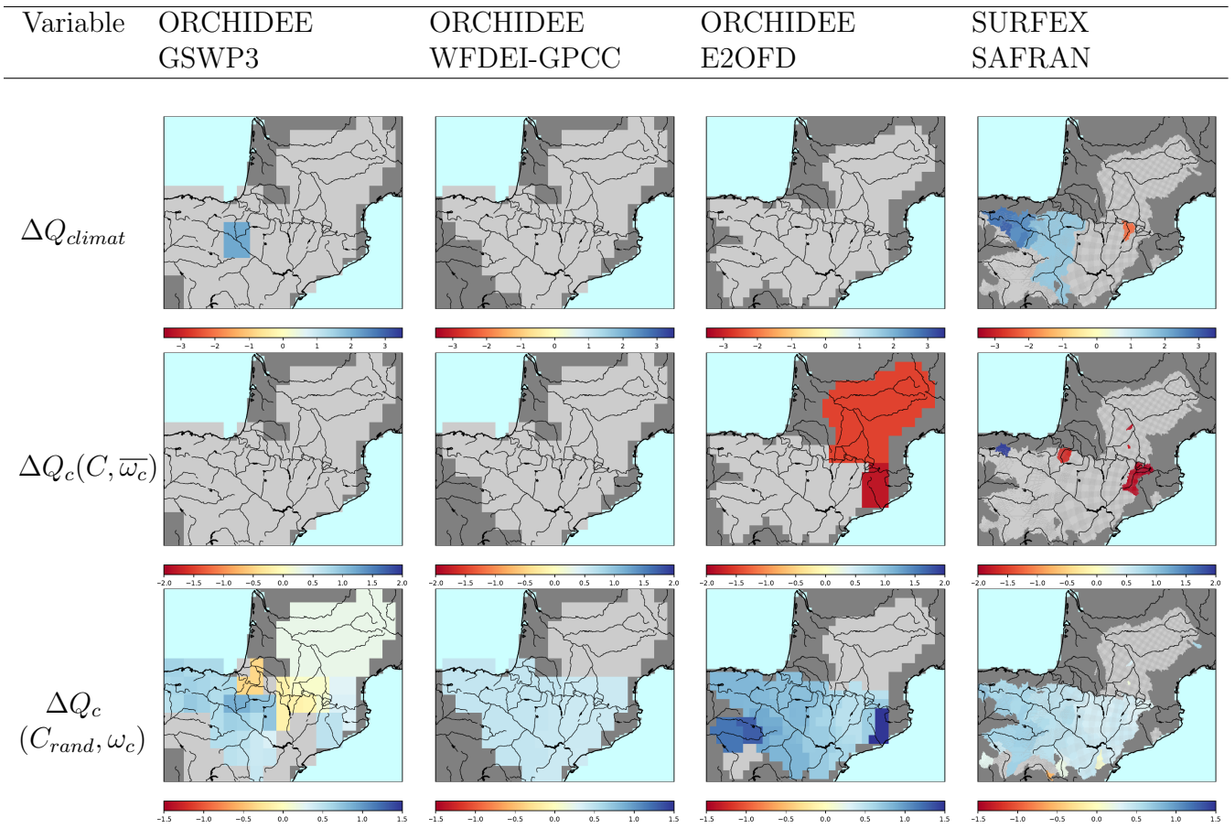


Figure 4.7 – Comparison of discharge relative trends (1979-2013) for the different forcings over the Ebro area for the "climatic" system.

This comparison only covers 30 years so most relative trends in discharge are not significant, due to the high climate variability (Fig. 4.7 and 4.8, first two rows). However, if we look at the isolated partial effect of change in evaporation efficiency ω for both systems ("climatic" and "actual"), we get similar order of magnitude and similar pattern for all forcings and both models tested (Fig. 4.7 and 4.8, third rows). This shows that our methods captures a similar effect of catchments behaviors for both LSM tested. Again here, the run with E2OFD is the one showing less similarities with the other forcings tested, probably because of its lowest precipitation average in the Pyrennees (Fig. 4.5). SURFEX forced with SAFRAN runs at higher resolution so we better separate the effects over small catchments, otherwise covered at broader resolution.

We conclude that the results of our method is not really sensitive to either the forcing or the LSM used since we get similar trends. This analysis could be further improved if we had access to other forcings covering longer periods and if we could run other LSM over broader areas.

Results presented in the following sections are obtained with the forcing dataset GSWP3 which covers the longest time period 1901-2012 and is thus most relevant for evaluating the long-term driver of river discharge trends.

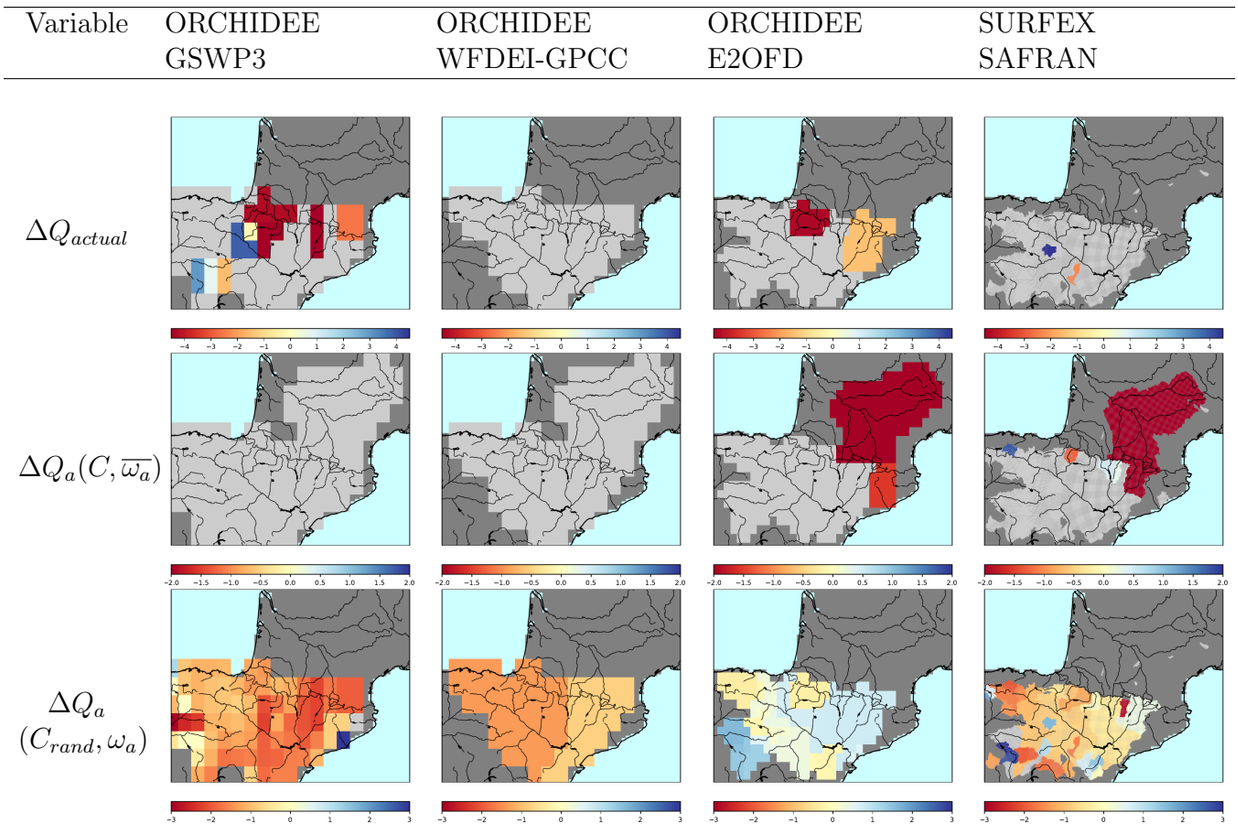


Figure 4.8 – Comparison of discharge relative trends (1979-2013) for the different forcings over the Ebro area for the "actual" system.

4.4 Decomposing the evolution of Q over the past century

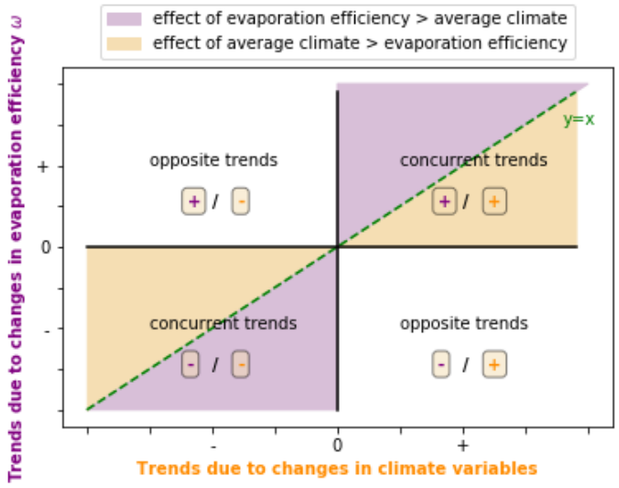
We now look at the decomposition of Q trends, in each system, the "climatic" system and the "actual" system. In each system, our methodology allows to separate the trend due to a change in the annual mean of P and PET from the changes due to an evolution of the evaporation efficiency. In the "climatic" system, the latter is mainly due to changes at the sub-annual time-scale in the distribution of climate variables (changes in the synchronization between the energy demand and the water availability), as seen in chapter 3. In the actual system, the changes in evaporation efficiency also include all the effects of land surface changes and human water management.

4.4.1 General changes in Q and in its components for the "climatic" and the "actual" system

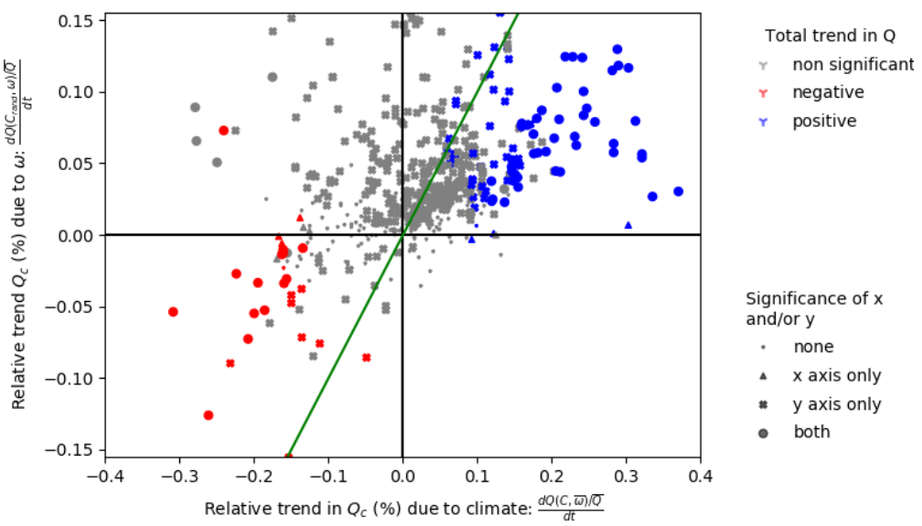
The relative trend in Q can be separated into a part explained by the evolution of climate (annual mean of P and PET) and another to changes in the catchment which also affect evaporation efficiency (see Fig. 4.2). Figure 4.9 summarizes and compare for each system, for all the catchments studied, the relative trends in Q and in these two components.

It is presented in the form of a graph, one graph for each system. The colored points represent the catchments for which the overall trend in Q is significant. The quadrant of the graphs represent different pattern of behavior in the different components of Q changes:

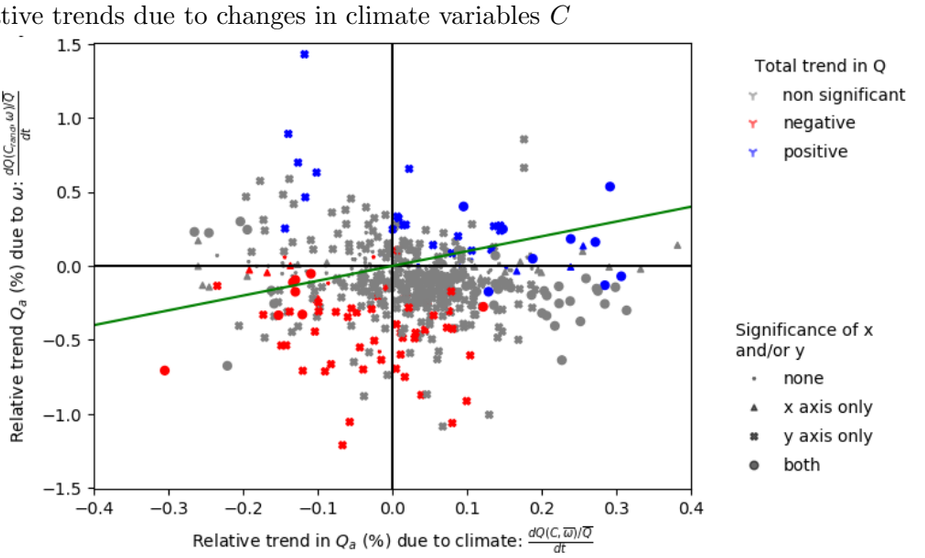
- The top-right and the bottom-left quadrants include the catchments for which both components of Q changes have a concurrent/complementary effect.
  - In the top-right quadrant, both of them tend to an increase in discharge.



(a) Interpretation scheme: comparing significant trends due to climate variables or due to ω. The graphs have four quadrants: the top right and the bottom left ones correspond to area of the graph where the climatic trend and the area due to ω are complementary/ going into the same direction. The top left one contains the basins for which the trends due to ω are positive and the trends due to climate variables are negative and the bottom right quadrant the opposite.



(b) "Climatic" system Q_{climat}: relative trends (% per year over the century) due to changes in ω_c versus relative trends due to changes in climate variables C



(c) "Actual" system Q_{actual}: relative trends (% per year over the century) due to changes in ω_a versus relative trends due to changes in climate variables C

Figure 4.9 – Comparing the relative trends ($\frac{dQ}{dt}/\overline{Q}$) due to a change in climate variables or due to a change in evaporation efficiency ω in the evolution of discharge, for both system Q_{climat} and Q_{actual} (detailed equations in Fig. 4.2). One point corresponds to a basin with at least 50 years of river discharge observations over Europe. The scale of trends due to ω_a in the "actual" system is ten times larger than the one for trends due to ω_c in the "natural" system. The green line is the line $y = x$. The color scale represents the significance of the trend in Q when all factors are considered. The markers stand for whether the partial trends, due to changes in C (x-axis), due to changes in ω (y-axis), or both, are significant.

  - In the bottom-left, both tend to a decrease in discharge.

- the two other quadrants contain catchment for wich the two components have opposite effects.

  - In the top-left, it contains catchments where the effect of the evolution of the annual mean of P and PET tend to a decrease of Q but where the decrease in the evaporation efficiency tend to counteract that effect.
  - In the bottom-right, both components counteract in the other direction, with the evolution of annual mean of climate variables tending to an increase in Q while the increase in the evaporation efficiency tends to a decrease in runoff and Q.

In the case of the "climatic" system (Fig. 4.9b), almost all catchments where Q has significantly changed have concurrent trends in both components. More generally, the partial trends due to changes in annual mean P and PET are larger than the one due to a change in evaporation efficiency.

There is a dominance of the effect of the annual mean in climate variables P, PET, amplified by the responses of catchments and changes in evaporation efficiency. These cases corresponds to catchments where an increase in P and/or a decrease in PET tends to an increase in Q, or inversely for a decrease. For instance, if an increase in P is also asymmetrical over the year, with an even stronger increase in winter precipitation, where the partitioning towards runoff is usually higher, it translates in a decreased evaporation efficiency and an even stronger increase of Q. Therefore in that example, the increase the average Q is not only due to an increase in annual mean P, but is amplified due to a more contrasted seasonality, captured in the changes of evaporation efficiency.

There are almost no catchments in the fourth quadrant: where an increase in Q due to changes in the annual mean of P and PET is associated to an increase in evaporation efficiency. More generally, there are fewer catchments in the lower half of the graph, meaning that the changes in the evaporation efficiency mostly tend to increase Q in the "climatic" system, except when they concur with a high decrease in Q due to a decrease in P and/or an increase in PET (bottom-left quadrant).

There are also catchments with no significant trends in the overall Q because of opposite trends in the two components (top-left quadrant). All these points are in the top-left quadrant, corresponding to a decrease in discharge due to changes in the annual mean of P and PET and evolution of the evaporation efficiency tending toward higher runoff and Q. It corresponds for instance to catchments where P decreases and/or PET increases, tending to lower Q, while increasing extreme precipitation events and/or winter precipitation ratio would decrease the evaporation efficiency and therefore tend to increase Q.

In the case of the "actual" system (Fig. 4.9c), the repartition of the catchments looks very different. Most catchments are this time in the lower half: contrary to the effects of climate alone, land surface changes and human water management tend to increase the evaporation efficiency of catchments and therefore decrease Q. This is coherent with activities such as irrigation which aims at optimizing the evapotranspiration over catchments.

Moreover, the partial trends due to changes in evaporation efficiency are larger than the partial trends due to annual mean in P and PET by a factor of 3. More generally, the overall trends (colored points) follow the partial trends due to changes in the catchments' evaporation efficiency. Therefore in the "actual" system, the trends in Q are mainly due to changes in catchments behavior due to non climatic factors.

4.4.2 Map of the general trends and of its components

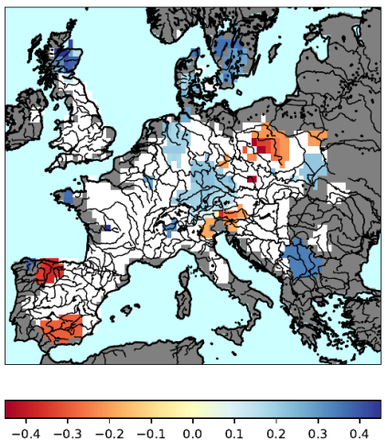
Figure 4.10 shows the same results with their spatial distribution, for the catchments where the trends are significant. These trends are spatially coherent, which also attests of the robustness of the method.

These results for the "climatic" system corresponds to the left side maps. The significant relative trends in Q_{climat} are shown in the top left map 4.10a. Some basins are getting dryer such as in Spain (-0.2% to >-0.4% per year over the past century) and in eastern Europe while other in center and northern Europe see an increase in their "climatic" discharge (+0.2% to <0.4% per year over the past century). Maps 4.10c and 4.10e show the two components of Q_{climat} changes: respectively the partial relative trends due to changes in average climate variables C and partial relative trends due to changes in evaporation efficiency ω_c. In that system, the trends in discharge Q_{climat} (Fig. 4.10a) mostly relates to the partial trend due to changes in average climate variables C (Fig. 4.10c).

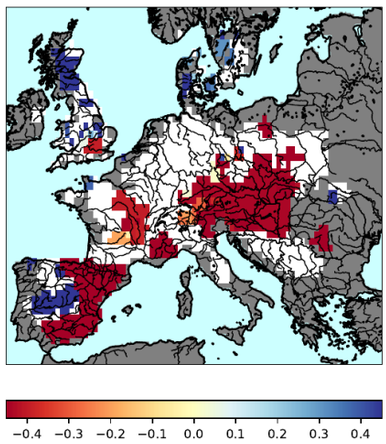
We also get significant partial trends due to changes in the evaporation efficiency in the "climatic" system (Fig. 4.10e). However these partial trends are small (-0.1% to +0.1% per year over the past century) and their effect is mostly hidden and non significant when looking at the total trends in discharge. It can however amplify the partial trend due to changes in C. It corresponds to the top-right and bottom-left quadrant in Fig. 4.9b, and to catchments such as for the Duero basin in north-western Spain where both partial trends concur to a decrease in Q. It can also cancel them out as for the Tiber river in Italy where the decrease in Q due to changes in C is not significant in the overall changes in Q_{climat}. This corresponds to the top-left quadrant Fig. 4.9b.

The results for the "actual" system corresponds to the right side maps in Fig. 4.10. In that system, we see that the discharge trends (Fig. 4.10b) mostly relates to the effect of changes in the evaporation efficiencies (Fig. 4.10f). Here the changes in the evaporation efficiency ω_a encompass all changes in catchment's evaporative behaviours. Similarly to the results of Vicente-Serrano et al. (2019) over western Europe, we find the highest negative trends over Southern Spain. The scales of the Fig. 4.10 are fixed for comparison purposes so for Fig. 4.10d and Fig. 4.10f the scales are saturated, not showing that the trends are a lot higher in Spain, with the highest absolute trends, below < -0.5 % change per year over the past century in Fig. 4.9c. The south of Spain corresponds to an area where both climate changes and changes related to human activities led to a significant decrease in river discharge over the past century. Over Spain, the Guadiana river stands out in our maps. It seems that over that catchment, the overall effect of human water management and land surface changes tend to increase Q, contrary as in the rest of Spain. However, mainly, the changes in evaporation efficiency result in decreasing trends in Q (Fig. 4.10f). Over the rest of Europe, trends are lower and less significant, with positive trends rather in northern Europe, Great Britain and Sweden and negative trends over the center of Europe. However not all basins where partial trends due to changes in evaporation efficiency are significant also have an overall significant trend in Q_{actual}. This goes for instance for western France, northern Germany, Serbia. For these areas, the trends induced by changes in the evaporation efficiency loose their significance when the climate variability is taken into account in the reconstructed discharge.

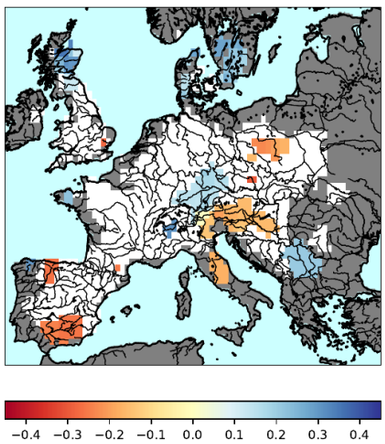
Interestingly when we draw similar maps for sub-periods of 10 years, the respective weight of the two components of discharge changes is inverted for the "actual" system. At decadal



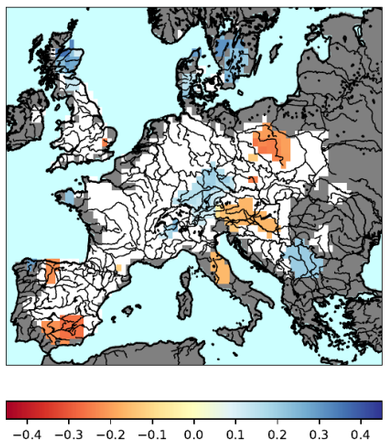
(a) Relative changes in Q_{climat}



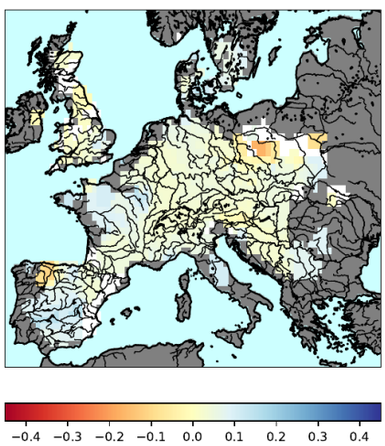
(b) Relative changes in Q_{actual}



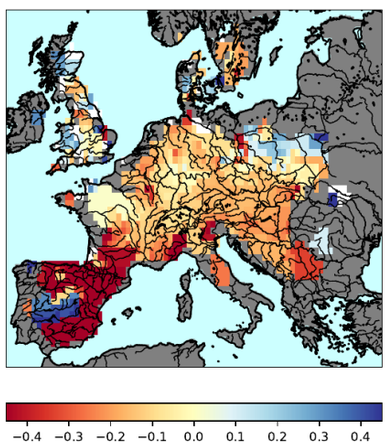
(c) Relative trend in $Q_c(C, \overline{\omega_c})$: trends in relative Q_{climat} changes due to climate variables C.



(d) Relative trend in $Q_a(C, \overline{\omega_a})$: trends in relative Q_{actual} changes due to climate variables C.



(e) Relative trends in $Q_c(C_{rand}, \omega_c(t))$: trends in relative Q_{climat} changes due to changes in the evaporation efficiency ω_c



(f) Relative trends in $Q_a(C_{rand}, \omega_a(t))$: trends in relative Q_{actual} changes due to changes in the evaporation efficiency ω_a.

Figure 4.10 – Significant trends in the relative annual mean river discharge $Q/\overline{Q}$ over the time period 1901-2012 (% per year over the century). A positive trends corresponds to a significant increase of the river discharge over the century while a negative trends corresponds to a significant decrease of the streamflow. The left side presents the relative trends in Q_{climat} with the associated partial trends due to changes in climate variables C or in evaporation efficiency ω_c. The right side presents the relative changes in Q_{actual} with the associated partial trends due to changes in climate variables C or in evaporation efficiency ω_a. The scales have been forced to be the same for all maps for comparison purposes but the extrema can go higher or lower. The construction of the colormaps are detailed in infobox 5.

scale, the climatic variability is high. This climatic noise covers the effect of changes in the catchment's evaporation efficiency in discharge trends. At the scale of the century, the signal-to-noise ratio is higher, allowing to bring to light the long-term effect of changes in catchment's evaporation efficiency and catchment's behavior on discharge.

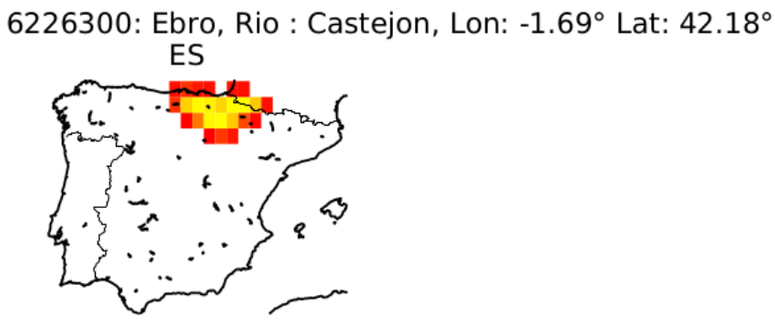
4.5 Evolution of the evaporation efficiency related to anthropic drivers and land surface changes

Due to a high climatic noise, the decadal trends in discharge are not significant over most catchments. However, we can extract trends in evaporation efficiency changes. This allows to look at shorter term changes in catchments characteristics and try to find possible correlations with specific potential drivers of change.

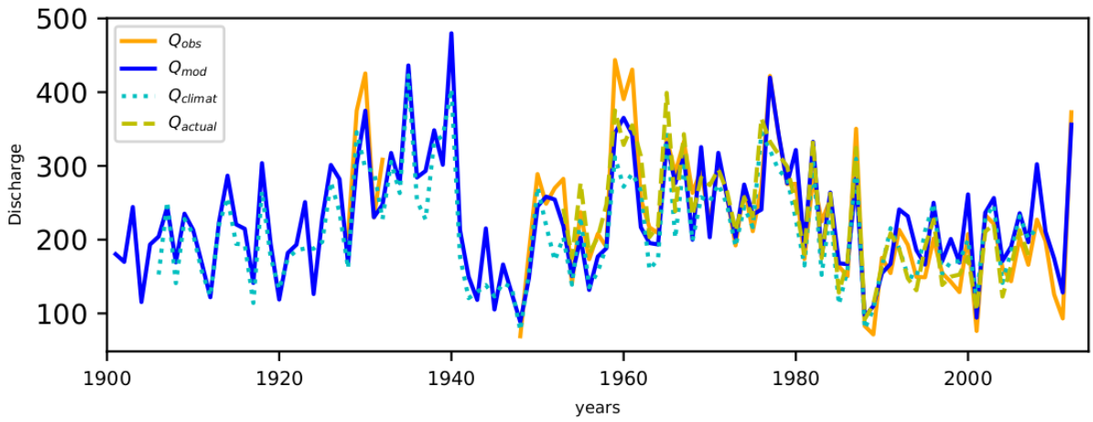
We therefore focus on the analysis of the evaporation efficiency ω. We here compare ω_c to ω_a. Similarly to fingerprint methodologies (Hegerl & Zwiers, 2011; Douville et al., 2021), we identify trends in evaporation efficiency ω that can't be explained by climate variability and trend only. In order to determine if the observed changes of evaporation efficiency (ω_a) are beyond what would be possible within a natural system, we use the statistics of all the climatic trends observed (ω_c). This comparison is done over all the 11-year trends which are computed in our method. Thus if the 11-year trend of ω_a is below or above or the 5-95% quantile of the distribution of all ω_c trends in this catchment, we can consider that it is not caused by climate change at 90% probability. We can also correlate ω_a with possible drivers such as development of human activities as a first step towards attributing changes.

4.5.1 Illustration of the analysis at the catchment level

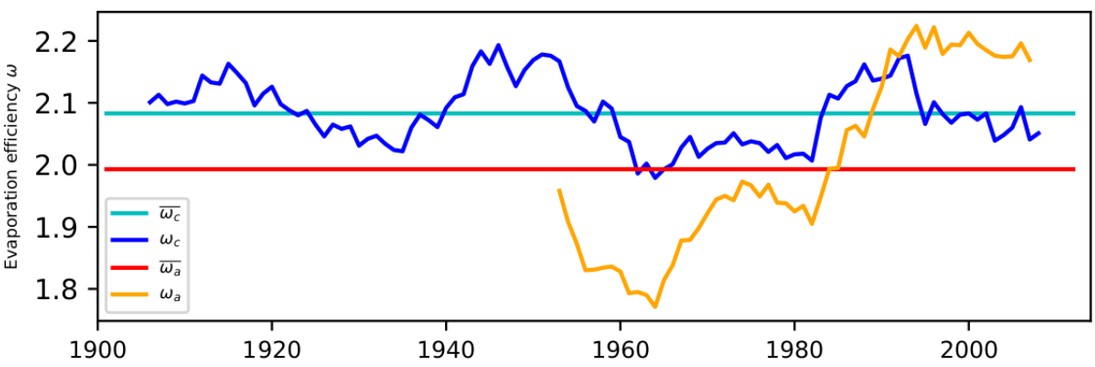
We focus here on two examples to illustrate the results available at the station level. We chose the station of Castejon, upstream of the Ebro river in Spain (Fig. 4.11a) and the station Pontelagoscuro for the Po river in Italy (Fig. 4.12a). The discharge (Fig. 4.11b) at the station level has continuous observations from the 1950's (Q_{obs}) for Castejon and from 1930 to 1990 for Pontelagoscuro. In the case of Castejon (Fig. 4.11b), we see that if the variability of Q_{obs} and Q_{mod} are very similar, we see that over the observation period covered by the observation, at the beginning of the period (1950-1970), $Q_{obs} > Q_{mod}$ while at the end of the period (1990-2010), $Q_{obs} < Q_{mod}$. Both tend to decrease but Q_{obs} has a steeper decrease. Looking at the variations of ω in both systems help to explain that difference. ω_c is not constant over time but its variability is smaller than the variability of ω_a. This remains true for all forcing datasets tested. There are other non climatic factors inducing higher trends. For the particular case of Castejon, there are two time periods at the end of the 1960's end in the 1985-1995 period where there are trends in ω_a with a slope which is higher than 90% of all of ω_c slopes over the entire century (Fig. 4.11d). Therefore there is a very high probability that these slopes can not be explained by climatic phenomena only. They are positive trends: non-climatic factors tend to increase evaporation efficiency (associated with a decrease in discharge, not significant however at the decadal scale). Similarly for the period 1930-1950 for Pontelagoscuro (Fig. 4.12d). More details on the interpretation of the results over the Po river are detailed later on.



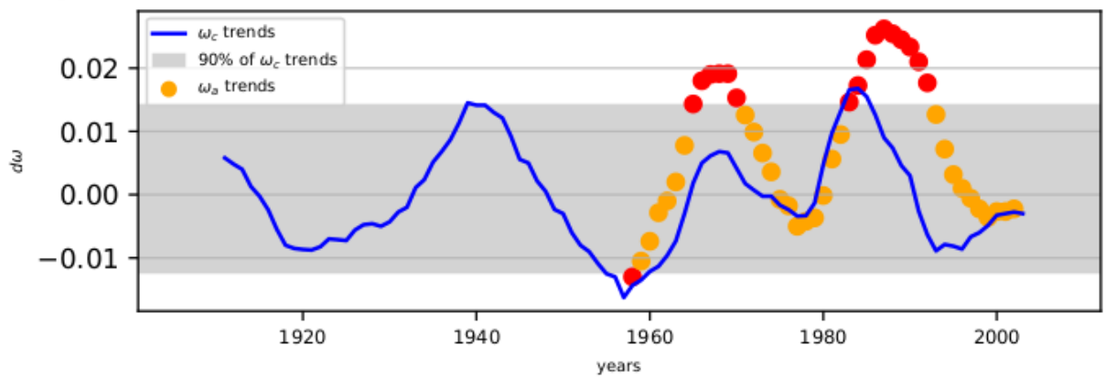
(a) Watershed of the gauging station Castejon on the Ebro river



(b) Annual mean discharge at the station outlet: Observed discharge Q_{obs} (orange), modeled discharge from the LSM ORCHIDEE Q_{mod} (blue) and from the Budyko framework fitted on the model Q_{climat} (dotted blue) and on the observations Q_{actual} (dashed orange)

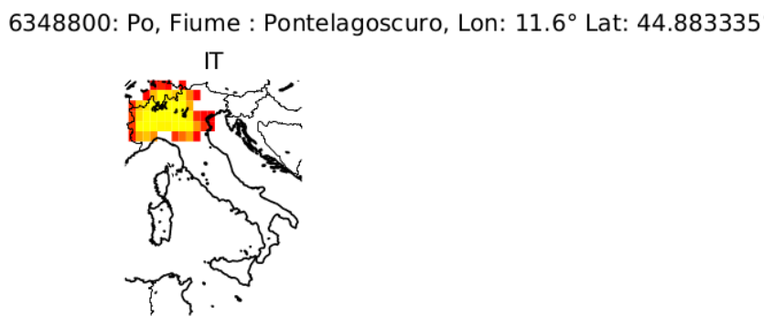


(c) ω fitted on the model outputs (ω_c (blue) corresponding to the "climatic" ω, compared to ω_a (orange) fitted on the observations).

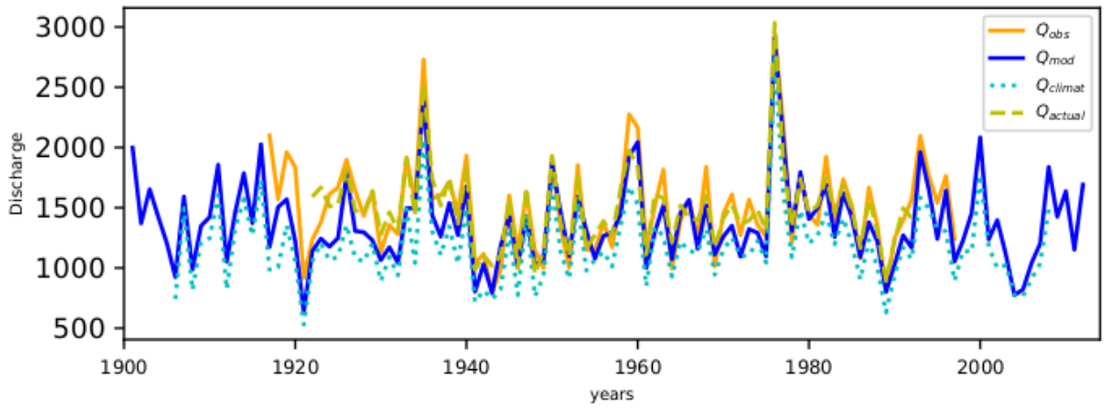


(d) Slopes of ω calculated with an 11-year time moving window (slope calculated over 11 years, 5 years prior and after the referenced year), for ω_c (blue) and for ω_a (orange). The red points corresponds to years for which the absolute slope of ω_a is different from 90% of all ω_c slopes (grey area).

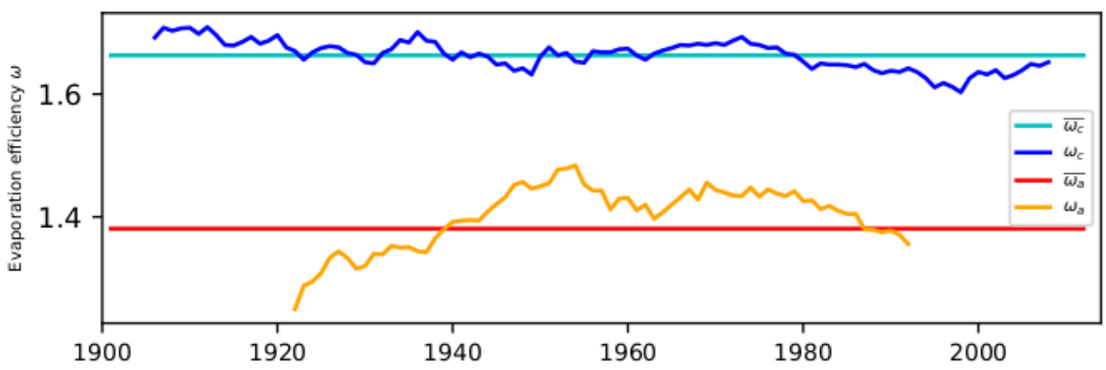
Figure 4.11 – Example of the results at the station level for the gauging station Castejon on the Ebro river in Spain



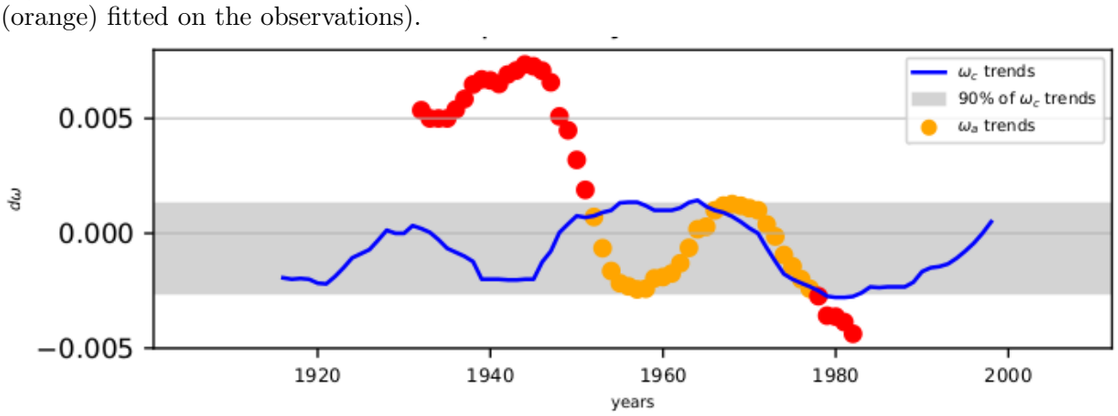
(a) Watershed of the gauging station Pontelagoscuro on the Po river



(b) Annual mean discharge at the station outlet: Observed discharge Q_{obs} (orange), modeled discharge from the LSM ORCHIDEE Q_{mod} (blue) and from the Budyko framework fitted on the model Q_{climat} (dotted blue) and on the observations Q_{actual} (dashed orange)



(c) ω fitted on the model outputs (ω_c (blue) corresponding to the "climatic" ω, compared to ω_a (orange) fitted on the observations).



(d) Slopes of ω calculated with an 11-year time moving window (slope calculated over 11 years, 5 years prior and after the referenced year), for ω_c (blue) and for ω_a (orange). The red points corresponds to years for which the absolute slope of ω_a is different from 90% of all ω_c slopes (grey area).

Figure 4.12 – Example of the results at the station level for the gauging station Pontelagoscuro on the Po river in Italy

4.5.2 Hypotheses behind a change in evaporation efficiency

We come back to the water-balance equation 3.1: $P = E + Q + \Delta S$. The Budyko framework assumes no variation in the water storage $\Delta S = 0$ and a closed system, at the catchment level. These hypotheses are verified in the modeled "climatic" system (see infobox 1), while in the observed catchments this hypothesis could be violated. In the latter, actual changes in ΔS and/or transfers of water from or toward other catchments would result in a change of the apparent evaporation efficiency due to sources or sinks of water unaccounted for.

A negative trends corresponds to a decrease of the apparent E (Fig. 4.1) or to an increase in the apparent river discharge. It can match an decrease in vegetation cover (which decreases transpiration and therefore E), an increase in relative runoff due to factors such as soil sealing. It can also corresponds to rising water input, not accounted for by a change in average P and PET only. It could be the expected effect of glaciers melting or groundwater uptakes. These factors correspond to additional sources of water not accounted for. Therefore the assumption $\Delta S = 0$ is wrong due to such factors, and it results in a change of the apparent evaporation efficiency. Similarly water transfer from another catchment would be a source of water unaccounted for.

A positive trend in the evaporation efficiency corresponds to an increase of the apparent E (Fig. 4.1) or to a decrease in the apparent Q. It corresponds to rising water uptakes (including rising evapotranspiration). It would be the expected effect of non climatic factors such as the development of irrigation and infrastructures such as dams which artificially enhance evapotranspiration. It can also match an increase in vegetation cover (which increases transpiration) or to a transfer removing water from the catchment.

4.5.3 Decadal changes in catchments' evaporation efficiency

We expand the analysis over all basins to have an overview of these results over all Europe. We map, for 10-year successive time-periods from 1920 to 2010, the average trend in the actual evaporation efficiency ω_a (Fig. 4.13 with red for positive trends, blue for negative trends). Blue areas correspond to rising water intake not explained by climate phenomena included in the model only. Red areas match rising water uptakes (including rising evapotranspiration). We also map where the trends in ω_a are larger than 90% of all ω_c slopes over the century. For each sub-decade, Fig. 4.13 (second column) show the proportion of ω_a trends larger than 90% of all ω_c slopes. It corresponds to the fraction of points outside of the grey zone at the catchment level in Fig. 4.11d. The larger that fraction (the closer to 1), the more likely it is that the these trends can't be explained by climate variability only, represented by ω_c. It therefore corresponds to the significance of non-climate-induced trends in evaporation efficiencies. Values close to 1 logically corresponds to areas where the starkest trends in ω_a are.

Since the beginning of the century (1920-1940), the trends in actual evaporation efficiency ω_a are significantly different from ω_c especially in the South of Europe. These trends are for the most part increasing trends in evaporation efficiency, which can correspond to an increase in the relative evapotranspiration due to the development of irrigation. These increasing trends are especially strong for the period 1920-1940 (Spain (especially the Ebro river), Italy (especially the Po river), south of France), 1960-1980 where they extended towards the Netherlands and Germany and 1980-2000 towards central France and southern Spain. It matches areas where irrigation infrastructures were developed over the century (Fig. 4.14,

Decade	Average $\Delta\omega_a$ (unit per year)	Fraction of $\Delta\omega_a$ 90% different from $\Delta\omega_c$ (0-1)

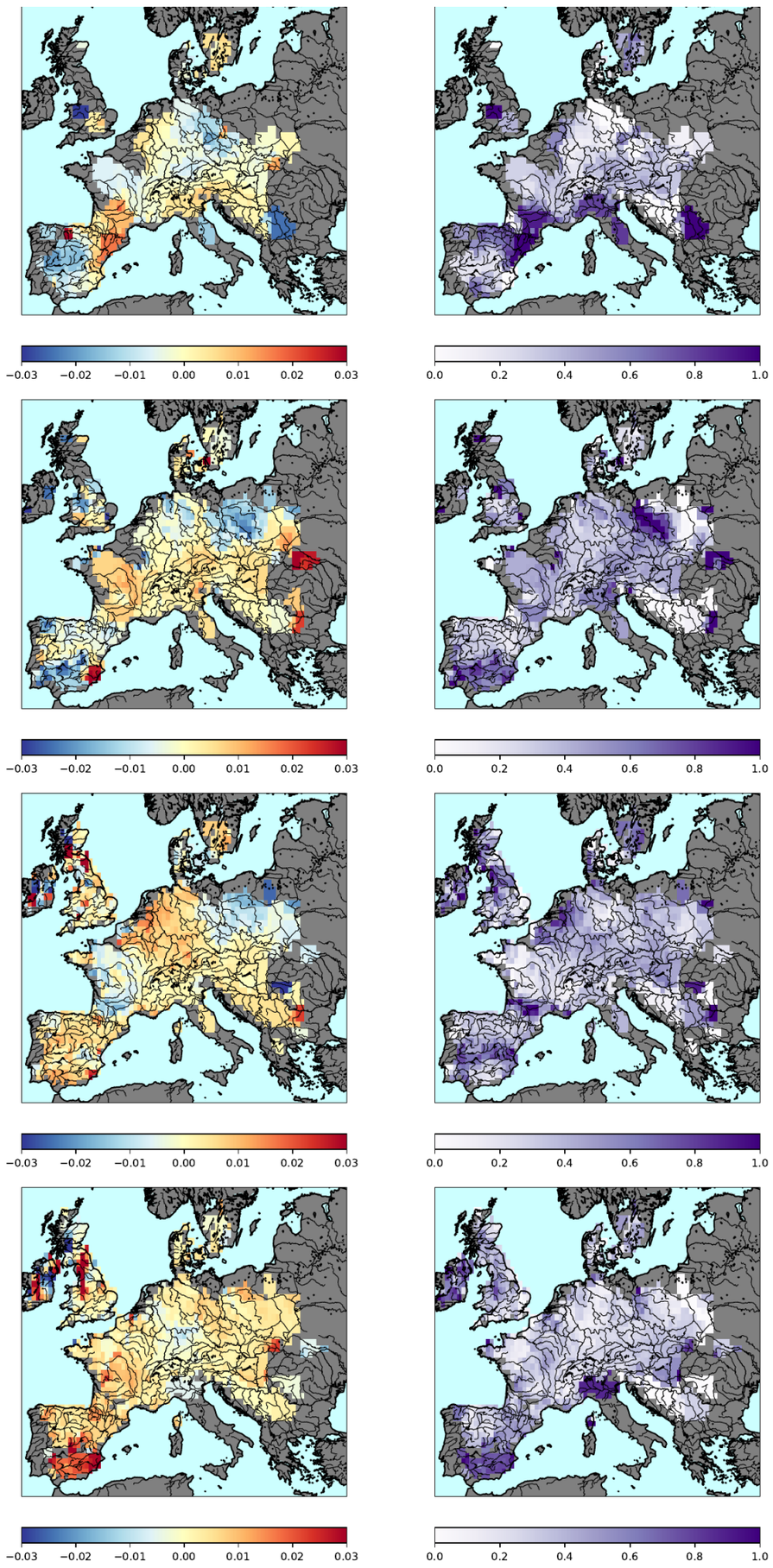


Figure 4.13 – Average decadal trends in evaporation efficiency over 20 years periods in Europe from **1920 to 2000**. The first column shows the average trends over each decade in evaporation efficiency ω_a, blue for negative trends corresponding to an increase in water relative intake and red to positive trends corresponding to an increase in relative water uptake. The second column is the comparison between ω_c and ω_a. For each sub-decade, we map the proportion of ω_a trends larger than 90% of all ω_c slopes. A value close to 1 means that over that decade most changes in the evaporation efficiency ω_a can't be explained by climate variability only.

from the dataset (Siebert et al., 2015)).

There are also decreasing trends, significantly different from climate variability. We were able to make fewer hypotheses to explain these trends. The significant decrease in the Center/South of Spain (source of the Guadiana river/Jucar basin, Fig. 4.16a) in the decade 1960-2000 coincides with intensive groundwater mining in that area (Holtz & Pahl-Wostl, 2012; Llamas et al., 2015; Esteban & Albiac, 2012). This could lead to a decreasing evaporation efficiency due to an additional sources of water in the system, unaccounted for. However groundwater mining is associated to the development of irrigation which would have an opposite effect by increasing the relative evapotranspiration. Therefore it is difficult to know which effect would dominate, and the sign of the trend in evaporation efficiency over that area varies from one decade to another.

In a different area, we can raise a different hypothesis. Looking at the Po river in Italy (Fig. 4.12, which is a major irrigated area in Europe (Fig. 4.14), we would expected strong positive trends due to the development of irrigation. It is indeed what we get for the period 1920-1940 and to a lesser extend up to 1950-1960 (Fig. 4.13). However, even though irrigation has continued to develop in the area (Fig. 4.14b), the positive trends in ω_a are getting less significant for the period 1960-1980 and turn to significantly negative trends for the last period (1980-2000) (Fig. 4.12, 4.13). To explain that change, along with the slight negative trends north of the Alps for the period 1980-2000, one hypothesis would be glacier melting. It would also be an additional source of water because of an unaccounted for unbalance to the system. In that area, glaciers can contribute up to 50% to streamflow (Schaner et al., 2012) and suffers from a regionally consistent acceleration of mass lover over recent decades (Vincent et al., 2017). It lead to increasing trends over the past decades over the Po river. With glaciers depletion, that increasing trend is however projected to stop by the end of the century, once glaciers depletion is so that melting runoff will decrease again (Huss & Hock, 2018).

An additional difficulty at comparing periods to each other is the stations with available discharge observation and the area covered by the upstream catchment of these stations can vary from on decade to another. This analysis is therefore only conjectures with high limitations. Furthermore here we analyze changes in the evaporation efficiency parameter ω, but since it does not have a truly defined physical meaning, it would be more interesting to look at discharge trends (which here have very few significant trends at the decade level).

4.5.4 Correlation with land surface and anthropic drivers

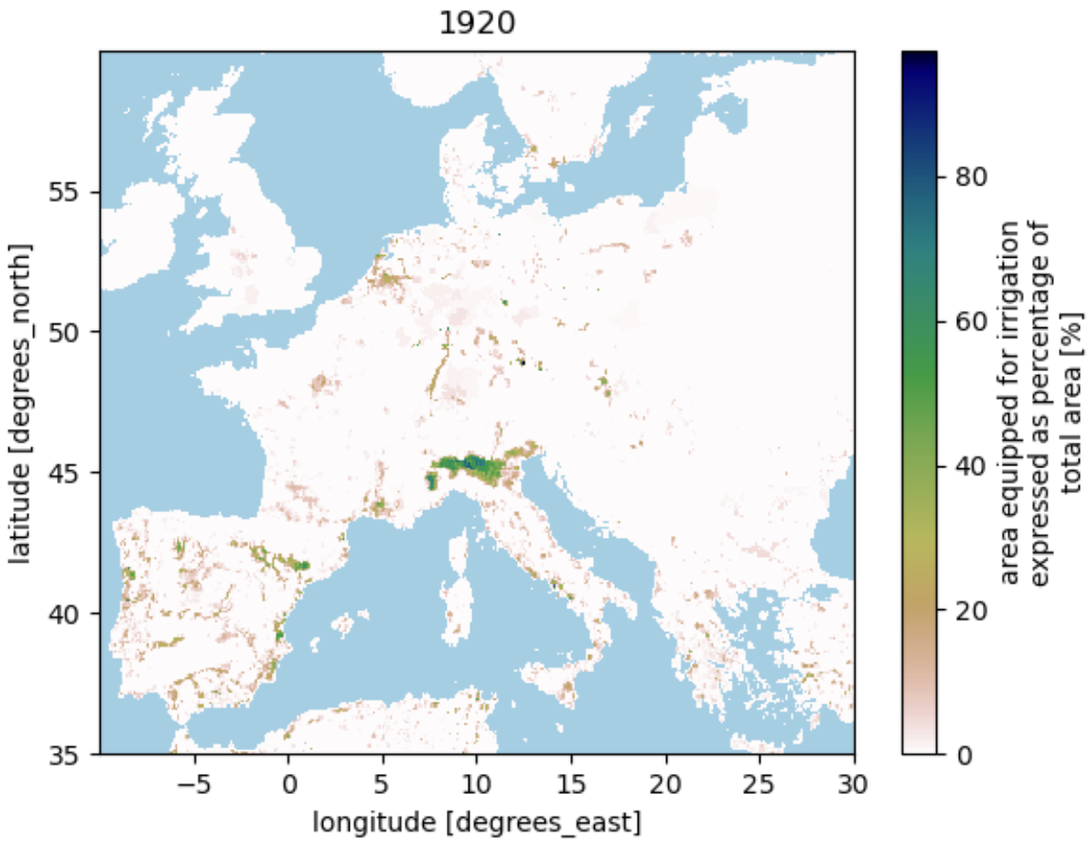
To better associate trends in evaporation efficiency to specific drivers, we can look at correlation between these changes and the evolution of given drivers. This however is not an attribution method.

4.5.4.1 Proxy to irrigation development

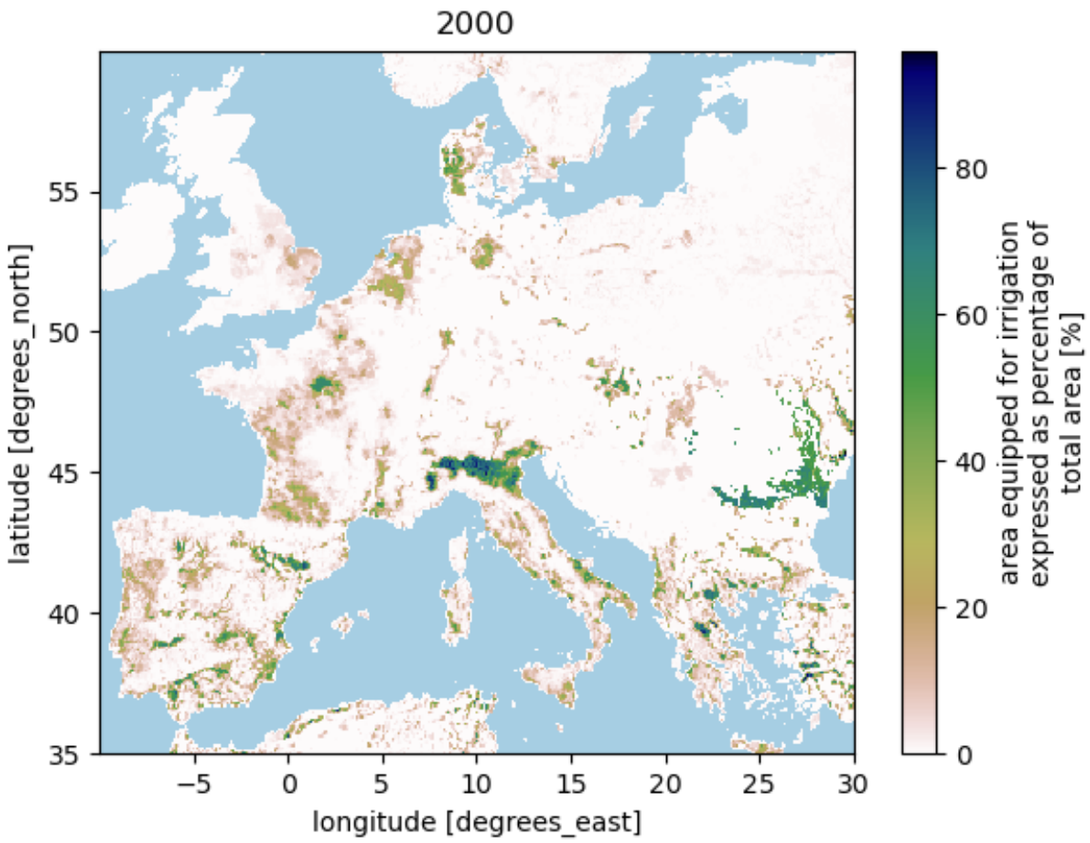
We hypothesize that the changes in evaporation efficiency should relate to the development of irrigation, and especially over Spain. We do not know a database of the amount of water use for irrigation, so we used proxy variables:

- **Surfaces equipped for irrigation**, from the database (Fig. 4.14) (Siebert et al., 2015). We can integrate over each catchment to get the change over time at the catchment level of percentage of area equipped for irrigation.

- **Evolution of water stored in dams**: over Spain, we had access to a database,

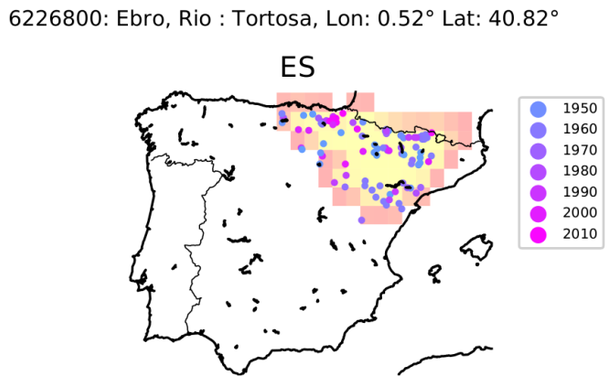


(a) Area equipped for irrigation (%) in 1920

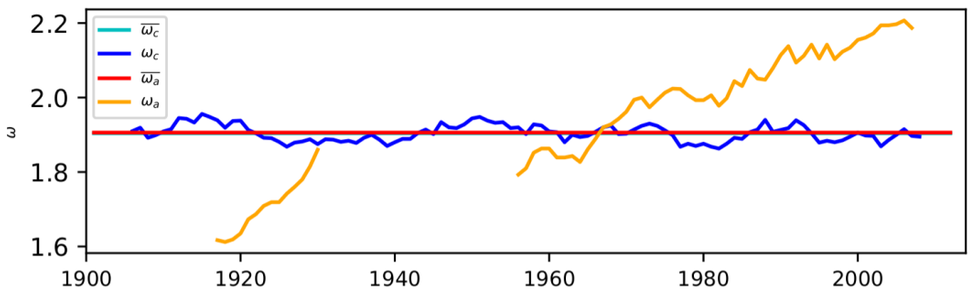


(b) Area equipped for irrigation (%) in 2000

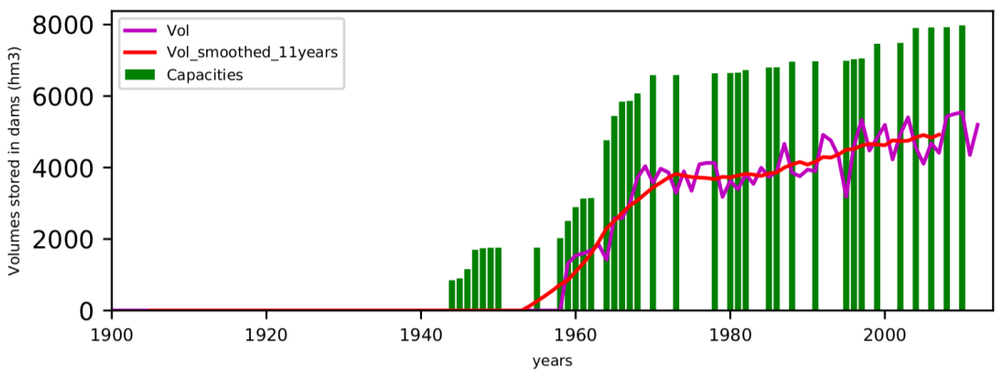
Figure 4.14 – Area equipped for irrigation (%): comparison between the equipment level in Europe in 1920 and in 2000, data from (Siebert et al., 2015).



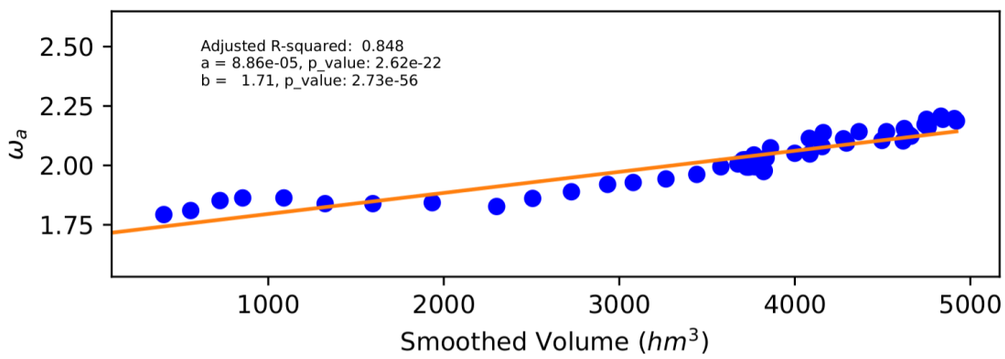
(a) Watershed of the gauging station Tortosa on the Ebro river. The points represent the dams reported on the catchment, coloured according to their date of implementation. The font represents the catchment limits, with the proportion of each points within the catchment (yellow for points fully in the catchment).



(b) ω fitted on the model outputs (ω_c (blue) corresponding to the "climatic" ω, compared to ω_a (orange) fitted on the observations (see Fig. 4.11c).

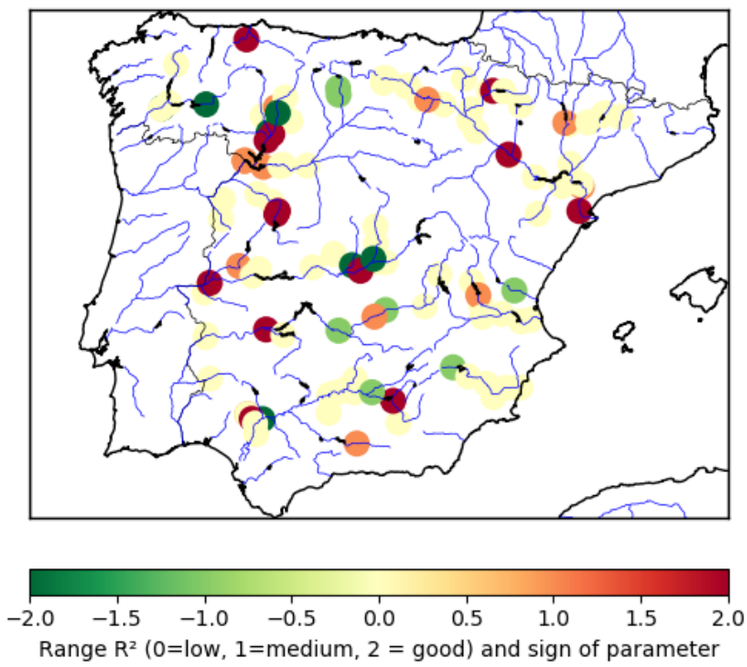


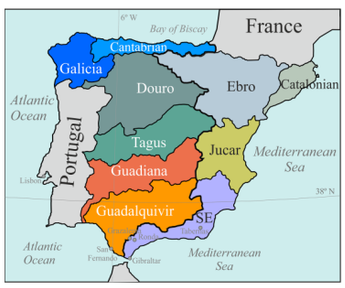
(c) Evolution of the volume stored in dams per year (purple line). Green bars represent the evolution of storage capacities when a new dam is implemented. The red line corresponds to the smoothed volume considering an 11-year sliding period.



(d) Result of the linear regression between ω_a and the smoothed volumed stored in dams over the catchment.

Figure 4.15 – Example of the results linking changes in evaporation efficiency and dams' water storage for the gauging station Tortosa on the Ebro river in Spain





(a) Main Spain rivers (García-Barrón et al., 2013)

Figure 4.16 – Performance of the linear regression between ω_a and the smoothed volumed stored in dams over Spain's catchments. R^2 is considered low ($R^2 < 0.5$), medium (R^2 0.5-0.7) and good ($R^2 > 0.7$). Then the sign corresponds to wether the correlation is positive or negative.

associated with the position of dams. Associating reservoirs to catchments, it gives us over Spain the evolution of water stored in dams over the century for each catchment.

For each catchment and each variable, we tested the correlation with the evaporation efficiency ω_a. To test correlation, we use the adjusted R^2 of a linear regression. We find weak correlations with the evolution of area equipped for irrigation. This could be that confounding effects take place over most catchments and/or that the factor chosen is not a good proxy to represent the main drivers of efficiency changes. Indeed, areas equipped for irrigation may not be used and not match the effectively irrigated areas. It could also be because we have very few points over time, with only values every ten years from 1910 to 1980 and every five years up to 2005.

Over some catchments in Spain however, we find a good correlation between the water stored in dams and the evolution of evaporation efficiency for 20% of the basins (example for Tortosa station Fig. 4.15d and summary of the results over Spain Fig. 4.16). The correlation is considered good for $R^2 > 0.7$. Over the entire Ebro basin (gauging station of Tortosa, Fig. 4.15), the linear regression between the evaporation efficiency ω_a trends and the volume of water stored in dams over the entire basin has a significant coefficient and an adjusted R^2 of 0.84 (Fig. 4.15). It is not an isolated result and in northern Spain (Ebro, dowstream of the Duero and Tagus river, Fig. 4.16), the increase of ω_a in the "actual" system is strongly positively correlated with the development of water storage in dams. In these areas, the development of dams seems to be a good indicator of the impact of human management on water resources impacting the evaporation efficiency of watersheds. It could be due to the direct effects of water storage in dams leading to an increase of evaporation and/or to indirect effects, dams development being an indicator of expansion of water use for irrigation.

The results are less clear in the South and East of the Peninsula. It corresponds to areas with large and highly exploited groundwater aquifers (La Mancha aquifer, over the catchment of the Jucar and upstream of the Guadiana river (Esteban & Albiac, 2012; Holtz

& Pahl-Wostl, 2012), aquifers over the Segura catchment (Custodio et al., 2016)). It could be that in such areas, the dominant effect on evaporation efficiency changes is the effect of groundwater uptake (increasing apparent runoff) and not the increase of relative evapotranspiration due to increasing irrigation. Both dams development and increase in groundwater uptake are both associated to a development of irrigated crops (Holtz & Pahl-Wostl, 2012). It would explain the negative correlation in that area: dams development is also correlated to the increasing groundwater uptakes which tend to reduce the evaporation efficiency by increasing the water supply. It would also explain the variability of the result over the area, dependent on which effect dominates.

This complementary study only analyses correlations and allows to raise assumptions. It however doesn't allows to definitely attribute the meaningful drivers behind the trends in evaporation efficiency, since these hypotheses are not tested with a model including the potential drivers.

4.5.4.2 Vegetation cover changes

We also hypothesize that changes in vegetation cover changes vegetation dynamics and therefore may impact evaporation efficiency of catchments. We use the HILDA dataset from Wageningen University (Fuchs et al., 2015) to represent changes in **different vegetation covers (forest, grasslands, croplands)**. We integrate each cover at the catchment level. Here also, most factors tested (changes in forest/grassland/cropland areas) have weak correlation to ω_a.

As a first step towards attribution, we tested the land cover changes effects by adding changing Plant Functional Type (PFT) maps into the LSM ORCHIDEE (Lawrence et al., 2016). The changes on evaporation efficiency time series where negligible with less than 3% of change. The current implantation of land cover changes in the LSM ORCHIDEE has a non significant effect on the hydrological behavior of the watersheds at the time scale of this study.

To continue testing attribution and the hypotheses introduced in this section, we would need to test the other factors (irrigation, dams development, groundwater uptake, glacier melting) and therefore introduce these possible drivers in the LSM.

4.6 Discussion and conclusion

Discharge is the result of integrated flows over a catchment. The functioning of the catchments and its evaporative behavior respond to climate (Troch et al., 2015; García-Ruiz et al., 2011). It is also sensitive to all human activities and other surface changes such as glacier melting, impacting the balance between evapotranspiration and runoff. In general, all of these phenomena have confounding effects which makes it difficult to attribute trends over basins where there is not a clear dominating factor.

In order to understand and reproduce discharge models are used. Physical-based models such as LSM are constructed to reproduce known physical processes and study their effect. They reproduce well the effect of climate processes but are incomplete and therefore fail to reproduce the full complexity of reality. Empirical calibrated models are better at reproducing the actual discharge but there is no clear understanding of what phenomena stand behind the calibrated parameters and their climate dependency (Coron et al., 2014; Nicolle et al., 2021), which needs to be accounted for to fully comprehend the role of climate change on discharge trends. This is true for any parsimonious hydrological model including the

parametric equations of the Budyko framework and there is a need to quantify this climate dependency before evaluating the role of non-climatic changes on the modeled discharge.

The usage of a state-of-the-art LSM to obtain a "climatic" reference is key in the methodology as it provides an estimate of the evolution of the empirical Budyko parameter, and thus river discharge, without anthropogenic pressure on the catchment. Then when comparing to our "actual" system with Budyko model fitted on observation, we quantify the effect of modifications of the evaporation efficiency due to non-climatic factors. In principle other parsimonious hydrological models could be used. Their empirical parameter could be adjusted over the LSM outputs and then over observations to decompose the contribution of climatic and non-climatic factors.

Similarly to Milly et al. (2005); Zhang et al. (2023); Gudmundsson et al. (2017b), we found when we use a LSM externally forced to represent the "climatic" behavior of the system that climate has a significant impact on discharge trends. The changes in annual mean discharge due to climate processes are mainly driven by annual mean climate variables P and PET but climate induced changes in the evaporation efficiency of the catchment can also play a role in discharge trends (chapter 3).

Over the "actual" system, we find that over the last century, annual mean discharge of European rivers has significantly changed (Fig. 4.10b). We find the largest decreasing trends in Spain, south of France and center Europe and the largest increasing trends in northern countries and northern England, similarly to the results of Vicente-Serrano et al. (2019); Yang et al. (2018). At decadal level, the high climatic variability tends to hide discharge trends. However at the scale of the century, the signal-to-noise ratio is higher and the effects of non climatic factors influencing catchments' evaporation efficiency emerges from the climatic noise. It dominates the trends over most of Europe, especially in Spain (Fig. 4.10d and 4.10f). The south of Spain corresponds to an area where both climate changes and changes related to human activities led concurrently to a significant decrease of the river discharge over the past century.

With this methodology, we can only estimate the non-climatic trends but not attribute them to specific factors. In some areas where we can hypothesize that some factors, such as irrigation, are dominant, we can correlate these trends to some specific factors to look for probable causes. For instance, over Spain, we correlate the strong increase in evaporation efficiency and dry up of rivers to the development of dams. Dams water storage is an indicator of human management on water resources impacting the evaporation efficiency of watersheds. More generally, we see that the changes in the evaporation efficiency intensify over the second part of the century where areas equipped for irrigation have been developing (Angelakιs et al., 2020; Siebert et al., 2015). However the correlation with that latter factor is less clear probably due to confounding effects and/or because that indicator doesn't account well for the effective amount of water use for irrigation.

Glaciers melting, groundwater pumping, can explain positive trend (or confounding effects) in discharge due to additional sources of water not accounted for in the "climatic" system. Other phenomena such as soil sealing and river management would be expected to have similar effects, due to a decrease of E or to an artificial enhancement of runoff. Changes in vegetation cover is shown to have little effect over the studied period and area but could have a more significant effect at local scale over small catchments.

Attributing discharge changes to specific factors remains challenging because most fac-

tors have concurring and competing effects. Detection and attribution method have been developed in climate studies to assess anthropogenic climate change. They have allowed to determine the role of various factors by reproducing them first in GCMs (Hegerl & Zwiers, 2011; Douville et al., 2021). Similarly, we would need to simulate water management and other missing phenomena in the LSMs so that their impact on the evaporation efficiency can be identified and their contribution to the non-climatic trend quantified. Matching the results of Gudmundsson et al. (2017b), our results highlight the fact that not accounting for them leads to high under-estimation of discharge changes in the LSM used and therefore to high uncertainties in projections for future trends.

This would allow to better understand how to balance the different factors and how their respective effects may change in the future. Changes in climatic variables are expected to increase (Gudmundsson et al., 2017b; Alkama et al., 2013). Concurrently in Europe, human water management is expected to evolve to adapt to climate change and other constraints such as changes in water and energy demand and regulation schemes (Arheimer et al., 2017). For instance, the extend of irrigated land in Europe has peaked at the end of the 20th century and the future irrigation evolution are expected to follow new goals and mostly rely on improved efficiency (Adeyeri et al., 2020). Therefore the balance between the different terms influencing catchments' evaporation efficiency and discharge may change. If non climatic factors dominated over the past century to explain discharge trends it may not be the same in the future.

KEY POINTS TO REMEMBER

- Our method uses Fu's equation to separate the effect of average annual climate change from the effect of changes in catchments' evaoprative behavior.
- The state-of-the-art LSM allows for a "climatic" reference, to extract the changes in evaporative behavior linked with climate (see chapter 3).
- When comparing that "climatic" system to the "actual" system, we can compare the changes in evaporative behavior due to non-climatic factors, not taken into account in the system modeled with the LSM.
- Over the past century in Europe, our method highlights that the changes in discharge are mainly due to non-climatic changes, especially the negative trends in the Iberian Peninsula. In the North Europe, streamflow is rather increasing, with a higher role of climate variables (P) at play.
- We can hypothesized that the non-climatic negative trends in streamflow are due to the development of irrigation and dams which increases evaporation efficiency of catchments, especially over the Mediterranean are. However, this hypothesis does not apply everywhere, with detected increasing non-climatic trends in highly irrigated areas. We can raise other hypotheses to explain these confusing trends, such as groundwater pumping for the South of Spain and glaciers increased malting rates for the Po river in Italy. These are however just correlation and hypothesis, no true attribution of discharge non-climatic drivers.
- For true attribution of streamflow changes to specific human activities and better understanding of their feedback on the atmosphere, these activities need to be included into LSMs, at a higher resolution than the current one.

CHAPTER 5

Towards better regional scale modeling: create a km-scale resolution forcing

This chapter introduces the first steps, constructed during the last month of the thesis, of a project to run different land surface models at kilometric scale with a common high resolution atmospheric forcing. This project is part of the LIAISE project and aims at testing model performances to reproduce the hydrological dynamics and other land surface processes at kilometric scale, over a specific contrasted area, known for the impact of climate change, human water use, contrasted topography and where highly qualitative data have been gathered.

Contents

5.1 Introduction

As shown in the previous chapters, human activities and related processes are missing from most of LSMs, while the effect of these missing processes is non negligible and even predominant in discharge changes in Europe over the past century. There is therefore a need to include regional processes such as irrigation, dams and human water management in LSMs, to get discharge and more generally the components of the water cycle correct for the right reasons. Running models at kilometric scale will improve finer scale processes representation more relevant to include human water use, associated to the correct water bodies (Wang et al., 2018; Stephens et al., 2023).

Human water managements is highly related to the river networks and to river flows. Dams are constructed over the river, usually in mountainous catchment, and store water to release it strategically. Irrigation subtract water to the available reservoirs (groundwater or near-by river), sometimes associated to artificial reservoirs and canals to take water from reservoirs further away, where it is available. It therefore relies on and changes the full regional dynamic of water fluxes. Before including the human disturbances into the models, it is necessary to test the models' ability to reproduce that regional dynamic at high resolution and improve the processes represented if needed. Several studies and projects aim at running models at higher resolution, and more recently at the kilometric scale (Lucas-Picher et al., 2021; Stephens et al., 2023). Due to the higher resolution, the atmospheric moisture transport and, in the case of the km-scale, deep convection are better simulated with finer processes reproduced (Lucas-Picher et al., 2021; Douville et al., 2021). More specifically, higher resolution allows to reproduce topography more accurately and therefore to better model the contrasts linked to altitude and slopes (Müller et al., 2021; Lucas-Picher et al., 2021; Ban et al., 2021; Zhao & Li, 2015; Stephens et al., 2023; Fan et al., 2019). Mountainous hydrology is important as water fallen over high elevation areas can sustain downstream ecosystem through surface and sub-surface flows converging towards valleys (Riedel & Weber, 2020; Fan et al., 2019), and remains a challenge to model and understand (Stephens et al., 2023). Müller et al. (2021) estimate that orographic precipitation accounts for 40% of global land precipitation but produces 50% of the global runoff. Underestimating it should therefore lead to overly dry systems (Fan et al., 2019). Improving the topography precision and the atmospheric processes represented at high resolution allows to better represent orographic precipitation (Müller et al., 2021; Ban et al., 2021; Lucas-Picher et al., 2021), more intense and more localized. It also improves the timing in snow-melt and evapotranspiration due to a better representation of terrain slope and orientation, impacting the exposition to radiation and local surface energy fluxes (Zhao & Li, 2015; Fan et al., 2019; Decharme et al., 2019). Overall, higher resolution models should perform better at reproducing small catchments with a strong influence of orographic precipitation (Müller et al., 2021).

However, if higher resolution runs are better at reproducing streamflow in small mountainous catchments, the observations in general are still in better agreement with low resolution models (Müller et al., 2021). This could be due to several reasons. First, despite their importance for downstream ecosystem dynamics and water availability, observations covering mountainous areas and complex terrain are limited. Weather stations are less dense in such areas (Poschlod et al., 2020; Lucas-Picher et al., 2021; Ban et al., 2021; Quintana-Seguí et al., 2020; Zhao & Li, 2015) and the altitude, slope and heterogeneity in the terrain complicates other types of measurements, such as radar measurements (Ban et al., 2021; Lucas-Picher et al., 2021). The majority of observations cover more accessible and lower areas and, in the case of streamflow, larger catchments where the share of orographic precipitation in the final

runoff is reduced. The good agreement of lower resolution models would be mainly due to a compensation at large scale of positive and negative biases (Müller et al., 2021). At higher resolution, the biases are not compensated and limitations appear due to missing important regional scale processes. Indeed LSMs were initially designed for coarser resolution, with the parametrization and the included processes validated accordingly (Kour et al., 2016; Zhao & Li, 2015). Processes currently at the sub-grid level and included in the parametrization would require an implementation of additional physics at finer scales (Wood et al., 2011). For instance, some interaction between vegetation rooting depth, slope and groundwater dynamics are still missing from models (Decharme et al., 2019; Fan et al., 2019). Depending on the depth of the water table and the infiltration capacity, both impacted by the altitude and the average slope of the terrain, vegetation can more or less access water, depending on its rooting system. It plays a role in vegetation growth and evapotranspiration dynamics, especially when the slopes contrasts are stronger, with a better representation of steeper slopes at higher resolution. Similarly, as shown in previous chapters, the impact of human water management is also still missing from models. The finer resolution improves the simulation of some processes, especially over mountainous areas, but highlights the limitations of models performances due to these missing processes. High resolution modeling needs to be improved to better account for local spatial heterogeneities and missing regional processes, to model more accurately streamflow and for the right reasons.

The current limitations to run and develop LSMs at high resolution (and km-scale) are the computational considerations (Wood et al., 2011; Kour et al., 2016; Lucas-Picher et al., 2021; Ban et al., 2021), especially in coupled mode, and the lack in high resolution data input. Since the end of the past century, remote sensing has been developed, allowing for time series of observation for land surface parameters (vegetation cover, leaf area index, albedo...) of increasing length and resolution (Zhao & Li, 2015). Spatial and temporal reanalysis of atmospheric variables based on observation data assimilation are constructed at coarser resolution. Different techniques are used to disaggregate these datasets and reach gridded products are km-scale. The first common technique is statistical downscaling. It relies on a network of observation stations and the data are disaggregated by spatial and temporal interpolation methods (Zhao & Li, 2015; Kour et al., 2016). Examples of such methodologies are multiple linear regression method, weather patterns generators, artificial neural networks (Khan et al., 2006; Zorita & von Storch, 1999; Hewitson et al., 2014). They all rely on using the observed relationship and pattern between large-scale and local scale climate to construct their disaggregation model (Zorita & von Storch, 1999). Most of these techniques use station measurements, their location and topographic information to adjust an interpolation function for each atmospheric variable and construct a high-resolution gridded product of these atmospheric variables (Zhao & Li, 2015; Quintana-Seguí et al., 2020; Kour et al., 2016). However statistical downscaling has limitations. The spatial and temporal disaggregation may not be physically coherent, depending on the interpolation function. Furthermore, such a technique necessitates a dense network in order to characterize the spatial contrasts at high resolution, which is an issue in the case of insufficient or low quality observation network, such as in mountainous areas. These downscaling techniques usually tend to smooth out extreme values and exaggerate the spatial extend of small scarce events, depending on which events are captured or not by the station network (Lucas-Picher et al., 2021; Ban et al., 2021; Quintana-Seguí et al., 2020). The second common technique is dynamical downscaling (Kour et al., 2016; Zhao & Li, 2015; Lucas-Picher et al., 2021). A regional climate model (RCM), consisting of a coupled atmospheric model and land surface model, is run nested in a coarser resolution dataset. The coarser dataset, which can be an

atmospheric reanalysis at coarser scale, is used to provide initial and boundary conditions to the RCM. Then the RCM runs at high resolution in these boundaries and spatially and temporally simulates a high resolution distribution of atmospheric variables. The distribution is physically consistent, reproducing coherent spatial and altitudinal contrasts and localized extremes. However these methods are highly demanding in data storage and computational power (Kour et al., 2016; Stephens et al., 2023). Furthermore, they are sensitive to the RCM biases and internal variability (Kour et al., 2016; Zhao & Li, 2015), which can lead to uncertainties and divergences over time in the nested area, drifting away from the real system it aims at reproducing.

This chapter introduces the method we implemented to construct a high resolution atmospheric forcing at km-scale. It relies on an observation-based atmospheric dataset (Système d'analyse fournissant des renseignements atmosphériques à la neige (SAFRAN), non gridded product covering geographical zones of about 1000 km^2) but combines it to the outputs of RCMs to disaggregate it at the sub-diurnal and km-scale. The latter have biases but present a physically consistent, high resolution distribution of atmospheric variables. We first introduce the prerequisites to the method, data and area of interest. Then we detail the method and the different choices and hypotheses that we made to construct a km-scale forcing. Finally, we present the first tests we had time to do. As it is a project started at the end of this PhD thesis, we did not have time to construct and test all the hypotheses we made when constructing such a forcing, nor to run LSMs using the final forcings produced. We therefore only present the current state of advancement and the future perspective and tests to be lead to continue the project.

5.2 Construction of a km-scale forcing: prerequisites

5.2.1 Problematic and objective

As introduced, there is a need to test and run LSMs at kilometric resolution as they were initially designed for coarser resolutions. Running LSMs at kilometric scales would allow to better assess water availability in mountainous areas and connect it better to lowlands hydrology. It would also allow to integrate local processes such as human water usage in LSMs. One limitation to running LSMs at higher resolution is the lack of data. Observation-based and reanalysis atmospheric forcing datasets are usually at broader scales, and can be limited at higher resolution especially in high altitude, due to sparser observation networks.

Therefore in this section, we present our methodology to construct a kilometric scale atmospheric forcing, using the output of on-line runs of regional climate models (RCMs) to spatially and temporally disaggregate an observation-based atmospheric dataset, keeping the sub-diurnal and spatial coherence of the processes involved in the conversion-permitting modeling. The final dataset needs to gather all the atmospheric variables needed to run a LSM, over a kilometric scale grid, and ideally cover a period long enough to cover natural long-term climate variability.

5.2.2 Area of interest

In order to construct a kilometric-scale forcing to run LSMs, we need to focus on a region, for computational purposes. We also need a data rich area, both for the construction of a forcing and later on to test the LSMs performances. Since there is also a need to include human water management in high resolution runs, we should choose an area interesting in this regard.

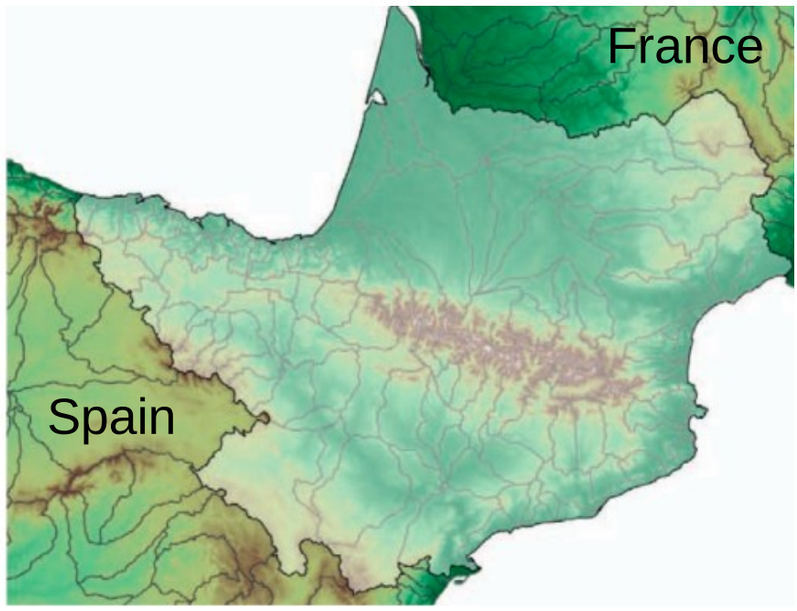


Figure 5.1 – Area of interest: South of France and North of Spain (Ebro basin), including the Pyrenees, from Project PIRAGUA https://www.opcc-ctp.org/fr/piragua (Accessed 02/09/2023).

We choose to work on the area defined North and South of the Pyrenees (Fig. 5.1) for the cross-boarder project PIRAGUA (Beguería et al., 2019). This project aims at characterizing the water cycle over the Pyrenees, over a territory vulnerable to climate change and gathers data related to the water cycle over the area (Palazón & Beguería, 2022).

Furthermore, this area also covers the area of interest in the Land surface Interactions with the Atmosphere over the Iberian Semi-arid Environment (LIAISE) project (Boone, 2019). This project focuses on the Ebro basin, to better understand the human impact on the water and energy cycle in a semi-arid environment and the limitations of models to represent all aspects of the terrestrial water cycle in such environment. This area is known as being largely impacted by human water management and irrigation and as having already suffered consequences of climate change with a decrease of snow-related runoff and increased irrigation needs. An intensive field campaign was conducted in summer 2021 to gather surface and atmospheric data over irrigated and non-irrigated lands, from multiple sources, from airborne to on-site measurements. This rich observational dataset can be used to test models run at high resolution over the area. Our work described in this chapter is part of that latter component of the LIAISE project. Ideally, the forcing constructed would cover 2021, so the outputs of the models can be compared to a rich qualitative set of field data.

In the end, this is a data rich region, ideal both to construct the forcing and to validate the later model runs. It is also of particular interest due to its semi-arid climate, contrasted with the mountainous sub-area covered by the Pyrenees where most of the water flow originates, and to its large impact of human water usage.

5.2.3 Data

5.2.3.1 SAFRAN observation-based dataset

Système d'analyse fournissant des renseignements atmosphériques à la neige (SAFRAN) was initially designed to provide meteorological information for snow models over the Alps (Durand et al., 1993). It was then extended and validated over France (Quintana-Seguí et al., 2008) and Spain (Quintana-Seguí et al., 2017). It uses an optimal interpolation method to

analyze most of the parameters, over defined climatically homogeneous zones (polygons, Fig. 5.2). These zones are based on "expert opinions" and are of irregular shape, of a surface usually smaller than 1000 km^2, where the spatial horizontal gradient (especially for precipitation) is supposed to be weak. Within each zone, the parameters are dependent on elevation (Quintana-Seguí et al., 2008). By an interpolation of the quality-controled observational data within the zone (and sometimes from neighboring zones), SAFRAN estimates one value for each parameter for each zone, at several altitude levels (every 300 meters) (Quintana-Seguí et al., 2020) (Fig. 5.3). As it is based on the observational network (Fig. 5.2), the number of meteorological stations used changes with time, and the quality of the dataset is better over more recent years. For many atmospheric variables with few observations such as radiative variables, a first guess is used, based on ERA-Interim outputs (infobox 12) (Quintana-Seguí et al., 2020).

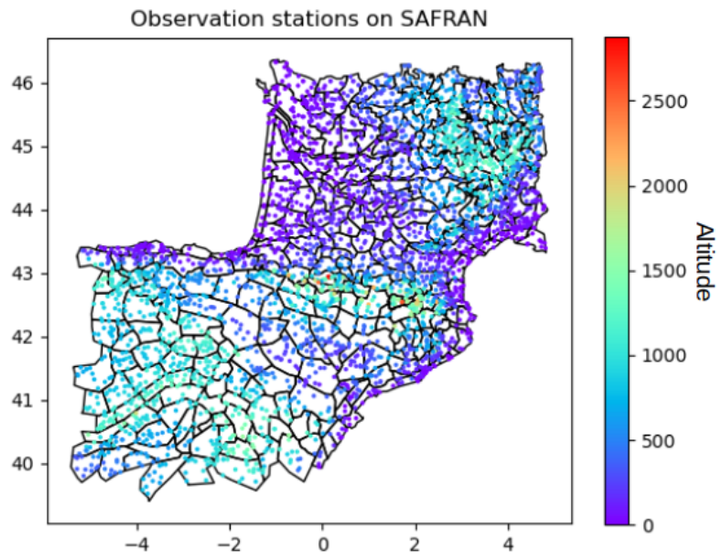


Figure 5.2 – Meteorological station network used in SAFRAN (year 1979) and SAFRAN climatically homogeneous zones (black polygons) over our area of interest. The color corresponds to the altitude of the station. We can see that the stations at high altitudes are scarce.

One downside of the climatically homogeneous zones is that they create artificial discontinuities at the borders of the zones (Quintana-Seguí et al., 2017) (Fig. 5.3). The main sources of error in the current downscaled gridded versions of SAFRAN originates with the interpolation to sub-diurnal values and spatial disaggregation within the geographical zones (Quintana-Seguí et al., 2008). It improperly captures spatial heterogeneities, with an underestimation of intense events (smoothing effect) and an low precipitation events either missed or overestimated, depending on whether they are locally captured and locally extended, especially in the areas with scarce observations such as in mountainous areas (Ban et al., 2021; Quintana-Seguí et al., 2020).

In this study, we use SAFRAN covering the period 1979-2014, at a daily time scale at the scale of the about 1000 km^2 geographical zones, later on referred as polygons. It is currently being extended to cover more recent years. We here present a method to disaggregate this dataset at the sub-daily time scale and within each polygon.

INFO BOX 12

Atmospheric reanalyses

Atmospheric reanalyses are multivariate datasets which provides a physically coherent, spatially and temporally complete reconstruction of past weather (Dee et al., 2011; Cucchi et al., 2020; Hersbach et al., 201904). They combine the use of models and assimilated observations (Hersbach, 2016), with increasing performances today with the development and growing availability of satellite data. The model allows to extrapolate the locally observed information to unobserved parameters and forward in time. Reanalyses are therefore impacted by biases of the model. Still, reanalyses are considered as a reference in the climate community, as the best possible reconstruction of the state of the atmosphere following observations (Cucchi et al., 2020).

- **ERA-Interim climate reanalysis (Dee et al., 2011)**

ERA-Interim is a global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is based on a forecast model and the assimilation of available observations. It is produced with a sequential data assimilation scheme, advancing forward in time using 12-hourly analysis cycles. Over each cycle, observations are combined with the forecast model for physical coherence and initialize the next cycle. It covers the period from 1979 onwards at a spatial resolution of about $\sim$ 79 km (T255) and 3-hourly temporal resolution.

- **ERA5 climate reanalysis (Hersbach et al., 201904)**

With the development of observations methods and forecast models, reanalyses techniques and data assimilation are continuously improved. ERA5 is the fifth generation of ECMWF atmospheric reanalyses, following ERA-Interim. ERA5 is developed since 2016 (Hersbach, 2016), with an increased assimilation of satellite and ground-based radar observations and beneficing from development in model physics and data assimilation. It tends to replace ERA-Interim which is not updated since 2019. ERA5 is available from 1979 onwards at higher temporal and horizontal resolution: 1 hour time step and about 31 km resolution. ERA5 product also includes information over uncertainties.

- **WFDE5 climate forcing (Cucchi et al., 2020)**

WFDEI (introduced in chapter 4) is a forcing constructed with the WATCH Forcing Data (WFD) methodology to bias-adjust ERA-Interim climate reanalysis. A sequential elevation and a monthly bias correction are applied to the dataset, based on monthly gridded observations from CRU and GPCC for precipitations totals (Weedon et al., 2014). WFDE5 is the result of the updated version of this methodology, with WFD applied to ERA5 (Cucchi et al., 2020). The final dataset has a higher temporal resolution (hourly) than WFDEI (3-hourly) and the same spatial resolution of 0.5°.

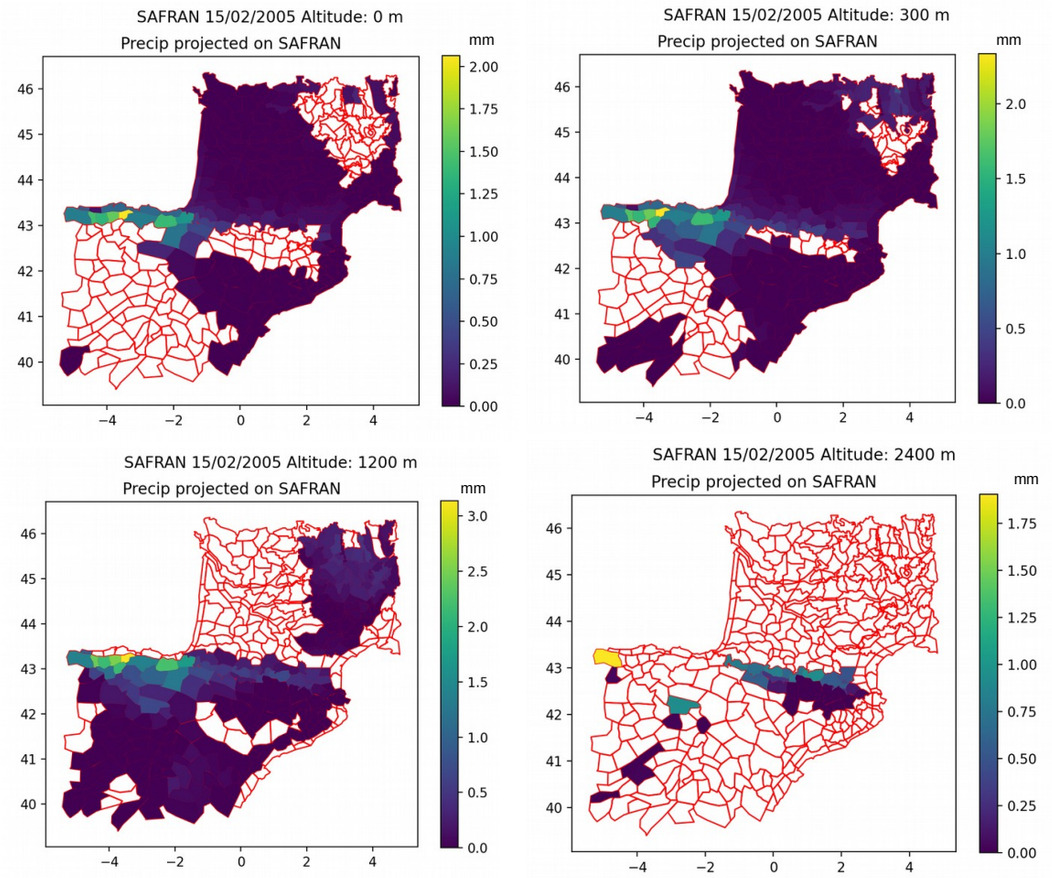


Figure 5.3 – Example of SAFRAN Precipitation data (15/02/2005) over the polygons and over different altitude classes (0m; 300m; 1200m; 2400m). Not all altitude classes are represented in each polygon, only the pertinent classes.

5.2.3.2 High resolution convection-permitting regional climate model (CPRCM) outputs

In the past decade, a new tool has emerged to produce fine-resolution (km-scale) decadal-long climate simulation: convection-permitting regional climate models (CPRCMs) (Lucas-Picher et al., 2021). They are designed similarly to regional climate model simulations (RCMs), with a dynamical downscaling, which consist in running a climate model at high spatial resolution over a limited area, using a coarser-resolution climate information to provide initial and lateral boundary conditions (Lucas-Picher et al., 2021). They can be run and evaluated in a "perfect framework", forced at their boundaries by reanalysis, which represent as best as possible the past atmospheric states at larger scale due to observation assimilations (Lucas-Picher et al., 2021). Compared to RCMs, CPRCMs are run at a scale finer than 4-km to be able to reproduce meso-scale atmospheric structures and explicitly resolve deep convection processes which are parameterized in RCMs. The explicit simulation of deep convection associated to the finer and more accurate topography improves the ability to model precipitation characteristics (diurnal cycle and hourly distribution, intensity, frequency, duration of precipitation events along with short localized extreme events) (Lucas-Picher et al., 2021; Ban et al., 2021) and temperature diurnal cycle and altitudinal gradients (Lucas-Picher et al., 2021).

CPRCMs and RCMs are run over limited domains due to limitations in computational power and storage capabilities. In the context of the European Climate Prediction system (EUCP) H2020 project (https://www.eucp-project.eu/) (Hewitt & Lowe, 2018) and COordinated Regional climate Downscaling Experiment (CORDEX), different European research

teams have performed CPRCMs simulations over different domains in Europe at km-scale (smaller than 4km). More specifically, an ensemble of 4 simulations covering the southern western Europe and forced by ERA-interim reanalysis (infobox 12) was performed for the period 2000-2009.

One instance of such simulations is the regional climate simulations with the coupled atmosphere (WRF)-land surface (ORCHIDEE) RegIPSL model 3 km (SWE3; convection-permitting/resolving) (Shahi et al., 2021). We will mainly use the outputs of this simulation in the following, but we have access to other simulations at similar km-scale, from other models run within the same framework for the same projects. The methodology developed later on can be applied using another one of them, to construct different products, whose quality can be compared. It would allow to test the sensitivity of the method to the model used (see section 5.5).

The CPRCMs performances are measured against high resolution gridded observation-based datasets, such as gridded products derived from SAFRAN or radar-based datasets. However, as we have seen for SAFRAN and it can be generalized to all gridded products derived from observation assimilation, high uncertainties remain around such gridded products. Indeed, especially in polar and mountainous regions where weather stations are sparsely distributed, the data gathered may not be sufficient to be representative of the area. For the same reason, precipitation extremes are undercatched due to their localized nature (Quintana-Seguí et al., 2020; Lucas-Picher et al., 2021; Ban et al., 2021). In the case of radar-based products, their quality is dependent on the distribution of radar in space since their efficiency is limited by distance and diminished over mountain ranges (Westrick et al., 1999; Trapero et al., 2009; Panziera et al., 2018). These limits question the quality of the spatial and temporal disaggregation at finer resolution of observation-based datasets, when most of the observation data are not available at the sub-diurnal time step and the observation network may not be dense enough (Quintana-Seguí et al., 2020; Lucas-Picher et al., 2021; Ban et al., 2021). At least in mountains, the skill of the CPRCMs in simulating rain and snow could exceed the skill of observation networks in measuring precipitation (Lundquist et al., 2019). The methodology presented in this chapter proposes a new method to get a spatial and sub-diurnal disaggregation of the observation-based product SAFRAN, favoring the spatial and sub-diurnal distribution of the CPRCM. We also construct different forcings, testing the sensitivity of our final outputs to whether we favor the altitudinal distribution of SAFRAN or of the CPRCMs outputs (see section 5.3.2.1).

INFO BOX 13

Compatibility between SAFRAN/CPRCM variables and associated conversions

SAFRAN available data do not match exactly the atmospheric output of the RCMs (Tab. 5.1). Therefore there is a step of conversion necessary before comparing the two datasets and launch the disaggregation methodology.

Table 5.1 – Available atmospheric variables for each dataset

	SAFRAN	CPRCM ouputs
Temperature	T (°C)	tas (K)
Wind speed	Wind ($m.s^{-1}$)	*uas*: Eastward Near-Surface Wind *vas*: Northward Near-Surface Wind
Relative humidity	Hum (%)	hurs (%)
Surface Downwelling Longwave Radiation (LWdown)	IR: Infra-red incident radiation ($W.m^{-2}$)	rlds ($W.m^{-2}$)
Surface Downwelling Shortwave Radiation (SWdown)	Soldr: Solar direct radiations Soldf: Solar diffuse radiations ($W.m^{-2}$)	rsds ($W.m^{-2}$)
Precipitation	RRtot (mm)	pr ($kg.m^{-2}.s^{-1}$)
Pressure	-	psl: Sea level pressure (Pa)
Altitude	-	Orography (m)

Pre-conversions necessary for comparison purposes:

- From °C to K for SAFRAN temperature data
- From *uas* and *tas* to wind speed in the CPRCM outputs: $Wind = \sqrt{uas^2 + vas^2}$
- Surface Downwelling Shortwave Radiation: SWdown = Soldr + Soldf = rsds
- Pressure is not corrected with the methodology since the variable is not in SAFRAN. We keep the initial outputs from the CPRCM.

The final version of the kilometric-scale dataset should include all variables necessary to run the LSM with the forcing. Therefore new conversions are needed.
The final variables needed are presented in Tab. 5.2.

Table 5.2 – Needed final atmospheric variables to run ORCHIDEE

Variable	Unit	Full name	Conversion
Tair	K	Near-Surface Air Temperature	-
ASurf	m	Surface Altitude	-
PSurf	Pa	Surface Air Pressure	1- Fom sea level pressure (psl), temperature (Tair) and altitude (ASurf)
Qair	$kg.kg^{-1}$	Near-Surface Specific Humidity	2- From relative humidity (hurs), temperature (Tair) and surface air pressure (PSurf)
Wind		Near-Surface Wind Speed	-
Rainf Snowf	$kg.m^{-2}.s^{-1}$ $kg.m^{-2}.s^{-1}$	Rainfall Flux Snowfall Flux	3- From precipitation variable and temperature
LWdown	$W.m^{-2}$	Surface Downwelling Longwave Radiation	-
SWdown	$W.m^{-2}$	Surface Downwelling Shortwave Radiation	-

Final conversions necessary:

1. Surface atmospheric pressure ($PSurf$) from the sea level pressure (psl), temperature ($Tair$) and altitude ($ASurf$):
   $PSurf = psl * exp(\frac{-g*ASurf*Ma}{Tair*R})$
   with g =9.81 $m.s^{-1}$ the gravitational constant; $Ma = 0.02897$ $kg.mol^{-1}$ the molar mass of dry air; $R = 8.314$ $J.mol^{-1}.K^{-1}$ the universal gas constant.
2. Surface specific humidity ($Qair$) from relative humidity, temperature and surface atmospheric pressure:
   - Partial pressure of water vapour at saturation (Ps) calculated from $Tair$ (Huang, 2018)
   - $Qair = \frac{hurs}{100} * eps * Ps$
     with eps = 0.622, ratio between the gas constant for dry air and water vapour.
3. Divide precipitation into rainfall and snowfall according to temperature:
   Precip = $Snowf$ if $Tair < 0$, $Rainf$ otherwise

5.2.3.3 Choice of target grid

We want a km-scale final atmospheric forcing which can be used to run different LSMs at high resolution and test their performance. We therefore needed the final forcing to have a km-scale grid, easy to use for all the LSMs which would want to. The outputs of the different CPRCMs we had available are gridded at the km-scale but they all have different types of grids and resolution. In order to have a defined final grid, which would be common to all the final products for comparison purposes and which would be easy to use for different types of LSMs, we constructed a new grid.

This grid is a regular lat/lon grid of $\sim$ 3 km resolution (infobox 14). All the outputs of the CPRCM used to construct the final forcing are first projected on this new grid.

INFO BOX 14

Regular latitude/longitude grid description

- Size of the final grid lat x lon : 263 x 286
- Region covered: latmin, latmax = 39.3°, 46.4°; lonmin, lonmax = -5.5°, 5°
- Regular lat/lon steps ($\sim$ 3 km over the selected region): lat: 0.0270°; lon: 0.0368°

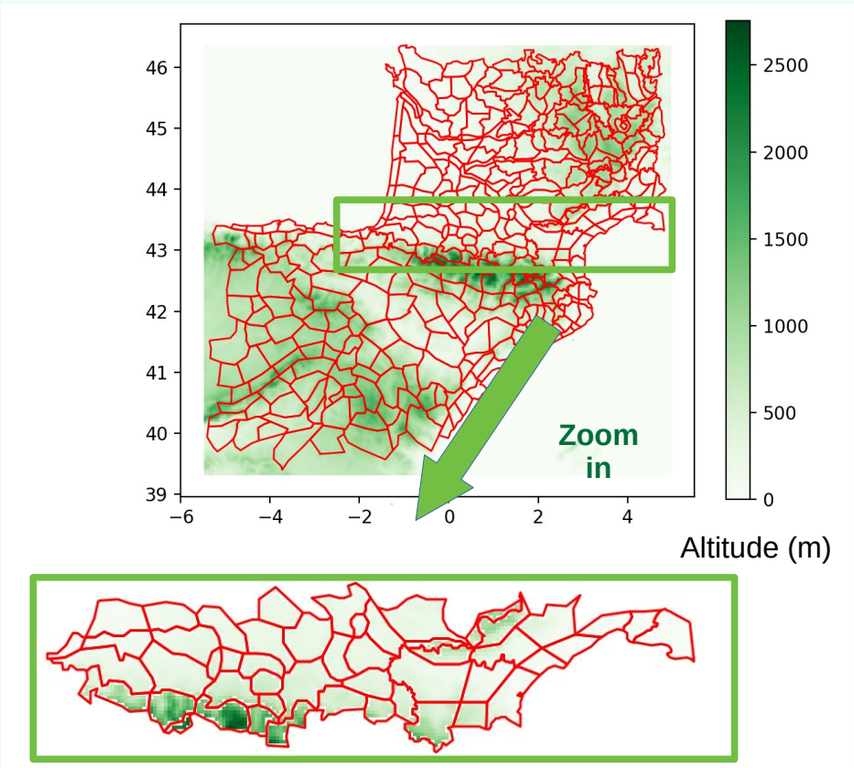


Figure 5.4 – Spatial subdivision of the polygons with the final grid (here colored by altitude)

5.3 Construction of a km-scale forcing: methodology

We have:

- an observation-based dataset SAFRAN at the scale of polygons, with daily values of atmospheric data for 1979 to 2014, with information over different altitude classes.

- the outputs of a CPRCM (here RegIPSL but other models can be used), at high resolution and hourly time step, for 2000 to 2009.

We want:

A full atmospheric forcing covering the longest period as possible, favoring the data of SAFRAN at the daily scale and polygon level but favoring the spatial and sub-diurnal coherence of the CPRCMs outputs.

5.3.1 Calculation of a Bias

As a first step, we focus on the time period common to the two datasets: 2000-2009. There is an initial step to convert the data from SAFRAN and the CPRCM so they can be compared (infobox 13).

5.3.1.1 Daily bias over each polygon

We want to keep the daily average of SAFRAN data in the final dataset. Therefore we calculate a bias to correct the CPRCM data to match SAFRAN daily average. To do so, we start by averaging the CPRCM data at a daily step and over each polygon (step 1, Fig. 5.5). At this stage, there is a choice to make, in how to consider the altitude classes. More details will be given in section 5.3.2.1.

To average at the polygon level (step 1, Fig. 5.5), first we calculate for each grid point p, its portion in each polygon:

$$Portion_p = \frac{Area_p \cap Polygon}{Area_p}, \forall p \tag{5.1}$$

Then we can calculate the spatial average of the daily averages for each atmospheric variable at the polygon level (complement in section 5.3.2.1):

$$\forall \text{ day and polygon, } Value_{model_{poly}} = \sum_{p \text{ in } poly} Portion_p * Value_p \tag{5.2}$$

Then we calculate a linear bias, ratio of the daily average of each dataset (Teutschbein & Seibert, 2012), for each polygon (step 2, Fig. 5.5):

$$\forall \text{ day and polygon, } Bias = \frac{Value_{SAFRAN}}{Value_{Model}} \tag{5.3}$$

5.3.1.2 Spatial disaggregation of the bias

Once we have a bias at the polygon level, we can extend that bias at all the points within that polygon (step 3, Fig. 5.5), proportionally to their portion within the polygon (Equ. 5.1).

Some issues remain at that stage (step 4, Fig. 5.5):

- Some points of the km-scale grid have an altitude not covered by SAFRAN altitude classes (depending on the choice made concerning altitude classes to calculate the average bias, detailed later in section 5.3.2.1, point 1). Therefore there is no bias extended to these points. For them, we apply the average bias of the neighboring points.

- For some atmospheric variables which can have a null value such as precipitation, the bias may not be defined if $Value_{Model} = 0$ for a given day and polygon. In

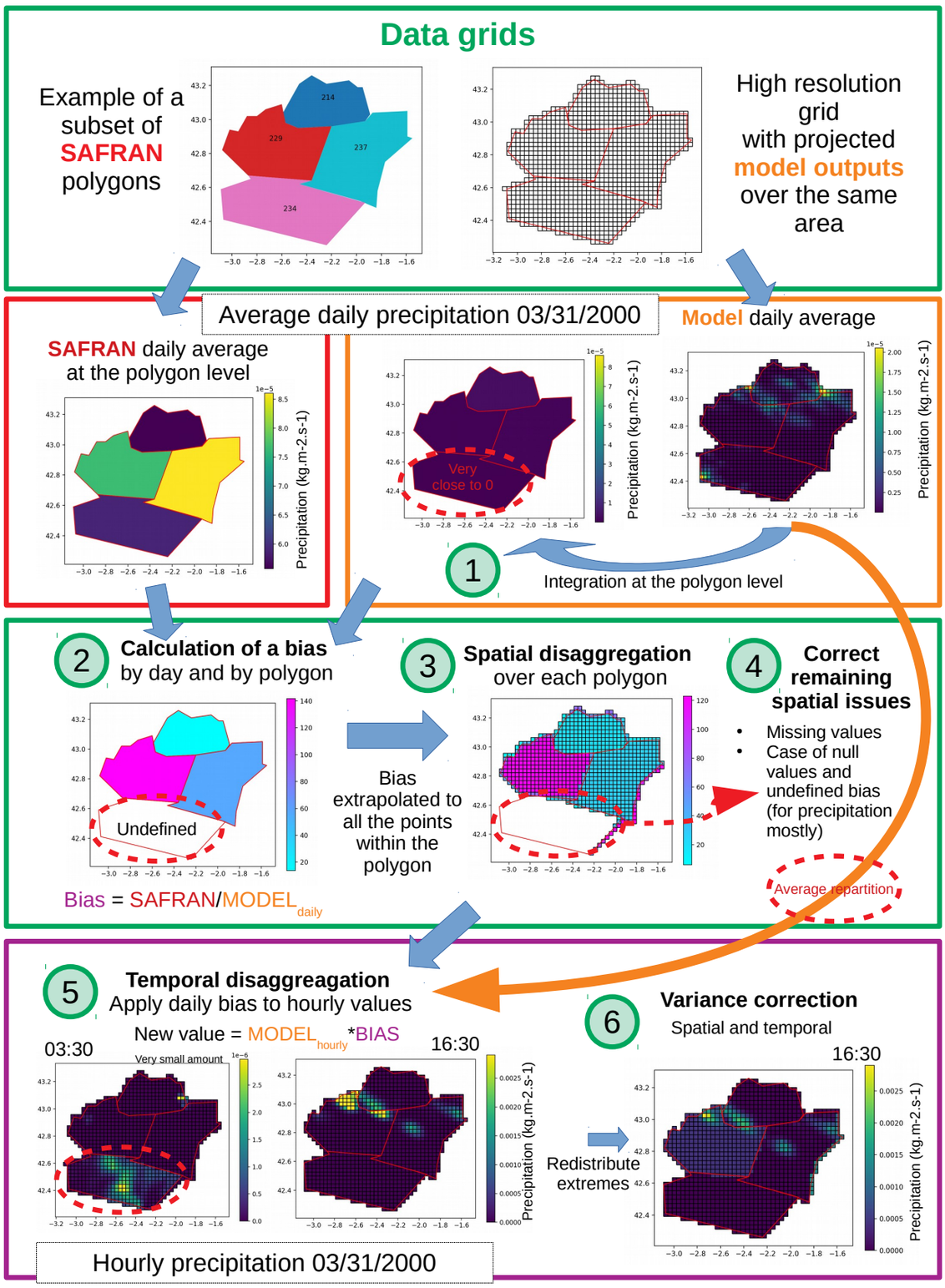


Figure 5.5 – Method to construct a high resolution forcing, favoring the daily values of SAFRAN and the spatial and sub-diurnal disaggregation of the outputs of the CPRCM model projected on a regular lat/lon grid. Example over a subset of polygons for precipitation over a given day (31/03/2023).
1- Aggregation of daily average of the output of the CPRCM for comparison with SAFRAN
2- Calculation of a bias between SAFRAN and the CPRCM
3- Spatial disaggreagtion of the daily bias over the high resolution grid
4- Correction of remaining issues : missing values and undefined bias
5- Application of the bias to correct the hourly values of the CPRCM
6- correction of the temporal and spatial variance to smooth out extreme values due to the disaggregation

the case of precipitation, we consider that a daily average below 0.01 mm/day (or 1.10^{-7} kg.m^{-2}.s^{-1}) is null, to get rid of drizzling in the model outputs. If we also have $Value_{SAFRAN} = 0$ then there is no point in calculating a bias. However, if $Value_{SAFRAN} \neq 0$, then the points within that polygon have to be treated differently later on (see infobox 15).

5.3.1.3 Temporal disaggregation: bias correction

Once we have a disaggregated bias, we apply it to correct the hourly values of the CPRCM outputs (step 5, Fig. 5.5):

$$\forall \text{ hour and point } p,\ NewValue_p = Value_p * Bias_p \tag{5.4}$$

Therefore, we keep the relative spatial and sub-diurnal distribution from the CPRCM within each polygon but scaled so the daily average of each polygon matches SAFRAN data.

5.3.1.4 Last intrinsic issue: variance correction

At that stage, the produced dataset matches our objectives of respecting the daily average of SAFRAN and the relative spatial and temporal distribution of the CPRCM. However one strong remaining issues is inherent to the methodology and to the calculation of a proportional bias. A positive proportional bias applied to all the points within a polygon increases the spatial contrasts over the polygon. Indeed, since the bias is applied to all the points in the polygon, the average at the polygon level is proportionally increased, but also the standard deviation (Teutschbein & Seibert, 2012) of the ensemble of all points within the polygon. Same goes for the daily average and standard deviation for an hourly series of a given point. It is partially what we wanted, favoring the contrasts of the CPRCM outputs to distribute variables and keep a representation of extreme events. However it can lead to overly contrasted values and highly exaggerated extreme values, especially for precipitation and wind which are highly contrasted spatially and radiations variables which are highly contrasted temporally. Therefore we need a final step to correct the final spatial and temporal variance and set a maximum acceptable variance to avoid overly contrasted spatial and sub-diurnal distributions.

First we define that maximum spatial and temporal variance acceptable over each month by taking the maximum variance for that month modeled with the CPRCM, both spatially (over each polygon) and temporally (over hourly series for each given grid point).

Then, for each polygon at each time step, we check if the spatial variance is higher than the corresponding maximum and if so, we correct it in three steps (Teutschbein & Seibert, 2012):

- First we shift the average of the polygon to center the values around the null value for that time step:

$$\forall p \text{ within the polygon},\ ShiftValue_p = Value_p - Average_{polygon} \tag{5.5}$$

- Then we correct the standard deviation around that shifted average:

$$\forall p \text{ within the polygon},\ NewShiftValue_p = ShiftValue_p * \frac{\sigma_{max}}{\sigma_{shift}} \tag{5.6}$$

At this stage, the standard deviation of the shifted series is corrected.

INFO BOX 15

Handeling the issue of undefined bias

There is an issue when a bias can't be defined, when $Value_{Model} = 0$. It needs to be addressed.

Along with the spatial disaggregation of the bias over the km-scale grid, we construct in parallel a mask to identify the points with undefined bias (0 if $Value_{Model} = 0$ and $Value_{SAFRAN} = 0$, -1 if $Value_{Model} = 0$ and $Value_{SAFRAN} \neq 0$, 1 if the bias is defined).

In the case of the mask = 0, then the value is set to 0, since the daily average over the polygon is also null for SAFRAN. There is no adjustment and bias correction needed.

In the case of the mask < 0, we have to make a choice. Indeed in that case, there is a non-zero value for SAFRAN at the daily scale but none for the corresponding day in the CPRCM outputs. Therefore we cannot distribute SAFRAN value using the relative distribution in the model since it does not exist. To solve this issue, for each month, we calculate an average spatial and temporal distribution of each variable in the CPRCM outputs (Fig. 5.6). When we are in that case of mask = −1, we disaggregate SAFRAN data spatially and temporally using that average distribution of the corresponding month (Fig. 5.6d). This is not ideal since it smooths out most of the contrasts. This is an important limitation of the method for precipitation disaggregation since it is a highly localized and contrasted phenomenon, sensitive to the internal variability of the model.

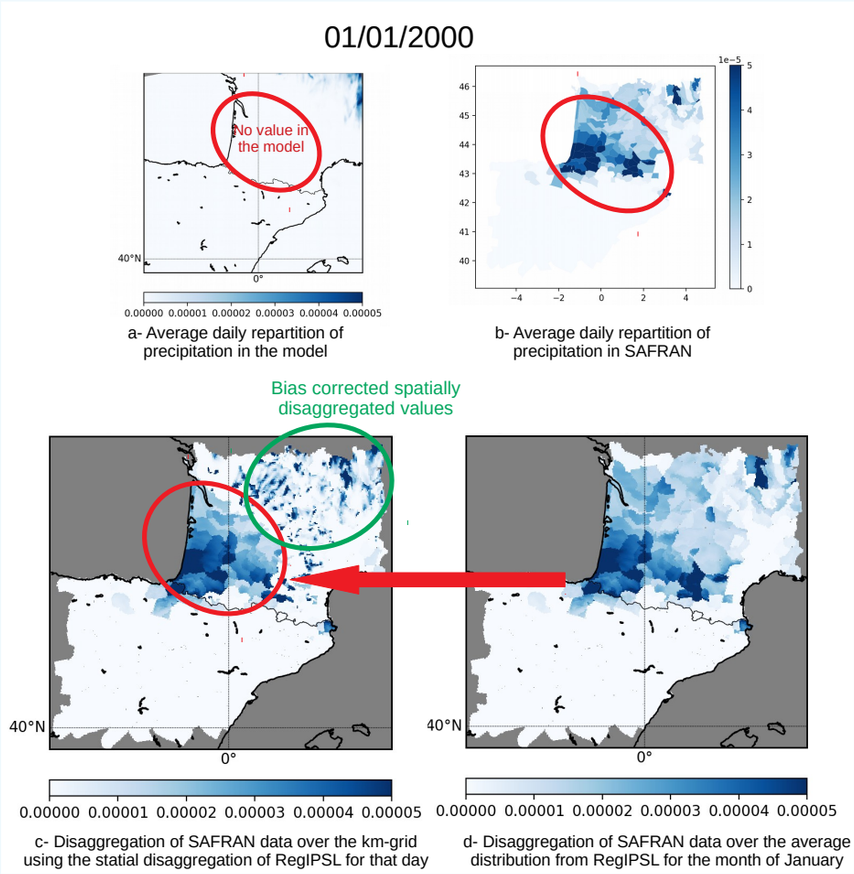


Figure 5.6 – Example of undefined bias at the daily scale for precipitation ($kg.m^{-2}.s^{-1}$) of 01/01/2000. In the outputs of RegIPSL for that day (a- red circle), there is an area with no precipitation, while SAFRAN reports precipitation there that day (b). Therefore, we construct an average disaggregation for the corresponding month (January) and spatially distribute SAFRAN data (d). Over the area with no values in the model, we fill the gap with that average spatial distribution of SAFRAN (c- red circle). For the rest of the area, we apply a bias correction normally (c-green circle).

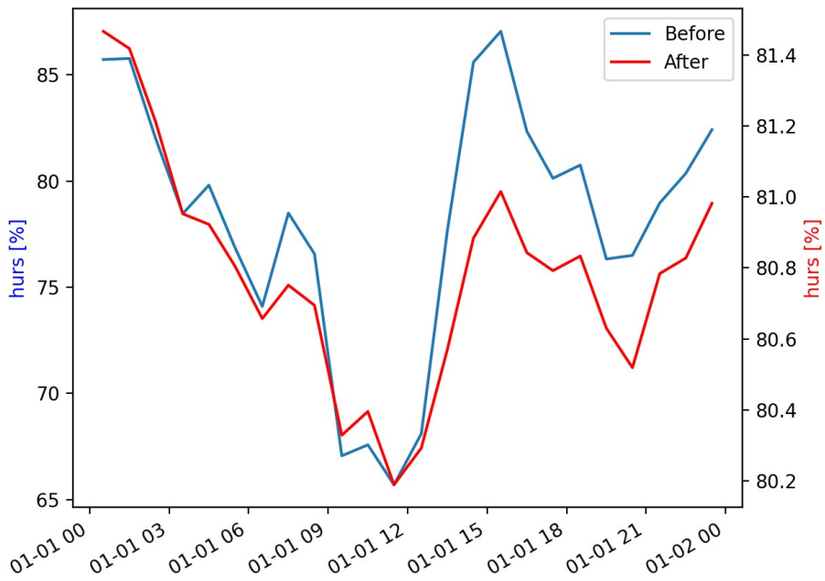


Figure 5.7 – Example of daily distribution before and after variance correction for relative humidity (hurs, %) over a given grid point (lat=42,73°, lon=-0.7896°) for a given day (01/0/2000). The shape of the daily variation is preserved but the scale after the variance correction (red axis) is dampened, with extremes reduced, which avoids an excessive contrast over the day. The shape of the daily distribution is not exactly the same, since a spatial variance correction was also applied to some peaks may also have been smoothed out and spatially redistributed over neighboring points because of it.

- Last, we shift the average back:

$$\forall p \text{ within the polygon, } NewValue_p = NewShiftValue_p + Average_{polygon} \tag{5.7}$$

In the end, the spatial series over the polygon keep the same average value but the standard deviation has been corrected and the exaggerated extremes smoothed out and redistributed over neighboring points (step 6, Fig. 5.5). Similarly, we correct the temporal standard deviation for hourly series over each grid points so the daily average is conserved but the hourly peaks are smoothed out when necessary (Fig. 5.7).

5.3.2 Construction of different forcings: towards sensitivity tests

We have described a general methodology, but we have left out some details which open different options and choices. Depending on these choices, we can construct different forcings to compare to each other. First, there are different choices and options possible to calculate the average of a variable at the polygon level before calculating a bias (Fig. 5.8, options 1, 2 and 3), see section 5.3.2.1. Then for the moment we have presented a methodology which works to construct a forcing over the period common to both initial datasets (SAFRAN and RegIPSL run). However, we introduce another method in section 5.3.2.2 (Fig. 5.8, option B), which would allow to extend the methodology to the full period covered by SAFRAN (1979-2014). This latter method would also attenuate the issues of undefined bias.

5.3.2.1 Choice of integration and bias calculation at the polygon level

SAFRAN variables are given by polygon and by altitude classes (corresponding to every 300m). For the model km-scale outputs, we have access to the orography and to the altitude associated to each points. We can make different choices, favoring either the altitudinal dis-

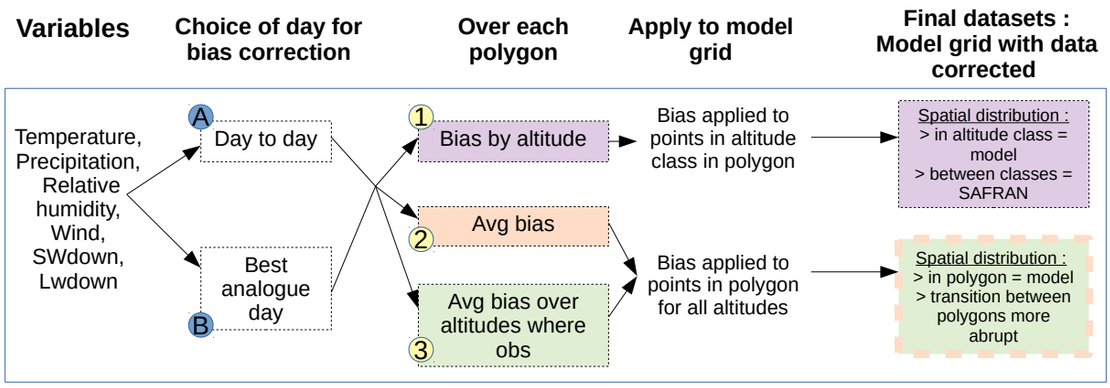


Figure 5.8 – Different options to construct a km-scale forcing with our methodology:
A-B: Choose the pattern of spatial disaggregation (section 5.3.2.2).
1-2-3: Choose how to calculate the bias at the polygon level, favoring the altitudinal distribution from SAFRAN or from the model (section 5.3.2.1).
With the different combinations of choices, we can create different high-resolution forcings and test the sensitivity of the final dataset to each of these choices.

tribution of SAFRAN or the one from the model outputs. This choice only impacts the part of the methodology associated to the bias calculation and its spatial disaggregation (steps 1, 2 and 3, Fig. 5.5). The rest of the methodology is not impacted.

Table 5.3 – Altitude classes in SAFRAN and in the model: each grid points of the model is attributed to a given class

Classes	class 0	class 1	class 2	class 3	...
SAFRAN	0m	300m	600m	900m	...
Model	[0-150m[	[150-450m[	[450-750m[	[750:1050m[	...

1. ***Bias by altitude class***
   In this case, we calculate one bias by polygon and by altitude class. We first attribute each grid points of the model to a given altitude class (Tab. 5.3). Then the model values are integrated at the polygon level with the average calculated for the point within the altitude class (Fig. 5.9). The bias between SAFRAN and the model (step 2, Fig. 5.5) is calculated for each classes and for each polygons.

   Finally when we spatially disaggregate the bias (section 5.3.1.2) for each polygon and altitude class, applying a given bias to all the points within the polygon and the altitude class. With a correction by altitude class, this choice favors the altitudinal distribution in SAFRAN and accentuate a contrast between altitude classes (Fig. 5.10a).

   Since the model orography may not be exactly similar to the one used to construct SAFRAN altitude classes, some points in the model grid are associated to an altitude not covered by the altitude classes of SAFRAN. Therefore these points don't have a calculated bias and are handled as described in section 5.3.1.2 for missing values.

2. ***Average bias over the polygon***
   In this case we calculate one bias per polygon, covering all altitude classes. The model values are integrated at the polygon level and averaged over all altitude classes, weighted by the relative ratio of each class within the polygon (Fig. 5.9). The bias between SAFRAN and the model (step 2, Fig. 5.5) is calculated once on this average for each polygons.

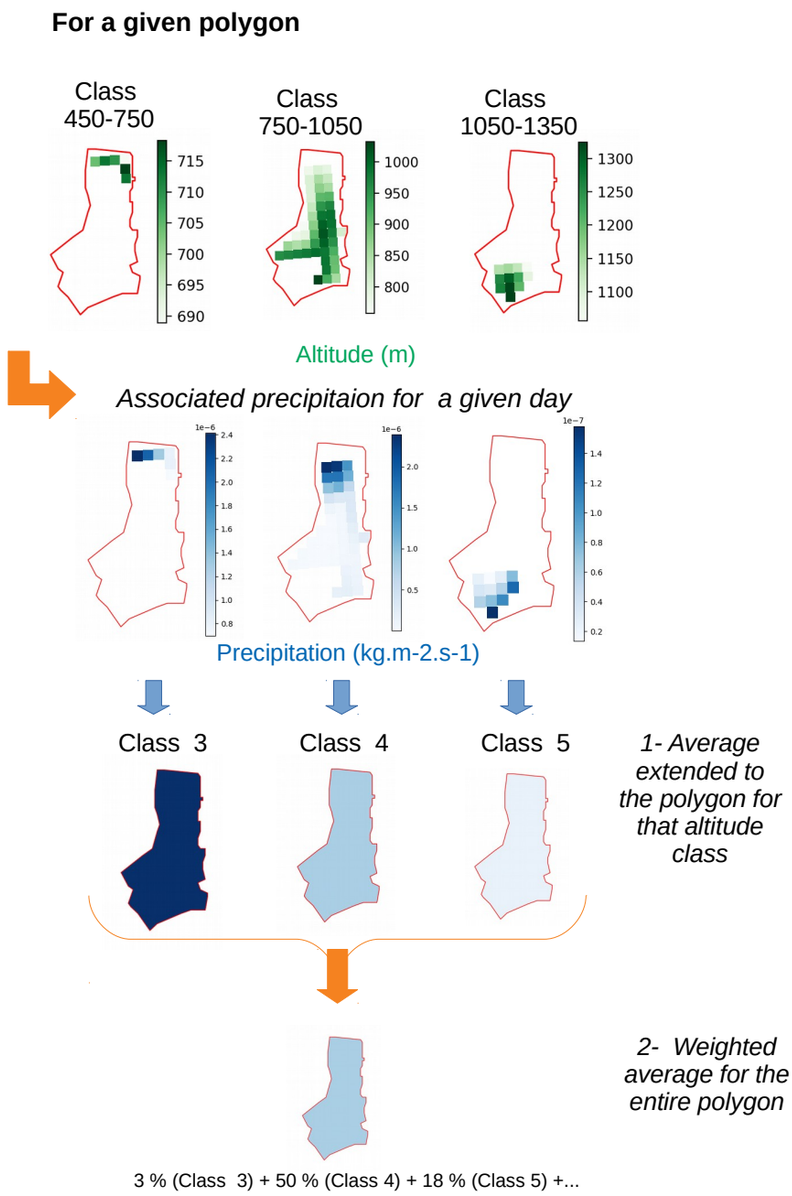


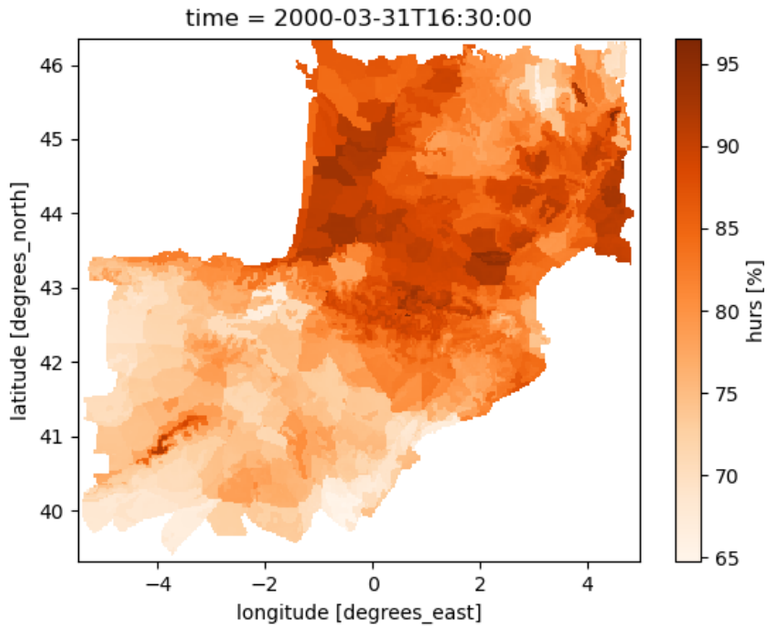
Figure 5.9 – Example over a given polygon of 1 - the integrated average by altitude class, 2 - the weighted average over the entire polygon, prior to the bias calculation. Illustrated in the case of precipitation for a given day.

Finally when we spatially disaggregate the bias (section 5.3.1.2), we apply the bias to all the points within the polygon, regardless of their altitude class. With an average correction per polygon, this choice favors the altitudinal distribution of the model within each polygon. It however increases the artificial discontinuities at the boarder of the polygons in the final product, especially visible over topographically homogeneous areas such in the southern part of the domain (Fig. 5.10b).

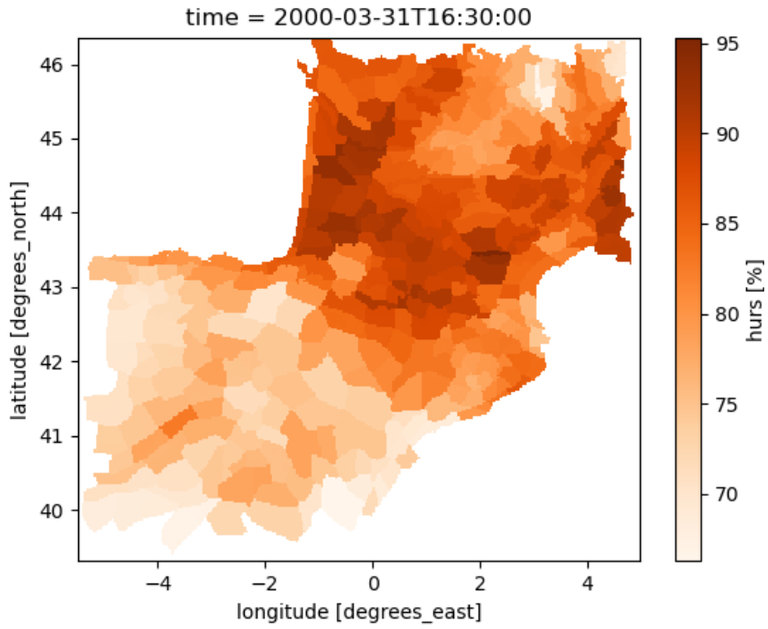
3. ***Average over selected altitude classes***

As explained earlier in the data description, due to the scarcity of weather stations, SAFRAN may be less accurate and reliable at high altitudes. Therefore to calculate the bias, we may want to put more emphasis on SAFRAN data in lower altitudes. We can for instance apply the previous method (***Average bias***) but by calculating the average bias only over lower altitude classes and then applying it to all the points.

For each year, we also have access to the number and location and altitude of all weather stations used to construct SAFRAN. We can therefore filter the altitude classes with and without any weather station for each polygon, and again apply the previous



(a) Relative humidity after the bias calculation for the option 1, bias by altitude class.



(b) Relative humidity after the bias calculation for the option 2, average bias by polygon.

Figure 5.10 – Corrected relative humidity (%) on the km-scale grid, after the full application of the method, for the day 03-31-2000, 16:30. The biases used to correct RegIPSL and disaggregate SAFRAN's data have been calculated with two different options:
1- with a bias by altitude class. This method favors the altitudinal disaggregation of SAFRAN and contrasted transition from one class to the next. It smoothes out the limits between the different polygons, especially when they contain contrasted altitudes classes;
2- with an average bias at the polygon level. This method favors the altitudinal distribution of RegIPSL within polygons, which smoothes the altitudinal transitions in each polygon, but highligths the transitional boarders between polygons.

method (***Average bias***) but by calculating the average bias only over the altitude classes including weather stations.

Comparing those three different way of proceeding will allow to test the sensitivity of the final result to the favored altitudinal distribution (either from SAFRAN or from the model) and to SAFRAN's extrapolation over higher altitudes where weather stations are scarcer.

5.3.2.2 Analogue Days

So far we have been calculating a bias day by day, taking one day in SAFRAN and using it to rescale the same day in RegIPSL outputs. However this method only allows to cover the common period for the two datasets (2000-2009). If we want to construct a forcing to run and test LSM performances, we need it to cover at least a period of 20 years, to cover a period significant enough to include long-term climate variability and slower hydrological processes. We therefore introduce here a method that, if validated, would allow to extrapolate the methodology out of the common period between SAFRAN and the models' runs.

After having integrated the models values at the polygon level for each altitude class (bias by altitude, section 5.3.2.1), we calculate, for each SAFRAN's day, each altitude class and for each variable a spatial correlation between SAFRAN's day and all the model's days (correlation matrix, example of a subset Fig. 5.11).

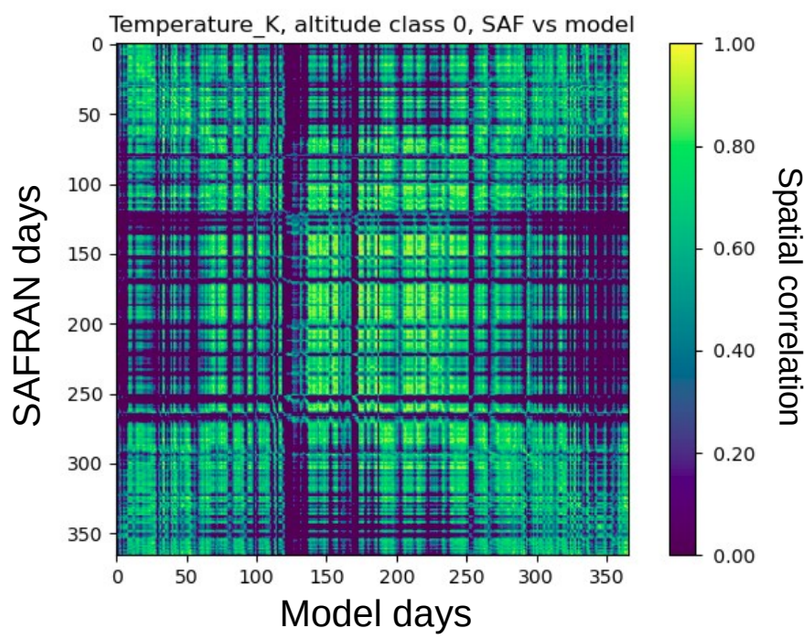


Figure 5.11 – Illustration of a subset of the correlation matrix (from 01/01/2000 to 12/31/2000), for temperature. Spatial correlation day by day for the altitude class 0. The visible diagonal shows that the correlation between one SAFRAN day and its matching day in the model is good. The full correlation matrix extend to 1979-2014 (SAFRAN) x 2000-2009 (Model).

Once we have the correlation matrix, instead of using the day in the model matching SAFRAN's day temporally, we can instead look for the day in the model best matching SAFRAN spatially. For each SAFRAN day, through the correlation matrix, we identify the model day, across all years covered by the model run, in the same month (to be sure to match the season), with the best spatial correlation. We select the model day with the best average spatial correlation, considering the spatial correlation for precipitation and temperature over the first four altitude classes, which cover most of the area of interest and where SAFRAN's data are probably most accurate. We call that selected model day

the "analogue day" corresponding to SAFRAN's day. Then, to temporally and spatially disaggregated SAFRAN's variables of a given day, instead of using the same day in the model, we instead use the "analogue" day. The methodology is developed the same way except the bias and all the rest are calculated comparing SAFRAN's day to the "analogue" day in the model. The "analogue" day is only selected once for each SAFRAN's day and is the same for all variables, to keep the coherence between all atmospheric variables.

This "analogue" day method allows to extend the forcing construction outside of the common period. For each SAFRAN's day before or after the period covered by the model run, there is a possibility to look for its best "analogue" day in the period covered by the model, and therefore use that day to disaggregate at high-resolution for each of SAFRAN's days. The "analogue" day method should also reduce the issue of undefined bias (infobox 15). Indeed, since we choose the best spatially correlated day at the polygons level between SAFRAN and the model, it should highly limit the occurrences of polygons where $Value_{Model} = 0$ and $Value_{SAFRAN} \neq 0$.

Applying that "analogue" day method over the common period and comparing it to the "day by day" method should allow to test and validate the method prior to extending it to outside of the common period.

This "analogue" day technique relates more generally to the commonly used statistical downscaling analogue method (Zorita & von Storch, 1999), where the large scale atmospheric patterns are compared to historical observations to look for the most similar distributions. The main limitation is often that the period over which the analogue is defined needs to be long enough so that a reasonable analogue can be found for each type of possible atmospheric pattern. In our case, if we had access to CPRCM simulations longer than 10 years, it would be interesting to test whether looking for "analogue" days over a longer period significantly changes the results compared to over a ten-year period. More generally, all statistical downscaling methods aim at finding a relationship between the compared datasets over a common period of test, to extrapolate this relationship outside of the studied period where only the coarser scale dataset is available. For instance, to extend CPRCMs outputs without having to run a computationally demanding simulation, Doury et al. (2023) developed an hybrid downscaling approach, combining CPRCMs runs to a neural network. They construct a regional climate emulator using deep learning, relating the spatial disaggregation from the CPRCM to the dataset used for boundary conditions such as ERA-Interim. Then that emulator is used to reconstruct the CPRCMs outputs for the full period covered by ERA-Interim. It would be interesting to compare or eventually combine our "analog day" method introduced here to other products, such as gridded SAFRAN products created with other statistical dowscalling methods or such as the product from Doury et al. (2023), to see how it changes or impacts our disaggregation of SAFRAN outside of the common period.

5.4 First tests

At this stage, we constructed forcings with the different options with the full methodology for (see options in Fig. 5.8):

- option A1: day by day, bias by altitude (Bias Alt IPSL)
- option A2: day by day, average bias (Avg Alt IPSL)
- option B2: analogue day over the common period, average bias (Analogue Avg Alt

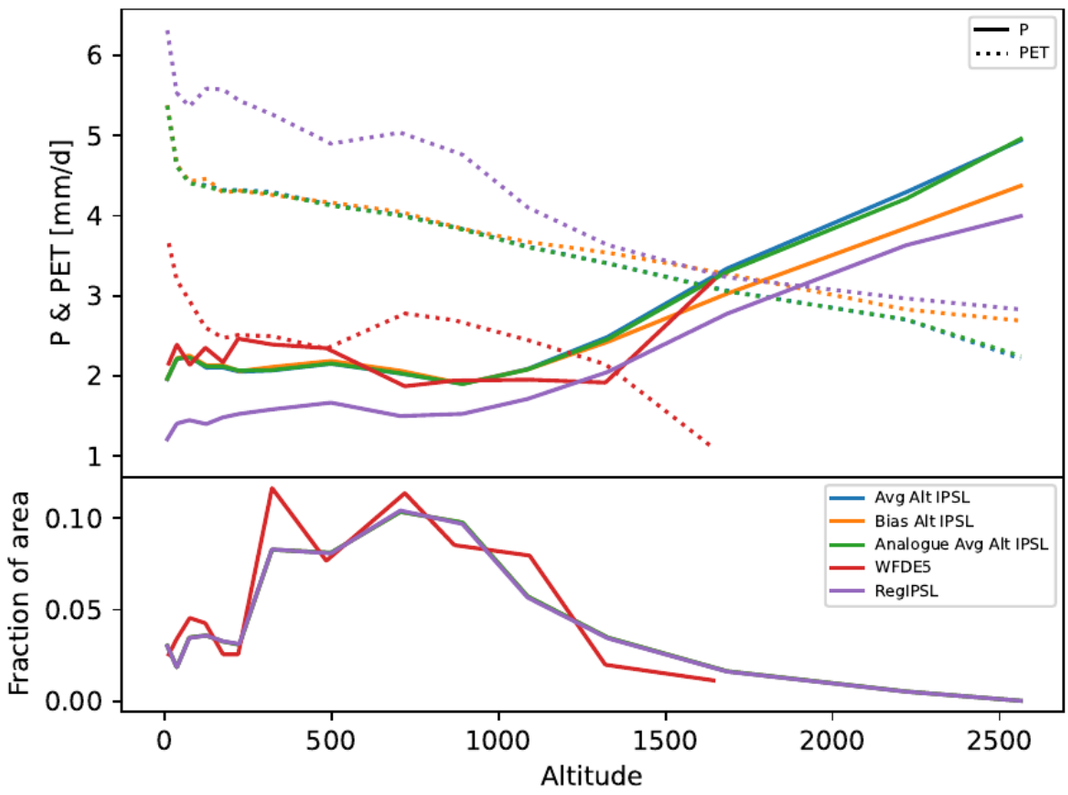


Figure 5.12 – Altitudinal distribution of P and PET for all the constructed forcings (Avg Alt IPSL, Bias ALt IPSL, Analogue Avg Alt IPSL), the initial model outputs (RegIPSL) and a coarser resolution reanalysis (WFDE5). Annual average over the entire area. The bottom graph shows the distribution of the grid points with altitude. The constructed forcings have the same grid point altitudinal distribution than RegIPSL.

IPSL). For this option, we only have the spatial disaggregation and a correction at the daily scale. The sub-daily disaggregation is still left to do.

As a first test, we look at the altitudinal distribution of P and PET, on average over the full area (Fig. 5.12), comparing the different forcing constructed (Avg Alt IPSL, Bias ALt IPSL, Analogue Avg Alt IPSL), the distribution of the model before correction (RegIPSL) and another coarser scale reanalysis, WFDE5 (see infobox 12). Here PET is calculated with the Penman-Monteith equation recommended by the Food and Agriculture Organization (FAO) (Allen et al., 1998).

First we compare the different forcings we created. We can verify that the altitudinal distribution for the methodology applied day by day (option A) of by analogue day (option B) is similar on average (Avg Alt IPSL and Analogue Avg IPSL, Fig. 5.12). The choice of one of these methodology changes the daily spatial distribution used for disaggregation, but it should not and does not impact its yearly average and differences are not detected with this type of graph. The altitudinal distribution is rather dependent on the choice of bias correction (option 1, 2 or 3). Indeed, when we compare the forcing Bias Alt IPSL (option 1) and Avg Alt IPSL (option 2) (Fig. 5.12), we can see a difference at higher altitude in the average P between the two forcings. The contrast between low and high altitude is more marked for Avg Alt IPSL than for Bias Alt IPSL. Bias Alt IPSL favors the relative altitudinal distribution of SAFRAN while Avg Alt IPSL favors the relative altitudinal distribution of RegIPSL. Therefore, this result is in agreement with the hypothesis that the CPRCM model represent a more contrasted distribution of P and high P events at higher altitudes, while SAFRAN data tend to maybe underestimate P at high altitudes. That effect is less marked for the other atmospheric variables involved in the calculation of PET. Over lower altitudes

(<∼1400m), all constructed forcings agree on the altitudinal repartition of P and PET.

When comparing the constructed forcings to the initial RegIPSL outputs, it shows that the bias correction with SAFRAN data overall increased the yearly average of P at every altitude and decreased the yearly average of PET over lowlands. So RegIPSL seems to, on average, underestimate P and overestimate PET over low lands compared to SAFRAN (and therefore the constructed forcings). Looking at that same altitudinal distribution but on monthly average (not shown), we can better assess that this underestimation of P in the initial RegIPSL outputs is mostly over the drier months (May to September) and over lower altitudes, while in wetter months at high altitudes, the constructed forcings tend to have lower P than RegIPSL. Our bias correction method based on SAFRAN allows to correct the dry bias over low lands and summer months in RegIPSL. Looking at the shape of the distribution, as expected, the slope of RegIPSL is closer from the slope of Avg Alt IPSL than from the slope of Bias Alt IPSL, since the first one favors the relative altitudinal distribution of RegIPSL and the other the one from SAFRAN. Overall, Bias Alt IPSL (and to a lesser extend Avg Alt IPSL) tends to have a lower slope, with a less contrasted relative altitudinal distribution. Over wetter month especially, the altitudinal contrast in RegIPSL is a lot higher than in the constructed forcings. Our bias correction method based on SAFRAN tends to reduce the altitudinal contrasts in P compared to RegIPSL, especially for the forcing Bias Alt IPSL.

It would have been interesting to compare our final forcings to the original SAFRAN distribution. However, since it is not a gridded product but is based on geographical zones of very different shapes and altitude classes, it is more difficult to use and we had no time to construct such a calculation. We are currently looking into gridded product of SAFRAN from statistical downscaling methods, used at least over the french portion of our area of interest, to test how it compares to our results but it is still on going work.

Now we compare the constructed forcings to the coarser atmospheric reanalysis WFDE5. This is a reference atmospheric reanalysis dataset at coarser scale (infobox 12). We can first see that at coarser resolution, high altitudes are smoothed out and are missing from the forcing. Here, there is no points in WFDE5 above 1600m (bottom graph Fig. 5.12). Therefore we can only make a comparison for lower altitudes. The distribution is also more noisy due to the fewer grid points covering the area of interest. The constructed forcings are in better agreement with WFDE5 for P over lower altitudes than the initial RegIPSL outputs. Therefore, our bias correction methodology using SAFRAN data effectively improves the average P distribution from RegIPSL outputs over low lands compared to the observation-based datasets. PET however is still very different, comparing WFDE5 to the other forcings. This variable is more difficult to assess since it relies on non-linear relationships between a number of atmospheric variables. A closer look to the different variables involved would be needed here.

These first results are in agreement with the hypothesis we make in 5.3.2, for the different options of forcings construction (Fig. 5.8) and the objectives we had for constructing these forcings.

5.5 Next steps and conclusion

The forcings constructed still need to be further tested, to better look at the contrasts, extreme values, spatial distribution and diurnal cycle. It would also be interesting to finish

INFO BOX 16

Remaining issues

- **Issue of extremes**

The disaggregation method we present aims at contrasting the values over the polygon. Sometimes that contrast is exaggerated and we correct it by rescaling the standard deviation. However that correction is not always sufficient, and some extreme values remain. It is one limit of this methodology.
One way to avoid it would be to impose a maximum limit to the variables to avoid unphysical values. It is what we do for the relative humidity (hurs) which can't be above 100%. This choice however tends to slightly reduce the average at the polygon level.

- **Issue of transition from one day to the next**

The bias is calculated at the daily scale and then applied over all the sub-time steps of the associated day. Therefore the transition is quite abrupt from one day to the next. It could be interesting in the future to smooth the transition from one day to the next, using for instance a sliding window covering neighboring time steps. Since it is in the middle of the night, in this region, such a forcing should not highly impact the outputs of the LSM.

to construct the forcings with the other combinations and options not tested yet, such as the option to filter SAFRAN data depending on the altitudes with higher or low concentration in weather stations (option 3, 5.3.2.1). Also, we only tested the disaggregation method using RegIPSL while we had other CPRCM outputs available. It would be interesting to see how sensitive our final results are to the chosen CPRCM. We also want to run a LSM with the constructed forcings to see how sensitive the LSM outputs are to each of our hypothesis, choices and options.

We also need to run a LSM with these forcings to see how the LSM performances changes using a km-scale resolution dataset compared to results with coarser reanalysis such as WFDE5. First tests to run ORCHIDEE with these constructed forcings have been started but are not completed yet. It would be particularly interesting to test the LSM run at high resolution to reproduce streamflow over small mountainous catchments, if possible little impacted by human activities, since we assume that due to the increased topographic contrast, its performance should increase. Again for that step, several LSM could be run with these forcings to test the performance of the ensemble.

Finally, we need to further look at the compared performance of the forcings constructed using a day by day constructed bias or based on the "analogue day" methodology over the common period (2000-2009). If the final LSM performance is similar in both methods (or even maybe improved by the "analogue" method), it would validate the "analogue" method and we will be able to extend the application of that method to construct a forcing covering the full extend of SAFRAN data (1979-2014, still being extended towards 2021). This will enable to run LSMs over a period long enough to cover long-term variability, and to test their performances against data collected during the LIAISE field experiment.

This disaggregation method seems promising to obtain km-scale atmospheric forcings to

run LSMs over the area. It should enable to better represent the contrast between high and steep mountainous areas and the hydrological redistribution of water towards valleys. If the LSMs are shown to perform well over catchments little impacted by human activities, it will also enable to test the implementation of anthropic processes such as dams and irrigation in LSMs and test their performances over highly managed areas.

KEY POINTS TO REMEMBER

- There is a need for km-scale atmospheric forcings to run and test LSMs at high resolution and improve the representation of regional scale processes, altitudinal contrasts and include human water use.
- Current reanalysis based on data assimilation have been validated at coarser scale and their spatial and sub-diurnal disaggregation are limited by the interpolation methods, which tend to attenuate extreme events and are biased in regions with scarce observations.
- High resolution convection-permitting regional climate model (CPRCM) allows for a physically coherent spatial and sub-diurnal distribution at km-scale. However, the internal variability of the model introduces a divergence between the model outputs and the real distribution of atmospheric variables during the period covered by the run.
- We construct a methodology combining the two types of data to construct a km-scale forcing: a rather coarse observation-based dataset (SAFRAN, 1979-2014) and the km-scale outputs of a CPRCM (RegIPSL, 2000-2009). We keep the spatial and sub-diurnal physical coherence of the model outputs but scale it to match the average of the observation-based dataset.
- We test different options to test the sensitivity of the final results to how the rescaling is done. For instance, favoring the relative altitudinal distribution from SAFRAN or from the model, showing a higher altitudinal contrast of precipitation in the latter case.
- The methodology could be extended to construct a full forcing covering more than 20 years using "analogue" days.
- Still further tests are needed but the method and the results seem promising.

CHAPTER 6

Conclusion and perspectives

During this thesis work, I have explored the knowledge we have about measuring and modeling the water cycle, with a focus on streamflow. I reviewed the observed and predicted changes in discharge found in the literature, along with the different types of models used to describe, reproduce, attribute and project these changes. In the following chapters, I combined different types of models, to highlight their respective strengths and limits, and identify and quantify the changes in streamflow, related to both climate change and surface characteristic changes. To better decompose, understand and attribute the drivers included in the surface characteristic changes related to streamflow dynamics, process-based modeling needs to be improved to include new and more detailed processes, at higher resolution. This led us to the next chapter towards land surface modeling at km-scale. In this final chapter, I summarize the main conclusions about my work, draw up the final picture I have on these problematics, along with my understanding of the perspectives, next steps and challenges, as I understand them at the end of this thesis.

Contents

6.1 Conclusion

6.1.1 Main results and messages

- The Earth system is complex and hydrological processes are highly related to both atmospheric and land surface dynamics. Furthermore, these dynamics are evolving, both with natural and anthropogenic climate change and with more direct influences of human intervention, modifying land use, using and redistributing water. This was our main problematic, how to better understand and detect the role of each climatic and anthropogenic drivers in discharge evolution.

- Different types of model exist to represent at best this complexity, with different functions and different uses. We used a complex LSM to represent the changes in hydrological dynamics in response to climate change, to attribute the effects of different climate characteristics (changes in averages, in intra-annual distribution dynamics...). We combine it with a parsimonious model. The LSM allows to attribute the effects of climate variability projected on to the parameter of the parsimonious Budyko framework. In turn, this framework gives a structure for interpreting the outputs and compare it to observations, highlighting and quantifying the limitations of the LSM due to missing or ill-representation of surface processes.

- Over Europe, we diagnose that the main climate-related changes in discharge are attributed to changes in average precipitation, which tend to increase discharge in Northern and Western Europe and decrease it in Southern and eastern Europe. Then secondary climatic drivers identified are potential evapotranspiration and the intra-annual distribution of precipitation. The first one, with and increase of evaporative demand all over Europe, tends to decrease streamflow, especially in Northern Europe where water is abundant. In water-limited areas such as around the Mediterranean area, changes in evaporative demand are less impactful and changes in precipitation distribution plays a more important role in discharge generation.

  However, these effects of climate change on discharge are small compared to the effects of non-climatic drivers, about ten times larger, especially over the South of Europe and over the Iberian Peninsula. Our method is only a diagnostic, quantifying the total effect of these non-climatic factors but does not allow to attribute specific changes to specific drivers, since these factors and the related processes are not represented in the LSM. We can only hypothesize that irrigation, groundwater uptakes, dams management and glacier melting are some probable drivers of surface dynamics changes impacting discharge.

- Land surface models tend towards more and more complexity to represent the Earth system and towards higher resolution (km-scale), to work at a scale more relevant to decision makers taking into account human activities and better represent hydrological dynamics. This is still a great challenge, with more tests and developments needed to use LSM at such resolutions. We contributed here with our work, with first steps towards a high resolution forcing to test LSM at km-scale over a highly contrasted area, covering from mountainous catchments to highly managed and irrigated valleys in a semi-arid climate.

6.1.2 Reflexion around earth system modeling and hydrology

Models are used to decompose a problem and express it in simpler, more understandable elements and links. It is a useful tool to address a question otherwise too complex to be fully apprehended. And if any, the earth system is a complex system, where a myriad of variables are embedded and connected, at different spatial and temporal scales. With the development of computer systems, over the past fifty years, more and more complex models have been constructed to try to reproduce and understand the main drivers of the earth system and its associated natural cycles. It allows to test the sensitivity of the modeled system to specific drivers, and comparing the resulting patterns of change to observed trends, to attribute the main sources of changes.

More and more complex models have been developed in many scientific fields (atmospheric physics, agronomy, hydrology...), starting from more simple parsimonious models based on few relationships and empirical parameters and going towards complex models driven by physical processes. These models are getting more and more successful to reproduce relatively complex interactions for "natural" systems, at the scale at which they were designed. For instance, atmospheric and climate models have shown that over the past century, climate has been changing at a rate that natural variability and long-term cycles only cannot explain, addressing the responsibilities of human activities and carbon emissions in these changes. In the complex earth system, these detected atmospheric changes impact the full energetic, carbon and water cycles. Focusing on water, managing the resource has been a stake for human societies almost at all times, water being essential for domestic use and many other activities, with irrigation being developed since prehistorical times. Therefore hydrological models have been developed early on, mostly aiming at reproducing streamflow at best and predict short term evolution, to best operate water management infrastructure, estimate flood risks and water available for irrigation and other water demands. With improvement of modeling tools and in the context of climate change and development of water management, hydrological models are also getting more and more complex, towards physical-based models including more and more relationships with atmospheric processes, vegetation... Generally, all process-based models started from a given scientific field (i.e. hydrology or climatology) and tend to improve by including more and more processes (representation of atmosphere, vegetation, soil processes, surface hydrological schemes...) from other scientific fields (atmospheric physics, chemistry, agronomy, biology, geology, pedology, hydrology...), trying to address the complexity of the full system through multi-discipline integration.

However, it is important not to get lost in this increasing complexity and understand its limitations. Complex models will always not be complex enough to be a perfect representation of the Earth system. Adding processes also inserts new sources of uncertainty and often increases the need in inputs. We need to understand for each model, the temporal and spatial scales and eventual environment over which they perform best, often depending on how they were designed. Furthermore, an increased complexity also complicates the interpretation of the outputs. This is where more parsimonious and less computer intensive models can be very interesting to use, as simpler operational and diagnostic tools. Indeed, these parsimonious models are efficient at reproducing the integrated dynamics of complex systems (modeled or observed) but with no or few details on the underlying processes. They can therefore give a framework to understand and compare observations and the outputs of complex models as we have shown in our work. We project the complex model outputs with numerous degree of freedom on to the parsimonious and more restricted framework,

which facilitates its understanding and the comparison to observations. It helps to show the shortcomings of complex process-based models. In turn, process-based models are the only way to understand the causes of detected changes and therefore develop predictions and adaptation scenarios.

There is a necessity to keep track of the objectives, and too much complexity may not be the solution if it is associated with too much uncertainties. We need to assess and quantify these uncertainties and to not forget simpler tools and their respective qualities, which may be more adapted or good complementary tools in given cases.

6.2 Perspectives

6.2.1 Reflexion around regional scale modeling and high resolution

Models continue to improve, with better representation of processes, either explicitly by an increased resolution or through improved parametrization and coupling. The current dynamics goes toward increased resolution to run process-based models, allowing to represent more spatial contrasts and decompose smaller scale processes. Regional models and kilometric-scale simulations also reach a resolution more adapted to address local issues and be more appealing to decision-makers for operational uses. They enable us to develop the representation of processes such as human water management at a relevant scale, directly addressing regional problematics and operational issues. This is one important step to get a full representation of the water cycle and a more accurate estimation and prediction of water availability, since these missing processes can highly impact regionally the amplitude and the variability of streamflow, as we have shown in our work.

Hyper-resolution (km-scale) solves but also raises challenges. It has been addressed over the past decade, such as in the paper of Wood et al. (2011) and its answer by Beven & Cloke (2012). Physical-based models have been developed and validated at large scale, with a parametrization designed accordingly. There is a need to adapt them at a regional scale. In some cases, the higher resolution allows for a direct improvement of processes representation. For instance, it improves the representation of convective processes in climate models, replacing the parametrization needed at coarser scale (Lucas-Picher et al., 2021). However, in the case of hydrological dynamics, further tests are still needed to assess the performance of models at higher resolution. Many processes are still missing, or implicitly covered by the large scale interpretation, such as sub-terrain and lateral flows and root water uptakes, which effects are likely to be more impactful in the final equilibrium between evapotranspiration and streamflow at higher resolution with larger spatial and altitudinal contrasts. This raises the question of whether the current physical representation of processes and the parametrization related to it are sufficient to address the high heterogeneities at regional and km-scale. Therefore there is a need to test whether missing processes have a higher impact on model performances than uncertainties related to parametrization. If not, then there is first a need to refine the representation of processes already accounted for to reduce that uncertainty. There is a compromise to be found between constructing a model flexible enough to cover spatial heterogeneities and local specificities and keeping a model robust enough to be generalized in space and time. If the processes are rely too much on specific parametrization to match regional conditions, it will raise similar limits as for adjusted parsimonious models: such models are accurate to represent a specific area and period but may not be adapted to be used in a different context where we don't have

access to the same level of data, such as for another region or for projection scenarios. It can be useful for short- and medium-term diagnostics but it is not the main goal when physical-based models were initially designed to complement climate projection scenarios. The potentialities and model parametrization and performances at km-scale therefore need to be tested, with the hope that progress in data acquisition and assimilation, in computational efficiency and in the understanding and representation of the underlying physical processes will reduce the uncertainties and limits associated to parametrization. Ideally, the more we understand the processes and their complexity, the less parametrization will be needed and the more the models will be efficient and flexible. There will also be more demanding in computational power, the more we complexify the represented processes and we increase the resolution.

For other processes, it has already been shown that high resolution can improve our estimation, even compared to observational networks, such as for precipitation in mountainous areas (Lundquist et al., 2019), where weather stations are scarce and field measurements and detection more difficult. This raises another challenge linked to higher resolution: the lack of high resolution quality data, both in input of the model and to validate it. As we have discussed in chapter 5, reanalysis datasets present global scale gridded data derived from most of available observations, answering at best the need for a large scale homogeneous dataset. They are widely used, improving with each generation of data-assimilation models and increasing remote-sensing data from new satellite missions and ground-based radars. And still these datasets are limited in resolution, with the questions of optimal disaggregation towards kilometric scale still to be answered. Also data assimilation are not homogeneous in time, with more and more data assimilated over recent period, which improves the quantitative estimations for more recent periods but complicates the representation of trends, such as for snow cover, which is a variable of interest in water management (Monteiro & Morin, 2023). To validate models, when data are not available, such as measurements of evapotranspiration over large areas, or when we suspect that the model can be better than the observations, such as precipitation in mountainous areas, we need to use proxy variables and simpler models to test the outputs. For instance, to test the modeling of precipitation in mountainous areas against observations, we can test the effectiveness to reproduce streamflow in small and steep catchments where evapotranspiration is small, with a simple water balance model using one or the other precipitation product (Behrangi et al., 2014). Otherwise, in order to test and validate models, there is a need to improve data acquisition or at least to have access to rich observational datasets such as the one gathered during LIAISE field campaign, to test models over a rich set of different data, and ideally over different contrasted areas to attest of model robustness.

The case of adding a physical representation of human activities and water managements further expand these two challenges. First, as for many types of observations and measurements, data are not available equally everywhere, with some data more accessible than others, with a quality which may also be dependent on the area. In the case of information related to water consumption and water management, there are added political and economic considerations and confidentiality issues, increasing the difficulty to get shared and qualitative data. Therefore, adding human activities faces both the question of how to represent and parameterize these new processes to be included, and the limitation of uneven data availability. Both questions are actually related. With new processes to be added, new data are needed, such a water quantities for irrigation, water demand or dams operations. Again here not all of these data exist or are easily accessible everywhere. Should we choose

(and is it possible to create) a representation as generic as possible, relying only on widely accessible data and a generic parametrization? Or should we construct options to adapt to local specificities and data availability? The first case would allow to model human water usage in places where data are scarce and would be more suitable for projections. However, the second case would be more interesting for regional stake-holders and decision makers and to test specific scenarios. If for physical processes we can imagine that we can improve our understanding of the system towards finding the essential and generalizable processes, there is no certainty that it will be possible to find a rationalized representation of human behaviors, generalizable in space and time. In the end, I believe there is a need to construct different options, which can be activated depending on the data availability and scenarios tested, allowing for more flexibility to account for societal and economical diversities. In places where data are scarce or not accessible and for projection scenarios, new hypotheses and modeling sub-processes for water demand, water consumption, installation performances (irrigation, dams...) are needed to palliate to this lack of information (example in Neverre & Dumas (2015)). This involves new scientific fields such as agronomy, economy and social sciences, to reproduce decision-makers water regulation and laws along with water demand and users behaviors and to predict adaptation pathways.

To be met, all of these challenges necessitate to have different communities working together, to pool resources for computing power and expertises and construct common frameworks for consistent validation, sensitivity tests and comparison purposes.

6.2.2 Next steps and challenges

- Test LSMs performances at km-scale
  - Define common framework and performance indices to compare models to each other.
  - Make the best use of available data, ideally using rich and diversified field observations over contrasted areas.
  - Test models sensitivities and uncertainties, to see if the internal biases and uncertainties linked to parametrization are higher or lower than the effect of missing processes on model performances.
  - Improve the representation of hydrological processes, accounting for higher spatial and altitudinal contrasts at km-scale.
- Incorporate human activities in models
  - Need for different communities to work together: integrate economic and social sciences to agronomic, hydrological sciences and atmospheric physics.
  - Construct water demand and management models to add to physical models, along with improving the attribution of water uptakes from the hydrological system.
  - Associate to these models a social and economic reflexion, to account for the diversity of behaviors and management choices, and construct projection scenarios. Reflect on the balance between generalization and relevant representation.

APPENDIX A

Article submitted to *Water Resources Research*

AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR034509

Budyko Framework Based Analysis of the Effect of Climate Change on Watershed Evaporation Efficiency and Its Impact on Discharge Over Europe

Julie Collignan[1], Jan Polcher[1], Sophie Bastin[2], and Pere Quintana-Segui[3]

[1]Laboratoire de Météorologie Dynamique/IPSL, Ecole Polytechnique/CNRS, Paris, France, [2]Laboratoire Atmosphères, Observations Spatiales/IPSL, CNRS, Paris, France, [3]Observatori de l'Ebre, Universitat Ramon Llull, CSIC, Roquetes, Spain

Key Points:

- Evaluate the role of climate change on the evolution of the discharge of European rivers
- Using the Budyko model as a detection tool for changes in hydrological behaviors of watersheds
- Analyze the effects of changes in the intra-annual distribution of precipitation on the evolution of the annual discharge

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Collignan,
julie.collignan@lmd.ipsl.fr

Citation:

Collignan, J., Polcher, J., Bastin, S., & Quintana-Segui, P. (2023). Budyko framework based analysis of the effect of climate change on watershed evaporation efficiency and its impact on discharge over Europe. *Water Resources Research*, *59*, e2023WR034509. https://doi.org/10.1029/2023WR034509

Received 5 APR 2023
Accepted 22 AUG 2023

Author Contributions:

Conceptualization: Julie Collignan, Jan Polcher
Formal analysis: Julie Collignan
Methodology: Julie Collignan, Jan Polcher, Sophie Bastin
Resources: Pere Quintana-Segui
Supervision: Jan Polcher, Sophie Bastin
Writing – original draft: Julie Collignan
Writing – review & editing: Jan Polcher, Sophie Bastin, Pere Quintana-Segui

Abstract In the context of climate change, the stakes surrounding water availability are rapidly intensifying. Decomposing and quantifying the effects of climate on discharge allows us to understand their impact on water resources better. We propose a methodology to separate the effect of change in the annual mean of climate variables from the effect of the intra-annual distribution of precipitation. It combines the Budyko framework with land surface model (LSM) outputs. The LSM is used to reproduce the behavior of 2,134 reconstructed watersheds across Europe between 1902 and 2010, with climate inputs as the only source of change. We fit a one-parameter approximation of the Budyko framework to the LSM outputs. It accounts for the evolution of the annual mean in precipitation (P) and potential evapotranspiration (PET). We introduce a varying parameter in the equation, representing the effect of long-term variations in the intra-annual distribution of P and PET. To better assess the effects of changes in annual means or intra-annual distribution of P, we construct synthetic forcings fixing one or the other. European results show that the trends in the annual averages of P dominate the trends in discharge due to climate. The second main climate driver is PET, except over the Mediterranean area, where changes in intra-annual variations of P have a higher impact on discharge than trends in PET. Therefore, the effects of changes in the intra-annual distribution of climate variables are to be addressed when looking at changes in annual discharge.

Plain Language Summary Water availability is a challenge for all of society. Various competing activities rely on this resource, and its scarcity can lead to social, economic, and environmental conflicts. With climate change, river discharge and, more generally, the full water cycle is impacted. Furthermore, multiple human actions such as dams and irrigation concurrently change the balance of the water cycle over watersheds. To comprehensively understand the dynamics of discharge, it is essential to analyze the potential influence of direct human activities alongside the impacts of natural climatic factors. Models are a way to represent reality with an understanding of the physical phenomena included. They can be used to represent the behavior of watersheds without human intervention. In light of this, we have developed a methodology to highlight the climate factors impacting discharge. Annual discharge changes are driven mainly by changes in annual precipitation over Europe. The increasing temperature leads to an ever-growing evaporative demand and is the second most impacting factor over most of Europe. However, in the Mediterranean area, where water is more limited, changes in the seasonality of precipitation have a higher impact than changes in the evaporative demand.

1. Introduction

Water is a key resource for the whole of society and both its excess and scarcity can lead to challenging economic, environmental, and social issues. Understanding the hydrological cycle and how it evolves due to a changing climate is a significant challenge of this century.

Over the past century, several studies have shown the impact of climate change on climate variables in Europe. Annual precipitation increased between 1901 and 2005 over most of Europe except the Mediterranean area, where they tended to decrease (Christidis & Stott, 2022; Douville et al., 2021; Knutson & Zeng, 2018). Trends per decade are less significant due to the high inter-annual variability of precipitation P (Douville et al., 2021). Trends in potential evapotranspiration (PET) are linked to an increase in the energy available at the surface, which is highly correlated to rising temperatures (Douville et al., 2021; Vicente-Serrano et al., 2014). Few studies have directly examined European PET trends, except over the Mediterranean area, where studies have shown a

significant increase in PET over the end of the century (Kitsara et al., 2013; Vicente-Serrano et al., 2014, 2019). The intra-annual variations of climatic variables are more difficult to assess, and only a few indices exist to measure the inter-annual changes in the distribution of climate variables. For example, García-Barrón et al. (2018) defined indices to assess the evolution of the intra-annual cycle of *P* over time throughout the Iberian Peninsula. At the end of the century, they identified a shift of the main rainfall periods toward autumn, especially over the Atlantic basins, and an increase in the inter-annual variability of the intra-annual cycle, especially over the Mediterranean basins. For precipitation, studies have shown that not only is the annual average of *P* changing, but there are differences between summer and winter, depending on the area (Christidis & Stott, 2022; Zveryaev, 2004). Moreover, over the past few decades, extreme precipitation events affecting this area have significantly intensified (Ribes et al., 2019). Therefore, it is important to investigate the effects of changes in the annual averages of climate variables along with the effect of changes in seasonality and intra-annual distribution of these variables. The distribution of *P* within the year and its coupling or decoupling from the atmospheric demand PET will influence water partitioning between evapotranspiration and discharge on the annual scale.

Transformations in different climate variables governing the water cycle alter the equilibrium in the water balance over the different watersheds, thus impacting the discharge of rivers. Milly et al. (2005) showed that worldwide discharge trends are and will continue to be significantly impacted by changes in climatic factors. Over Europe, statistically significant trends in discharge are observed in historical records (positive in the northern region and negative in the south and east). These trends are spatially coherent with precipitation changes (Stahl et al., 2010; Vicente-Serrano et al., 2019). Y. Yang et al. (2018) show that discharge is less sensitive to PET changes than to changes in *P*.

The effects of intra-annual variations of *P* on discharge are primarily considered in the literature through the study of seasonality and annual extremes of *P* and PET in order to examine their impact on floods (Douville et al., 2021; Milly et al., 2002; Rottler et al., 2020), drought events (Douville et al., 2021; Vicente-Serrano et al., 2014), and more generally, on discharge peaks (Bouwer et al., 2008; Tuel et al., 2022). Stahl et al. (2010) found the trends in discharge over the end of the century were disconnected between summer flows and winter flows for an ensemble of small near-natural catchments in Europe. Blöschl et al. (2019) showed that increasing autumn and winter rainfall led to increased floods in northwestern Europe.

However, rivers are also highly managed, and human activities are an important driver of change in how watersheds function (Ficklin et al., 2018; Riedel & Weber, 2020). A significant difficulty in analyzing the effect of climate on historic discharge changes is decomposing the effects of the different drivers of change and isolating them from each other to better understand their relative importance (Stahl et al., 2010; Vicente-Serrano et al., 2019). Several studies have concentrated on catchments that are regarded as near-natural or unimpaired in order to investigate the effects of climatic changes on discharge (Stahl et al., 2010; Y. Yang et al., 2018). However, this highly limits the areas studied (Vicente-Serrano et al., 2019), especially in Europe, where the high population density and long history of water management limit the study to small catchments (Stahl et al., 2010).

Another approach is to use models to separate the factors involved in discharge changes; different types of models have been developed. Models relying on a few variables and adjusted parameters are favored for their simplicity of use and interpretation. One example is using statistical models fitted over specific areas, such as linear regressions (Bouwer et al., 2008; Ficklin et al., 2018; Vicente-Serrano et al., 2019). More complex models integrate nonlinear relationships and physical boundaries. However, all these parsimonious models are empirical: they rely on adjusted parameters over the area and the time period studied and lack a clear physical meaning. The parameters often cannot be generalized and transposed to other areas or future climates (Coron et al., 2014; Reaver et al., 2020).

Other methods are physical-based hydrological and land surface models (LSMs). They require more data and computational power. They do not always accurately represent a whole real hydrological system depending on which processes are included in them but allow a meaningful assessment of hydrologic aridity (Y. Yang et al., 2018). They have grown increasingly complex and are able to reproduce the behavior of watersheds and to model "natural flow" regimes (Decharme et al., 2019; Gudmundsson et al., 2017; Schneider et al., 2013; F. Wang et al., 2018). However, due to their complexity, it is more difficult to decompose the effects of individual climate factors and to interpret their outputs than with other simpler models.

In light of this, we propose a tool that combines the simplicity of the more empirical model with the heightened performance and complexity of the physical-based model to better understand the phenomena encapsulated behind the adjusted parameters.

We use here the well-known and widely used empirical Budyko framework (Mianabadi et al., 2020). It is predicated upon utilizing the annual mean of water and energy balances at the watershed scale (Tian et al., 2018), taking into account the water and energy limitations of the physical system. It was initially conceived over multiple catchments. Parametric equations were developed to introduce an empirical parameter adjusting the framework to the specific evaporation efficiency of each catchment over an equilibrium period (H. Yang et al., 2008; L. Zhang et al., 2004). However, equilibrium disruptions, due to climate change or any other direct human activities and vegetation change, highlight limitations to the model. Moreover, most disruptive features are concurrent. The parameter introduced has no evident physical meaning and is just a well-adjusted proxy to E/P over a specific catchment and period. There is no straightforward method to attribute changes in the adjusted parameters to specific climatic or nonclimatic features (Berghuijs et al., 2020; Reaver et al., 2022), as for any parsimonious model with calibrated parameters.

To focus on the effects of climate change, the present study applies the Budyko framework to the outputs of a state-of-the-art LSM. The latter represents the constant physical behavior of watersheds. The only source of change in the dynamics of the modeled watersheds is the evolving climate variables introduced. Using LSM outputs also allows for adjusting the near-surface atmospheric variables to more adequately decompose the effects of the different elements of climate variability and change.

This article is organized as follows: Section 2 covers the methodology developed. It describes the Budyko framework with its underlying hypothesis and limitations and the state-of-art LSM. Then, we describe how we apply the framework to the chosen LSM. Here, we create Synthetic forcings to test if our methodology yields an optimal analysis of the effects of different aspects of climate change. We also explain the use of the time-moving window to examine temporal trends in the different climatic effects. In Section 3, we present the results of the effect of different elements of climate change across Europe (changes in annual averages against changes in the intra-annual distribution of climate variables) on discharge trends over the past century. Section 4 provides a comprehensive analysis of the advantages of our findings, while also highlighting the present constraints and areas for further investigation. Finally, in Section 5, we summarize our conclusions.

2. Methodology

2.1. The Budyko Framework

2.1.1. General Presentation

Over watersheds considered as closed systems, the water balance Equation 1 applies when explaining the equilibrium between the variables of the hydrological cycle: the river discharge (Q), the evapotranspiration (E), the precipitation (P) and the change in the water storage over the watershed between two-time steps (ΔS).

$$P - \Delta S = Q + E \tag{1}$$

Long-term, ΔS can be negligible. Ideally, this hypothesis should be applied over a long enough period that the system's equilibrium is reached (L. Zhang et al., 2008). It also supposes no external disturbances impact the water budget, such as groundwater mining or water transfers to or from other basins.

The Budyko framework, which is frequently used in hydrological research to study the partitioning of P into E and Q, draws from this long-term equilibrium of water balance over a catchment coupled with the energy balance. It postulates that the partition of the annual water budget between runoff and evapotranspiration over catchments, represented by the evapotranspiration E, is a function of the relative water supply (rainfall P) and the atmospheric water demand (PET) (Tian et al., 2018; Xing et al., 2018; D. Yang et al., 2007). The latter depends on both available energy and aerodynamic resistance (Barella-Ortiz et al., 2013). Therefore, this framework considers the system's water and energy limitations, which cannot evaporate more than the atmospheric demand allows and more water than the catchment receives from the water source (P). In short, it defines the "Budyko space" (Berghuijs et al., 2020; Reaver et al., 2022).

This framework relies on a closed water budget in time and space, neglecting ΔS. Therefore, it must be applied over a closed watershed and fitted on a long-term equilibrium. To be freed from seasonal water storage variations, we use a time series of a yearly resolution (hydrological year) in this study. For the region considered, the hydrological year starts in September, at the end of the dry season, when the reservoirs are supposedly at their lowest. It

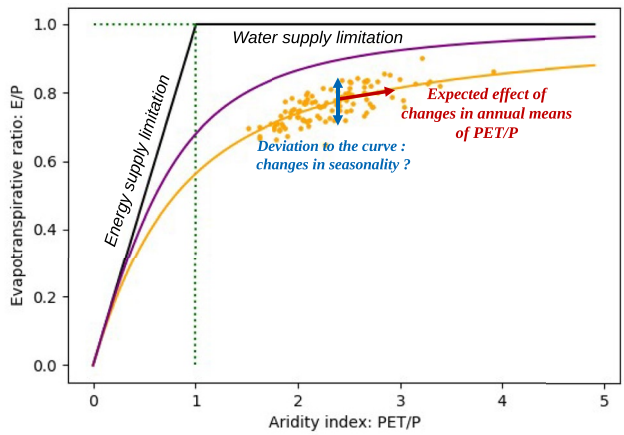


Figure 1. Budyko framework: relationship between evapotranspiration ratio (E/P) and aridity index (PET/P) (Fu's equation). E, PET, P are annual averages. ω associated with the purple curve is larger than ω associated with the orange curve and translates into a higher evaporation efficiency above the watershed. For a given watershed with constant characteristics, there is still a dispersion around the curve of the dots for a given year due to intra-annual variations of the climate cycle (orange dots). The curve and its associated ω represent the average behavior of the watershed. The framework includes trends in annual climate variables by a displacement along the curve (red arrow). However, it does not include trends that could impact the way water is partitioned over the catchment such as long-lasting trends in the intra-annual distribution of P and PET (blue arrows).

minimizes the differences in ΔS from year to year. Later on, unless specified otherwise, the variables P, E, and Q represent the annual averages over the hydrological year. We then apply the framework over minimum periods of 11 years, considered a long enough period for ΔS to be negligible over most catchments, dependent on the area (Han et al., 2020). We tested this hypothesis with the outputs of the LSM, and we found that ΔS is about a hundred times smaller than Q when 11-year sub-periods are considered (not shown).

2.1.2. One Parameter Equation

The original Budyko framework was empirically constructed over a set of catchments to define a curve followed, on average, by catchments in the Budyko space. Different analytical approximations to this hypothesis (Budyko curves) have been developed, expressing the evapotranspiration rate (E/P) as a function of the aridity index (PET/P) over a catchment (Figure 1).

More specifically, the framework was extended to analyze individual catchments over a stable period. Parametric equations were developed which introduced an empirical parameter representing the specific position of the catchment within the Budyko space (H. Yang et al., 2008).

Two of the most widely used are the Fu equation (Equation 2) (Ning et al., 2019; Simons et al., 2020; L. Zhang et al., 2004, 2008; Zheng et al., 2018) and the Mezentsev-Choudhury-Yang equation (Equation 3) (Luo et al., 2020; Roderick & Farquhar, 2011; W. Wang et al., 2020; Xing et al., 2018; Xiong et al., 2020; H. Yang et al., 2008). These can be found under different names in the literature, such as the Tixeront-Fu equation for Equation 2 or Turc-Mezentsev for Equation 3 (Andréassian & Sari, 2019).

$$\frac{E}{P} = 1 + \frac{PET}{P} - \left(1 + \left(\frac{PET}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} \tag{2}$$

$$E = \frac{P * PET}{(P^n + PET^n)^{\frac{1}{n}}} \tag{3}$$

The two parameters derived from Equations 2 and 3 are linearly correlated, implying that both equations are almost equivalent (Andréassian & Sari, 2019; Du et al., 2016; Roderick & Farquhar, 2011; H. Yang et al., 2008). We examine the sensitivity of the results to the parametric equation used. We obtain very similar results for the methodology with either equation used. We conclude that we could use either equation. For the rest of the study, we use results obtained with Fu's equation (Equation 2).

E measurements are not available over large spatial and temporal scales. Therefore, most studies work from the analysis of Q, which can be calculated from the water balance Equation 1, where ΔS has been neglected. With Fu's equation (Equation 2) used to express E in Equation 1, it yields (Equation 4):

$$Q = P * \left(1 + \left(\frac{PET}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} - PET = f(P, PET, \omega) \tag{4}$$

2.1.3. Discussion of the Watershed Parameter

The watershed parameter is empirical; it is obtained by fitting data from a specific catchment during a period of assumed equilibrium state. It determines the position of the catchment in the Budyko space.

The specificity of the parameter relates to all factors impacting the evaporation efficiency of the watershed other than changes in the average aridity index (Donohue et al., 2012; Padrón et al., 2017; L. Zhang et al., 2004). The most common hypothesis is that it reflects the various hydrological characteristics of the watershed, such as topography, vegetation coverage, and soil properties, which play a part in the annual partitioning of P into E and Q over the catchment (Gudmundsson et al., 2017; Reaver et al., 2022). Some are considered time-invariant

(soil type, topography, etc.), while others are possibly affected by long-lasting changes. These can occur in the hydrological properties of the surface water system, most likely due to direct anthropogenic activities such as river management, irrigation, and land cover changes. It leads to the "catchment trajectory conjecture" (Reaver et al., 2022), which suggests that the watersheds would follow an average Budyko-curve (Figure 1, red arrow) if it were not for changes in hydrological properties independent of changes in the average aridity index.

Several studies attempted to analyze the evolution of watershed behavior between two equilibrium states, a period of reference and a period of post-changes (Jiang et al., 2015; Luo et al., 2020; W. Wang et al., 2020; Zhao et al., 2018; Zheng et al., 2018) and then fit the parameter independently over each period. Two distinct curves (Figure 1) were acquired using distinct watershed parameters to characterize the pre- and post-change equilibrium states. As a first hypothesis, they then considered that deviation from the initial curve (period of reference) is only due to changes in the land surface, such as the effect of anthropogenic activities and land cover variations. Assuming ω to be climate invariant, the changes due to climate are considered in the framework only through the modifications of the average P/PET (Figure 1, red arrow). It follows the hypothesis that watersheds follow their Budyko curve if the catchment's surface characteristics remain constant.

However, studies have shown that not all catchments under climate change exhibit this behavior. There is a climate dependence of the deviation to the initial curve. Reaver et al. (2022) showed that reference catchments with the long-term stability of land use did not always follow their Budyko-curve. With the previous hypothesis, this deviation could be misinterpreted as a change in the land-surface characteristics. Padrón et al. (2017) found that the variability in the parameter is highly correlated to climate features such as snow fraction precipitation and the storm arrival rate. Over their extensive global database, the correlation between vegetation indices and direct anthropogenic influence factors is only secondary. Jaramillo et al. (2022) used CMIP6 multi-models ensemble to fit Budyko curves over several basins for the period 1901–1950 and to calculate ET/P PET/P for 2051–2100. They compared the results of the ensemble to those obtained with the hypothesis that catchments should follow their initial Budyko curve. Most basins will not follow the curve under climate change, showing a climate dependence of the deviation from the initial curve.

To circumvent the limitation due to the hypothesis of ω being climate invariant, several studies have tried to locate an expression of the watershed parameter as a function of pertinent factors. It would allow us to express the evolution of ω over time and decompose the effects of climate and human activities through the different factors chosen. If valid, it would also allow transposing the expression to unmonitored catchments where ω cannot be directly fitted or to future catchment conditions. Different methods, such as step-wise regressions and neural networks, were used to identify pertinent factors. Such methods require enough information on the chosen factors; strong hypotheses stand behind the expression.

Most studies construct their function across several basins, accounting for spatially different human, climate, and land characteristics (D. Li et al., 2013; S. Li et al., 2022; Ning et al., 2019; Tian et al., 2018; Xing et al., 2018; X. Zhang et al., 2019). A variety of factors were selected: environmental factors such as soil moisture, seasonality of P and PET, aridity index (S. Li et al., 2022; Ning et al., 2019), vegetation fraction and routing depth (Gentine et al., 2012; D. Li et al., 2013; Ning et al., 2019), relief ratio, drought severity index, seasonality of P, and synchronicity between P and PET (Xing et al., 2018); direct human factors such as irrigated surfaces (Tian et al., 2018), the amount of water applied for irrigation (D. Li et al., 2013), land use, land cover change in highly managed areas (Tian et al., 2018), and even gross domestic product per capita (X. Zhang et al., 2019). The chosen factors are highly dependent on the area studied.

Another strong hypothesis is that such a relationship defined over spatial differences is applicable to explain temporal differences (Luo et al., 2020). Other studies (Jiang et al., 2015; Zhao et al., 2018) looked at time-varying human activities and climate change to construct expressions, using a time-moving window to fit the evolution of the catchment parameter over a basin. This approach faces another limitation due to the availability of information on the different factors' time evolution. Ning et al. (2019) used a mixed technique, applying their fit across 30 basins at different time scales using moving time windows and found that the impact of vegetation cover and climate seasonality on the watershed parameter was stronger over longer time steps, showing that the weight of different factors varies with the time scale. Ultimately, the pertinent factors highly differ among studies, regions, time periods, and climate types (Padrón et al., 2017).

Moreover, recent studies question the hypothesis underlying these studies and "the catchment trajectory conjecture" (Berghuijs et al., 2020; Reaver et al., 2020, 2022). The study demonstrates that the parameter exhibits a lack of independence from climate but also depends on the biophysical characteristics of the catchment directly due to

the dependence of E, PET, or P on those features (Reaver et al., 2020). The highly nonlinear relationships between all the features involved in the evaporation efficiency of the catchment and the average P and PET contradict the hypothesis that the parameter ω can be expressed as a simple function of independent parameters. It also explains why previous studies were so different from one another. The catchment parameter is, therefore, a mathematical tool to represent the evaporation efficiency of a catchment over a given period and has no physical meaning in itself. It is not transferable through time and space. It only positions the catchment within the Budyko space (Reaver et al., 2022). It can still be used to study the position of the catchment in the Budyko space. Studying the deviation from the curve may provide insight into how factors besides aridity affect the water balance (Berghuijs et al., 2020).

2.2. Simulations With a LSM

To isolate the climate change effect from other factors that could affect watersheds, we work with the outputs of a LSM. The model constructs watersheds with constant hydrological properties and represents an idealized watershed without any direct changes from human activities and other nonclimatic disturbances. Therefore, the only source of long-term change would be due to a difference in response to an evolving climate.

2.2.1. A "Natural Reference" Simulation

This study uses the LSM Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) from the Institut Pierre Simon Laplace. It includes biophysical and biogeochemical processes to simulate the global carbon cycle and quantify terrestrial water and energy balance. It runs coupled to an atmospheric model or in stand-alone conditions with an independent data set to force the atmospheric conditions. Here, we use the model in stand-alone conditions, forced with the data set GSWP3 covering 1901–2013 (Hyungjun, 2017) at the resolution of 0.5° for all climate variables.

The hydrological network of the ORCHIDEE LSM is constructed from the hydrological elevation model HydroSHEDS (Lehner et al., 2008), which covers the area studied with the resolution of 30 arc s (approximately 1 km at the equator). The hydrological information is then upscaled to the resolution of the atmospheric grid, the hydrological coherence being preserved by the construction of hydrological transfer units at the subgrid level (Polcher et al., 2022). From a database of gauging stations, upstream basins are reconstituted on the hydrological elevation model grid and then projected on the atmospheric grid during the process. We have access to 2,134 stations over the area studied for which the LSM calculates a discharge and for which we have the reconstituted upstream basin (Figure S1 in Supporting Information S1).

The LSM ORCHIDEE, more specifically the Schématisation des EChanges Hydriques à l'Interface Biosphère-Atmosphère module, uses the USEB (unstressed surface energy balance) method to model PET (details in Barella-Ortiz et al., 2013). This method relies on the turbulent diffusion equation to calculate the potential soil evaporation PET_{soil}, obtained from the air density, the aerodynamic resistance, and the humidity gradient. The USEB method estimates the virtual surface temperature from an unstressed surface-energy balance, computing a new energy balance considering a saturated surface (Barella-Ortiz et al., 2013). Potential transpiration is driven by PET_{soil} but limited by vegetation resistance, calculated in LSM ORCHIDEE and based on plant functioning types maps and LAI calculations (Guimberteau et al., 2012). Then, the total PET is calculated by summing the potential evaporation and the potential transpiration. PET is reduced to the actual evapotranspiration E by a "moisture availability function" (Barella-Ortiz et al., 2013).

Over the course of several years, the model has been tested and validated on many aspects of the land surface processes (hydrology, vegetation, and carbon cycle processes). This attests to its quality to reproduce the water and energy balance and also discharge over different areas over the globe (Guion et al., 2022; Nguyen-Quang et al., 2018; Polcher et al., 2022; Tafasca et al., 2020; F. Wang et al., 2018). Comparing the LSM outputs directly to observations for discharge is challenging, mainly due to the absence of certain processes in the models, including those resulting from direct human activities and the extensive water and river management (F. Wang et al., 2018), as it is the case in our area of interest. Based on previous literature, we can assume the model proficiently emulates the mechanisms underlying actual evaporation, thereby effectively replicating the "natural" response of watersheds with persistent physical attributes to the past climate conditions prevalent in Europe. We study Q variations and not the absolute value of Q since we know that the output of the LSM does not represent the complete processes over real catchments. We focus here on the impact of the changes in atmospheric parameters

Table 1
Synthetic Forcings Created

	Forcing name	Average P	Intra-annual variation of P	Description[a]
1	*ref*	–	–	Reference forcing: GSWP3 (1901–2012)
2	*f*2000	Fixed	Fixed	P has been entirely fixed for each year, equal to the precipitation and the seasonality of the year 2000.
3	*cstmean*	Fixed	–	Only the average value of P has been fixed for every year to the one of the year 2000
4	*cstintravar*	–	Fixed	Only the intra-annual variations of P have been fixed for every year to the one of the year 2000

[a]For forcings 2–4, P has been modified compared to the reference: the average value of P over the year and/or the distribution of precipitation over the year have been fixed for each year to the value of the year 2000.

on land surface responses with constant characteristics. The modeled watersheds react to the climate data input at each time step (30 min time step). Therefore, the LSM output depends on both the evolving annual average and the evolving distribution over the year of the climate variables.

For consistency in the calculation of E and PET, we take both from the output of ORCHIDEE forced with GSWP3. The gridded outputs (PET, E) are at the resolution of the forcing data set (0.5°). P is the sum of rainfall and snowfall in GSWP3. Then, we consider the annual mean P, PET, and E over hydrological years, integrated over each catchment. The catchments' shape has been reproduced at a finer resolution and then projected on the 0.5°grid.

2.2.2. Synthetic Forcings to Analyze the Effect of Variation of Seasonality

In order to better understand the effect of inter and intra-annual climate variations on the Budyko framework and on discharge Q, we construct synthetic climate forcings, fixing one or the other.

The calculation of PET includes many related climate variables and nonlinear relationships, making it very difficult to anticipate how a change in a given climate variable may influence its behavior. It is, therefore, too complicated to create synthetic forcings for which we can modify climate variables to fix PET seasonality, for instance. Therefore, we only modify the precipitation P in the synthetic forcings to see how it impacts our results compared to the reference forcing.

The reference forcing is the GSWP3 data set from September 1901 to September 2012 (3 hr time step). Then we constructed three forcings, which were modified over hydrological years (Table 1, Figure 3a):

- *f*2000: A forcing where all 3 hr values of P are set to the values of the year 2000 (September 1999 to September 2000) for each year. Therefore, all components of P (average and intra-annual variations) are set constant.
- *cstmean*: A forcing for which we keep the relative intra-annual distribution of P of each year, but where the average P of each year is set constant. The 3 hr values of P are scaled so the hydrological year average is set to the one of the year 2000 (September 1999 to September 2000).
- *cstintravar*: A forcing for which we keep the annual average of P for each year, but where the relative intra-annual distribution of P is set constant. The 3 hr values of P are set to the values of the year 2000 (September 1999 to September 2000) for each year and then scaled over each hydrological year so the yearly average is set to the one of the corresponding years in the reference forcing.

2.3. Combining Both Models

In this study, we apply the Budyko framework to the output of an LSM to explore the sensitivity of the empirical parameter to climate change and the resulting effect on discharge. The watersheds in the LSM have constant

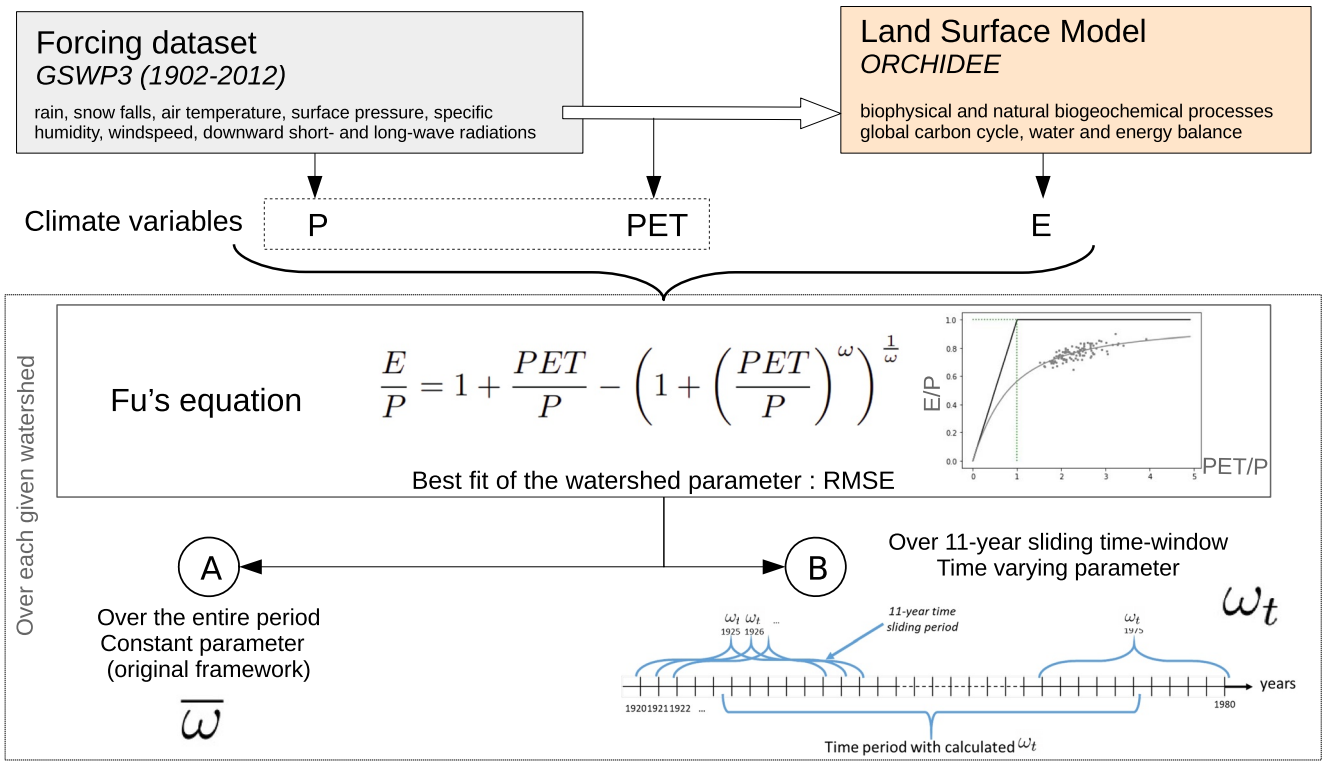


Figure 2. Scheme of the method: the land surface model (LSM) is obligated with the forcing data set to calculate *E*. The LSM is considered to represent the "climatic reality" over a catchment without any changes in the watershed characteristics. We then average *P*, PET, and *E* and integrate them over each watershed to get annual averages for all catchments. Then, we fit the Fu equation. (a) The fit of the equation over the entire century results in the calculation of an empirical parameter $\overline{\omega}$, which represents the average catchment characteristics. (b) To have an evolution of ω_t over time, the fit was then successively applied over an 11-year sliding time period.

biophysical characteristics. The LSM then reproduces the interaction of the land surface with climate parameters. It is affected by climate change and no other source of change. We use the changes in ω as a proxy for changes in the partitioning of *P* into *Q* and *E* other than direct changes in average PET/*P*; it focuses on the deviation from the initial curve and attempts to decompose its dependence on climate. In this case, any deviation to the curve is only due to climate effects. Since ω has no clear physical meaning, we do not analyze directly the changes in ω but rather how they impact the evolution of discharge.

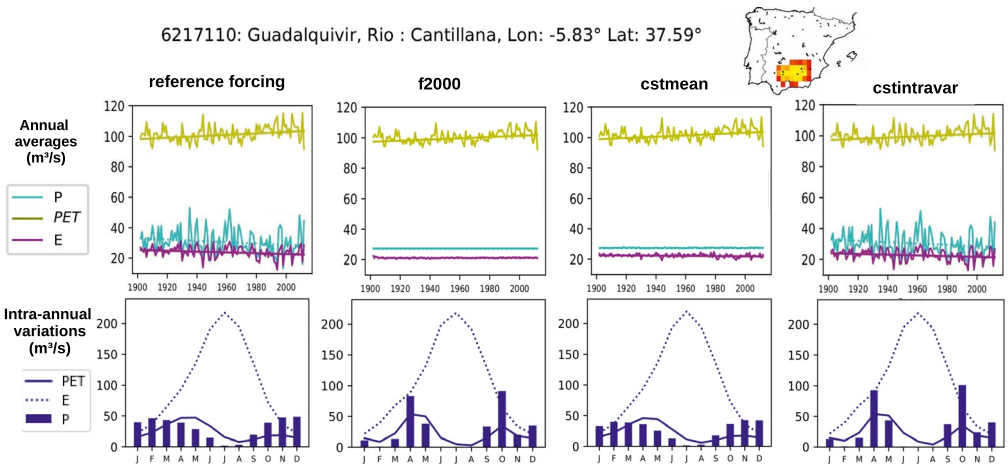
Using an LSM, we can also change various climate parameters to better address how they weigh in the modeled changes. We develop a varying ω_t to capture part of the change in the evaporation efficiency of the watersheds due to climate. We compare its effects to the magnitude of change in discharge already captured with the traditional framework, which only considers changes in annual averages of PET/*P*.

2.3.1. Fit of the Evaporation Efficiency Parameter ω

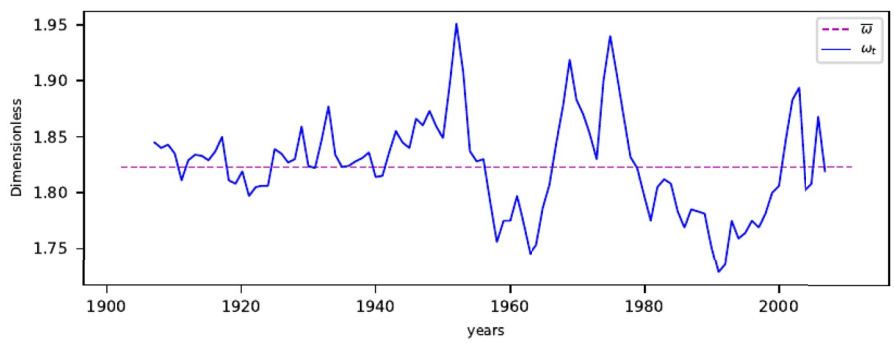
The watershed parameter of the Budyko curve is calculated over each catchment with a fit of the equation curve $E/P = f(\text{PET}/P)$ (Equation 2), using the minimum root mean square error for a given set of annual averages of evapotranspiration *E*, precipitation *P* and PET data (Jiang et al., 2015; D. Yang et al., 2007). We fit the parameter once with all points over the entire period covered by the climate data set to obtain $\overline{\omega}$ representing the average behavior for each catchment (Figure 2a).

For a watershed with constant hydrological properties (which is the case when considering modeled watershed in ORCHIDEE), if we consider the "catchment trajectory conjecture," ω is independent of climate, and the catchment follows its initial curve. However, ω varies for a given watershed because of climate. For instance, over an equilibrium state, intra-annual variations of the climate cycle induce a variability of the annual values (*E*/*P*, PET/*P*) around the fitted curve. The distribution of rain changes the covariance between *P* and PET over the year. A difference in storm depth over a catchment can change the capacity of the soil to store water, the response of vegetation, and change the dynamic of the water partition into runoff and evaporation even if the annual amount

(a) Example for give watershed. Map: the colors show the share of the grid point within the watershed. Yellow points are completely within it while bordering grid points are red.
Modified forcings over a given basin: the first row of graphs shows the inter-annual variability of P (cyan), PET (green) and E (purple) for each forcings. The second row shows the average seasonal distribution of P over the catchment for each forcings over the entire century. The average monthly distributions of P shown here are computed over the century. It however varies from one year to another for the reference forcing and the forcing *cstmean* which is not illustrated here.



(b) Watershed parameter $\overline{\omega}$ fitted over the entire time period (dashed purple line) and ω_t fitted successively over a sliding 11-year time-window (blue line) for the reference forcing.



(c) Discharge estimated with Budyko for the reference forcing: $Q = f(P, PET, \omega_t)$ (blue line), $Q_c = f(P, PET, \overline{\omega})$ (purple line), $Q_\omega = f(P_{rand}, PET_{rand}, \omega_t)$ (black line) with their associated trends. Unsignificant trends are dashed. Here all trends are significant.

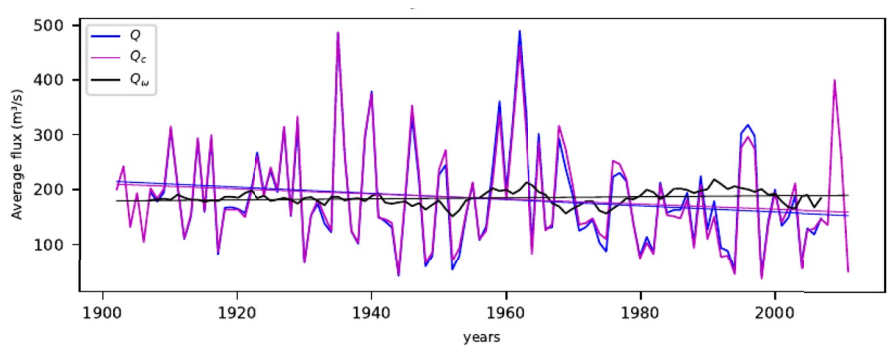


Figure 3.

of precipitation stays constant (Donohue et al., 2012). More generally, a change in synchronization between P (water available) and PET (energy demand from the atmosphere) will change E/P for the same average climate (Abatzoglou & Ficklin, 2017; S. Li et al., 2022). In an equilibrium state, the intra-annual variations should be without trends and only result in white noise around that equilibrium. The fitted parameter ω represents the average behavior of the basin. For a catchment under climate change, however, this variability could lead to a significant permanent deviation from the initial curve if this intra-annual distribution tends to have a trend.

2.3.2. Introducing a Varying Watershed Parameter ω_t

With its simple framework, the Budyko model does not cover possible changes at intra-annual time scales. The average effect of this synchronization is included in the adjustment parameter ω, which is, therefore, not completely independent of climate. Therefore, long-term changes in seasonality should induce a climatic time dependence which is not accounted for in the framework with a constant ω. Therefore, considering a varying parameter should improve the Budyko model to reproduce E/P and its climatic evolution.

To obtain a varying parameter ω_t for each catchment, we carry out several fits over successive 11-year time-sliding sub-periods (Figure 2b). We chose 11 years as the smallest time length to apply the Budyko framework relevantly, considering that each 11-year sub-period is stationary ($\Delta S = 0$). This allows us to focus on long-term changes and to minimize the impact of year-to-year "transient" effects (e.g., soil storage and groundwater changes) (Y. Yang et al., 2018). Tian et al. (2018) found that below a certain time length, the fit of the ω parameter was too unstable to be relevant.

2.3.3. Decomposing the Impact of Climate on Discharge Trends

The watershed parameter ω is a conceptual variable that provides little insight into the magnitude of discharge changes. Thus, we examine the impact of ω_t changes on the river discharge Q and compare these changes to the impact of annual averages of climate variables (P and PET) changes on Q over time. To simplify the discussion, we gather the annual averages of P and PET in a "climate" variable $C = (P, \text{PET})$.

Following our previous hypothesis (Equation 4), Q can be estimated with the Budyko framework using C and ω: $Q = f(C, \omega_t)$.

Q can be decomposed with first-order partial derivatives (Equation 5), with the first term of the right-hand side representing the partial derivative due to climate variables C and the second term for the partial derivative due to changes in the watershed parameter ω_t. We then estimate the partial derivatives due to C and due to ω independently.

$$\frac{dQ}{dt} = \frac{\delta Q}{\delta C}\frac{dC}{dt} + \frac{\delta Q}{\delta \omega}\frac{d\omega}{dt} \qquad with \quad C = (P, PET) \tag{5}$$

To independently estimate the partial derivative due to climate variables C, we must cancel the second term (Equation 5, left side). To do so, we calculate the discharge $Q_c = f\left(C, \overline{\omega}\right)$, with a constant value of ω. The trend of that discharge $\frac{dQ_c}{dt}$ matches the term with the partial derivative due to C in Equation 5.

To estimate the partial discharge trend due to ω_t, we need to eliminate the trends in annual averages of P and PET over the century to cancel the first term (Equation 5, left side). We randomly draw P and PET pairings for each year. We do so several times and average the results for each year. It gives us a random climate without trends over the century. We then apply Fu's equation (Equation 2) with the resulting random annual averages of P and PET and the varying ω_t calculated with the forcing before the random drawing. It gives $Q_\omega = f(C_{\text{rand}}, \omega_t)$ for which the climate trends are only due to variations captured by the time-varying parameter ω_t. The trend $\frac{dQ_\omega}{dt}$ matches the term with the partial derivative due to ω in Equation 5. In the end, we get:

Figure 3. Time series obtained through the full application of our methodology for a given basin in Spain. (a) shows the inter-annual variability of annual averages of climate variables P, potential evapotranspiration (PET), and E modeled by the land surface model, forced with the different synthetic forcings. E mostly relates to P. (b and c) are results for the reference forcing. (b) shows the varying ω_t resulting from the time-sliding window calculation (blue curve), compared to $\overline{\omega}$ calculated with one fit over the entire century (dashed purple line). (c) shows the decomposition of the discharge, comparing the full discharge to partial discharges and their respective trends. The full discharge Q is modeled with Fu's equation with annual averages of P and PET from the reference forcing and ω_t. The first partial discharge Q_C is the one calculated with the constant parameter $\overline{\omega}$. It covers most Q variations for the given basins. The second partial discharge Q_ω covers some of the missing variations of Q and some of the missing trends due to deviations to the average curve. From that figure, we can conclude that most variations and trends of the discharge in this basin are explained by $C = (P, \text{PET})$.

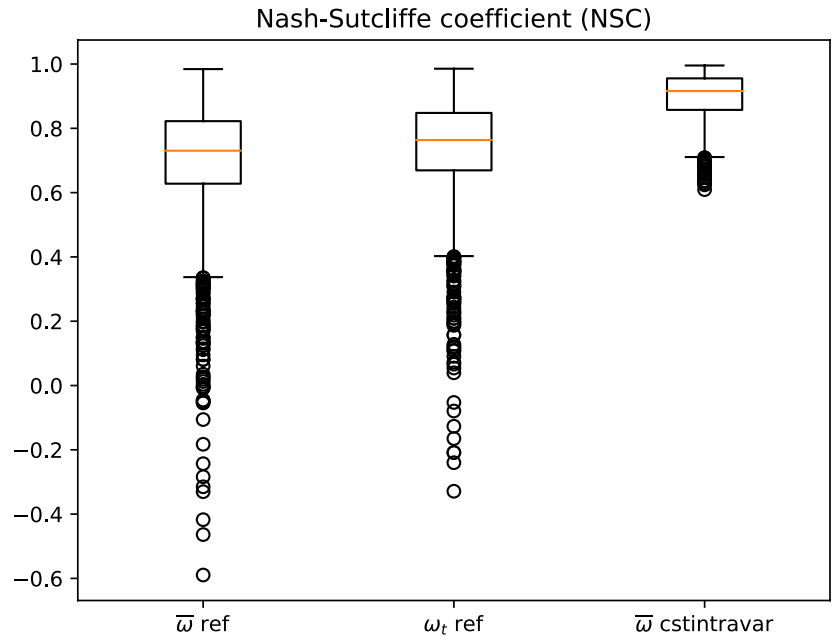


Figure 4. Boxplot of Nash-Sutcliffe coefficient (NSC) for all watersheds: for the forcing of reference with the constant parameter $\overline{\omega}$, with the varying parameter ω_t and for the forcing *cstintravar* (where the seasonal distributions of P have been fixed over the entire time period) with a constant $\overline{\omega}$. It represents how well the Budyko model reproduces the discharge output from Organizing Carbon and Hydrology In Dynamic Ecosystems. A value above 0.5 is considered satisfactory. Very similar results are found when looking at R^2 from a linear regression.

$$\frac{dQ}{dt} = \frac{dQ_c}{dt} + \frac{dQ_\omega}{dt} \tag{6}$$

We calculate the trends of each term and their significance using the Mann-Kendall nonparametric test, associated with the Thiel-Sen slope estimator. It gives us time series and associated trends for each studied watershed. Figure 3 shows an example of a watershed in southern Spain.

3. Results

3.1. Performance of Budyko With or Without a Variant Parameter ω

We hypothesize that for watersheds with constant hydrological properties, the dispersion of annual points around the curve is due to intra-annual variations of climate. If these variations did not exist, catchments would follow their Budyko curve, and we could use it to model the discharge almost perfectly.

To test this hypothesis, we examine the performance of the Budyko curve with a constant parameter $\overline{\omega}$ to reproduce the discharge from the LSM for the reference forcing compared to the forcing *cstintravar*. For that latter forcing, we removed the intra-annual variations of P from 1 year to another, which should render the performance of the Budyko curve model close to perfect if the hypothesis is valid.

We use the Nash-Sutcliffe coefficient (NSC) as a performance indicator (Equation 7, Figure 4). We consider an NSC >0.5 to be satisfactory (Moriasi et al., 2007).

$$NSC = 1 - \frac{\sum_{i=0}^{years}(Ql_i - Qb_i)^2}{\sum_{i=0}^{years}\left(Ql_i - \overline{Ql}\right)^2} \left\{ \begin{array}{l} \text{with } Ql = \text{dicharge from the LSM} \\ \text{and } Qb = \text{Result from the methodology with Fu's equation} \end{array} \right. \tag{7}$$

We obtain NSC values above 0.5 for 89.9% of all 2,134 watersheds tested for the Budyko curve with a constant parameter (Q_c, calculated with a constant $\overline{\omega}$) applied with the reference forcing (boxplot on the left, Figure 4). Therefore, the average curve model is rather effective in reproducing the annual discharge over watersheds with constant hydrological properties reacting to an evolving climate.

For the forcing *cstintravar*, NSC for Q_c increases to above 0.6 for all watersheds (boxplot on the right, Figure 4). It confirms our hypothesis: the average Budyko curve model is even more effective if there are no intra-annual variations of P from 1 year to another. Therefore, most of the variability that is not captured by the average Budyko curve over the past century is due to the intra-annual variability of P and the covariance of P and PET.

When looking at NSC for the framework applied to the reference forcing with a varying parameter $Q(ref) = f(C(ref), \omega_t)$, we gain up to 0.26 points of NSC for the tested watershed and reach 94.1% of all watersheds with NSC >0.5 (boxplot on the center, Figure 4). It does not reach the performance to reproduce Q_c with the forcing *cstintravar*. However, it allows to catch some of the deviation to the curve due to intra-annual trends of climate variables. We capture long-term trends following our choice of the 11-year time-moving window. It validates our hypothesis that introducing a varying watershed parameter ω_t improves the framework to better encompass climate variability and the effect of climatic trends on discharge, including the effect of climate change on the intra-annual distribution and covariance of climate variables (P and PET).

To sum up, for watersheds with constant hydrological properties under historical climate, most of the deviation to the average curve model (i.e., changes in the evaporation efficiency of catchments) is due to variations in the intra-annual distribution of climate variables (P and PET). Our varying parameter improves the framework by allowing us to capture the long-term trends of these variations. We now analyze their effect on the discharge and compare them to the direct effect of trends in the annual average of climate variables.

3.2. Comparing the Effects of Intra-Annual Variations of *P* on Discharge *Q* to the Effects of Variations in Annual Averages of *P* in Europe

We consider our area of study, western Europe (2,134 watersheds modeled) (Figure 5). To better illustrate our results, we also take two contrasted basins: one in Italy (Figure 6) and another in England (Figure 7).

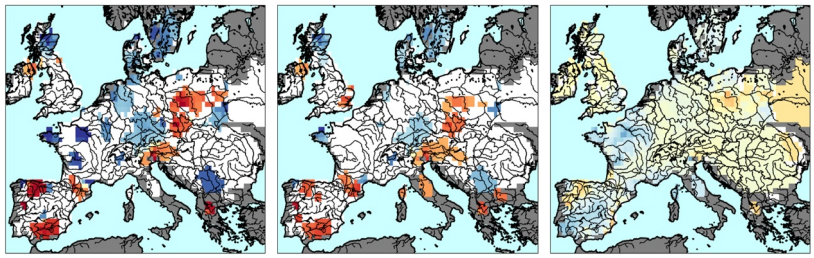
Figures 5a–5c show the relative trends over each basin for the reference forcing, respectively, of Q, Q_c, and Q_ω. There are significant decreases in the total discharge Q (Figure 5a) (−0.3% to −0.4% per year over the past century) over sparse basins in Spain, the Pyrenees, Italy, Slovenia, Greece, and Eastern Europe. There are significant increases (Figure 5a) (+0.2% to +0.4% per year over the past century) over sparse basins in France, Germany, Denmark, Sweden, Northern UK, and Serbia. These trends are primarily due to changes in the annual averages $C = (P, \text{PET})$ since the average Budyko curve model Q_c captures most of the signal (Figure 5b). The inter-annual variability of C is high, making the trends less than 95% significant over most basins for Q and Q_c. Both selected catchments better illustrate it (Figures 6b and 7b): for the reference forcing (top left), the dominant effect in the variations of annual discharge Q (blue line) is due to the annual mean of climate variables C (purple line). Clearly, both curves have very similar high inter-annual variations and trends.

Changes in C are the dominant factors explaining the climatic trends in Q over the past century in Europe. The results obtained with the forcing *cstintravar* (bottom right for Figures 6b and 7b and maps Figures 5j–5l) confirm it. It shows that without inter-annual changes in P distribution (in other words, with a maximum reduction of the inter-annual changes in the annual covariance of P and PET), the discharge Q obtained and the associated relative trends are very similar to the results obtained with the reference forcing. Therefore, the effects of changes in the annual covariance of P and PET are minor compared to the effects of changes in the annual mean of climate variables C in most of Europe.

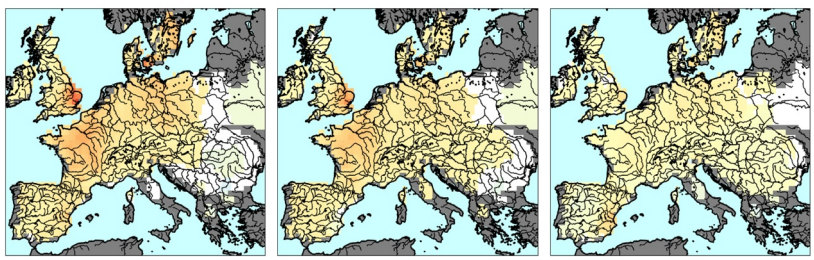
However, in some areas, the effects of the intra-annual distribution of P should be addressed. If we look at the Tiber River in Italy (Figure 6b), the trend in Q_c (purple line) is significant for both the reference forcing and the forcing *cstintravar*. However, the total discharge Q (blue line) trend is only significant for the forcing *cstintravar*. For the reference forcing, the decreasing trend in the discharge due to C (Q_c) is counteracted by the increasing trend due to changes in the evaporation efficiency (Q_ω), making the final trend in discharge Q insignificant.

More generally, over Europe, when we erase the inter-annual variability of C, we capture the effect of trends in the intra-annual distribution of P and PET, through changes in the evaporation efficiency, in Q_ω (Figure 5c). It tends to increase discharge, especially in southwestern Spain, Italy, and the west of France (+0.1% per year over the century). It corresponds to the increasing trend of the black line in Figure 6b, top left graphs for the Tiber River. It has an opposite trend toward decreasing discharge in eastern Europe and has a relatively neutral effect in the rest of the continent (Figure 5c and, in the example of the English basin, Figure 7b, top left graph, black line). It amplifies the trends due to annual averages C changes over certain watersheds such as the Duero basin (north-western Spain, decrease in discharge), western France, and northern Germany. Indeed, we note a significant increase in discharge over certain watersheds where the effect of changes in C alone was insignificant.

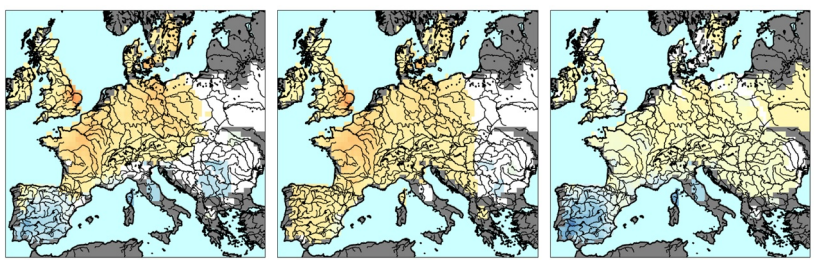
(a) Relative trends $\frac{dQ/\overline{Q}}{dt}$: ref **(b)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: ref **(c)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: ref



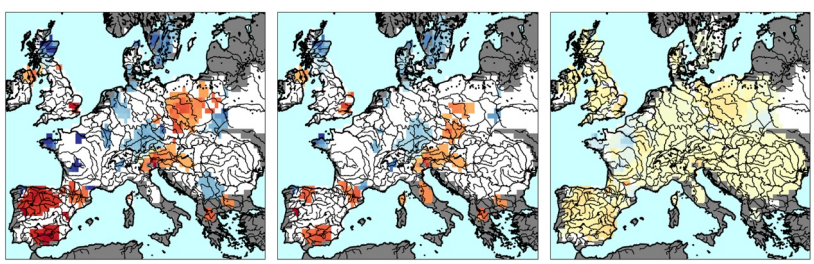
(d) Relative trends $\frac{dQ/\overline{Q}}{dt}$: f2000 **(e)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: f2000 **(f)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: f2000



(g) Relative trends $\frac{dQ/\overline{Q}}{dt}$: cstmean **(h)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: cstmean **(i)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: cstmean



(j) Relative trends $\frac{dQ/\overline{Q}}{dt}$: cstintravar **(k)** Relative trends $\frac{dQ_c/\overline{Q}}{dt}$: cstintravar **(l)** Relative trends $\frac{dQ_\omega/\overline{Q}}{dt}$: cstintravar



(m) Colorbar: significant relative trends (% of change per year over the century)

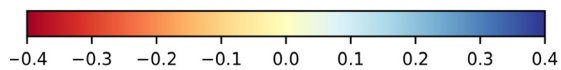


Figure 5.

In other areas, such as the Tiber River in Italy, or in southern UK, the intra-annual variability of P and PET counteracts the effect of C, making the relative total Q trends lose their significance due to opposite signals. We note the decreasing trend is due to the evolution of C, while the effect of the change in the intra-annual distribution of the climate variables tends to increase the discharge.

In order to investigate the impacts of intra-annual variations of P on discharge, we analyze the results of the synthetic forcing $f2000$ and *cstmean* (respectively top right and bottom left Figures 6b and 7b and maps Figures 5d–5f and 5g–5i). For the synthetic forcing $f2000$ (Figures 5d–5f), P have been entirely set for each year to P of the year 2000. Therefore, this only yields the trends due to changes in PET, both for changes due to annual climate variables and changes in the evaporation efficiency of the catchment. For the synthetic forcing *cstmean*, only the annual mean of P has been set. In this case, the trends are due to PET and changes in the intra-annual distribution of P.

For the forcing $f2000$, the effect of PET is toward a decrease in discharge over all of Europe (less than -0.1% to -0.2% per year over the century) (Figure 5d). For both the chosen examples, the effect of PET (top right graphs) tends to decrease discharge (purple line, Q_c when P has been fixed). It is consistent with the significant increase in PET (Figures 6a and 7a, top right). The effect of intra-annual variations of PET on changes in the evaporation efficiency (Figure 5f and black lines, top right graph Figures 6b and 7b) has the same order of magnitude, if not a little smaller (less than -0.1% per year over the century), than the effect of inter-annual change of the annual average of PET (Figure 5e or purple line top right graph Figures 6b and 7b). It tends to amplify the latter's effect, especially over western France and southern UK. It has a slightly opposite effect toward increasing trends in Q (less than $+0.08\%$ per year over the century) over the east of Europe, west of Spain, and for the Tiber river. The effect of changes in the annual mean of PET, in this specific case, is canceled in the total discharge (blue line) by the effect of the changes in the intra-annual distribution of PET captured in Q_ω (black line) (Figure 6b).

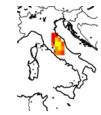
For the forcing *cstmean*, we now add the effect of changes in the intra-annual covariance of P and PET due to changes in the intra-annual distribution of P. Depending on the area, there are two different responses. The two basins chosen in the example each correspond to one type of response. In the case of the basin in England (Trent River), the results obtained for the forcing *cstmean* (Figure 7b, bottom left) are very similar to the results obtained for $f2000$ (Figure 7b, top right). This means that the effect is due to changes in intra-annual synchronicity of P and PET has little impact compared to the effect of the annual mean of PET over that particular basin. It matches the results over northern Europe, especially over France, Germany, and southern UK, where the trends in Q (Figure 5g) are mainly driven by changes in the annual mean of PET (Figure 5h). However, over the Tiber River in Italy, the results obtained for the forcing *cstmean* (Figure 6b, bottom left) shows that the changes in the total discharge Q (blue line) match the changes due to the evolution of ω_t (Q_ω, black line). In this latter case, the effect of the intra-annual variations of P is dominant compared to the effect of changes in PET. This matches the results over southern Europe (Spain, Italy) where for the forcing *cstmean*, the trends in Q (Figure 5g) are driven mainly by changes in the evaporation efficiency (Figure 5i). This increase in discharge diverges from the trends due to changes in C in the area (reference forcing and forcing $f2000$, purple lines).

The discharge trends for both forcings, namely $f2000$ and *cstmean*, are statistically significant across multiple watersheds, independent of the high inter-annual variability observed in the annual mean of P. Trends are significant for 1,883 basins with the forcing $f2000$ and 1,756 for the forcing *cstmean* against only 352 basins with significant trends in Q out of 2,134 for the reference forcing. However, the magnitude of these trends is also quite small. Comparing the discharge obtained with the reference forcing shows that the main factor driving Q is the annual mean of P since the discharge trends look entirely different, when free of its variations.

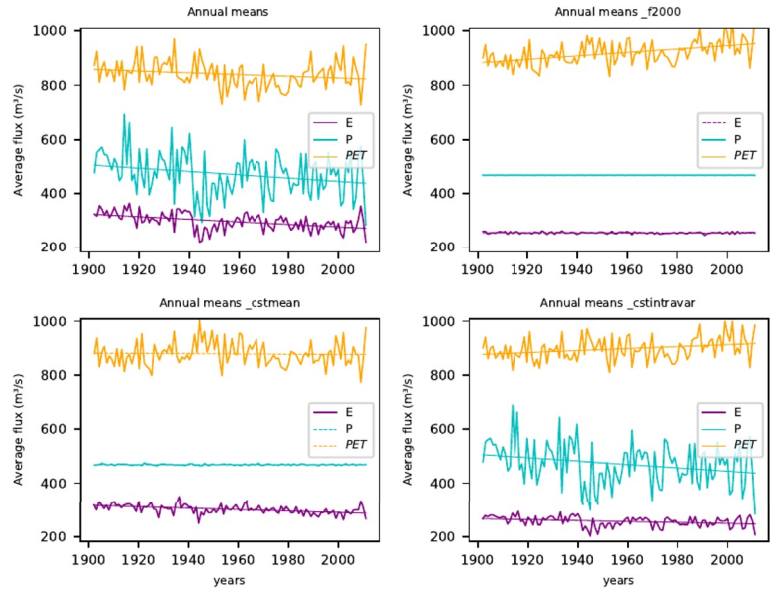
To sum up, the results obtained with the synthetic forcings, the annual mean of P is the first driver of change in the annual discharge over all of Europe. However, its high inter-annual variability tends to hide the trends in most areas. The second most important climatic driver of discharge change depends on the area. Over southern Europe (Italy, Spain), where water is the limiting factor to evapotranspiration, the second most important climatic factor driving discharge changes is the intra-annual distribution of P. Over the rest of Europe, where water is less limiting, the second most important factor driving discharge changes is the increasing PET.

Figure 5. Decomposition of significant relative Q trends (% of change per year over the century) for all the tested forcings: the first line is the reference forcing. The first column is the total change in Q, the second is the partial change due to trends in the annual average of P and potential evapotranspiration (PET), and the last column is the partial change due to changes in the watershed parameter, mostly due to trends in the intra-annual distribution of P and PET. For the modified forcings: $f2000$ has the annual average and intra-annual distribution of P fixed for every year to their value for the year 2000. *cstmean* has only the annual average of P fixed. *cstintravar* has only the intra-annual distribution of P fixed. White areas do not have significant trends.

Tiber River:
Roma



(a) Annual average of climate variables P (light blue line), PET (yellow line) and E (purple line) modeled with the LSM for each academic forcing. Not shown here, the intra-annual distribution of P has been fixed for the forcings *f2000* and *cstintravar*.



(b) Discharge estimated with Budyko and their respective trends: full discharge modeled $Q = f(P, PET, \omega_t)$ (blue line), discharge $Q_c = f(C, \overline{\omega})$ (purple line) with only variations in C accounted for, and $Q_\omega = f(P_{rand}, PET_{rand}, \omega_t)$ (black line) with only the variations of evaporation efficiency, mostly due to the intra-annual covariance of P and PET, accounted for. Unsignificant trends are dashed. Here the results for each academic forcing are shown. We represented the normalized anomaly of discharge $((Q - Q_{mean})/Q_{mean})$ in order to better compare the plots to each other. The scale of the y-axis changes and is divided by 5 for the focings *f2000* and *cstmean*.

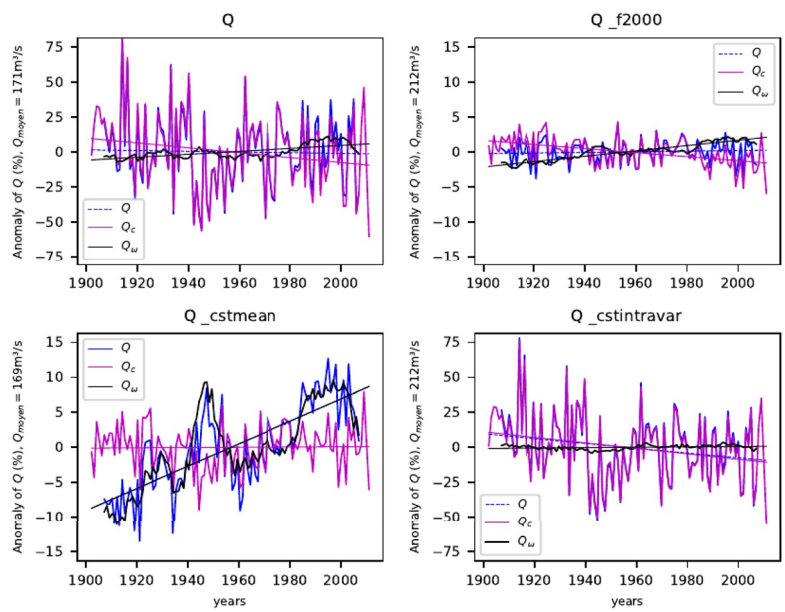
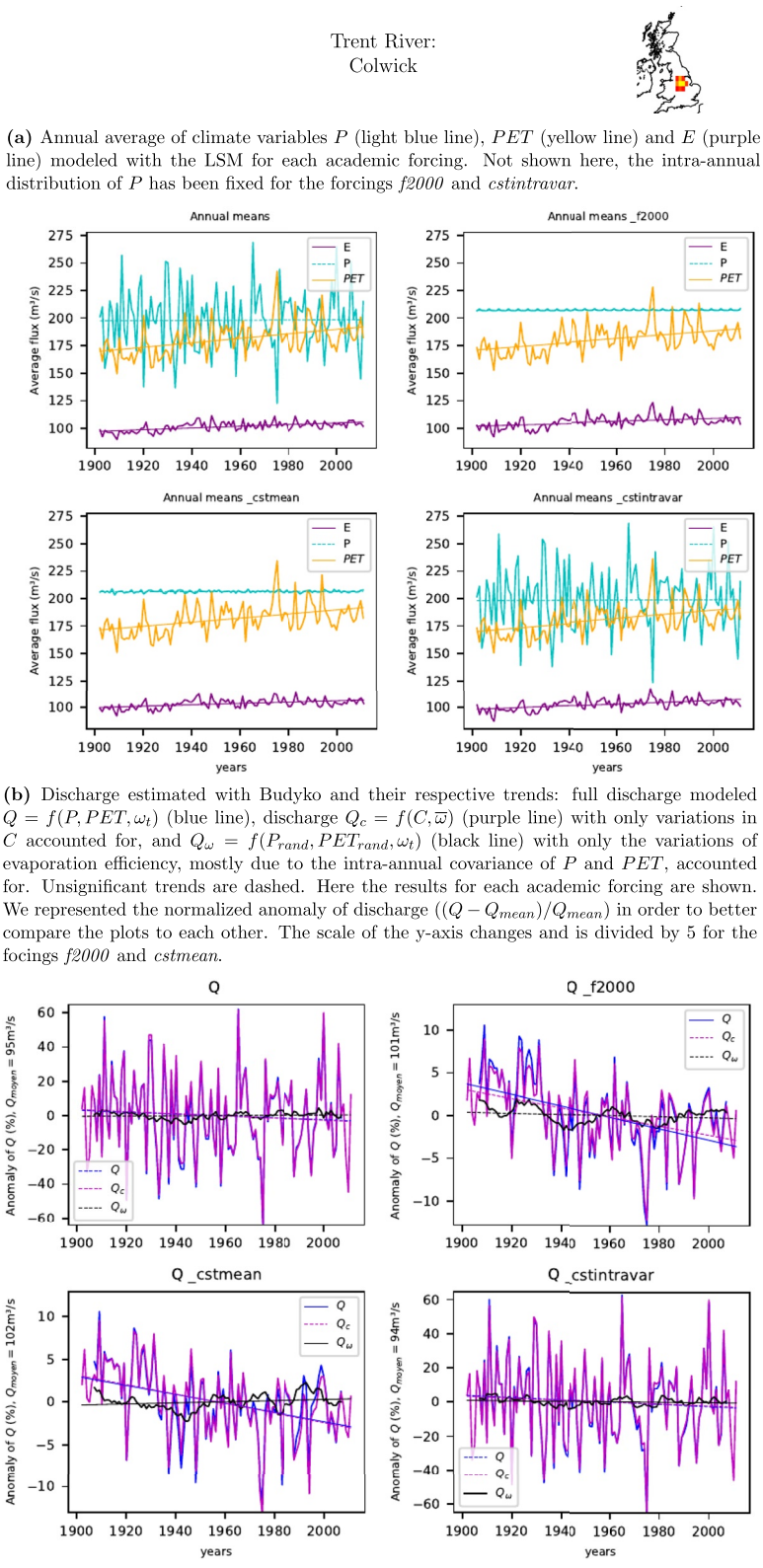


Figure 6. Example 1: time series obtained through the full application of our methodology for a given basin in Italy.



(a) Annual average of climate variables P (light blue line), PET (yellow line) and E (purple line) modeled with the LSM for each academic forcing. Not shown here, the intra-annual distribution of P has been fixed for the forcings *f2000* and *cstintravar*.

(b) Discharge estimated with Budyko and their respective trends: full discharge modeled $Q = f(P, PET, \omega_t)$ (blue line), discharge $Q_c = f(C, \overline{\omega})$ (purple line) with only variations in C accounted for, and $Q_\omega = f(P_{rand}, PET_{rand}, \omega_t)$ (black line) with only the variations of evaporation efficiency, mostly due to the intra-annual covariance of P and PET, accounted for. Unsignificant trends are dashed. Here the results for each academic forcing are shown. We represented the normalized anomaly of discharge $((Q - Q_{mean})/Q_{mean})$ in order to better compare the plots to each other. The scale of the y-axis changes and is divided by 5 for the focings *f2000* and *cstmean*.

Figure 7. Example 2: time series obtained through the full application of our methodology for a given basin in England.

4. Discussion

Similar to the results of several studies (Abatzoglou & Ficklin, 2017; Jaramillo et al., 2022; S. Li et al., 2022; Padrón et al., 2017; Reaver et al., 2022; Xing et al., 2018), we find that the average Budyko curve model with a constant watershed parameter $\overline{\omega}$ does not capture climate-related changes in the watershed behavior impacting its evaporation efficiency. Even with constant hydrological land surface characteristics, most catchments do not follow their average curve over the past century. The deviation to the curve can significantly affect Q's long-term trends over the past century if we free our analysis from the high inter-annual variability of P. It is in accordance with the results of Reaver et al. (2022) criticizing the "catchment trajectory conjecture."

Parameter ω has no direct physical meaning but is a proxy to represent the evaporation efficiency of catchments (Berghuijs et al., 2020; Reaver et al., 2022). However, since it cannot be expressed as a function of clearly defined factors, it is difficult to attribute the changes in the evaporation efficiency to specific climatic features (Berghuijs et al., 2020). Y. Yang et al. (2018) assume that further reductions in Q declining trends due to changes in catchment properties are likely associated with elevated atmospheric CO_2 concentration or increased rainfall intensity. Other studies find correlations between changes in the evaporation efficiency of catchments and storm depth, the portion of precipitation such as snow (Donohue et al., 2012; Padrón et al., 2017; Xing et al., 2018). Using the outputs of an LSM, our studies allow us to test a selection of hypotheses by adjusting climate parameters. We find that the climatic deviation to the average Budyko curve over the historical records is mainly due to variations in the intra-annual distribution of P.

We introduce a time-varying window to fit the parameter of Fu's equation in order to capture trends in the deviation to the average curve in the Budyko space. The choice of window size determines the temporal scale of accounted-for trends. This functions as a frequency filter and only captures the effect of variations over periods the size of the window or larger. We must balance the length of our data set and the appropriate length of the trends we choose to analyze. Since our aim is to investigate the effects of climate change, we do not need to capture the high inter-annual variability and can focus on decadal trends or longer (Y. Yang et al., 2018). Furthermore, a shorter time window would not be adapted to the hypothesis of the Budyko framework, which needs a long enough period to be fitted. So, the window cannot be shorter (Tian et al., 2018). An exploration of an extended time window could be conducted to investigate the limited duration of time that captures the most significant impact on discharge. However, the longer the time window, the fewer points we will have to evaluate the trends.

In our methodology, we decompose the trends due to climatic changes in evaporation efficiency and the trends due to changes in average climate variables P and PET. One limitation in our decomposition method is that the variations in the evaporation efficiency captured in the deviation to the average curve are not entirely independent from the variations of average P and PET. The relationship between P, PET, and evaporation efficiency is complex and relies on many interrelated factors (Reaver et al., 2020). We find in our study that the changes captured in the varying ω_t are mostly due to changes in the covariance of the intra-annual distribution of P and PET. However, the effect of the intra-annual distribution of climate variables on discharge is not completely independent from the annual mean of P and PET because of the difference in sensitivity of the system to a change in water availability. It can impact the magnitude of the identified trends. It is shown by the slight differences observed in Q_ω between the reference forcing and the forcing *cstmean* (Figures 5c and 5i) and between the forcing $f2000$ and the forcing *cstintravar* (Figures 5f and 5l). For each pair of forcings, the intra-annual distribution of P is the same, but the inter-annual mean of P differs. The difference in Q_ω for each pairing is due to a link between the annual mean and the intra-annual distribution of P. Therefore, the amplitude of the effect of the intra-annual distribution of P and PET quantified here may depend on the choice of the fixed average P (again, P from the year 2000 in this study). The observed differences were found to be comparatively insignificant in light of the identified trends, indicating that the fundamental findings regarding Europe would remain unchanged; therefore, we opted to disregard them. When studying specific basins, it could be interesting to choose specific pairings of intra-annual distributions/annual averages of P to construct synthetic forcings, to compare how specific associations combine.

Furthermore, we cannot simply fix PET or its intra-annual variations in our synthetic forcings due to its nonlinearity dependence on a number of climate variables. Therefore, we are unable to decompose the effects of PET as easily as for the effects of P, which would be interesting to do, especially in the areas where P is less limiting, such as in western France or northern Europe.

Our methodology allows for the separation of the effect of primary and secondary climatic drivers on discharge trends. We look at the trends in *P*, and PET for the forcing GSWP3. Our results concur with those in the literature, validating that this forcing reasonably reproduces the climatic trends of the past century over Europe. The trends in PET are significantly (95% level) increasing over Europe. However, the trends in *P* are most often nonsignificant because of its high inter-annual variability, with a significant trend in the annual average of *P* for 413 catchments out of 2,134 selected. The present study finds that the main driver of annual discharge *Q* (trends and inter-annual variability) is the annual mean of *P*. As expected with the increase in *P* over western Europe and the decrease in *P* observed in the Mediterranean area (Christidis & Stott, 2022; Douville et al., 2021; Knutson & Zeng, 2018), the trends in *Q* have followed the same direction. It concurs with the finding of Stahl et al. (2010) and Vicente-Serrano et al. (2019), who found strong spatial consistency between streamflow changes and global rainfall changes.

H. Yang et al. (2008) show that *Q* is universally more sensitive to changes in *P* than to changes in PET, for a fixed land surface condition. Similarly, we find that over most of Europe, the second most important climatic factor on discharge changes is PET, which leads to a decrease in discharge due to the increasing evaporative demand by the atmosphere. Over the Iberian Peninsula and the Mediterranean area, however, PET trends have a lesser impact. There, the water limit is the prevailing factor, having been attained by the end of spring and persisting throughout the entirety of summer. Therefore, a warmer summer does not have a strong impact. The evolution of intra-annual variations of *P* is the second most important factor impacting the changes in the annual discharge, with a higher effect on discharge than the increase of PET over the past century. The intra-annual covariance of *P* and PET impacts the annual behavior of the catchment and the annual balance between evapotranspiration and discharge since it changes the timing between water and energy available throughout the year. The evolution of the intra-annual cycle of *P* tends towards decreasing discharge in the Mediterranean area. It partially counteracts the effect of decreasing *P* and increasing PET on discharge. Therefore, the intra-annual distribution of *P* deserves more attention when studying the evolution of annual discharge. In most studies, it is only considered to look at changes in discharge peaks, floods, or droughts (Douville et al., 2021; Milly et al., 2002; Rottler et al., 2020; Tuel et al., 2022; Vicente-Serrano et al., 2014). We calculate the indices defined by García-Barrón et al. (2013) to evaluate the trends in the intra-annual cycle of *P* for the forcing GSWP3. Similarly to the authors' findings, in Spain, we identify a shift over the end of the century towards a more bimodal distribution of precipitation throughout the year. However, the trends in the intra-annual cycle are mostly qualitative. The tendencies of the annual cycle to have an increasingly marked seasonality, concentrating rain events in fewer but more extreme events over the year, can explain the increasing runoff and relative discharge. Our methodology allows to identify these effects despite the only qualitative trends observed in the indices that measure the intra-annual distribution of *P*.

We apply our parametric model to LSM outputs to isolate the discharge variations due to changes in climate factors. This methodology relies on the capacity of the chosen LSM to reproduce the "natural" response of a catchment to climate, such as its behavior and response to changes in the intra-annual distribution of *P*. The amplitude of our results could depend on the choice of the LSM or the forcing data. We tested the use of other forcing data sets: WFDEI (Weedon et al., 2014), which covers the period from 1979 to 2010, with the same resolution as GSWP3, and E2OFD (Beck et al., 2017), while also covering 1979–2010 but at a lower resolution. We also tested another model, SURFEX (Quintana-Seguí et al., 2020), forced with SAFRAN (Quintana-Seguí et al., 2017), over the Ebro river. This yielded similar results over the overlapping period with little differences in the trends' significance and amplitude. This indicates that the resolution of the forcing exerts a greater influence on the results compared to a specific forcing or model employed. This confirms the suitability of utilizing an LSM as a climatic reference in accordance with our methodology. In the future, when looking at specific basins, it would be interesting to use higher resolution forcings to obtain a more accurate picture of the effects of climate change on discharge. In this case, the diversity of behaviors exhibited among subbasins within a given catchment could be elaborated upon by distinguishing the behavior of upstream subbasins within mountainous regions from that of the downstream portion, which may display differential responses to climate change.

5. Conclusion

Our methodology combines a physical-based model to a parsimonious model. The first allows to identify the climatic changes in the empirical parameter of the second. The second allows for a simple decomposition of the relative changes in discharge. In this case, the Budyko framework and a one-parameter equation: the deviation

from the average curve corresponds to a change in the evaporation efficiency of the catchment. The state-of-the-art LSM was used to simulate changes in the evaporation efficiency under the climate of the past century, independent from any other disruptive process. The successive fit of the parametric equation allows us to find the climatic dependence of the deviation to the average curve in the Budyko space over time.

For a given catchment, we quantify its effect on Q. Over the past century, the primary climatic source of deviation to the average curve is the change in the intra-annual distribution of P. We compare the impact of that deviation on changes in the average annual discharge compared to the change due to average climate variables P and PET. Over Europe for the past century, the main climatic driver of change in the average Q is the change in the average P. The second main driver of discharge change is PET over most of Europe except the Mediterranean area, where a change in the intra-annual distribution of P weighs more on Q changes than PET. Therefore, the effect of the intra-annual distribution of P should be addressed when studying the evolution of the average discharge and water availability under climate change, especially over the Mediterranean.

If we were to work from observations instead of model outputs, there would be other non-climate-related sources of variability, such as direct human activities or vegetation changes which would modify watershed behavior. Our next step is to apply the methodology to quantify these human-induced changes and compare their magnitude to those attributed to climate change in the present study's responses.

Data Availability Statement

The forcing data set GSWP3 used to grid P and other climate data and run the LSM over Europe between 1901 and 2010 in the study is freely available upon registration (Hyungjun, 2017). The LSM used to calculate PET and model the discharge in this study is ORCHIDEE (IPSL [Institut Pierre Simon Laplace], 2017), available on their website. The outputs used for this study at the annual time step for each catchment are gathered in a file freely available on Zenodo.org (Collignan et al., 2023). Stations used in the study come from the Global Runoff Data Centre (GRDC) (2022), completed with the Geoportal of Spain Ministerio (Ministerio para la Transición Ecológica y el Reto Demográfico, 2020) and over France with data from the database HYDRO (Ministere de l'ecologie, du developpement durable et de l'energie, 2021), where the data are freely accessible but have to be gathered region by region and station by station. The file on Zenodo.org (Collignan et al., 2023) also includes the list of the stations used in the study and their main related metadata: their location and the size of the upstream area used to position the station on the grid. The upstream watersheds are reconstructed using the hydrological elevation model HydroSHEDS (Lehner et al., 2008) to construct the routing graphs for rivers on the LSM grid.

Acknowledgments
We would like to acknowledge the support of the Agence Nationale de la Recherche under contract HLIAISE (ANR-19-CE01-0017-02). The lead author would like to thank Institut Polytechnique de Paris for the Gaspard Monge fellowship which funded her Ph.D. thesis.

References

Abatzoglou, J. T., & Ficklin, D. L. (2017). Climatic and physiographic controls of spatial variability in surface water balance over the contiguous United States using the Budyko relationship. *Water Resources Research*, *53*(9), 7630–7643. https://doi.org/10.1002/2017WR020843

Andréassian, V., & Sari, T. (2019). Technical note: On the puzzling similarity of two water balance formulas—Turc–Mezentsev vs. Tixeront–Fu. *Hydrology and Earth System Sciences*, *23*(5), 2339–2350. https://doi.org/10.5194/hess-23-2339-2019

Barella-Ortiz, A., Polcher, J., Tuzet, A., & Laval, K. (2013). Potential evaporation estimation through an unstressed surface-energy balance and its sensitivity to climate change. *Hydrology and Earth System Sciences*, *17*(11), 4625–4639. https://doi.org/10.5194/hess-17-4625-2013

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, *21*(1), 589–615. https://doi.org/10.5194/hess-21-589-2017

Berghuijs, W. R., Gnann, S. J., & Woods, R. A. (2020). Unanswered questions on the Budyko framework. *Hydrological Processes*, *34*(26), 5699–5703. https://doi.org/10.1002/hyp.13958

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., et al. (2019). Changing climate both increases and decreases European river floods. *Nature*, *573*(7772), 108–111. https://doi.org/10.1038/s41586-019-1495-6

Bouwer, L. M., Vermaat, J. E., & Aerts, J. C. J. H. (2008). Regional sensitivities of mean and peak river discharge to climate variability in Europe. *Journal of Geophysical Research: Atmospheres*, *113*(D19), D19103. https://doi.org/10.1029/2008JD010301

Christidis, N., & Stott, P. A. (2022). Human influence on seasonal precipitation in Europe. *Journal of Climate*, *35*(15), 5215–5231. https://doi.org/10.1175/JCLI-D-21-0637.1

Collignan, J., Polcher, J., Bastin, S., & Quintana-Segui, P. (2023). Output of the land surface model ORCHIDEE over river catchments in Europe, run with GSWP3 and synthetic forcings where the precipitation is modified [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.8211025

Coron, L., Andréassian, V., Perrin, C., Bourqui, M., & Hendrickx, F. (2014). On the lack of robustness of hydrologic models regarding water balance simulation: A diagnostic approach applied to three models of increasing complexity on 20 mountainous catchments. *Hydrology and Earth System Sciences*, *18*(2), 727–746. https://doi.org/10.5194/hess-18-727-2014

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., et al. (2019). Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1207–1252. https://doi.org/10.1029/2018MS001545

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2012). Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. *Journal of Hydrology*, *436*(437), 35–50. https://doi.org/10.1016/j.jhydrol.2012.02.033

Douville, H., Raghavan, K., Renwick, J., Allan, R., Arias, P., Barlow, M., et al. (2021). Water cycle changes. In *Climate change 2021: The physical science basis* (pp. 1055–1210). Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/9781009157896.010

Du, C., Sun, F., Yu, J., Liu, X., & Yaning, C. (2016). New interpretation of the role of water balance in an extended Budyko hypothesis in arid regions. *Hydrology and Earth System Sciences*, *20*(1), 393–409. https://doi.org/10.5194/hess-20-393-2016

Ficklin, D. L., Abatzoglou, J. T., Robeson, S. M., Null, S. E., & Knouft, J. H. (2018). Natural and managed watersheds show similar responses to recent climate change. *Proceedings of the National Academy of Sciences*, *115*(34), 8553–8557. https://doi.org/10.1073/pnas.1801026115

García-Barrón, L., Aguilar-Alba, M., Morales, J., & Sousa, A. (2018). Intra-annual rainfall variability in the Spanish hydrographic basins. *International Journal of Climatology*, *38*(5), 2215–2229. https://doi.org/10.1002/joc.5328

García-Barrón, L., Morales, J., & Sousa, A. (2013). Characterisation of the intra-annual rainfall and its evolution (1837–2010) in the southwest of the Iberian Peninsula. *Theoretical and Applied Climatology*, *114*(3), 445–457. https://doi.org/10.1007/s00704-013-0855-7

Gentine, P., D'Odorico, P., Lintner, B., Sivandran, G., & Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. *Geophysical Research Letters*, *39*(19), 19404. https://doi.org/10.1029/2012GL053492

Global Runoff Data Centre (GRDC). (2022). In situ river discharge [Dataset]. Retrieved from https://www.bafg.de/GRDC/EN/02_srvcs/21_tmsrs/riverdischarge_node.html

Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Correspondence: Flawed assumptions compromise water yield assessment. *Nature Communications*, *8*(1), 14795. https://doi.org/10.1038/ncomms14795

Gudmundsson, L., Seneviratne, S. I., & Zhang, X. (2017). Anthropogenic climate change detected in European renewable freshwater resources. *Nature Climate Change*, *7*(11), 813–816. https://doi.org/10.1038/nclimate3416

Guimberteau, M., Laval, K., Perrier, A., & Polcher, J. (2012). Global effect of irrigation and its impact on the onset of the Indian summer monsoon. *Climate Dynamics*, *39*(6), 1329–1348. https://doi.org/10.1007/s00382-011-1252-5

Guion, A., Turquety, S., Polcher, J., Pennel, R., Bastin, S., & Arsouze, T. (2022). Droughts and heatwaves in the Western Mediterranean: Impact on vegetation and wildfires using the coupled WRF-ORCHIDEE regional model (RegIPSL). *Climate Dynamics*, *58*(9), 2881–2903. https://doi.org/10.1007/s00382-021-05938-y

Han, J., Yang, Y., Roderick, M. L., McVicar, T. R., Yang, D., Zhang, S., & Beck, H. E. (2020). Assessing the steady-state assumption in water balance calculation across global catchments. *Water Resources Research*, *56*(7), e2020WR027392. https://doi.org/10.1029/2020WR027392

Hyungjun, K. (2017). Global soil wetness project phase 3 atmospheric boundary conditions (Experiment 1) [Dataset]. Data Integration and Analysis System (DIAS), *5*. https://doi.org/10.20783/DIAS.501

IPSL (Institut Pierre Simon Laplace). (2017). Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) [Software]. Retrieved from https://orchidee.ipsl.fr/

Jaramillo, F., Piemontese, L., Berghuijs, W. R., Wang-Erlandsson, L., Greve, P., & Wang, Z. (2022). Fewer basins will follow their Budyko curves under global warming and fossil-fueled development. *Water Resources Research*, *58*(8), e2021WR031825. https://doi.org/10.1029/2021WR031825

Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., & Xu, C.-Y. (2015). Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters. *Journal of Hydrology*, *522*, 326–338. https://doi.org/10.1016/j.jhydrol.2014.12.060

Kitsara, G., Papaioannou, G., Papathanasiou, A., & Retalis, A. (2013). Dimming/brightening in Athens: Trends in sunshine duration, cloud cover and reference evapotranspiration. *Water Resources Management*, *27*(6), 1623–1633. https://doi.org/10.1007/s11269-012-0229-4

Knutson, T. R., & Zeng, F. (2018). Model assessment of observed precipitation trends over land regions: Detectable human influences and possible low bias in model trends. *Journal of Climate*, *31*(12), 4617–4637. https://doi.org/10.1175/JCLI-D-17-0672.1

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data [Dataset]. Eos, Transactions American Geophysical Union, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Li, D., Pan, M., Cong, Z., Zhang, L., & Wood, E. (2013). Vegetation control on water and energy balance within the Budyko framework. *Water Resources Research*, *49*(2), 969–976. https://doi.org/10.1002/wrcr.20107

Li, S., Du, T., & Gippel, C. J. (2022). A modified Fu (1981) equation with a time-varying parameter that improves estimates of inter-annual variability in catchment water balance. *Water Resources Management*, *36*(5), 1645–1659. https://doi.org/10.1007/s11269-021-03057-1

Luo, Y., Yang, Y., Yang, D., & Zhang, S. (2020). Quantifying the impact of vegetation changes on global terrestrial runoff using the Budyko framework. *Journal of Hydrology*, *590*, 125389. https://doi.org/10.1016/j.jhydrol.2020.125389

Mianabadi, A., Davary, K., Pourreza-Bilondi, M., & Coenders-Gerrits, A. M. J. (2020). Budyko framework; towards non-steady state conditions. *Journal of Hydrology*, *588*, 125089. https://doi.org/10.1016/j.jhydrol.2020.125089

Milly, P. C. D., Dunne, K. A., & Vecchia, A. V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, *438*(7066), 347–350. https://doi.org/10.1038/nature04312

Milly, P. C. D., Wetherald, R. T., Dunne, K. A., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, *415*(6871), 514–517. https://doi.org/10.1038/415514a

Ministere de l'ecologie, du developpement durable et de l'energie. (2021). Hydro [Dataset]. Retrieved from https://www.hydro.eaufrance.fr/rechercher/entites-hydrometriques

Ministerio para la Transición Ecológica y el Reto Demográfico. (2020). Geoportal [Dataset]. Retrieved from https://www.miteco.gob.es/es/cartografia-y-sig/ide/geoportal.html

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*, 16.

Nguyen-Quang, T., Polcher, J., Ducharne, A., Arsouze, T., Zhou, X., Schneider, A., & Fita, L. (2018). ORCHIDEE-ROUTING: Revising the river routing scheme using a high-resolution hydrological database. *Geoscientific Model Development Discussions*, *11*(12), 4965–4985. https://doi.org/10.5194/gmd-11-4965-2018

Ning, T., Zhou, S., Chang, F., Shen, H., Li, Z., & Liu, W. (2019). Interaction of vegetation, climate and topography on evapotranspiration modelling at different time scales within the Budyko framework. *Agricultural and Forest Meteorology*, *275*, 59–68. https://doi.org/10.1016/j.agrformet.2019.05.001

Padrón, R. S., Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Large-scale controls of the surface water balance over land: Insights from a systematic review and meta-analysis. *Water Resources Research*, *53*(11), 9659–9678. https://doi.org/10.1002/2017WR021215

Polcher, J., Schrapffer, A., Dupont, E., Rinchiuso, L., Zhou, X., Boucher, O., et al. (2022). Hydrological modelling on atmospheric grids; using graphs of sub-grid elements to transport energy and water. *EGUsphere*, 1–34. https://doi.org/10.5194/egusphere-2022-690

Quintana-Seguí, P., Barella-Ortiz, A., Regueiro-Sanfiz, S., & Miguez-Macho, G. (2020). The utility of land-surface model simulations to provide drought information in a water management context using global and local forcing datasets. *Water Resources Management*, *34*(7), 2135–2156. https://doi.org/10.1007/s11269-018-2160-9

Quintana-Seguí, P., Turco, M., Herrera, S., & Miguez-Macho, G. (2017). Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-Interim. *Hydrology and Earth System Sciences*, *21*(4), 2187–2201. https://doi.org/10.5194/hess-21-2187-2017

Reaver, N. G. F., Kaplan, D. A., Klammler, H., & Jawitz, J. W. (2020). Technical note: Analytical inversion of the parametric Budyko equations. *Hydrology and Earth System Sciences Discussions*, 1–19. https://doi.org/10.5194/hess-2020-585

Reaver, N. G. F., Kaplan, D. A., Klammler, H., & Jawitz, J. W. (2022). Theoretical and empirical evidence against the Budyko catchment trajectory conjecture. *Hydrology and Earth System Sciences*, *26*(5), 1507–1525. https://doi.org/10.5194/hess-26-1507-2022

Ribes, A., Thao, S., Vautard, R., Dubuisson, B., Somot, S., Colin, J., et al. (2019). Observed increase in extreme daily rainfall in the French Mediterranean. *Climate Dynamics*, *52*(1), 1095–1114. https://doi.org/10.1007/s00382-018-4179-2

Riedel, T., & Weber, T. K. D. (2020). Review: The influence of global change on Europe's water cycle and groundwater recharge. *Hydrogeology Journal*, *28*(6), 1939–1959. https://doi.org/10.1007/s10040-020-02165-3

Roderick, M. L., & Farquhar, G. D. (2011). A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. *Water Resources Research*, *47*(12). https://doi.org/10.1029/2010WR009826

Rottler, E., Francke, T., Bürger, G., & Bronstert, A. (2020). Long-term changes in central European river discharge for 1869–2016: Impact of changing snow covers, reservoir constructions and an intensified hydrological cycle. *Hydrology and Earth System Sciences*, *24*(4), 1721–1740. https://doi.org/10.5194/hess-24-1721-2020

Schneider, C., Laizé, C. L. R., Acreman, M. C., & Flörke, M. (2013). How will climate change modify river flow regimes in Europe? *Hydrology and Earth System Sciences*, *17*(1), 325–339. https://doi.org/10.5194/hess-17-325-2013

Simons, G. W. H., Bastiaanssen, W. G. M., Cheema, M. J. M., Ahmad, B., & Immerzeel, W. W. (2020). A novel method to quantify consumed fractions and non-consumptive use of irrigation water: Application to the Indus Basin Irrigation System of Pakistan. *Agricultural Water Management*, *236*, 106174. https://doi.org/10.1016/j.agwat.2020.106174

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., et al. (2010). Streamflow trends in Europe: Evidence from a dataset of near-natural catchments. *Hydrology and Earth System Sciences*, *14*(12), 2367–2382. https://doi.org/10.5194/hess-14-2367-2010

Tafasca, S., Ducharne, A., & Valentin, C. (2020). Weak sensitivity of the terrestrial water budget to global soil texture maps in the ORCHIDEE land surface model. *Hydrology and Earth System Sciences*, *24*(7), 3753–3774. https://doi.org/10.5194/hess-24-3753-2020

Tian, L., Jin, J., Wu, P., & Niu, G.-y. (2018). Quantifying the impact of climate change and human activities on streamflow in a semi-arid watershed with the Budyko equation incorporating dynamic vegetation information. *Water*, *10*(12), 1781. https://doi.org/10.3390/w10121781

Tuel, A., Schaefli, B., Zscheischler, J., & Martius, O. (2022). On the links between sub-seasonal clustering of extreme precipitation and high discharge in Switzerland and Europe. *Hydrology and Earth System Sciences*, *26*(10), 2649–2669. https://doi.org/10.5194/hess-26-2649-2022

Vicente-Serrano, S. M., Lopez-Moreno, J.-I., Beguería, S., Lorenzo-Lacruz, J., Sanchez-Lorenzo, A., García-Ruiz, J. M., et al. (2014). Evidence of increasing drought severity caused by temperature rise in southern Europe. *Environmental Research Letters*, *9*(4), 044001. https://doi.org/10.1088/1748-9326/9/4/044001

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., et al. (2019). Climate, irrigation, and land cover change explain streamflow trends in countries bordering the northeast Atlantic. *Geophysical Research Letters*, *46*(19), 10821–10833. https://doi.org/10.1029/2019GL084084

Wang, F., Polcher, J., Peylin, P., & Bastrikov, V. (2018). Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles. *Hydrology and Earth System Sciences*, *22*(7), 3863–3882. https://doi.org/10.5194/hess-22-3863-2018

Wang, W., Zhang, Y., & Tang, Q. (2020). Impact assessment of climate change and human activities on streamflow signatures in the Yellow River Basin using the Budyko hypothesis and derived differential equation. *Journal of Hydrology*, *591*, 125460. https://doi.org/10.1016/j.jhydrol.2020.125460

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, *50*(9), 7505–7514. https://doi.org/10.1002/2014WR015638

Xing, W., Wang, W., Shao, Q., & Yong, B. (2018). Identification of dominant interactions between climatic seasonality, catchment characteristics and agricultural activities on Budyko-type equation parameter estimation. *Journal of Hydrology*, *556*, 585–599. https://doi.org/10.1016/j.jhydrol.2017.11.048

Xiong, M., Huang, C.-S., & Yang, T. (2020). Assessing the impacts of climate change and land use/cover change on runoff based on improved Budyko framework models considering arbitrary partition of the impacts. *Water*, *12*(6), 1612. https://doi.org/10.3390/w12061612

Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., & Lei, Z. (2007). Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. *Water Resources Research*, *43*(4). https://doi.org/10.1029/2006WR005224

Yang, H., Yang, D., Lei, Z., & Sun, F. (2008). New analytical derivation of the mean annual water-energy balance equation. *Water Resources Research*, *44*(3). https://doi.org/10.1029/2007WR006135

Yang, Y., Zhang, S., McVicar, T. R., Beck, H. E., Zhang, Y., & Liu, B. (2018). Disconnection between trends of atmospheric drying and continental runoff. *Water Resources Research*, *54*(7), 4700–4713. https://doi.org/10.1029/2018WR022593

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., & Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, *40*(2). https://doi.org/10.1029/2003WR002710

Zhang, L., Potter, N., Hickel, K., Zhang, Y., & Shao, Q. (2008). Water balance modeling over variable time scales based on the Budyko framework—Model development and testing. *Journal of Hydrology*, *360*(1), 117–131. https://doi.org/10.1016/j.jhydrol.2008.07.021

Zhang, X., Dong, Q., Costa, V., & Wang, X. (2019). A hierarchical Bayesian model for decomposing the impacts of human activities and climate change on water resources in China. *Science of the Total Environment*, *665*, 836–847. https://doi.org/10.1016/j.scitotenv.2019.02.189

Zhao, J., Huang, S., Huang, Q., Wang, H., & Leng, G. (2018). Detecting the dominant cause of streamflow decline in the Loess Plateau of China based on the latest Budyko equation. *Water*, *10*(9), 1277. https://doi.org/10.3390/w10091277

Zheng, Y., Huang, Y., Zhou, S., Wang, K., & Wang, G. (2018). Effect partition of climate and catchment changes on runoff variation at the headwater region of the Yellow River based on the Budyko complementary relationship. *Science of the Total Environment*, *643*, 1166–1177. https://doi.org/10.1016/j.scitotenv.2018.06.195

Zveryaev, I. I. (2004). Seasonality in precipitation variability over Europe. *Journal of Geophysical Research: Atmospheres*, *109*(D5), D05103. https://doi.org/10.1029/2003JD003668

Bibliography

Abatzoglou, J. T., & Ficklin, D. L. (2017). Climatic and physiographic controls of spatial variability in surface water balance over the contiguous United States using the Budyko relationship. *Water Resources Research*, *53*, 7630–7643.

Adeyeri, O. E., Laux, P., Lawin, A. E., & Arnault, J. (2020). Assessing the impact of human activities and rainfall variability on the river discharge of Komadugu-Yobe Basin, Lake Chad Area. *Environmental Earth Sciences*, *79*(6), 143.

Ahn, K.-H., & Merwade, V. (2014). Quantifying the relative impact of climate and human activities on streamflow. *Journal of Hydrology*, *515*, 257–266.

Alkama, R., Decharme, B., Douville, H., & Ribes, A. (2011). Trends in Global and Basin-Scale Runoff over the Late Twentieth Century: Methodological Issues and Sources of Uncertainty. *Journal of Climate*, *24*(12), 3000–3014.

Alkama, R., Kageyama, M., & Ramstein, G. (2010). Relative contributions of climate change, stomatal closure, and leaf area index changes to 20th and 21st century runoff change: A modelling approach using the Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) land surface model. *Journal of Geophysical Research: Atmospheres*, *115*(D17).

Alkama, R., Marchand, L., Ribes, A., & Decharme, B. (2013). Detection of global runoff changes: Results from observations and CMIP5 experiments. *Hydrology and Earth System Sciences*, *17*(7), 2967–2979.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Chapter 4 - Determination of ETo. In *Crop Evapotranspiration - Guidelines for Computing Crop Water Requirements*. Rome: Food and Agriculture Organization of the United Nations (FAO).

Andréassian, V., Coron, L., Lerat, J., & Le Moine, N. (2016a). Climate elasticity of streamflow revisited – an elasticity index based on long-term hydrometeorological records. *Hydrology and Earth System Sciences*, *20*(11), 4503–4524.

Andréassian, V., Mander, Ü., & Pae, T. (2016b). The Budyko hypothesis before Budyko: The hydrological legacy of Evald Oldekop. *Journal of Hydrology*, *535*, 386–391.

Andréassian, V., & Sari, T. (2019). Technical Note: On the puzzling similarity of two water balance formulas – Turc–Mezentsev vs. Tixeront–Fu. *Hydrology and Earth System Sciences*, *23*(5), 2339–2350.

Angelakıs, A. N., Zaccaria, D., Krasilnikoff, J., Salgot, M., Bazza, M., Roccaro, P., Jimenez, B., Kumar, A., Yinghua, W., Baba, A., Harrison, J. A., Garduno-Jimenez, A., & Fereres, E. (2020). Irrigation of World Agricultural Lands: Evolution through the Millennia. *Water*, *12*(5), 1285.

Arheimer, B., Donnelly, C., & Lindström, G. (2017). Regulation of snow-fed rivers affects flow regimes more than climate change. *Nature Communications*, *8*(1), 62.

Arnell, N. W. (1999). The effect of climate change on hydrological regimes in Europe: A continental perspective. *Global Environmental Change*, *9*(1), 5–23.

Ban, N., Caillaud, C., Coppola, E., Pichelli, E., Sobolowski, S., Adinolfi, M., Ahrens, B., Alias, A., Anders, I., Bastin, S., Belušić, D., Berthou, S., Brisson, E., Cardoso, R. M., Chan, S. C., Christensen, O. B., Fernández, J., Fita, L., Frisius, T., Gašparac, G., Giorgi, F., Goergen, K., Haugen, J. E., Hodnebrog, Ø., Kartsios, S., Katragkou, E., Kendon, E. J., Keuler, K., Lavin-Gullon, A., Lenderink, G., Leutwyler, D., Lorenz, T., Maraun, D., Mercogliano, P., Milovac, J., Panitz, H.-J., Raffa, M., Remedio, A. R., Schär, C., Soares, P. M. M., Srnec, L., Steensen, B. M., Stocchi, P., Tölle, M. H., Truhetz, H., Vergara-Temprado, J., de Vries, H., Warrach-Sagi, K., Wulfmeyer, V., & Zander, M. J. (2021). The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: Evaluation of precipitation. *Climate Dynamics*, *57*(1), 275–302.

Baratgin, L., Polcher, J., Quirion, P., & Dumas, P. (2023). Simulating French hydropower operations in a land surface model. Tech. Rep. EGU23-3396, Copernicus Meetings.

Barella-Ortiz, A., Polcher, J., Tuzet, A., & Laval, K. (2013). Potential evaporation estimation through an unstressed surface-energy balance and its sensitivity to climate change. *Hydrology and Earth System Sciences*, *17*(11), 4625–4639.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, *21*(1), 589–615.

Beguería, S., López-Moreno, J. I., Lorente, A., Seeger, M., & García-Ruiz, J. M. (2003). Assessing the Effect of Climate Oscillations and Land-use Changes on Streamflow in the Central Spanish Pyrenees. *AMBIO: A Journal of the Human Environment*, *32*(4), 283–286.

Beguería, S., Luis, J., Lambán, J., Dewandel, B., Desprats, J.-F., Zabaleta, A., Antigüedad, I., Caballero, Y., Lanini, S., Cointe, P. L., Béranger, S., Arnaud, L., Pinson, S., & Lechevalier, J. (2019). Groundwater resources of the Pyrenees in the global change context -The PIRAGUA Project. In *46th IAN Congress Malaga - IAH 2019*.

Behrangi, A., Andreadis, K., Fisher, J. B., Turk, F. J., Granger, S., Painter, T., & Das, N. (2014). Satellite-Based Precipitation Estimation and Its Application for Streamflow Prediction over Mountainous Western U.S. Basins. *Journal of Applied Meteorology and Climatology*, *53*(12), 2823–2842.

Berghuijs, W. R., Gnann, S. J., & Woods, R. A. (2020). Unanswered questions on the Budyko framework. *Hydrological Processes*, *34*(26), 5699–5703.

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The Relative Importance of Different Flood-Generating Mechanisms Across Europe. *Water Resources Research*, *55*(6), 4582–4593.

Beven, K. J., & Cloke, H. L. (2012). Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al.: COMMENTARY. *Water Resources Research*, *48*(1).

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT Mission and Its Capabilities for Land Hydrology. In A. Cazenave, N. Champollion, J. Benveniste, & J. Chen (Eds.) *Remote Sensing and Water Resources*, Space Sciences Series of ISSI, (pp. 117–147). Cham: Springer International Publishing.

Blöschl, G., Bierkens, M. F., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Allen, S. T., Amin, A., Andréassian, V., Arheimer, B., Aryal, S. K., Baker, V., Bardsley, E., Barendrecht, M. H., Bartosova, A., Batelaan, O., Berghuijs, W. R., Beven, K., Blume, T., Bogaard, T., Borges de Amorim, P., Böttcher, M. E., Boulet, G., Breinl, K., Brilly, M., Brocca, L., Buytaert, W., Castellarin, A., Castelletti, A., Chen, X., Chen, Y., Chen, Y., Chifflard, P., Claps, P., Clark, M. P., Collins, A. L., Croke, B., Dathe, A., David, P. C., de Barros, F. P. J., de Rooij, G., Di Baldassarre, G., Driscoll, J. M., Duethmann, D., Dwivedi, R., Eris, E., Farmer, W. H., Feiccabrino, J., Ferguson, G., Ferrari, E., Ferraris, S., Fersch, B., Finger, D., Foglia, L., Fowler, K., Gartsman, B., Gascoin, S., Gaume, E., Gelfan, A., Geris, J., Gharari, S., Gleeson, T., Glendell, M., Gonzalez Bevacqua, A., González-Dugo, M. P., Grimaldi, S., Gupta, A. B., Guse, B., Han, D., Hannah, D., Harpold, A., Haun, S., Heal, K., Helfricht, K., Herrnegger, M., Hipsey, M., Hlaváčiková, H., Hohmann, C., Holko, L., Hopkinson, C., Hrachowitz, M., Illangasekare, T. H., Inam, A., Innocente, C., Istanbulluoglu, E., Jarihani, B., Kalantari, Z., Kalvans, A., Khanal, S., Khatami, S., Kiesel, J., Kirkby, M., Knoben, W., Kochanek, K., Kohnová, S., Kolechkina, A., Krause, S., Kreamer, D., Kreibich, H., Kunstmann, H., Lange, H., Liberato, M. L. R., Lindquist, E., Link, T., Liu, J., Loucks, D. P., Luce, C., Mahé, G., Makarieva, O., Malard, J., Mashtayeva, S., Maskey, S., Mas-Pla, J., Mavrova-Guirguinova, M., Mazzoleni, M., Mernild, S., Misstear, B. D., Montanari, A., Müller-Thomy, H., Nabizadeh, A., Nardi, F., Neale, C., Nesterova, N., Nurtaev, B., Odongo, V. O., Panda, S., Pande, S., Pang, Z., Papacharalampous, G., Perrin, C., Pfister, L., Pimentel, R., Polo, M. J., Post, D., Prieto Sierra, C., Ramos, M.-H., Renner, M., Reynolds, J. E., Ridolfi, E., Rigon, R., Riva, M., Robertson, D. E., Rosso, R., Roy, T., Sá, J. H., Salvadori, G., Sandells, M., Schaefli, B., Schumann, A., Scolobig, A., Seibert, J., Servat, E., Shafiei, M., Sharma, A., Sidibe, M., Sidle, R. C., Skaugen, T., Smith, H., Spiessl, S. M., Stein, L., Steinsland, I., Strasser, U., Su, B., Szolgay, J., Tarboton, D., Tauro, F., Thirel, G., Tian, F., Tong, R., Tussupova, K., Tyralis, H., Uijlenhoet, R., van Beek, R., van der Ent, R. J., van der Ploeg, M., Van Loon, A. F., van Meerveld, I., van Nooijen, R., van Oel, P. R., Vidal, J.-P., von Freyberg, J., Vorogushyn, S., Wachniew, P., Wade, A. J., Ward, P., Westerberg, I. K., White, C., Wood, E. F., Woods, R., Xu, Z., Yilmaz, K. K., & Zhang, Y. (2019a). Twenty-three unsolved problems in hydrology (UPH) – a community perspective. *Hydrological Sciences Journal*, *64*(10), 1141–1158.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G. T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Frolova, N., Ganora, D., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Volpi, E., Wilson, D., Zaimi, K., & Živković, N. (2019b). Changing climate both increases and decreases European river floods. *Nature*, *573*(7772), 108–111.

Boone, A. A. (2019). Land surface Interactions with the Atmosphere over the Iberian Semiarid Environment (LIAISE).

Bouwer, L. M., Vermaat, J. E., & Aerts, J. C. J. H. (2008). Regional sensitivities of mean and peak river discharge to climate variability in Europe. *Journal of Geophysical Research: Atmospheres*, *113*(D19).

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., & Fernández-Prieto, D. (2018). How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. *International Journal of Applied Earth Observation and Geoinformation*, *73*, 752–766.

Budyko, M. I. M. I. (1974). Climate and life. *Academic Press, New York*, (p. 508pp.).

Christidis, N., & Stott, P. A. (2022). Human Influence on Seasonal Precipitation in Europe. *Journal of Climate*, *35*(15), 5215–5231.

Coch, A., & Mediero, L. (2016). Trends in low flows in Spain in the period 1949–2009. *Hydrological Sciences Journal*, *61*(3), 568–584.

Collignan, J., Polcher, J., Bastin, S., & Quintana-Segui, P. (2023a). Budyko Framework Based Analysis of the Effect of Climate Change on Watershed Evaporation Efficiency and Its Impact on Discharge Over Europe. *Water Resources Research*, *59*(10), e2023WR034509.

Collignan, J., Polcher, J., Bastin, S., & Quintana-Seguí, P. (2023b). Output of the Land Surface Model ORCHIDEE over river catchments in Europe, run with GSWP3 and synthetic forcings where the precipitation is modified.

Coron, L., Andréassian, V., Perrin, C., Bourqui, M., & Hendrickx, F. (2014). On the lack of robustness of hydrologic models regarding water balance simulation: A diagnostic approach applied to three models of increasing complexity on 20 mountainous catchments. *Hydrology and Earth System Sciences*, *18*(2), 727–746.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., & Buontempo, C. (2020). WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, *12*(3), 2097–2120.

Custodio, E., Andreu-Rodes, J. M., Aragón, R., Estrela, T., Ferrer, J., García-Aróstegui, J. L., Manzano, M., Rodríguez-Hernández, L., Sahuquillo, A., & del Villar, A. (2016). Groundwater intensive use and mining in south-eastern peninsular Spain: Hydrogeological, economic and social aspects. *Science of The Total Environment*, *559*, 302–316.

Dai, A. (2016). Historical and Future Changes in Streamflow and Continental Runoff. In *Terrestrial Water Cycle and Climate Change*, chap. 2, (pp. 17–37). American Geophysical Union (AGU).

Dai, A., Qian, T., Trenberth, K. E., & Milliman, J. D. (2009). Changes in Continental Freshwater Discharge from 1948 to 2004. *Journal of Climate*, *22*(10), 2773–2792.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., & Voldoire, A. (2019). Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1207–1252.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., & Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

Dezsi, Ş., Mîndrescu, M., Petrea, D., Rai, P. K., Hamann, A., & Nistor, M.-M. (2018). High-resolution projections of evapotranspiration and water availability for Europe under climate change. *International Journal of Climatology*, *38*(10), 3832–3841.

Dirmeyer, P., Gao, X., & Oki, T. (2003). The Second Global Soil Wetness Project (GSWP-2).

Dirmeyer, P. A., Dolman, A. J., & Sato, N. (1999). The Pilot Phase of the Global Soil Wetness Project. *Bulletin of the American Meteorological Society*, *80*(5), 851–878.

Do, H. X., Zhao, F., Westra, S., Leonard, M., Gudmundsson, L., Boulange, J. E. S., Chang, J., Ciais, P., Gerten, D., Gosling, S. N., Müller Schmied, H., Stacke, T., Telteu, C.-E., & Wada, Y. (2020). Historical and future changes in global flood magnitude – evidence from a model–observation investigation. *Hydrology and Earth System Sciences*, *24*(3), 1543–1564.

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2012). Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. *Journal of Hydrology*, *436–437*, 35–50.

Doury, A., Somot, S., Gadat, S., Ribes, A., & Corre, L. (2023). Regional climate model emulator based on deep learning: Concept and first evaluation of a novel hybrid downscaling approach. *Climate Dynamics*, *60*(5), 1751–1779.

Douville, H., Raghavan, K., Renwick, J., Allan, R., Arias, P., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T., Gergis, J., Jiang, D., Khan, A., Pokam Mba, W., Rosenfeld, D., Tierney, J., & Zolina, O. (2021). Water Cycle Changes. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. (pp. 1055–1210).

Du, C., Sun, F., Yu, J., Liu, X., & Yaning, C. (2016). New interpretation of the role of water balance in an extended Budyko hypothesis in arid regions. *Hydrology and Earth System Sciences*, *20*, 393–409.

Durand, Y., Brun, E., Merindol, L., Guyomarc'h, G., Lesaffre, B., & Martin, E. (1993). A meteorological estimation of relevant parameters for snow models. *Annals of Glaciology*, *18*, 65–71.

Esteban, E., & Albiac, J. (2012). The problem of sustainable groundwater management: The case of La Mancha aquifers, Spain. *Hydrogeology Journal*, *20*(5), 851–863.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., van Verseveld, W., Volk, J., & Yamazaki, D. (2019). Hillslope Hydrology in Global Change Research and Earth System Modeling. *Water Resources Research*, *55*(2), 1737–1772.

Fang, H., Hrubiak, P., Kato, H., Rodell, M., Teng, W. L., & Vollmer, B. E. (2008). Global Land Data Assimilation System (GLDAS) Products from NASA Hydrology Data and Information Services Center (HDISC). In *American Society for Photogrammetry and Remote Sensing (ASPRS)*. Portland, OR.

Federal Institute of Hydrology (BfG), K., Germany (2023). The GRDC - Global Runoff Database. https://www.bafg.de/GRDC/EN/01_GRDC/13_dtbse/database_node.html.

Fenta, A. A., Yasuda, H., Shimizu, K., & Haregeweyn, N. (2017). Response of streamflow to climate variability and changes in human activities in the semiarid highlands of northern Ethiopia. *Regional Environmental Change*, *17*(4), 1229–1240.

Ficklin, D. L., Abatzoglou, J. T., Robeson, S. M., Null, S. E., & Knouft, J. H. (2018). Natural and managed watersheds show similar responses to recent climate change. *Proceedings of the National Academy of Sciences*, *115*(34), 8553–8557.

FitzHugh, T. W., & Vogel, R. M. (2011). The impact of dams on flood flows in the United States. *River Research and Applications*, *27*(10), 1192–1215.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., Suzuki, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., van Vliet, M., Biber, M. F., Betts, R. A., Bodirsky, B. L., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., & Yamagata, Y. (2017). Assessing the impacts of 1.5 ${}\^{}\circ$ C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, *10*(12), 4321–4345.

Fuchs, R., Herold, M., Verburg, P., Clevers, J., & Eberle, J. (2015). Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010. *Global Change Biology*, *21*(1), 299–313.

García-Barrón, L., Aguilar-Alba, M., Morales, J., & Sousa, A. (2018). Intra-annual rainfall variability in the Spanish hydrographic basins. *International Journal of Climatology*, *38*(5), 2215–2229.

García-Barrón, L., Morales, J., & Sousa, A. (2013). Characterisation of the intra-annual rainfall and its evolution (1837–2010) in the southwest of the Iberian Peninsula. *Theoretical and Applied Climatology*, *114*(3), 445–457.

García-Ruiz, J. M., López-Moreno, J. I., Vicente-Serrano, S. M., Lasanta–Martínez, T., & Beguería, S. (2011). Mediterranean water resources in a global change scenario. *Earth-Science Reviews*, *105*(3), 121–139.

Gardner, L. R. (2009). Assessing the effect of climate change on mean annual runoff. *Journal of Hydrology*, *379*(3), 351–359.

Gentine, P., D'Odorico, P., Lintner, B., Sivandran, G., & Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. *Geophysical Research Letters*, *39*, 19404.

Greve, P., Gudmundsson, L., Orlowsky, B., & Seneviratne, S. I. (2015). Introducing a probabilistic Budyko framework. *Geophysical Research Letters*, *42*(7), 2261–2269.

Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017a). Correspondence: Flawed assumptions compromise water yield assessment. *Nature Communications*, *8*(1), 14795.

Gudmundsson, L., Seneviratne, S. I., & Zhang, X. (2017b). Anthropogenic climate change detected in European renewable freshwater resources. *Nature Climate Change*, *7*(11), 813–816.

Guimberteau, M., Laval, K., Perrier, A., & Polcher, J. (2012). Global effect of irrigation and its impact on the onset of the Indian summer monsoon. *Climate Dynamics*, *39*(6), 1329–1348.

Guion, A., Turquety, S., Polcher, J., Pennel, R., Bastin, S., & Arsouze, T. (2022). Droughts and heatwaves in the Western Mediterranean: Impact on vegetation and wildfires using the coupled WRF-ORCHIDEE regional model (RegIPSL). *Climate Dynamics*, *58*(9), 2881–2903.

Han, J., Yang, Y., Roderick, M. L., McVicar, T. R., Yang, D., Zhang, S., & Beck, H. E. (2020). Assessing the Steady-State Assumption in Water Balance Calculation Across Global Catchments. *Water Resources Research*, *56*(7), e2020WR027392.

Han, S., Hu, H., Yang, D., & Liu, Q. (2011). Irrigation impact on annual water balance of the oases in Tarim Basin, Northwest China. *Hydrological Processes*, *25*(2), 167–174.

Hanasaki, N., Kanae, S., & Oki, T. (2006). A reservoir operation scheme for global river routing models. *Journal of Hydrology*, *327*(1), 22–41.

Hannerz, F., & Destouni, G. (2006). Spatial Characterization of the Baltic Sea Drainage Basin and Its Unmonitored Catchments. *AMBIO: A Journal of the Human Environment*, *35*(5), 214–219.

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1), 109.

Hegerl, G., & Zwiers, F. (2011). Use of models in detection and attribution of climate change. *WIREs Climate Change*, *2*(4), 570–591.

Hersbach, H. (2016). The ERA5 Atmospheric Reanalysis. *2016*, NG33D–01.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Raluca, R., Schepers, D., Simmons, A., Soci, C., & Dee, D. (201904). Global reanalysis: Goodbye ERA-Interim, hello ERA5. https://www.ecmwf.int/en/elibrary/81046-global-reanalysis-goodbye-era-interim-hello-era5.

Hewitson, B. C., Daron, J., Crane, R. G., Zermoglio, M. F., & Jack, C. (2014). Interrogating empirical-statistical downscaling. *Climatic Change*, *122*(4), 539–554.

Hewitt, C. D., & Lowe, J. A. (2018). Toward a European Climate Prediction System. *Bulletin of the American Meteorological Society*, *99*(10), 1997–2001.

Holtz, G., & Pahl-Wostl, C. (2012). An agent-based model of groundwater over-exploitation in the Upper Guadiana, Spain. *Regional Environmental Change*, *12*(1), 95–121.

Huang, J. (2018). A Simple Accurate Formula for Calculating Saturation Vapor Pressure of Water and Ice. *Journal of Applied Meteorology and Climatology*, *57*(6), 1265–1272.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., & Zhang, X. (2020). Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geoscientific Model Development*, *13*(11), 5425–5464.

Huss, M., & Hock, R. (2018). Global-scale hydrological response to future glacier mass loss. *Nature Climate Change*, *8*(2), 135–140.

Hyungjun, K. (2017). Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1). *Data Integration and Analysis System (DIAS)*, (p. 5).

ISIMIP (2012). The Inter-Sectoral Impact Model Intercomparison Project. https://www.isimip.org/.

Jaramillo, F., Piemontese, L., Berghuijs, W. R., Wang-Erlandsson, L., Greve, P., & Wang, Z. (2022). Fewer Basins Will Follow Their Budyko Curves Under Global Warming and Fossil-Fueled Development. *Water Resources Research*, *58*(8), e2021WR031825.

Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., & Xu, C.-Y. (2015). Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters. *Journal of Hydrology*, *522*, 326–338.

Khan, M. S., Coulibaly, P., & Dibike, Y. (2006). Uncertainty analysis of statistical downscaling methods. *Journal of Hydrology*, *319*(1), 357–382.

Kirby, J. M., Mainuddin, Md., Ahmad, M. D., & Gao, L. (2013). Simplified Monthly Hydrology and Irrigation Water Use Model to Explore Sustainable Water Management Options in the Murray-Darling Basin. *Water Resources Management*, *27*(11), 4083–4097.

Kitsara, G., Papaioannou, G., Papathanasiou, A., & Retalis, A. (2013). Dimming/brightening in Athens: Trends in Sunshine Duration, Cloud Cover and Reference Evapotranspiration. *Water Resources Management*, *27*(6), 1623–1633.

Knutson, T. R., & Zeng, F. (2018). Model Assessment of Observed Precipitation Trends over Land Regions: Detectable Human Influences and Possible Low Bias in Model Trends. *Journal of Climate*, *31*(12), 4617–4637.

Kour, R., Patel, N., & Krishna, A. P. (2016). Climate and hydrological models to assess the impact of climate change on hydrological regime: A review. *Arabian Journal of Geosciences*, *9*(9), 544.

Labat, D. (2010). Cross wavelet analyses of annual continental freshwater discharge and selected climate indices. *Journal of Hydrology*, *385*(1), 269–278.

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., de Noblet-Ducoudré, N., Pongratz, J., Seneviratne, S. I., & Shevliakova, E. (2016). The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6:rationale and experimental design. *Geoscientific Model Development*, *9*(9), 2973–2998.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New Global Hydrography Derived From Spaceborne Elevation Data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94.

Li, D., Pan, M., Cong, Z., Zhang, L., & Wood, E. (2013). Vegetation control on water and energy balance within the Budyko framework. *Water Resources Research*, *49*(2), 969–976.

Li, S., Du, T., & Gippel, C. J. (2022). A Modified Fu (1981) Equation with a Time-varying Parameter that Improves Estimates of Inter-annual Variability in Catchment Water Balance. *Water Resources Management*, *36*(5), 1645–1659.

Llamas, M. R., Custodio, E., de la Hera, A., & Fornés, J. M. (2015). Groundwater in Spain: Increasing role, evolution, present and future. *Environmental Earth Sciences*, *73*(6), 2567–2578.

Lucas-Picher, P., Argüeso, D., Brisson, E., Tramblay, Y., Berg, P., Lemonsu, A., Kotlarski, S., & Caillaud, C. (2021). Convection-permitting modeling with regional climate models: Latest developments and next steps. *WIREs Climate Change*, *12*(6), e731.

Lundquist, J., Hughes, M., Gutmann, E., & Kapnick, S. (2019). Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. *Bulletin of the American Meteorological Society*, *100*(12), 2473–2490.

Luo, Y., Yang, Y., Yang, D., & Zhang, S. (2020). Quantifying the impact of vegetation changes on global terrestrial runoff using the Budyko framework. *Journal of Hydrology*, *590*, 125389.

McDermid, S., Nocco, M., Lawston-Parker, P., Keune, J., Pokhrel, Y., Jain, M., Jägermeyr, J., Brocca, L., Massari, C., Jones, A. D., Vahmani, P., Thiery, W., Yao, Y., Bell, A., Chen, L., Dorigo, W., Hanasaki, N., Jasechko, S., Lo, M.-H., Mahmood, R., Mishra, V., Mueller, N. D., Niyogi, D., Rabin, S. S., Sloat, L., Wada, Y., Zappa, L., Chen, F., Cook, B. I., Kim, H., Lombardozzi, D., Polcher, J., Ryu, D., Santanello, J., Satoh, Y., Seneviratne, S., Singh, D., & Yokohata, T. (2023). Irrigation in the Earth system. *Nature Reviews Earth & Environment*, *4*(7), 435–453.

Mianabadi, A., Davary, K., Pourreza-Bilondi, M., & Coenders-Gerrits, A. M. J. (2020). Budyko framework; towards non-steady state conditions. *Journal of Hydrology*, *588*, 125089.

Milliman, J. D., Farnsworth, K. L., Jones, P. D., Xu, K. H., & Smith, L. C. (2008). Climatic and anthropogenic factors affecting river discharge to the global ocean, 1951–2000. *Global and Planetary Change*, *62*(3), 187–194.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? *Science*, *319*(5863), 573–574.

Milly, P. C. D., & Dunne, K. A. (2016). Potential evapotranspiration and continental drying. *Nature Climate Change*, *6*(10), 946–949.

Milly, P. C. D., Dunne, K. A., & Vecchia, A. V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, *438*(7066), 347–350.

Milly, P. C. D., Wetherald, R. T., Dunne, K. A., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, *415*(6871), 514–517.

Ministere de l'ecologie, du developpement durable et de l'energie (2021). HYDRO. https://www.hydro.eaufrance.fr/rechercher/entites-hydrometriques.

Ministerio para la Transición Ecológica y el Reto Demográfico (2020). GeoPortal. https://www.miteco.gob.es/es/cartografia-y-sig/ide/geoportal.html.

Montanari, A. (2012). Hydrology of the Po River: Looking for changing patterns in river discharge. *Hydrology and Earth System Sciences*, *16*(10), 3739–3747.

Monteiro, D., & Morin, S. (2023). Multi-decadal analysis of past winter temperature, precipitation and snow cover data in the European Alps from reanalyses, climate models and observational datasets. *The Cryosphere*, *17*(8), 3617–3660.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). MODEL EVALUATION GUIDELINES FOR SYSTEMATIC QUANTIFICATION OF ACCURACY IN WATERSHED SIMULATIONS. *TRANSACTIONS OF THE ASABE*, *50*, 16.

Müller, O. V., Vidale, P. L., Vannière, B., Schiemann, R., & McGuire, P. C. (2021). Does the HadGEM3-GC3.1 GCM Overestimate Land Precipitation at High Resolution? A Constraint Based on Observed River Discharge. *Journal of Hydrometeorology*, *22*(8), 2131–2151.

Nazemi, A., & Wheater, H. S. (2015a). On inclusion of water resource management in Earth system models – Part 1: Problem definition and representation of water demand. *Hydrology and Earth System Sciences*, *19*(1), 33–61.

Nazemi, A., & Wheater, H. S. (2015b). On inclusion of water resource management in Earth system models – Part 2: Representation of water supply and allocation and opportunities for improved modeling. *Hydrology and Earth System Sciences*, *19*(1), 63–90.

Neverre, N., & Dumas, P. (2015). Projecting and valuing domestic water use at regional scale: A generic method applied to the Mediterranean at the 2060 horizon. *Water Resources and Economics*, *11*, 33–46.

Nguyen-Quang, T., Polcher, J., Ducharne, A., Arsouze, T., Zhou, X., Schneider, A., & Fita, L. (2018). ORCHIDEE-ROUTING: Revising the river routing scheme using a high-resolution hydrological database. *Geoscientific Model Development Discussions*, *11*(12), 4965–4985.

Nicolle, P., Andréassian, V., Royer-Gaspard, P., Perrin, C., Thirel, G., Coron, L., & Santos, L. (2021). Technical note: RAT – a robustness assessment test for calibrated and uncalibrated hydrological models. *Hydrology and Earth System Sciences*, *25*(9), 5013–5027.

Ning, T., Zhou, S., Chang, F., Shen, H., Li, Z., & Liu, W. (2019). Interaction of vegetation, climate and topography on evapotranspiration modelling at different time scales within the Budyko framework. *Agricultural and Forest Meteorology*, *275*, 59–68.

Nohara, D., Kitoh, A., Hosaka, M., & Oki, T. (2006). Impact of Climate Change on River Discharge Projected by Multimodel Ensemble. *Journal of Hydrometeorology*, *7*(5), 1076–1089.

Oki, T., KIM, H., Ferguson, C. R., Dirmeyer, P., & Seneviratne, S. I. (2013). Global water balances reconstructed by multi-model offline simulations of land surface models under GSWP3 (Invited). *2013*, H53J–01.

Oldekop, E. (1911). Evaporation from the surface of river basins (in Russian). *Collection of the Works of Students of the Meteorological Observatory*, (p. 209).

Padrón, R. S., Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis. *Water Resources Research*, *53*(11), 9659–9678.

Palazón, L., & Beguería, S. (2022). Water cycle and water resources of the Pyrenees under climate change: The PIRAGUA datasets. (pp. EGU22–7825).

Palmer, M. A., Reidy Liermann, C. A., Nilsson, C., Flörke, M., Alcamo, J., Lake, P. S., & Bond, N. (2008). Climate change and the world's river basins: Anticipating management options. *Frontiers in Ecology and the Environment*, *6*(2), 81–89.

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Ottle, C., Poulter, B., & Zaehle, S. (2019). Evaluation of global terrestrial evapotranspiration by state-of-the-art approaches in remote sensing, machine learning, and land surface models. Preprint, Global hydrology/Modelling approaches.

Panziera, L., Gabella, M., Germann, U., & Martius, O. (2018). A 12-year radar-based climatology of daily and sub-daily extreme precipitation over the Swiss Alps. *International Journal of Climatology*, *38*(10), 3749–3769.

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, *279*(1), 275–289.

Polcher, J., Schrapffer, A., Dupont, E., Rinchiuso, L., Zhou, X., Boucher, O., Mouche, E., Ottlé, C., & Servonnat, J. (2022). Hydrological modelling on atmospheric grids; using graphs of sub-grid elements to transport energy and water. *EGUsphere*, (pp. 1–34).

Poschlod, B., Willkofer, F., & Ludwig, R. (2020). Impact of Climate Change on the Hydrological Regimes in Bavaria. *Water*, *12*(6), 1599.

Quintana-Seguí, P., Barella-Ortiz, A., Regueiro-Sanfiz, S., & Miguez-Macho, G. (2020). The Utility of Land-Surface Model Simulations to Provide Drought Information in a Water Management Context Using Global and Local Forcing Datasets. *Water Resources Management*, *34*(7), 2135–2156.

Quintana-Seguí, P., Moigne, P. L., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., & Morel, S. (2008). Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France. *Journal of Applied Meteorology and Climatology*, *47*(1), 92–107.

Quintana-Seguí, P., Turco, M., Herrera, S., & Miguez-Macho, G. (2017). Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-Interim. *Hydrology and Earth System Sciences*, *21*(4), 2187–2201.

Reaver, N. G. F., Kaplan, D. A., Klammler, H., & Jawitz, J. W. (2020). Technical Note: Analytical Inversion of the Parametric Budyko Equations. *Hydrology and Earth System Sciences Discussions*, (pp. 1–19).

Reaver, N. G. F., Kaplan, D. A., Klammler, H., & Jawitz, J. W. (2022). Theoretical and empirical evidence against the Budyko catchment trajectory conjecture. *Hydrology and Earth System Sciences*, *26*(5), 1507–1525.

Ribes, A., Thao, S., Vautard, R., Dubuisson, B., Somot, S., Colin, J., Planton, S., & Soubeyroux, J.-M. (2019). Observed increase in extreme daily rainfall in the French Mediterranean. *Climate Dynamics*, *52*(1), 1095–1114.

Riedel, T., & Weber, T. K. D. (2020). Review: The influence of global change on Europe's water cycle and groundwater recharge. *Hydrogeology Journal*, *28*(6), 1939–1959.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), 651–659.

Roderick, M. L., & Farquhar, G. D. (2011). A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. *Water Resources Research*, *47*(12).

Roderick, M. L., Sun, F., Lim, W. H., & Farquhar, G. D. (2014). A general framework for understanding the response of the water cycle to global warming over land and ocean. *Hydrology and Earth System Sciences*, *18*(5), 1575–1589.

Rottler, E., Francke, T., Bürger, G., & Bronstert, A. (2020). Long-term changes in central European river discharge for 1869–2016: Impact of changing snow covers, reservoir constructions and an intensified hydrological cycle. *Hydrology and Earth System Sciences*, *24*(4), 1721–1740.

Rudolf, B., Beck, C., Grieser, J., Schneider, U., & Deutscher Wetterdienst, D. (2005). The Global Precipitation Climatology Centre (GPCC).

Schaner, N., Voisin, N., Nijssen, B., & Lettenmaier, D. P. (2012). The contribution of glacier melt to streamflow. *Environmental Research Letters*, *7*(3), 034029.

Schneider, C., Laizé, C. L. R., Acreman, M. C., & Flörke, M. (2013). How will climate change modify river flow regimes in Europe? *Hydrology and Earth System Sciences*, *17*(1), 325–339.

Schrapffer, A., Sörensson, A., Polcher, J., & Fita, L. (2020). Benefits of representing floodplains in a Land Surface Model: Pantanal simulated with ORCHIDEE CMIP6 version. *Climate Dynamics*, *55*(5), 1303–1323.

Shahi, N. K., Polcher, J., Bastin, S., Pennel, R., & Fita, L. (2021). Assessment of the spatial variability of the added value on precipitation of convection-permitting simulation over the Iberian Peninsula using the RegIPSL regional earth system model. (pp. EGU21–9062).

Sherwood, S., & Fu, Q. (2014). A Drier Future? *Science*, *343*(6172), 737–739.

Siebert, S., Kummu, M., Porkka, M., Doell, P., Ramankutty, N., & Scanlon, B. (2015). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, *19*, 1521–1545.

Simons, G. W. H., Bastiaanssen, W. G. M., Cheema, M. J. M., Ahmad, B., & Immerzeel, W. W. (2020). A novel method to quantify consumed fractions and non-consumptive use of irrigation water: Application to the Indus Basin Irrigation System of Pakistan. *Agricultural Water Management*, *236*, 106174.

Slater, L. J., Singer, M. B., & Kirchner, J. W. (2015). Hydrologic versus geomorphic drivers of trends in flood hazard. *Geophysical Research Letters*, *42*(2), 370–376.

Sordo-Ward, A., Granados, I., Iglesias, A., & Garrote, L. (2019). Blue Water in Europe: Estimates of Current and Future Availability and Analysis of Uncertainty. *Water*, *11*(3), 420.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. a. J., Sauquet, E., Demuth, S., Fendekova, M., & Jódar, J. (2010). Streamflow trends in Europe: Evidence from a dataset of near-natural catchments. *Hydrology and Earth System Sciences*, *14*(12), 2367–2382.

Stephens, G., Polcher, J., Zeng, X., van Oevelen, P., Poveda, G., Bosilovich, M., Ahn, M.-H., Balsamo, G., Duan, Q., Hegerl, G., Jakob, C., Lamptey, B., Leung, R., Piles, M., Su, Z., Dirmeyer, P., Findell, K. L., Verhoef, A., Ek, M., L'Ecuyer, T., Roca, R., Nazemi, A., Dominguez, F., Klocke, D., & Bony, S. (2023). The First 30 Years of GEWEX. *Bulletin of the American Meteorological Society*, *104*(1), E126–E157.

Tafasca, S., Ducharne, A., & Valentin, C. (2020). Weak sensitivity of the terrestrial water budget to global soil texture maps in the ORCHIDEE land surface model. *Hydrology and Earth System Sciences*, *24*(7), 3753–3774.

Teng, J., Chiew, F. H. S., Vaze, J., Marvanek, S., & Kirono, D. G. C. (2012). Estimation of Climate Change Impact on Mean Annual Runoff across Continental Australia Using Budyko and Fu Equations and Hydrological Models. *Journal of Hydrometeorology*, *13*(3), 1094–1106.

Teutschbein, C., & Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, *456–457*, 12–29.

Tian, L., Jin, J., Wu, P., & Niu, G.-y. (2018). Quantifying the Impact of Climate Change and Human Activities on Streamflow in a Semi-Arid Watershed with the Budyko Equation Incorporating Dynamic Vegetation Information. *Water*, *10*(12), 1781.

Trapero, L., Bech, J., Rigo, T., Pineda, N., & Forcadell, D. (2009). Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network. *Atmospheric Research*, *93*(1), 408–418.

Troch, P. A., Lahmers, T., Meira, A., Mukherjee, R., Pedersen, J. W., Roy, T., & Valdés-Pineda, R. (2015). Catchment coevolution: A useful framework for improving predictions of hydrological change? *Water Resources Research*, *51*(7), 4903–4922.

Tuel, A., Schaefli, B., Zscheischler, J., & Martius, O. (2022). On the links between subseasonal clustering of extreme precipitation and high discharge in Switzerland and Europe. *Hydrology and Earth System Sciences*, *26*(10), 2649–2669.

Vicente-Serrano, S. M., Lopez-Moreno, J.-I., Beguería, S., Lorenzo-Lacruz, J., Sanchez-Lorenzo, A., García-Ruiz, J. M., Azorin-Molina, C., Morán-Tejeda, E., Revuelto, J., Trigo, R., Coelho, F., & Espejo, F. (2014). Evidence of increasing drought severity caused by temperature rise in southern Europe. *Environmental Research Letters*, *9*(4), 044001.

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., López-Moreno, J. I., Beguería, S., Noguera, I., Harrigan, S., & Vidal, J.-P. (2019). Climate, Irrigation, and Land Cover Change Explain Streamflow Trends in Countries Bordering the Northeast Atlantic. *Geophysical Research Letters*, *46*(19), 10821–10833.

Vincent, C., Fischer, A., Mayer, C., Bauder, A., Galos, S. P., Funk, M., Thibert, E., Six, D., Braun, L., & Huss, M. (2017). Common climatic signal from glaciers in the European Alps over the last 50 years. *Geophysical Research Letters*, *44*(3), 1376–1383.

Voisin, N., Hejazi, M. I., Leung, L. R., Liu, L., Huang, M., Li, H.-Y., & Tesfa, T. (2017). Effects of spatially distributed sectoral water management on the redistribution of water resources in an integrated water model. *Water Resources Research*, *53*(5), 4253–4270.

Wada, Y., van Beek, L. P. H., & Bierkens, M. F. P. (2012). Nonsustainable groundwater sustaining irrigation: A global assessment: NONSUSTAINABLE GROUNDWATER SUSTAINING IRRIGATION. *Water Resources Research*, *48*(6).

Wang, F., & Polcher, J. (2019). Assessing the freshwater flux from the continents to the Mediterranean Sea. *Scientific Reports*, *9*(1), 8024.

Wang, F., Polcher, J., Peylin, P., & Bastrikov, V. (2018). Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles. *Hydrology and Earth System Sciences*, *22*(7), 3863–3882.

Wang, W., Zhang, Y., & Tang, Q. (2020). Impact assessment of climate change and human activities on streamflow signatures in the Yellow River Basin using the Budyko hypothesis and derived differential equation. *Journal of Hydrology*, *591*, 125460.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, *50*(9), 7505–7514.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., & Best, M. (2011). Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *Journal of Hydrometeorology*, *12*(5), 823–848.

Westrick, K. J., Mass, C. F., & Colle, B. A. (1999). The Limitations of the WSR-88D Radar Network for Quantitative Precipitation Measurement over the Coastal Western United States. *Bulletin of the American Meteorological Society*, *80*(11), 2289–2298.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan,

M., Sheffield, J., Wade, A., & Whitehead, P. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research*, *47*(5).

Xing, W., Wang, W., Shao, Q., & Yong, B. (2018). Identification of dominant interactions between climatic seasonality, catchment characteristics and agricultural activities on Budyko-type equation parameter estimation. *Journal of Hydrology*, *556*, 585–599.

Xiong, M., Huang, C.-S., & Yang, T. (2020). Assessing the Impacts of Climate Change and Land Use/Cover Change on Runoff Based on Improved Budyko Framework Models Considering Arbitrary Partition of the Impacts. *Water*, *12*(6), 1612.

Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., & Lei, Z. (2007). Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. *Water Resources Research*, *43*(4).

Yang, H., Yang, D., Lei, Z., & Sun, F. (2008). New analytical derivation of the mean annual water-energy balance equation. *Water Resources Research*, *44*(3).

Yang, Y., Zhang, S., McVicar, T. R., Beck, H. E., Zhang, Y., & Liu, B. (2018). Disconnection Between Trends of Atmospheric Drying and Continental Runoff. *Water Resources Research*, *54*(7), 4700–4713.

Zanardo, S., Harman, C. J., Troch, P. A., Rao, P. S. C., & Sivapalan, M. (2012). Intra-annual rainfall variability control on interannual variability of catchment water balance: A stochastic analysis. *Water Resources Research*, *48*(6).

Zeng, L., Xiong, L., Liu, D., Chen, J., & Kim, J.-S. (2019). Improving Parameter Transferability of GR4J Model under Changing Environments Considering Nonstationarity. *Water*, *11*(10), 2029.

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., & Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, *40*(2).

Zhang, L., Potter, N., Hickel, K., Zhang, Y., & Shao, Q. (2008). Water balance modeling over variable time scales based on the Budyko framework – Model development and testing. *Journal of Hydrology*, *360*(1), 117–131.

Zhang, X., Dong, Q., Costa, V., & Wang, X. (2019). A hierarchical Bayesian model for decomposing the impacts of human activities and climate change on water resources in China. *Science of The Total Environment*, *665*, 836–847.

Zhang, Y., Zheng, H., Zhang, X., Leung, L. R., Liu, C., Zheng, C., Guo, Y., Chiew, F. H. S., Post, D., Kong, D., Beck, H. E., Li, C., & Blöschl, G. (2023). Future global streamflow declines are probably more severe than previously estimated. *Nature Water*, *1*(3), 261–271.

Zhao, J., Huang, S., Huang, Q., Wang, H., & Leng, G. (2018). Detecting the Dominant Cause of Streamflow Decline in the Loess Plateau of China Based onthe Latest Budyko Equation. *Water*, *10*(9), 1277.

Zhao, W., & Li, A. (2015). A Review on Land Surface Processes Modelling over Complex Terrain. *Advances in Meteorology*, *2015*, e607181.

Zheng, Y., Huang, Y., Zhou, S., Wang, K., & Wang, G. (2018). Effect partition of climate and catchment changes on runoff variation at the headwater region of the Yellow River based on the Budyko complementary relationship. *Science of The Total Environment*, *643*, 1166–1177.

Zhou, S., Yu, B., Huang, Y., & Wang, G. (2015). The complementary relationship and generation of the Budyko functions. *Geophysical Research Letters*, *42*(6), 1781–1790.

Zhou, X., Polcher, J., & Dumas, P. (2021). Representing Human Water Management in a Land Surface Model Using a Supply/Demand Approach. *Water Resources Research*, *57*(4), e2020WR028133.

Zorita, E., & von Storch, H. (1999). The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. *Journal of Climate*, *12*(8), 2474–2489.

Zveryaev, I. I. (2004). Seasonality in precipitation variability over Europe. *Journal of Geophysical Research: Atmospheres*, *109*(D5).

ECOLE DOCTORALE

Titre: Le cycle de l'eau continental: moteurs climatiques et non-climatiques des débits de rivières et évolution de la ressources en eau

Mots clés: Cycle de l'eau, Climat, Débits, Modélisation, Modèle de Budyko, Haute résolution

Résumé: Prévoir l'évolution des ressources en eau est un défi majeur dans un contexte de changement climatique et de rivières hautement anthropisées. Nous proposons une méthode innovante pour détecter et quantifier les changements dans le débit des rivières, climatiques et non climatiques. Un modèle de surface (LSM) est utilisé pour estimer la réponse "naturelle" de la surface continentale aux fluctuations climatiques. Le cadre conceptuel de Budyko est ensuite utilisé, pour décomposer l'évolution du débit en une réponse directe aux fluctuations climatiques, et une réponse indirecte, due aux changements de l'efficacité évaporative du bassin versant. Comparer l'application de ce cadre aux sorties du LSM et à des débits observés permet de mettre en évidence les zones où la réponse "naturelle" des bassins versants à la variabilité climatique est insuffisante pour expliquer les changements enregistrés. Les résultats obtenus en Europe montrent que la part de l'évolution des débits due au climat est dominée par la tendance sur les précipitations moyennes (P), avec en facteurs secondaires l'évapotranspiration potentielle (PET) dans la majeure partie de l'Europe et la répartition intra-annuelle de P en Méditerranée. Cependant, l'évolution générale des débits est dominée à l'échelle du siècle par des facteurs non pris en compte dans le système "naturel". Notre méthode permet d'identifier et de quantifier l'effet général de ces facteurs et de les corréler à certains vecteurs potentiels comme l'installation de barrages mais seul les futurs développements des LSM pour mieux intégrer les facteurs anthropiques permettrons d'attribuer les tendances non climatiques détectées. Or, la plupart des activités humaines qui influent sur le cycle de l'eau prennent place à petite échelle, celle des réservoirs ou des périmètres d'irrigation, et les forçages atmosphériques limitent la résolution d'exécution des LSM. La première étape consiste donc à construire un forçage atmosphérique à plus haute résolution. Pour aborder ce défi, nous combinons un jeu de données issu d'observations avec les résultats de modèles atmosphériques à l'échelle kilométrique. Ces derniers permettent de désagréger les observations selon des champs atmosphériques cohérent spatialement et en altitude.

Title: Continental water cycle: climatic and non-climatic drivers of river discharge and evolution of water resources

Keywords: Weter cycle, Climate, Discharge, Modeling, Budyko framework, High resolution

Abstract: Predict and manage the evolution of water resources is a key challenge in a context of climate change and highly managed rivers. We propose an innovative method to detect and quantify the changes in river discharge due to climate processes or to non-climatic factors. A land surface model (LSM) is used to estimate the "natural" response of the continental surface to climate fluctuations. Then the Budyko framework is used to decompose the streamflow response into a direct response to climate fluctuations and an indirect response to changes in evaporation efficiency of the watershed. Comparing the application of the framework to the LSM outputs and to observations allows to highlight the areas where the "natural" response of watersheds to climate variability is insufficient to explain the recorded changes in river discharge. Results over Europe show that, over the past century, changes in discharge due to climate processes are dominated by trends in annual mean precipitation (P). Secondary climatic factors are potential evapotranspiration (PET) over most of Europe and the intra-annual distribution of P for the Mediterranean area. However, the changes due to factors not accounted for in the "natural" system dominate over the century. Our method allows to quantify the overall effect of these non-climatic factors and correlate them to changes in potential specific drivers such as dams water storage but this is not trully an attribution. Future developments in LSMs will allow to better include human drivers of the hydrological cycle. Then they will be able to decompose and attribute the non-climatic changes detected. Yet, most human activities impacting the water cycle are at the regional scale. Since LSMs are limited to the resolution of the atmospheric forcing used, the first step is to construct a higher resolution atmospheric forcing, to later on test the performance of LSMs at the scales at which human activities modify the hydrological cycle. With that in mind, we use kilometric-scale outputs of atmospheric models to disaggregate an observation-based dataset. We reconstitute spatially and altitudinaly coherent atmospheric fields, with daily averages matching observations.

Institut Polytechnique de Paris
91120 Palaiseau, France