



# Agriculture's impact on water-energy balance varies across climates

Masoud Zaerpour<sup>a,1</sup> Shadi Hatami<sup>a</sup>, André S. Ballarin<sup>b</sup>, Simon Michael Papalexiou<sup>a,c</sup> Alain Pietroniro<sup>a</sup> and Ali Nazemi<sup>d</sup>

Affiliations are included on p. 10.

Edited by Andrea Rinaldo, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland; received May 28, 2024; accepted February 5, 2025

Agriculture is a cornerstone of global food production, accounting for a substantial portion of water withdrawals worldwide. As the world's population grows, so does the demand for water in agriculture, leading to alterations in regional water-energy balances. We present an approach to identify the influence of agriculture on the water-energy balance using empirical data. We explore the departure from the Budyko curve for catchments with agricultural expansion and their associations with changes in the water-energy balance using a causal discovery algorithm. Analyzing data from 1,342 catchments across three Köppen-Geiger climate classes—temperate, snowy, and others—from 1980 to 2014, we show that temperate and snowy catchments, which account for over 90% of stations, exhibit distinct patterns. Cropland percentage (CL%) emerges as the dominant factor, explaining 47 and 37% of the variance in deviations from the Budyko curve in temperate and snowy catchments, respectively. In temperate catchments, CL% shows a strong negative correlation with precipitation-streamflow (P-Q) causal strength (Spearman  $\rho = -0.75$ ), suggesting that cropland exacerbates precipitation-driven deviations. A moderate negative correlation with aridity-streamflow (AR-Q) causal strength ( $\rho = -0.42$ ) indicates additional influences of cropland through aridity-driven interactions. In snowy catchments, CL% is similarly influential, with a positive correlation with P-Q causal strength ( $\rho = 0.51$ ). However, the negative correlation with AR-Q causal strength ( $\rho = -0.45$ ) underscores the role of aridity as a secondary driver. While vegetation and precipitation seasonality also contribute to the deviations, their impacts are comparatively lower. These findings underscore the need for inclusion of agricultural activities in changing water-energy balance to secure future water supplies.

agriculture | water balance | irrigation | Budyko water balance

**1.1. Context.** Agriculture is crucial for meeting the world's growing food demands and accounts for approximately 72% of global water withdrawals (1–4). Global estimates of water utilized for crops span from 2,217 to 3,185 km<sup>3</sup> per year (5–8), while additional crop evapotranspiration encompasses a range of 927 to 1,530 km<sup>3</sup> per year (6, 8, 9). As the global population continues to increase, the need for water in agriculture has surged, resulting in often unknown consequences for the hydrological cycle. To sustain increasing food demand resulting from population growth and elevated living standards, global water withdrawals have surged nearly sixfold, escalating from approximately 500 km<sup>3</sup> per year in 1,900 to nearly 3,000 km<sup>3</sup> per year in 2000. During this period, crop yields have also substantially increased, reflecting significant advances in agricultural productivity (10); however, improved water use efficiency does not necessarily reduce total water consumption (11). Consequently, agriculture remains a predominant water user globally (8, 12, 13). While agricultural expansion sustains food security, at the same time, it has triggered severe water scarcity issues at regional to global levels (14, 15). For instance, the agriculture region in and around the Ogallala Aquifer, located in the Great Plains region of the United States, witnessed a dramatic increase in agricultural area, surging from 8,500 km<sup>2</sup> in 1949 to 63,000 km<sup>2</sup> by 2005 (16). Sustaining such a level of agricultural productivity, however, hinged on the ongoing extraction of water from the Ogallala Aquifer (17, 18).

Crop water needs can be fulfilled through three distinct sources: (1) green water, derived from local precipitation and temporarily stored in the soil, (2) blue water, encompassing surface water from local/nonlocal rivers, lakes, reservoirs, and renewable groundwater, and (3) nonrenewable groundwater, extracted from aquifers (19, 20). The term "nonrenewable groundwater abstraction" is used to underscore water extracted beyond recharge levels for a prolonged period, i.e., physically unsustainable groundwater use (21) that is unlikely to be replenished in years (22, 23).

## Significance

Agriculture plays a key role in global food security, intricately tied to water resources for crop growth. However, navigating the interplay between agriculture and water availability poses challenges, especially during the *Anthropocene*, where traditional perspectives often overlook agriculture's impacts on the water cycle. Understanding and integrating agriculture's influence on water dynamics becomes imperative in addressing contemporary challenges. Our study highlights the contrasting impacts of agricultural activities across temperate and snowy climates. In temperate catchments, agriculture weakens the precipitation-streamflow (P-Q) relationship, contributing to precipitation-driven deviations from the water-energy balance, while in snowy catchments, agricultural activities strengthen the P-Q relationship. These findings offer insights for shaping effective water management strategies, ensuring food security, and promoting sustainable development globally.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

Although PNAS asks authors to adhere to United Nations naming conventions for maps (<https://www.un.org/geospatial/mapsgeo>), our policy is to publish maps as provided by the authors.

<sup>1</sup>To whom correspondence may be addressed. Email: masoud.zaerpour@ucalgary.ca.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2410521122/-DCSupplemental>.

Published March 17, 2025.

Surface water is the primary source of freshwater for human use, with around 70% of the available surface resource used in agriculture (24). However, blue water alone does not fully represent the water needed for crop production (25–27), as rainfed crops primarily rely on rainfall infiltration, which constitutes a significant portion of water for agricultural croplands (28, 29). Additionally, actual agricultural water consumption is better measured as the difference between crop water use and what would be required by the natural vegetation that was replaced (30). As surface water sources become less reliable and predictable (31, 32), the importance of groundwater for agriculture is on the rise. In regions with limited access to surface water, groundwater can emerge as the primary source of agricultural water (33). However, unsustainable groundwater depletion can easily arise when extraction surpasses recharge over prolonged periods, resulting in declining groundwater levels and consequently reduced baseflow (20, 23, 33–39). Groundwater recovery can be gradual, with residence times spanning from several months in shallow aquifers to millions of years in deep aquifers. Hence, it is crucial to recognize unsustainable agricultural practices and their far-reaching consequences in altering the hydrological cycle.

A comprehensive understanding of the water cycle enhances our evaluation of freshwater resources and enables us to understand global water security threats more effectively. Also, on a global scale, climate change has been identified as a primary driver of hydrological cycle changes (40–42). Nonetheless, the hydrological cycle is not solely shaped by climate dynamics (43); it is also influenced by human activities like agriculture and groundwater extraction at the regional scale (44–46). Understanding the agricultural practices by source and their interplay with other drivers of changes in streamflow is imperative for advancing hydrological science at regional and global scales. For instance, surface water plays a pivotal role in flood dynamics (47–49), whereas groundwater is closely tied to baseflow changes and bears immense significance for environmental flow management and water supply in arid and semiarid regions (50–52). Two notable examples of groundwater use for agriculture in the United States are the Ogallala aquifer in High Plains and the aquifer system in California's Central Valley. These aquifers are extensively studied, with depletion estimates drawn from numerous well water level measurements and GRACE satellite data (36, 53, 54).

**1.2. Combining the Budyko Curve with Causal Analysis Is a Way to Provide Insight Into How Agriculture Affects the Water Cycle.** Despite the important role of agriculture and other anthropogenic water uses in shaping the hydrological cycle, assessing their global impact remains challenging due to uncertainties associated with factors such as agricultural area delineation, hydrological modeling, and critical local parameters such as soil hydraulic characteristics, the timing of the growing season, and its relation to water availability (3, 55).

By applying the Budyko water balance framework (56) and a causal discovery algorithm (57, 58), we explore the effect of agricultural water use on the hydrological cycle. The interplay of components of terrestrial freshwater ecosystems is uncertain in an era of human modifications and rapid global climate change. Although previous methods employed modeled streamflow data to identify causal connections between human-induced forcing and observed river flow patterns globally, uncertainties remain regarding the complex interactions between human activity, climate variables, and catchment-scale hydrological processes that impact water flow dynamics. Building on previous findings, we have used a large, measured dataset to better understand watershed responses to human changes in land surface caused by agricultural expansion.

The Budyko hypothesis offers a valuable tool for comparing normalized observations across a wide spectrum of climatic conditions, enabling the identification of secondary controls on a catchment's water balance (45, 59, 60). According to the Budyko curve, the ratio of mean annual evapotranspiration to mean annual precipitation (E/P, the evaporative fraction) is primarily influenced by the ratio of mean annual potential evapotranspiration to mean annual precipitation (Ep/P, the dryness index). Hereafter, the term water–energy balance refers to the streamflow–precipitation–aridity equilibrium described by the Budyko curve. Using this framework, we study the impact of agriculture by source on the water–energy balance at the country scale. Additionally, it is important to consider the role of precipitation seasonality and vegetation in influencing deviations from the Budyko curve. Studies have shown that seasonality of precipitation and potential evapotranspiration can substantially alter the timing and efficiency of water inputs to ecosystems, affecting evapotranspiration and runoff patterns (61, 62). Variations in vegetation type and density can further complicate these interactions, as they influence the partitioning of water between evapotranspiration and runoff (63, 64). Understanding these relationships is crucial for accurately interpreting the Budyko curve and the hydrological responses of different catchments.

It is essential to acknowledge that the Budyko curve, while highly informative, does not represent the complexity of natural hydrological processes. However, the Budyko curve follows a well-known characteristic curve when estimated for a series of natural basins. The spatial differences in the water balance between catchments can result from various factors beyond just agricultural water use. In addition, it is essential to acknowledge that a direct symmetry might not always exist in the temporal partitioning of precipitation into streamflow and evaporation within a catchment, in contrast to the observed spatial variations among catchments (65–67).

Because of these limitations, we also analyzed the interplay between main components of water–energy balance and their relationship with the cropland percentage (CL%; as an agriculture proxy), thus shedding light on the effect of agriculture on altering water–energy balance. Many Budyko studies often focus solely on spatial patterns of long-term averages (45, 66, 68, 69). Nevertheless, this methodology frequently omits a substantial amount of hydrological information and the interplay of various factors that could influence water balance at finer temporal scales (68). To address this challenge and gain insight into the intricate interplay of factors at finer time scales, we analyzed the control of agriculture on temporal changes of individual responses at the seasonal time scale (67, 69) using the PCMCI+ causal discovery method recently developed and applied in Earth system sciences (57, 58). By integrating the Budyko framework with the PCMCI+ causal discovery algorithm, we offer an approach to quantifying the equilibrium between water and energy balance in catchments. Our examination focuses on three key factors influencing streamflow (Q): precipitation (P), aridity (AR), and snow fraction (SF) (67, 70, 71). For calculating the SF, we assumed that precipitation falls as snow when the temperature is below 1 °C (see the *Materials and Methods* for further information on calculating the SF). We establish causal relationships between these drivers and streamflow and measure their strengths using the causal discovery method (58).

Here, we study the role of agriculture by source in altering the water–energy balance using 1,342 Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) catchments located across the contiguous United States and Great Britain (GB) from 1980 to 2014. The role of agriculture is studied by a spatial (between-catchment) comparison of the long-term partitioning of precipitation into evapotranspiration and streamflow

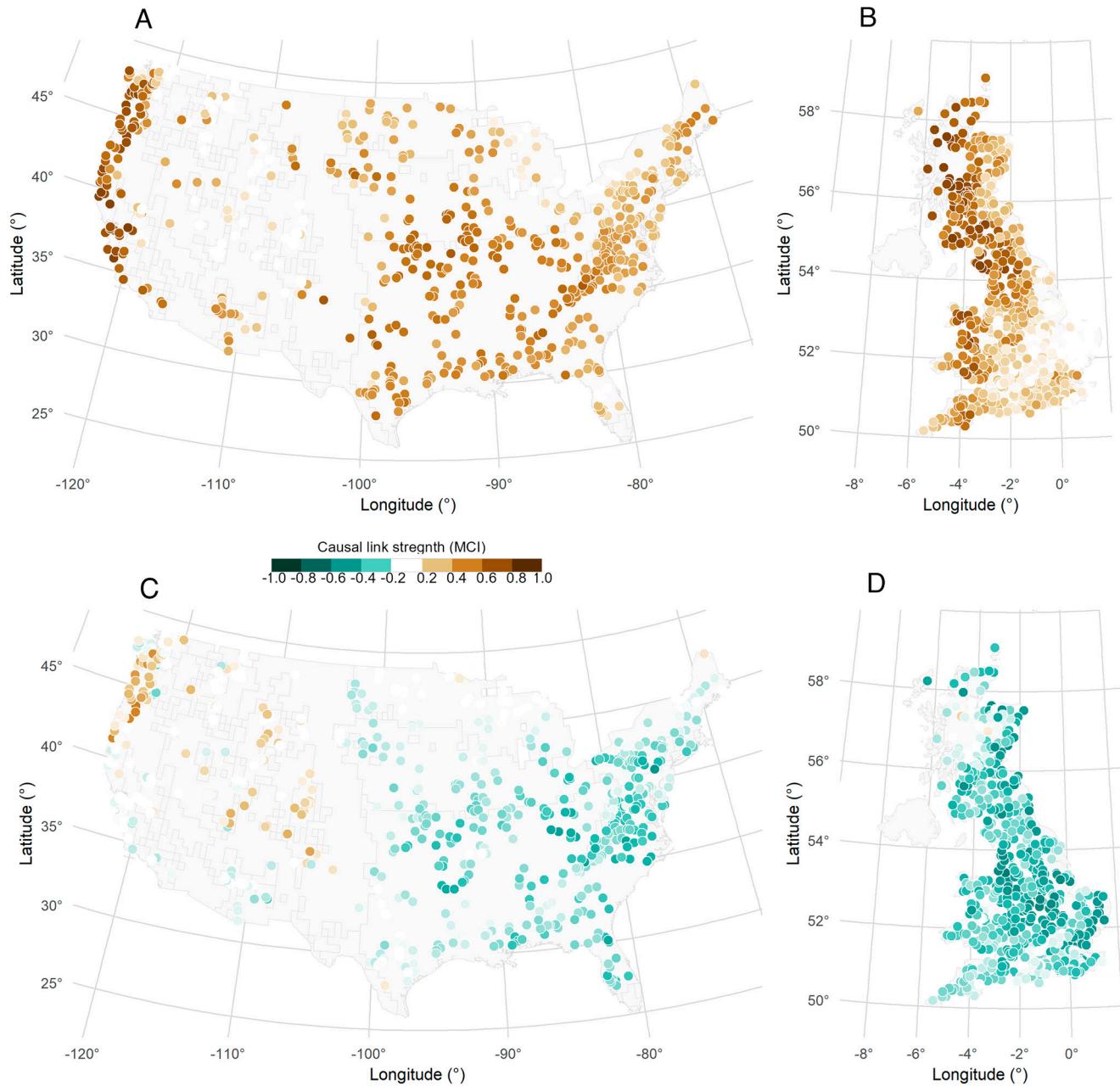
due to the lack of CL% time series. These long-term observations are then conceptualized within the Budyko framework (56).

## 2. Results

We first evaluate the causal strength between main components of water–energy balance including streamflow, precipitation, and AR, as depicted in Fig. 1. This analysis provides insights into the water–energy dynamics from the causal perspective. Causal link strength is quantified based on the momentary conditional independence (MCI) test statistic value (57, 58, 72, 73). The MCI test statistics yield a well-interpretable notion of a normalized causal strength (MCI test statistics value) that allows to measure the strength of causal links between variables (74, 75). An MCI value of +1 denotes the strongest positive link, while -1 represents

the strongest link with a negative impact. We acknowledge that the hydrological components are inherently interconnected through the principle of mass conservation, and deviations in one component are necessarily reflected in another. However, our analysis focuses on the temporal interactions and perturbations that drive catchments toward particular long-term balances, including deviations from the Budyko curve. These deviations are not a violation of mass balance but rather an indication of how external factors influence the partitioning and distribution of water–energy components.

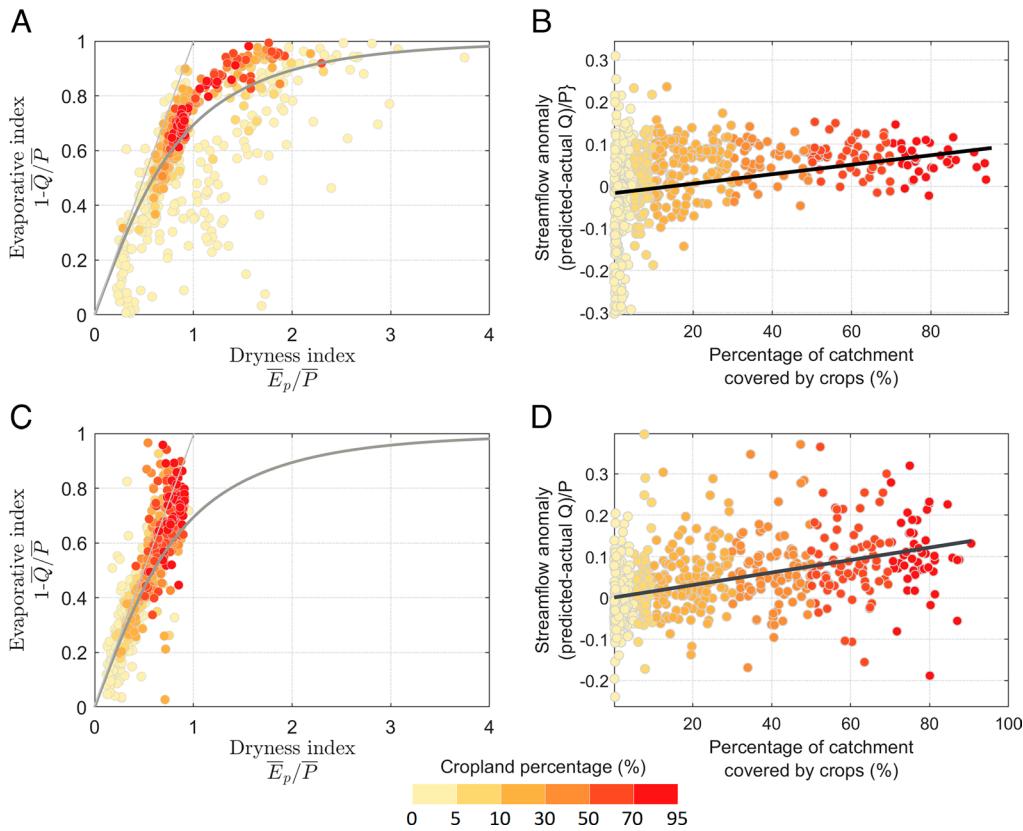
Panels (A) and (C) of Fig. 1 present the results for precipitation–streamflow (P–Q) and aridity–streamflow (AR–Q) links in the United States, respectively, while Panels (B) and (D) showcase corresponding results for GB. In the United States, precipitation emerges as the predominant factor influencing streamflow, demonstrating an



**Fig. 1.** Quantifying the causal link strength between main components of water–energy balance including streamflow, precipitation, and AR. The strength of these causal links is assessed through the MCI test statistic, where MCI values range from -1 to 1. Panels (A) and (B) illustrate the strength of causal links between streamflow–precipitation (P–Q) in the United States and Great Britain (GB), respectively. Panels (C) and (D) depict the strength of causal links between streamflow–aridity (AR–Q) in the United States and GB, respectively. White dots indicate catchments where the causal link strengths are not statistically significant.

expected MCI value of 0.43. AR also exerts control over streamflow, particularly in catchments located in Central and Coastal Plains; however, the expected MCI value of  $-0.12$  for all catchments in the United States is not statistically significant. In GB, similar to the US findings, precipitation predominantly control streamflow, with an expected MCI value of 0.41. AR also plays a significant role in streamflow, particularly evident in East England, with an expected MCI value stands at  $-0.31$  across all catchments. The results related to SF-Q can be found in *SI Appendix, Fig. S1*.

**2.1. The Percentage of Cropland May Explain Deviations from the Budyko Curve.** The long-term observations of the 671 CAMELS-US catchments in the context of the Budyko hypothesis are shown in Fig. 2A, stratified by the long-term percentage of cropland (CL%). Overall, the observed pattern in the United States aligns with the Budyko curve, with an average exceedance of the normalized mean streamflow ratio ( $Q/\bar{P}$ ) by 0.02. Additionally, the analysis reveals that higher CL% corresponds to a higher evaporative fraction ( $E/P$ ) and decreased runoff ratio ( $Q/P$ ). The CL% in the US catchments is significantly ( $P$ -value  $< 0.05$ ) correlated with the normalized streamflow anomaly, as shown in Fig. 2B (Spearman  $\rho = 0.56$ ), showing that the agricultural expansion is associated with higher streamflow anomalies. Hereafter, when we use the term “significant,” it specifically refers to  $P$ -value  $< 0.05$ . The reduction in variability of anomalies in Fig. 2B may be attributed to the moderating influence of agricultural practices in catchments with higher cropland coverage, where hydrological processes are more strongly governed by agricultural activities than by other natural mechanisms.

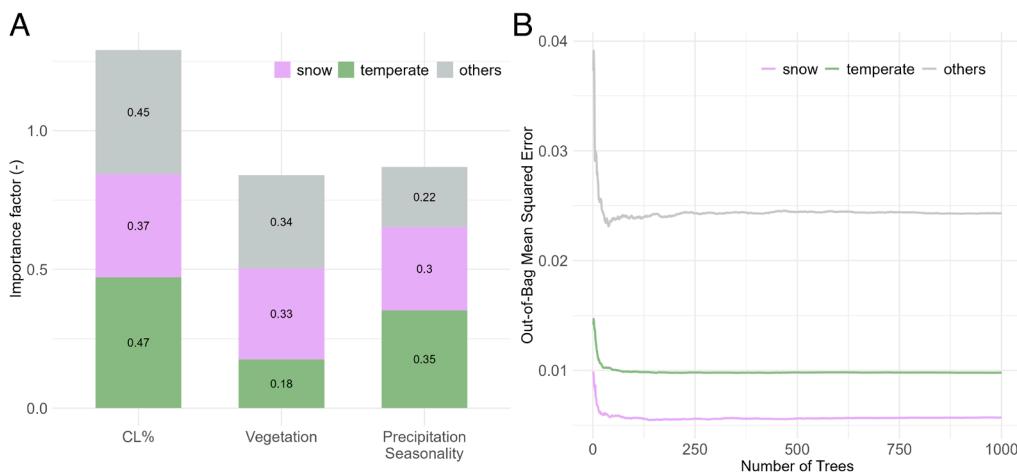


**Fig. 2.** The long-term observed streamflow and precipitation data are analyzed within the framework of the Budyko hypothesis (56). Panels (A) and (B) depict data for the United States, with (A) showing Budyko framework results and (B) illustrating the Spearman rank correlation between the percentage of cropland and the anomaly from the Budyko curve. Panels (C) and (D) present the corresponding results for Great Britain (GB), with (C) showing Budyko framework results and (D) illustrating the Spearman correlation. The colors indicate the percentage of cropland in each catchment. Spearman  $\rho$ s between the CL% and the normalized streamflow anomaly in the United States and GB are 0.56 and 0.51, respectively. The normalized streamflow anomaly is calculated as the difference between the simulated (i.e., Budyko-based) and actual streamflow divided by the long-term precipitation (simulated-actual  $\bar{Q}/\bar{P}$ ). These anomalies indicate how agricultural activities alter the natural partitioning of water between evaporation and runoff (See *Materials and Methods* for further information on the equations that were used).

Fig. 2C shows the results for the 671 CAMELS-GB catchments in GB. The pattern of observations deviates further from the Budyko curve compared to the US catchments with a mean exceedance of the normalized  $Q/\bar{P}$  of 0.04. Like the US catchments, the CL% is associated with normalized  $Q/\bar{P}$  anomaly (Spearman  $\rho = 0.51$ ), as shown in Fig. 2D. The analyses in both the United States and GB show that agricultural activities contribute to deviations from the Budyko curve.

However, spatial variations in the water balance can also be influenced by other factors, such as vegetation dynamics (61, 63, 76) and precipitation seasonality (62, 64). To investigate how these factors, alongside agricultural activities, contribute to deviations from the Budyko curve, we conducted a Random Forest (RF) analysis across hydrologically significant divisions, specifically three Köppen-Geiger climate classes: temperate, snow, and others—Please see *SI Appendix, Fig. S2*. In the analysis by climate class, temperate and snowy catchments—which together constitute over 90% of the stations (75.6% temperate and 15.2% snowy)—dominate the findings. These catchments are particularly relevant as they represent the majority of agricultural catchments in the studied regions. Conversely, the others have far fewer stations (9.2%), making their contribution to the overall patterns less significant. Fig. 3 presents the RF analysis, highlighting the contributions of cropland percentage (CL%), vegetation, and precipitation seasonality to deviations from the Budyko curve. The RF models were run with 1,000 trees, with the Out-of-Bag (OOB) error stabilizing after approximately 400 trees in all cases.

The analysis shows that CL% dominates vegetation and precipitation seasonality in its influence. In the temperate climate,



**Fig. 3.** RF model analysis of the relative importance of cropland percentage (CL%), vegetation, and precipitation seasonality in explaining deviations from the Budyko curve for catchments classified by the Köppen-Geiger climate classification. Panel (A) presents the factor importance for three climate classes. Panel (B) shows the OOB Error against the number of trees for each class. The analysis highlights that CL% consistently demonstrates the highest importance in temperate and snowy climates, surpassing the influence of vegetation and precipitation seasonality.

CL% demonstrates the highest importance (importance score = 0.47), notably exceeding the contributions of vegetation (importance score = 0.18) and precipitation seasonality (importance score = 0.35). In snow-dominated climates, CL% (importance score = 0.37) remains the most influential, although vegetation (importance score = 0.33) and precipitation seasonality (importance score = 0.3) show relatively higher contributions compared to temperate regions. In the others, which comprise less than 10% of the studied catchments, CL% (importance score = 0.45) again exhibits the greatest influence compared to vegetation (importance score = 0.34) and precipitation seasonality (importance score = 0.22).

Overall, CL% exhibits the highest importance (normalized importance score = 0.43), significantly exceeding the importance of vegetation (normalized importance score = 0.28) and precipitation seasonality (normalized importance score = 0.29). These results are consistent with the analysis presented in *SI Appendix*, Fig. S3 for US and GB catchments, where CL% emerged as the most dominant factor, reinforcing the significant role of agricultural activities in driving hydrological deviations across diverse climatic settings. While vegetation and precipitation seasonality are important, CL% stands out as the primary driver, particularly in temperate climate.

## 2.2. Causal Relationships Between Climatic Drivers and Streamflow Suggest Different Impacts of Agricultural Practices Across Diverse Climatic Settings.

To specifically investigate the influence of agriculture on the water–energy balance, hereafter we focus on the connection between CL% and the causal relationship between observed P–Q and AR–Q across climate classes as shown in Fig. 4 and the two countries in *SI Appendix*, Fig. S4. This dual focus provides insights into how agricultural intensity interacts with water–energy balance components to shape hydrological behavior over different climates and regions.

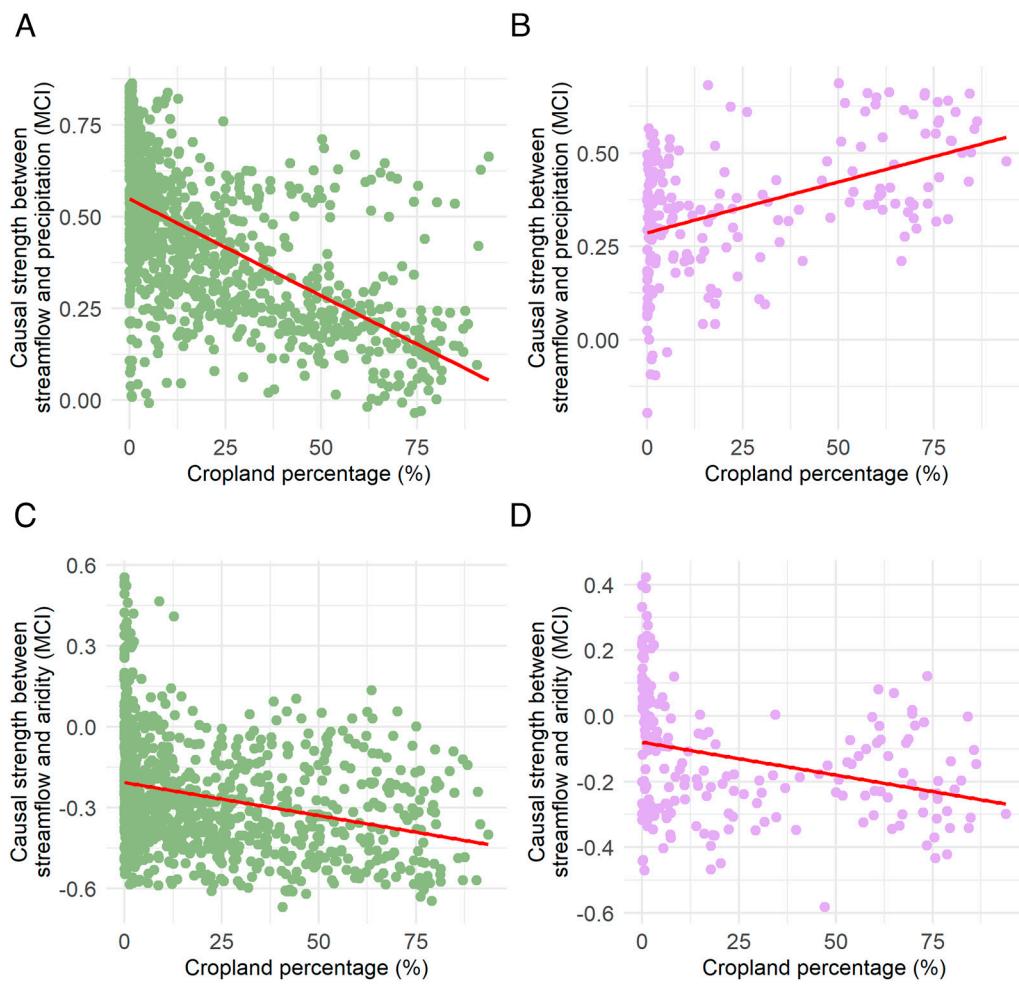
For temperate catchments, we found a strong negative correlation between CL% and P–Q causal strength (Spearman  $\rho = -0.75$ ), highlighting the role of cropland in exacerbating precipitation-driven deviations from the Budyko curve. In contrast, the relationship between CL% and AR–Q causality in temperate catchments was weaker but still significant ( $\rho = -0.42$ ), suggesting that cropland influences water–energy balance through complex interactions with precipitation and AR. However, snowy catchments exhibited a positive correlation between CL% and P–Q causality ( $\rho = 0.51$ ), indicating a stronger dependence on precipitation.

Snowy catchments showed a negative correlation between CL% and AR–Q causal strength ( $\rho = -0.45$ ), reinforcing the role of AR as a secondary driver. Fig. 4 further demonstrates that the relationships between P–Q and AR–Q are not independent but rather vary in their interactions across climate regions and external factors.

The analysis across countries in *SI Appendix*, Fig. S4 supports these class-specific findings while revealing regional nuances. In the United States where stations span a wide range of climates, the CL% was significantly negatively associated with AR–Q causal strength (Spearman  $\rho = -0.53$ ), consistent with the patterns observed in temperate catchments. However, the relationship between CL% and P–Q causality in the United States was weak ( $\rho = 0.12$ ), reflecting the variability in hydrological responses across diverse climates. In GB, a contrasting pattern emerged: CL% had a strong negative correlation with P–Q causality (Spearman  $\rho = -0.89$ ), suggesting an increased reliance on precipitation for streamflow in agricultural regions. The relationship between CL% and AR–Q causality in GB was weakly negative ( $\rho = -0.26$ ), aligning with its temperate climate, where AR plays a less dominant role in water–energy interactions.

## 2.3. Causal Relationships Suggest That Agriculture in the Temperate and Snow Catchments Relies on Different Water Sources.

The preceding analysis implies different impacts of agriculture on the interplay among streamflow, precipitation, and AR. In temperate catchments, causal analysis reveals a direct impact of agriculture on P–Q relationships, indicating that the P–Q causal relationship weakens as CL% increases in a catchment. That is, in crop catchments, P tends to exert a lower influence on Q. However, this is not observed in snowy catchments, where the impact of agriculture on the water balance differs significantly. To further investigate the sources of water for agriculture particularly in the snowy catchments, we partition streamflow into two components: direct flow (Qd; immediate runoff-fed component) and baseflow [Qb; groundwater-fed component (77, 78) utilizing a one-parameter low-pass filter (79–84) (See *Materials and Methods* for details)]. We examine the relationship between the rate of change in the annual baseflow and CL% to better understand the impact of CL% on the changes in the groundwater-fed component of streamflow (i.e., baseflow). We calculate the linear regression trend of normalized annual baseflows (TNB) over the study period and examine its relationship with the percentage of crop cover in the agricultural catchments.



**Fig. 4.** The scatter plots illustrate the correlation between changes in the causal link for P-Q, AR-Q, and cropland percentages (CL%) across temperate and snowy catchments. Panels (A) and (B) present the P-Q results for temperate and snowy catchments, respectively. Temperate catchments exhibit a strong negative correlation between CL% and P-Q causal strength (Spearman  $\rho = -0.75$ ), suggesting that cropland exacerbates precipitation-driven deviations from the Budyko curve. Conversely, snowy catchments show a positive strong dependence on precipitation in these regions ( $\rho = 0.51$ ). Panels (C) and (D) depict AR-Q results for temperate and snowy catchments, respectively. In temperate catchments, a moderate negative correlation ( $\rho = -0.42$ ) highlights the role of cropland in influencing streamflow through AR-driven evaporative demand. Similarly, snowy catchments show a negative correlation ( $\rho = -0.45$ ), underscoring AR's role as a secondary driver. These results reinforce the varying impacts of agricultural intensity on the water-energy balance components across climatic conditions.

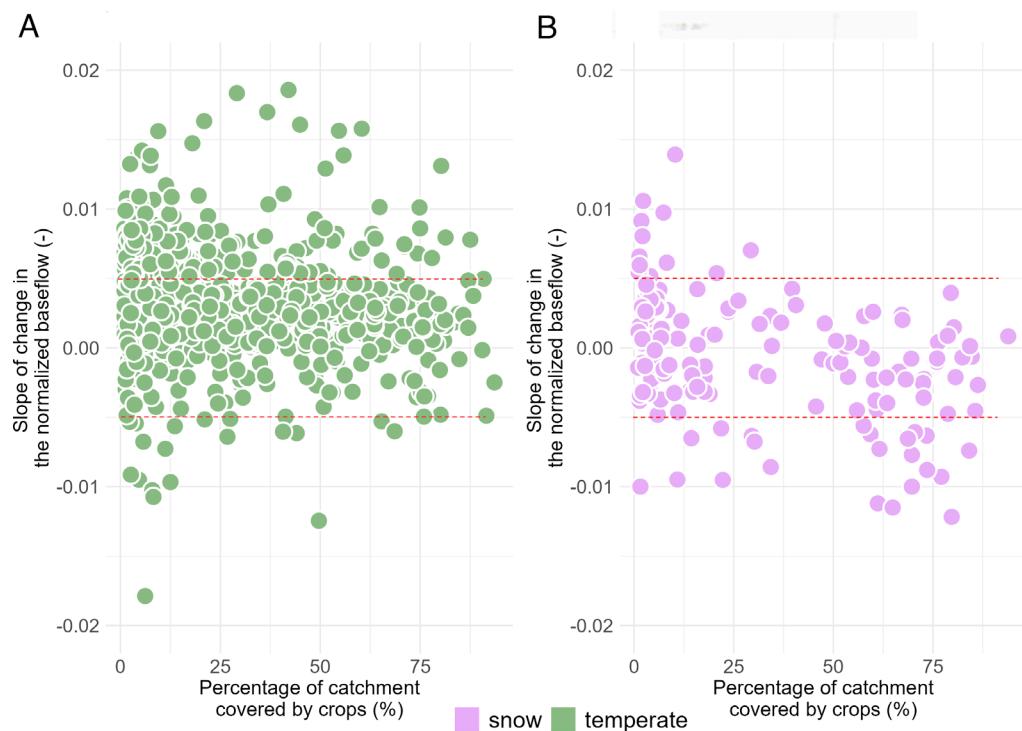
**2.4. Agricultural Catchments in a Snowy Climate in the United States Experience Decreases in Baseflow, Suggesting That Streams Are Increasingly Disconnected from Groundwater.** Fig. 5 summarizes findings of regression analysis of TNB, with the y-axis representing TNBs and the x-axis denoting CL%. Panels (A) and (B) display the results for the temperate and snowy catchments, respectively. The TNBs are determined through a simple linear regression trend, with dashed lines indicating the significance level (i.e., regression trend) at  $P$ -value = 0.05.

In the snowy catchments, a significant association is observed between the rate of baseflow change and crop cover, as evidenced by a Spearman rank correlation coefficient of  $-0.48$  (95% CIs spanning from  $-0.55$  to  $-0.41$ ). Remarkably, 38% of snowy catchments exhibit negative trends in baseflow, and among them, 10.3% display statistically significant changes. In contrast, the relationship between TNB and crop cover in temperate catchments appears comparatively weak, indicated by a Spearman correlation coefficient of  $-0.19$ . The 95% CI extends from  $-0.29$  to  $-0.1$ , underscoring the modest nature of this association. This difference can be attributed to the reliable availability of precipitation in temperate climates, making agriculture primarily dependent on precipitation rather than groundwater. Consequently, the correlation between baseflow changes and CL% is not as pronounced as it is in snowy

catchments. In temperate catchments, 13.6% of agricultural catchments demonstrate negative trends in baseflow, and among them, a mere 2.3% exhibit statistically significant changes unrelated to agricultural practices.

### 3. Discussion

The future of terrestrial freshwater ecosystems is uncertain in an era of human interventions and rapid global climate change (85–87). Although previous methods employed modeled streamflow data to identify causal connections between human-induced forcing and observed river flow patterns globally (88–90), uncertainties remain regarding the complex interactions between human activity, climate variables, and catchment-scale hydrological processes that impact water flow dynamics. These uncertainties increase the likelihood of unexpected outcomes as catchment modifications continue at the global scale, including urbanization (91, 92), increased water use (93), and alterations in the structure of land cover such as agricultural expansions (94, 95). Building on previous findings, we have used a large, measured dataset to better understand watershed responses to changes in land surface characteristics caused by agricultural expansion.



**Fig. 5.** Analysis of normalized baseflow changes in catchments across the temperature and snowy catchments, where cropland constitutes at least 1% of the total basin area, as depicted in Panels (A) and (B), respectively. Dashed lines represent the significance level (i.e., the threshold for regression trend) with a  $P$ -value of 0.05. The y-axis depicts the linear regression trend of normalized baseflow, while the x-axis represents the percentage of the catchment covered by crops. In the snowy catchments, the baseflow's regression trend strongly correlates with CL%, indicated by a Spearman correlation coefficient of  $-0.48$ . In temperature catchments, conversely, the linear regression trend in baseflow demonstrates a low correlation with CL% (Spearman  $\rho = -0.19$ ).

To advance our understanding of the water–energy balance and its association with human activities, it is imperative to identify driving forces. Earlier research primarily utilized the Budyko framework to evaluate the sensitivity of streamflow to certain variables. However, this framework, while effective in explaining spatial differences in streamflow across natural catchments, proves less adept at elucidating temporal changes within catchments, particularly concerning climate and human intervention. The intrinsic space–time asymmetry of the Budyko framework raises questions about its suitability for exploring water availability sensitivity at finer temporal scales (68, 69). To address this, we use empirical data and the PCMCI+ causal discovery method (57, 58) to shed light on how agriculture affects the long-term water–energy balance.

Our findings underscore that agricultural expansions induce alterations to the water–energy balance in catchments, contributing to deviations from the Budyko curve. Fig. 4 shows a significant association between agricultural expansion and the AR–Q causal relationship within these areas, indicating that more crop cover exacerbates the negative link between streamflow and AR in these areas. In catchments that rely primarily on groundwater for their water supply, this phenomenon is linked to increased crop cover, which increases evapotranspiration and, in turn, reduces surface runoff and streamflow (96–99) (Fig. 4). This finding is consistent with previous research that highlights the significant effects of groundwater-fed irrigation on evapotranspiration and alterations in the water–energy balance in the US High Plains (17, 33, 94, 100).

In particular, studies show that groundwater plays a critical role in sustaining agricultural activity in US locations with limited water availability. Groundwater-fed irrigation in the US High Plains increases evapotranspiration, which changes the water–energy balance (100). Other studies showed that groundwater levels are widely declined, especially in regions with extensive crop cover (101, 102). Our research supports these conclusions by

showing how groundwater-fed agriculture affects departures from the Budyko water–energy balance, reflecting the need of groundwater to meet the crop evapotranspiration (103, 104). This implies that groundwater gains or losses are more likely to have an impact on the theoretical relationship between dryness and the evaporative index (105, 106). Moreover, evapotranspiration and streamflow in the United States are significantly impacted by groundwater losses (94), with changes in evapotranspiration closely linked to crop development. This highlights the disruptive impact of irrigation expansion on the water–energy balance dynamics (107–109). While our spatial analysis offers valuable insights into the role of agriculture in shaping the water–energy balance, it is important to recognize that factors such as the length of the growing season, crop types, and vegetation replacement (110) can further influence temporal evapotranspiration dynamics, introducing complex interactions that may not be fully captured by our approach. Additionally, cropland greatly affects evapotranspiration and energy fluxes, further altering energy balance (106, 109). However, the P–Q relationship does not show a significant association with CL% in US agricultural catchments. This could be because groundwater is typically used instead of direct precipitation to meet agricultural water needs.

On the other hand, in GB, which is characterized by catchments with limited energy in which precipitation is the main source of water for agriculture, changes in the water–energy balance are primarily characterized by variations in the strength of causal relationships between streamflow and precipitation (P–Q). Results show a significant association between the agricultural expansion in temperate catchments and the P–Q relationship, which is not the case in the snowy catchments. As the percentage of CL% rises, the P–Q relationship decreases, suggesting that higher crop cover may result in a lower streamflow ratio as a consequence of changes in the P–Q dynamics. This phenomenon,

which is particularly apparent in areas where agriculture significantly depends on precipitation, can be explained by enhanced evapotranspiration and plant uptake of water (106, 107). The absence of correlation between CL% and the AR-Q relationship in temperate catchments can be ascribed to its relatively humid climate, where AR exerts less constraint on streamflow (111, 112). Studies conducted in GB suggest that elevated water usage by crops may result in a decreased distribution of precipitation, favoring evapotranspiration over runoff (111–113), which affects, in turn, the P-Q relationship. The reduction of the discrepancy between actual and potential evapotranspiration is largely dependent on the interplay between soil water storage and crop responses to water availability. This leads to a linear relationship between the evaporative ratio and AR (106, 107, 114) in these areas.

These distinct patterns underscore the nuanced interplay between agriculture and the water–energy balance, further emphasizing the impact of human activities as well as geographical and climatic variations on catchment dynamics. Understanding these intricacies is essential for effective water resource management, especially considering the potential implications of climate change and agricultural practices on streamflow patterns in energy-limited catchments like those in GB.

In conclusion, our study emphasizes the intricate relationship between agriculture by water source, the water–energy balance in a catchment, and deviations from the Budyko curve. A more profound understanding of these dynamics equips us to navigate water resource management adeptly, advocate for sustainable agricultural practices, and progress toward attaining global sustainable development goals. While our primary analysis centered on deviations from the Budyko curve due to alterations in the water–energy balance induced by agricultural practices, it is crucial to acknowledge additional influential factors shaping regional and global hydrological patterns. These factors encompass climatic variables such as climate variability and change (40, 115), human activities like drainage and groundwater extraction for domestic and industrial purposes, as well as catchment characteristics (37, 116, 117). For instance, in the Southeast region of GB, where extensive agricultural activities happen, rainfall-runoff events exhibit low runoff coefficient values. This is mainly due to the presence of chalk catchments characterized by high permeability (118). The occurrence of lateral subsurface flow in these catchments warrants consideration; however, its impact on the arguments presented in this paper is minimal,

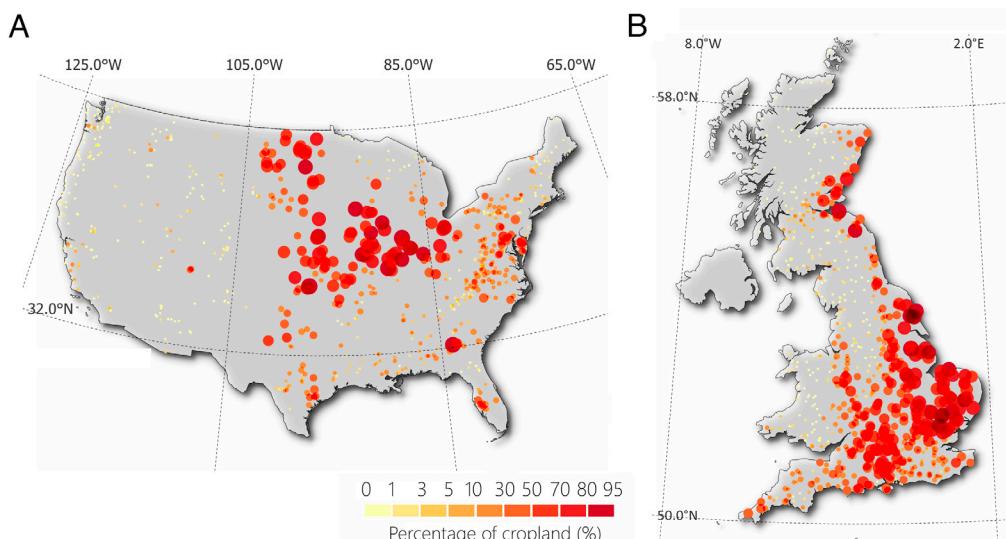
as the water involved typically travel the saturated zone swiftly and contributes to flow over the course of months (119, 120).

In attributing the change in the hydrological cycle to CL% as a proxy for agriculture, there is an opportunity for further research into the interplay of human and natural hydrological dynamics. For instance, a more comprehensive investigation of P-Q dynamic in GB can provide insights into the relationship between changes in the water–energy balance, CL%, and the implementation of field drainage in the 1970s and early 1980s (117). This, in turn, can unveil how agricultural practices aimed at reducing waterlogged conditions have impacted crop yield, transpiration, and water storage capacity within the upper soil profile, leading to changes in runoff changes (121–123). These multifaceted interactions warrant continued exploration and consideration in our ongoing efforts to address water-related challenges and secure a more sustainable future.

## 4. Materials and Methods

**4.1. Data.** The research focuses on analyzing watershed data from two comprehensive datasets, one from the United States and the other from GB, to understand hydrological variations across these regions. The US data are derived from the CAMELS dataset (124), which includes detailed information on 671 catchments across the contiguous United States. These catchments vary significantly in size, from 4 to 25,817 square kilometers, and provide a rich source of data including discharge, meteorological information, and various other attributes from 1980 to 2015. Key variables such as precipitation, temperature, potential evaporation, and streamflow are included, with streamflow data specifically sourced from the HCDN-2009 network (125). Additionally, the dataset incorporates information on cropland extent, calculated using a global cropland map (126), and groundwater-level data from the US Geological Survey (<https://waterdata.usgs.gov/nwis/inventory>).

In contrast, CAMELS-GB dataset (127), encompasses hydrometeorological and attribute data for 671 catchments across GB. These catchments also display a wide range of sizes, from 1.6 to 9,930 square kilometers, and include daily time series data for variables such as temperature, precipitation, potential evapotranspiration, and streamflow, covering the period from 1970 to 2015. The creation of the CAMELS-GB dataset aimed to synthesize existing data into a consistent, up-to-date collection that facilitates the comparability and reproducibility of hydrological studies across GB. This dataset represents a significant effort to provide a comprehensive view of the hydrological and meteorological characteristics of GB catchments, including aspects such as land cover, with crops and grassland being predominant in different regions, and the distribution of large reservoir capacities, especially in the more mountainous areas. Fig. 6 shows the spatial distribution of CAMELS catchments across the United States and GB, stratified with CL%.



**Fig. 6.** Geographical distribution of studied catchments. Panels (A) and (B) display the locations of CAMELS catchments in the United States and GB, respectively. The shades of red color represent the percentage of cropland. The size of dots is proportional to the magnitude of the percentage of cropland.

Both datasets cover a wide range of climatic, hydrological, landscape, and human management characteristics, underscores the complexity of watershed dynamics and the importance of large-sample studies in advancing our understanding of hydrological systems.

They serve as invaluable resources for researchers aiming to analyze regional variability in hydrologic model performance, assess the impact of climatic and land-use changes on water resources, and develop strategies for sustainable water management.

**4.2. Snow Estimation.** To employ the causal discovery algorithm and enhance our comprehension of the influences on water balance, it is imperative to pinpoint the primary drivers shaping variations in streamflow. Within this framework, our focus centers on three pivotal factors impacting streamflow (Q): precipitation (P), AR, and SF. Directly extracted from the CAMELS datasets, precipitation and AR data serve as foundational components. For the determination of the proportion of precipitation designated as snow (SF), a straightforward temperature-based threshold was applied. Specifically, when the average daily temperature dipped below 1 °C, all precipitation was categorized as snowfall (66, 128). Conversely, if temperatures surpassed 1 °C, the precipitation was classified as rainfall. Additionally, an alternative threshold of 0 °C (70) was explored, yielding consistent outcomes regardless of the chosen method for estimating snowfall.

**4.3. Budyko Framework.** Here, we employ the Budyko framework, a well-known approach to examine the complex relationships between streamflow elements, meteorological conditions, and catchment features in relation to the long-term water balance (41, 52, 73). Based on AR index ( $\phi$ ), this framework formalizes the partitioning of precipitation (P) into evapotranspiration (ET) and runoff (Q). The framework aids to evaluate the main variables affecting the components of water balance. Through an examination of the ways in which catchments with certain characteristics deviate from the Budyko curve, we may assess how much and how these features affect the components of the water balance beyond the AR index (66, 68). The Budyko curve is represented by the following equation (56):

$$\frac{1 - \bar{Q}/\bar{P}}{\bar{P}} = \sqrt{\frac{\bar{E}_p}{\bar{P}} \tanh\left(\frac{\bar{P}}{\bar{E}_p}\right) \left(1 - \exp\left(-\frac{\bar{E}_p}{\bar{P}}\right)\right)}, \quad [1]$$

where  $\bar{Q}$ ,  $\bar{P}$ , and  $\bar{E}_p$  denote the long-term mean values for streamflow, precipitation, and potential evaporation.

**4.4. RF for Analysis of Attribution.** To identify the factors contributing to deviations from the Budyko curve, it is essential to quantify the relative importance of each variable. Traditional approaches, such as linear or logistic regression models, can be used to evaluate the influence of explanatory variables on a response variable (129). However, multiple regression method is only applicable when there is a linear relationship between predictors and the response variable, and when the assumptions of normally distributed residuals, homoscedasticity, and independence are met (130).

Given the nonlinear interactions between deviations from the Budyko curve and contributing factors, we employed the RF algorithm. RF is well suited for handling complex, nonlinear relationships and interactions, offering robust performance in high-dimensional settings which enables the relative importance of influencing factors (131, 132). It uses an ensemble of Classification and Regression Trees (CARTs), enhancing predictive accuracy and reducing variance compared to single decision trees (133). The model employs bootstrap aggregation ("bagging") to create multiple subsets of the original data via random resampling with replacement, allowing each tree to train on unique subsets. Data points not selected in these samples form "OOB" samples, which are critical for assessing variable importance.

To evaluate the importance of a variable, RF permutes the variable in the OOB samples while keeping others constant and reruns the model. The variable's importance is calculated as the average difference in prediction accuracy between the original and permuted OOB samples. Mathematically, the variable importance (VI) of an explanatory variable  $x$  is defined as

$$VI(x) = \frac{1}{ntree} \sum_{i=1}^{ntree} (\tilde{y}_i - y_i), \quad [2]$$

where  $y$  is the prediction with the original OOB sample,  $\tilde{y}$  is the prediction with the permuted OOB sample, and  $ntree$  denotes the total number of trees in the forest. A larger value of VI indicates a higher importance of the variable.

In this study, RF was used to evaluate the relative importance of cropland percentage (CL%), vegetation, and precipitation seasonality in explaining deviations from the Budyko curve, offering insights into the key drivers of hydrological behavior across different catchments.

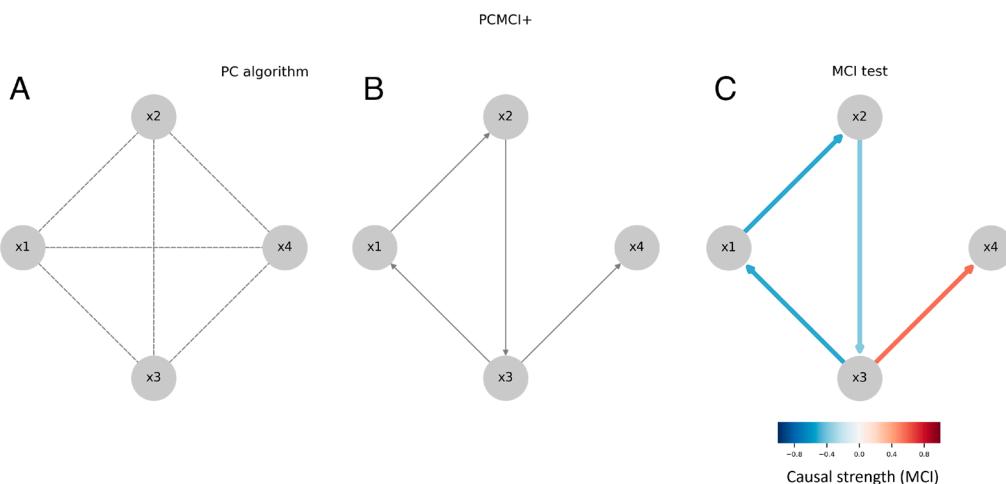
**4.5. Causal Discovery Algorithm.** In our study, we aim to understand the impact of agriculture on the water cycle across various catchments in the United States and GB by examining causal relationships between climate variables and streamflow. Traditional approaches, like the Budyko hypothesis, provide valuable insights into the spatial influence of agriculture on the water cycle; however, they fall short in explaining the causal mechanisms behind observed deviations in water balance. Most statistical attribution studies rely on qualitative reasoning, sensitivity-based analysis, or correlation-based techniques, which often focus on a single driver and do not adequately quantify causal relationships among multiple interacting factors (66, 69, 134, 135).

To address these limitations, we employ a causal discovery approach using the Peter-Clark MCI Plus (PCMCI+) algorithm. This algorithm allows us to explore the temporal interactions among hydrological components that shape deviations from the Budyko curve, without violating the principle of mass conservation ( $P - ET = Q$ ). The focus is on understanding how specific drivers, such as SF or cropland percentage (CL%), perturb these relationships over time, leading to deviations in long-term averages. PCMCI+ is specifically designed to detect both contemporaneous (simultaneous) and time-lagged causal relationships in high-dimensional time series data, which is common in environmental studies (136, 137). The PCMCI+ algorithm identifies causal connections among variables. It starts by constructing a Directed Acyclic Graph (DAG), which represents the system's causal structure (58, 138). In a DAG, variables are represented as nodes, and directed edges (arrows) illustrate causal relationships between them as shown in Fig. 7.

The PCMCI+ algorithm operates in two main stages (139). The first stage, known as the PC Stage (Conditional Independence Testing), uses the PC1 algorithm, a variant of the original Peter-Clark (PC) algorithm, to identify a set of variables that best explains the dependencies of that variable. This stage begins with a fully connected graph as shown in Fig. 7A and tests conditional independence between variables by incrementally increasing the number of variables in the conditioning set. By applying linear partial correlation tests, the algorithm determines which connections between variables can be removed while preserving the graph's validity. The process continues iteratively, aiming to converge on a subset of essential links at a significance level of 0.05, ensuring that only significant causal links remain as shown in Fig. 7B.

In the second stage, known as the MCI Stage (Causal Inference), the MCI test leverages the estimated conditions identified during the PC stage to infer causal relationships between variables. This stage assigns a causal strength value, measured by MCI partial correlation, to each detected causal link (74, 75) (Fig. 7C). The MCI test can capture both linear and nonlinear dependencies, providing a deeper understanding of causal interactions. Together, these two stages allow the PCMCI+ algorithm to effectively distinguish direct causal influences from indirect or spurious connections, enhancing the reliability of the resulting causal graph. The Python software for estimating the causal network is available at <https://jakobrunge.github.io/tigramite>.

**4.6. Decomposition of Streamflow.** To further explore the analysis of water sources for agriculture, we utilized a one-parameter low-pass filter method developed by Lyne and Hollick (79-84), commonly used in the literature (82, 84, 140, 141) to effectively decompose streamflow into its constituent components, namely baseflow and direct flow. This method relies on temporal filtering principles, wherein the streamflow hydrograph undergoes convolution with a smoothing function defined by a filter parameter, often referred to as the recession constant ( $\alpha$ ). We applied the recursive filter iteratively three times (forward-back-forward) to filter off flood peaks from the original streamflow time series, configuring the filter parameter [set as 0.925; a widely accepted value in hydrological studies (79, 80, 140, 142-144)]. The resulting separation let us to discern the slow, groundwater-derived baseflow from the rapid, rainfall-induced direct flow. To account for the inherent uncertainty of streamflow partitioning (81), we repeated our analysis using the Lyne Hollick one-pass configuration, which yielded quite similar results.



**Fig. 7.** Overview of the PCMI+ causal discovery algorithm. The algorithm consists of two main stages: the PC algorithm and MCI tests. In the first stage (Panel A), the PC1 algorithm begins with a fully connected graph. It then iteratively removes connections between variables by testing conditional independence with sets of increasing size (Panel B). In the second stage (Panel C), the MCI tests apply the conditions identified in the first step to determine strength of causal links.

To calculate baseflow, we used Lyne and Hollick's (79-84) one-parameter low-pass filter approach, which is a well-known method in the literature (82, 84, 140, 141). By applying temporal filtering principles, this approach effectively partitions streamflow into its two main components: baseflow and direct flow. Recession constant ( $\alpha$ ), a filter parameter, is used to build a smoothing function that is used to convolve the streamflow hydrograph. We applied the recursive filter three times (forward-back-forward), using a filter parameter set to 0.925, a widely accepted value in hydrological studies (79, 80, 140, 142-144). The original streamflow time series' flood peaks were effectively removed by this iterative procedure, making it possible to distinguish between rapid, rainfall-induced direct flow and slow, groundwater-derived baseflow. To account for the inherent uncertainty of streamflow partitioning (81), we repeated our analysis using the Lyne Holick one-pass configuration, which led to similar results.

**Data, Materials, and Software Availability.** Catchments in Great Britain (CAMELS-GB), daily hydro-meteorological time series and landscape features; data for the US (CAMELS); causal network estimation code data have been deposited in Environmental Information Data Centre; Geoscience Data Exchange Centre

of the NSF; Potsdam Institute for Climate Impact Research (PIK) (<https://catalogue.ceh.ac.uk/documents/8344e4f3-d2ea-44f5-8afa-86d2987543a9>; <https://dx.doi.org/10.5065/D6MW2F4D>; and <https://tocsy.pik-potsdam.de/tigramite.php>). Previously published data were used for this work (124, 127).

**ACKNOWLEDGMENTS.** Financial support for this study was provided by the Natural Sciences and Engineering Research Council of Canada. We would like to thank the Editor-in-Chief, Prof. May R. Berenbaum, and the Editor, Prof. Andrea Rinaldo, for handling our manuscript. We also extend our gratitude to the two anonymous reviewers for their constructive comments, which significantly improved our manuscript.

Author affiliations: <sup>a</sup>Department of Civil Engineering, Schulich School of Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada; <sup>b</sup>Department of Hydraulics and Sanitation, São Carlos School of Engineering, University of São Paulo, São Paulo 13566-590, Brazil; <sup>c</sup>Department of Water Resources and Environmental Modeling, Faculty of Environmental Sciences, Czech University of Life Sciences, Prague 165 00, Czech Republic; and <sup>d</sup>Department of Building, Civil, and Environmental Engineering, Concordia University, Montreal, QC H3G 2W1, Canada

Author contributions: M.Z., S.H., S.M.P., and A.P. designed research; M.Z. performed research; M.Z., S.H., and A.S.B. contributed new reagents/analytic tools; M.Z., S.H., and A.S.B. analyzed data; and M.Z., S.H., A.S.B., S.M.P., A.P., and A.N. wrote the paper.

- J. Zhang *et al.*, Sustainable irrigation based on co-regulation of soil water supply and atmospheric evaporative demand. *Nat. Commun.* **12**, 5549 (2021).
- A. Mrad *et al.*, Peak grain forecasts for the US high plains amid withering waters. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 26145–26150 (2020).
- D. Wisser *et al.*, Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets. *Geophys. Res. Lett.* **35**, 4–87 (2008).
- X. Cai, M. W. Rosegrant, Global water demand and supply projections. *Water Int.* **27**, 159–169 (2002).
- P. Döll *et al.*, Impact of water withdrawals from groundwater and surface water on continental water storage variations. *J. Geodyn.* **59–60**, 143–156 (2012).
- P. Döll *et al.*, Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resour. Res.* **50**, 5698–5720 (2014).
- J. Hoogeveen, J.-M. Faurès, L. Peiser, J. Burke, N. van de Giesen, GlobWat – a global water balance model to assess water use in irrigated agriculture. *Hydrol. Earth Syst. Sci.* **19**, 3829–3844 (2015).
- Y. Wada, D. Wisser, M. F. P. Bierkens, Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth Syst. Dyn.* **5**, 15–40 (2014).
- H. Hoff *et al.*, Greening the global water system. *J. Hydrol.* **384**, 177–186 (2010).
- M. D. Edgerton, Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiol.* **149**, 7–13 (2009).
- M. Maghrebi *et al.*, Iran's agriculture in the Anthropocene. *Earth's Future* **8**, e2020EF001547 (2020).
- I. A. Shiklomanov, Assessment of water resources and water availability in the world: Comprehensive assessment of the freshwater resources of the world (World Meteorological Organisation, Geneva, 1997). <https://digitallibrary.un.org/record/261669?ln=en&v=pdf>. Accessed 1 March 2024.
- P. Döll, S. Siebert, Global modeling of irrigation water requirements. *Water Resour. Res.* **38**, 8–1–8–10 (2002).
- R. Q. Grafton *et al.*, The paradox of irrigation efficiency. *Science* **361**, 748–750 (2018).
- D. Griggs *et al.*, An integrated framework for sustainable development goals. *Ecol. Soc.* **19**, 4 (2014).
- V. McGuire, *Water-Level and Recoverable Water in Storage Changes, High Plains Aquifer, Predevelopment to 2015 and 2013–15* (USGS Staff – Published Research, 2017).
- D. R. Steward *et al.*, Tapping unsustainable groundwater stores for agricultural production in the high plains Aquifer of Kansas, projections to 2110. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E3477–E3486 (2013).
- E. M. K. Haacker, A. D. Kendall, D. W. Hyndman, Water level declines in the high plains aquifer: Predevelopment to resource senescence. *Groundwater* **54**, 231–242 (2016).
- C. J. Vörösmarty, C. Lévéque, C. Revenga, Fresh water, in ecosystems and human well-being: Current states and trends, millennium ecosystem assessment. *Report*. (2005), [https://scholar.google.com/scholar\\_lookup?hl=en&publication\\_year=2005&pages=165-207&author=C.+J.+V%2C3%6f%C3%83%C3%A9&author=C.+Revenga&title=Ecosystems+and+Human+Well%2E%20%90Being%3A+Current+States+and+Trends](https://scholar.google.com/scholar_lookup?hl=en&publication_year=2005&pages=165-207&author=C.+J.+V%2C3%6f%C3%83%C3%A9&author=C.+Revenga&title=Ecosystems+and+Human+Well%2E%20%90Being%3A+Current+States+and+Trends).
- Y. Wada, L. P. H. van Beek, M. F. P. Bierkens, Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resour. Res.* **48**, 2055 (2012).
- J. Margat, S. Foster, A. Droulli, Concept and importance of non-renewable resources. *Non-Renew. Groundwater Resour.* **10**, 13–24 (2006).
- T. Gleeson *et al.*, Water balance of global aquifers revealed by groundwater footprint. *Nature* **488**, 197–200 (2012).
- M. F. P. Bierkens, Y. Wada, Non-renewable groundwater use and groundwater depletion: A review. *Environ. Res. Lett.* **14**, 063002 (2019).
- S. Rost *et al.*, Agricultural green and blue water consumption and its influence on the global water system. *Water Resour. Res.* **44**, W09405 (2008).
- X. Liu *et al.*, Multimodel uncertainty changes in simulated river flows induced by human impact parameterizations. *Environ. Res. Lett.* **12**, 025009 (2017).
- J. Liu, H. Yang, Spatially explicit assessment of global consumptive water uses in cropland: Green and blue water. *J. Hydrol.* **384**, 187–197 (2010).
- J. Schewe *et al.*, Multimodel assessment of water scarcity under climate change. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3245–3250 (2014).
- X. Liu *et al.*, Global agricultural water scarcity assessment incorporating blue and green water availability under future climate change. *Earth's Future* **10**, e2021EF002567 (2022).
- A. V. Veettil, A. K. Mishra, Water security assessment using blue and green water footprint concepts. *J. Hydrol.* **542**, 589–602 (2016).

30. R. L. B. Nóbrega *et al.*, Effects of conversion of native cerrado vegetation to pasture on soil hydro-physical properties, evapotranspiration and streamflow on the Amazonian agricultural frontier. *PLOS ONE* **12**, e0179414 (2017).

31. A. S. Richey *et al.*, Quantifying renewable groundwater stress with GRACE. *Water Resour. Res.* **51**, 5217–5238 (2015).

32. Z. W. Kundzewicz, P. Döll, Will groundwater ease freshwater stress under climate change? *Hydrol. Sci. J.* **54**, 665–675 (2009).

33. B. R. Scanlon *et al.*, Groundwater depletion and sustainability of irrigation in the US high plains and Central Valley. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9320–9325 (2012).

34. J. S. Perkin *et al.*, Groundwater declines are linked to changes in great plains stream fish assemblages. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7373–7378 (2017).

35. T. Gleeson *et al.*, Groundwater sustainability strategies. *Nat. Geosci.* **3**, 378–379 (2010).

36. J. S. Famiglietti *et al.*, Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophys. Res. Lett.* **38**, L03403 (2011).

37. M. GebreEgziabher, S. Jasechko, D. Perrone, Widespread and increased drilling of wells into fossil aquifers in the USA. *Nat. Commun.* **13**, 2129 (2022).

38. M. Rodell, I. Velicogna, J. S. Famiglietti, Satellite-based estimates of groundwater depletion in India. *Nature* **460**, 999–1002 (2009).

39. J. G. S. M. Uchôa *et al.*, Widespread potential for streamflow leakage across Brazil. *Nat. Commun.* **15**, 10211 (2024).

40. K. A. McColl *et al.*, The terrestrial water cycle in a warming world. *Nat. Clim. Chang.* **12**, 604–606 (2022).

41. R. P. Allan *et al.*, Advances in understanding large-scale responses of the water cycle to climate change. *Ann. N.Y. Acad. Sci.* **1472**, 49–75 (2020).

42. R. G. Taylor *et al.*, Ground water and climate change. *Nat. Clim. Chang.* **3**, 322–329 (2013).

43. L. E. Condon, A. L. Atchley, R. M. Maxwell, Evapotranspiration depletes groundwater under warming over the contiguous United States. *Nat. Commun.* **11**, 873 (2020).

44. Y. Wada *et al.*, Global depletion of groundwater resources. *Geophys. Res. Lett.* **37**, L20402 (2010).

45. D. Wang, M. Hejazi, Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States. *Water Resour. Res.* **47**, W00J12 (2011).

46. E. Kåresdotter *et al.*, Distinguishing direct human-driven effects on the global terrestrial water cycle. *Earth's Future* **10**, e2022EF002848 (2022).

47. W. R. Berghuijs *et al.*, The relative importance of different flood-generating mechanisms across Europe. *Water Resour. Res.* **55**, 4582–4593 (2019).

48. G. Blöschl *et al.*, Changing climate shifts timing of European floods. *Science* **357**, 588–590 (2017).

49. J.-F. Pekel, A. Cottam, N. Gorelick, A. S. Belward, High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).

50. I. E. M. de Graaf, T. Gleeson, L. P. H. van Beek, E. H. Sutanudjaja, M. F. P. Bierkens, Environmental flow limits to global groundwater pumping. *Nature* **574**, 90–94 (2019).

51. M. P. Miller, S. G. Buto, D. D. Susong, C. A. Rumsey, The importance of base flow in sustaining surface water flow in the upper Colorado River Basin. *Water Resour. Res.* **52**, 3547–3562 (2016).

52. M. Stoele, K. Stahl, A. Morhard, M. Weiler, Streamflow sensitivity to drought scenarios in catchments with different geology. *Geophys. Res. Lett.* **41**, 6174–6183 (2014).

53. L. Longuevergne, B. R. Scanlon, C. R. Wilson, GRACE hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA. *Water Resour. Res.* **46**, 11517 (2010).

54. W. Nie *et al.*, Groundwater withdrawals under drought: Reconciling GRACE and land surface models in the United States high plains Aquifer. *Water Resour. Res.* **54**, 5282–5299 (2018).

55. A. Puy *et al.*, The delusive accuracy of global irrigation water withdrawal estimates. *Nat. Commun.* **13**, 3183 (2022).

56. M. I. Budyko, *Climate and Life* (Academic, 1974), vol. 508.

57. J. Runge *et al.*, Inferring causation from time series in earth system sciences. *Nat. Commun.* **10**, 2553 (2019).

58. J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, D. Sejdinovic, Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci. Adv.* **5**, eaau4996 (2019).

59. L. Zhang, W. R. Dawes, G. R. Walker, Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resour. Res.* **37**, 701–708 (2001).

60. J. Liu, Y. You, The roles of catchment characteristics in precipitation partitioning within the Budyko framework. *J. Geophys. Res. Atmos.* **126**, e2021JD035168 (2021).

61. R. J. Donohue, M. L. Roderick, T. R. McVicar, Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. *J. Hydrol.* **436–437**, 35–50 (2012).

62. J. Fu, W. Wang, On the lower bound of Budyko curve: The influence of precipitation seasonality. *J. Hydrol.* **570**, 292–303 (2019).

63. D. Li, M. Pan, Z. Cong, L. Zhang, E. Wood, Vegetation control on water and energy balance within the Budyko framework. *Water Resour. Res.* **49**, 969–976 (2013).

64. P. Nasta, C. Allocà, R. Deidda, N. Romano, Assessing the impact of seasonal-rainfall anomalies on catchment-scale water balance components. *Hydrol. Earth Syst. Sci.* **24**, 3211–3227 (2020).

65. R. A. P. Perdigão, G. Blöschl, Spatiotemporal flood sensitivity to annual precipitation: Evidence for landscape-climate coevolution. *Water Resour. Res.* **50**, 5492–5509 (2014).

66. W. R. Berghuijs, R. A. Woods, M. Hrachowitz, A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nat. Clim. Chang.* **4**, 583–586 (2014).

67. W. R. Berghuijs, J. R. Larsen, T. H. M. van Emmerik, R. A. Woods, A global assessment of runoff sensitivity to changes in precipitation, potential evaporation, and other factors. *Water Resour. Res.* **53**, 8475–8486 (2017).

68. W. R. Berghuijs, S. J. Gnann, R. A. Woods, Unanswered questions on the Budyko framework. *Hydrol. Process.* **34**, 5699–5703 (2020).

69. W. R. Berghuijs, R. A. Woods, Correspondence: Space-time asymmetry undermines water yield assessment. *Nat. Commun.* **7**, 11603 (2016).

70. W. J. M. Knoben, R. A. Woods, J. E. Freer, A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resour. Res.* **54**, 5088–5109 (2018).

71. L. Stein, M. P. Clark, W. J. M. Knoben, F. Pianosi, R. A. Woods, How do climate and catchment attributes influence flood generating processes? A large-sample study for 671 catchments across the contiguous USA. *Water Resour. Res.* **57**, e2020WR028300 (2021).

72. E. Galytska *et al.*, Evaluating causal Arctic-Midlatitude teleconnections in CMIP6. *J. Geophys. Res. Atmos.* **128**, e2022JD037978 (2023).

73. M. Ombadi, P. Nguyen, S. Sorooshian, K. Hsu, Evaluation of methods for causal discovery in hydrometeorological systems. *Water Resour. Res.* **56**, e2020WR027251 (2020).

74. J. Runge, Quantifying information transfer and mediation along causal pathways in complex systems. *Phys. Rev. E* **92**, 062829 (2015).

75. J. Runge *et al.*, Identifying causal gateways and mediators in complex spatio-temporal systems. *Nat. Commun.* **6**, 8502 (2015).

76. P. Gentile, P. D'Odorico, B. R. Lintner, G. Sivandran, G. Salvucci, Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. *Geophys. Res. Lett.* **39**, 19404 (2012).

77. R. A. Freeze, Role of subsurface flow in generating surface runoff: 1. Base flow contributions to channel flow. *Water Resour. Res.* **8**, 609–623 (1972).

78. K. Eckhardt, How to construct recursive digital filters for baseflow separation. *Hydrol. Process.* **19**, 507–515 (2005).

79. V. Lyne, M. Hollick, Stochastic time variable rainfall-runoff modelling (1979), pp. 89–93. [https://scholar.google.com/scholar\\_lookup?hl=en&publication\\_year=1979&pages=82-92&author=V.+D.+Lyne&author=M.+Hollick&title=Hydrology+and+water+resources+symposium](https://scholar.google.com/scholar_lookup?hl=en&publication_year=1979&pages=82-92&author=V.+D.+Lyne&author=M.+Hollick&title=Hydrology+and+water+resources+symposium).

80. S. J. Gannan, R. A. Woods, N. J. K. Howden, Is there a baseflow Budyko curve? *Water Resour. Res.* **55**, 2838–2855 (2019).

81. A. A. Meira Neto, T. Roy, P. T. S. de Oliveira, P. A. Troch, An aridity index-based formulation of streamflow components. *Water Resour. Res.* **56**, e2020WR027123 (2020).

82. M. Sivapalan, M. A. Yaeger, C. J. Harman, X. Xu, P. A. Troch, Functional model of water balance variability at the catchment scale: 1. Evidence of hydrologic similarity and space-time symmetry. *Water Resour. Res.* **47**, W02522 (2011).

83. R. Trancoso, J. R. Larsen, T. R. McVicar, S. R. Phinn, C. A. McAlpine, CO<sub>2</sub>-vegetation feedbacks and other climate changes implicated in reducing base flow. *Geophys. Res. Lett.* **44**, 2310–2318 (2017).

84. A. S. Ballarin *et al.*, The impact of an open water balance assumption on understanding the factors controlling the long-term streamflow components. *Water Resour. Res.* **58**, e2022WR032413 (2022).

85. I. R. McDadden *et al.*, Linking human impacts to community processes in terrestrial and freshwater ecosystems. *Ecol. Lett.* **26**, 203–218 (2023).

86. R. I. McDonald *et al.*, Urban growth, climate change, and freshwater availability. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6312–6317 (2011).

87. G. Villarini, C. Wasko, Humans, climate and streamflow. *Nat. Clim. Chang.* **11**, 725–726 (2021).

88. L. Gudmundsson *et al.*, Globally observed trends in mean and extreme river flow attributed to climate change. *Science* **371**, 1159–1162 (2021).

89. Y. Zhang *et al.*, Future global streamflow declines are probably more severe than previously estimated. *Nat. Water* **1**, 261–271 (2023), 10.1038/s44221-023-00030-7.

90. X. Kuang *et al.*, The changing nature of groundwater in the global water cycle. *Science* **383**, eadf0630 (2024).

91. K. M. Suttles *et al.*, Assessment of hydrologic vulnerability to urbanization and climate change in a rapidly changing watershed in the Southeast U.S. *Sci. Total Environ.* **645**, 806–816 (2018).

92. N. K. Singh, N. B. Basu, The human factor in seasonal streamflows across natural and managed watersheds of North America. *Nat. Sustain.* **5**, 397–405 (2022).

93. S. Jasechko, H. Seybold, D. Perrone, Y. Fan, J. W. Kirchner, Widespread potential loss of streamflow into underlying aquifers across the USA. *Nature* **591**, 391–395 (2021).

94. L. E. Condon, R. M. Maxwell, Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion. *Sci. Adv.* **5**, eaav4574 (2019).

95. L. E. Condon, R. M. Maxwell, Groundwater-fed irrigation impacts spatially distributed temporal scaling behavior of the natural system: A spatio-temporal framework for understanding water management impacts. *Environ. Res. Lett.* **9**, 034009 (2014).

96. V. Andriassian, Waters and forests: From historical controversy to scientific debate. *J. Hydrol.* **291**, 1–27 (2004).

97. V. B. P. Chagas, P. L. B. Chaffé, G. Blöschl, Climate and land management accelerate the Brazilian water cycle. *Nat. Commun.* **13**, 5136 (2022).

98. D. R. Moore, S. M. Wondzell, Physical hydrology and the effects of forest harvesting in the Pacific Northwest: A review. *J. Am. Water Resour. Assoc.* **41**, 763–784 (2005).

99. M. J. Booij, T. C. Schipper, H. Marhafta, Attributing changes in streamflow to land use and climate change for 472 catchments in Australia and the United States. *Water* **11**, 1059 (2019).

100. R. Zeng, C. Yao, X. Cai, E. Haacker, Temporal and spatial pattern change in evapotranspiration over the High Plains: The impact of and guide on extensive groundwater-fed irrigation. *Water Resour. Res.* **59**, e2023WR035004 (2023).

101. D. S. Grogan, D. Wisser, A. Prusevich, R. B. Lammers, S. Frolking, The use and re-use of unsustainable groundwater for irrigation: A global budget. *Environ. Res. Lett.* **12**, 034017 (2017).

102. S. Jasechko *et al.*, Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature* **625**, 715–721 (2024).

103. Z. Liu, H. Chen, Z. Huo, F. Wang, C. C. Shock, Analysis of the contribution of groundwater to evapotranspiration in an arid irrigation district with shallow water table. *Agric. Water Manag.* **171**, 131–141 (2016).

104. D. Ketchum, Z. H. Hoylman, J. Huntington, D. Brinkerhoff, K. G. Jencso, Irrigation intensification impacts sustainability of streamflow in the Western United States. *Commun. Earth Environ.* **4**, 1–8 (2023).

105. L. Bouaziz *et al.*, Redressing the balance: Quantifying net intercatchment groundwater flows. *Hydrol. Earth Syst. Sci.* **22**, 6415–6434 (2018).

106. H. Chen, Z. Huo, L. Zhang, J. White, New perspective about application of extended Budyko formula in arid irrigation district with shallow groundwater. *J. Hydrol.* **582**, 124496 (2020).

107. L. Cheng, Z. Xu, D. Wang, X. Cai, Assessing interannual variability of evapotranspiration at the catchment scale using satellite-based evapotranspiration data sets. *Water Resour. Res.* **47**, W09509 (2011).

108. X. Wang, Z. Huo, S. Feng, P. Guo, H. Guan, Estimating groundwater evapotranspiration from irrigated cropland incorporating root zone soil texture and moisture dynamics. *J. Hydrol.* **543**, 501–509 (2016).

109. X. Wang *et al.*, Energy fluxes and evapotranspiration over irrigated maize field in an arid area with shallow groundwater. *Agric. Water Manag.* **228**, 105922 (2020).

110. A. P. Schreiner-McGraw *et al.*, Surface resistance controls differences in evapotranspiration between croplands and prairies in U.S. Corn Belt sites. *Water Resour. Res.* **60**, e2023WR035819 (2024).

111. E. K. Weatherhead, N. J. K. Howden, The relationship between land use and surface water resources in the UK. *Land Use Policy* **26**, S243–S250 (2009).

112. W. Stephens, T. M. Hess, J. W. Knox, Review of the effects of energy crops on hydrology (2001). <https://dspace.lib.cranfield.ac.uk/handle/1826/3368>. Accessed 1 March 2024.

113. P. Rameshwaran *et al.*, Use of abstraction and discharge data to improve the performance of a national-scale hydrological model. *Water Resour. Res.* **58**, e2021WR029787 (2022).

114. T. E. Huxman *et al.*, Convergence across biomes to a common rain-use efficiency. *Nature* **429**, 651–654 (2004).

115. S. M. Vicente-Serrano *et al.*, Climate, irrigation, and land cover change explain streamflow trends in countries bordering the Northeast Atlantic. *Geophys. Res. Lett.* **46**, 10821–10833 (2019).

116. D. Perrone, S. Jasechko, Deeper well drilling an unsustainable stopgap to groundwater depletion. *Nat. Sustain.* **2**, 773–782 (2019).

117. S. Harrigan *et al.*, Attribution of detected changes in streamflow using multiple working hypotheses. *Hydroclim. Earth Syst. Sci.* **18**, 1935–1952 (2014).

118. Y. Zheng *et al.*, Controls on the spatial and temporal patterns of rainfall-runoff event characteristics—A large sample of catchments across Great Britain. *Water Resour. Res.* **59**, e2022WR033226 (2023).

119. L. J. E. Lee, D. S. L. Lawrence, M. Price, Analysis of water-level response to rainfall and implications for recharge pathways in the Chalk aquifer SE England. *J. Hydrol.* **330**, 604–620 (2006).

120. A. G. Hughes *et al.*, Flood risk from groundwater: Examples from a Chalk catchment in southern England. *J. Flood Risk Manage.* **4**, 143–155 (2011).

121. K. W. King, N. R. Fausey, M. R. Williams, Effect of subsurface drainage on streamflow in an agricultural headwater watershed. *J. Hydrol.* **519**, 438–445 (2014).

122. M. Robinson, D. W. Rycroft, "The impact of drainage on streamflow" in *Agricultural Drainage* (John Wiley & Sons Ltd, 1999), pp. 767–800, 10.2134/agronmonogr38.c23.

123. R. W. Skaggs, M. A. Brevé, J. W. Gilliam, Hydrologic and water quality impacts of agricultural drainage. *Crit. Rev. Environ. Sci. Technol.* **24**, 1–32 (1994).

124. N. Addor *et al.*, The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydroclim. Earth Syst. Sci.* **21**, 5293–5313 (2017).

125. H. F. Lins, USGS Hydro-climatic data network 2009 (HCDN-2009). <https://pubs.usgs.gov/fs/2012/3047/> (2009). Accessed 1 March 2024.

126. N. Ramankutty *et al.*, Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob. Biogeochem. Cycles* **22**, 1003 (2008).

127. G. Coxon *et al.*, CAMELS-GB: Hydroclimatic time series and landscape attributes for 671 catchments in Great Britain. *Earth Syst. Sci. Data* **12**, 2459–2483 (2020).

128. R. Hock, Temperature index melt modelling in mountain areas. *J. Hydrol.* **282**, 104–115 (2003).

129. H. Fan, D. He, Temperature and precipitation variability and its effects on streamflow in the upstream regions of the Lancang–Mekong and Nu–Salween Rivers. *J. Hydrometeorol.* **16**, 2248–2263 (2015).

130. E. Vittinghoff *et al.*, *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. (Springer, US, Boston, MA, 2012).

131. Y. Wu, N. Li, Nonlinear control of climate, hydrology, and topography on streamflow response through the use of interpretable machine learning across the contiguous United States. *J. Water Clim. Change* **14**, 4084–4098 (2023).

132. L. Schoppa, M. Disse, S. Bachmair, Evaluating the performance of random forest for large-scale flood discharge simulation. *J. Hydrol.* **590**, 125531 (2020).

133. V. F. Rodriguez-Galiano *et al.*, Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **121**, 93–107 (2012).

134. G. Zhou *et al.*, Global pattern for the effect of climate and land cover on water yield. *Nat. Commun.* **6**, 5918 (2015).

135. B. Merz *et al.*, HESS Opinions 'More efforts and scientific rigour are needed to attribute trends in flood time series'. *Hydroclim. Earth Syst. Sci.* **16**, 1379–1387 (2012).

136. M. Zaerpour *et al.*, Climate shapes baseflows, influencing drought severity. *Environ. Res. Lett.* **20**, 014035 (2024).

137. D. Delforge *et al.*, Detecting hydrological connectivity using causal inference from time series: Synthetic and real karstic case studies. *Hydroclim. Earth Syst. Sci.* **26**, 2181–2199 (2022).

138. J. Runge "Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets" in *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)* (PMLR, 2020), pp. 1388–1397.

139. P. Spirtes *et al.*, *Causation, Prediction, and Search* (MIT Press, 2000).

140. S. Lee, H. Ajami, Comprehensive assessment of baseflow responses to long-term meteorological droughts across the United States. *J. Hydrol.* **626**, 130256 (2023).

141. S. J. Gnann *et al.*, Including regional knowledge improves baseflow signature predictions in large sample hydrology. *Water Resour. Res.* **57**, e2020WR028354 (2021).

142. A. R. Ladson *et al.*, A standard approach to baseflow separation using the Lyne and Hollick filter. *Aust. J. Water Resour.* **17**, 25–34 (2013).

143. W. R. Berghuijs, L. J. Slater, Groundwater shapes North American river floods. *Environ. Res. Lett.* **18**, 034043 (2023).

144. J. Zhang *et al.*, Evaluating relative merits of four baseflow separation methods in Eastern Australia. *J. Hydrol.* **549**, 252–263 (2017).