Young Scientists Summer Program

Global archetypes of groundwater interactions in social-ecological systems

Xander Huggins
xanderhuggins@uvic.ca

Approved by:

Mentor(s): Taher Kahil, Amanda Palazzo
Program: WAT
Date: 30 September 2022

This report represents the work completed by the author during the IIASA Young Scientists Summer Program (YSSP) with approval from the YSSP mentor.

It was finished by 30 September 2022 and has not been altered or revised since.

Mentor signatures:

Taher Kahil

Amanda Palazzo

Table of contents

List of Figures

List of Tables

ZVR 524808900

Abstract

Groundwater resources are deeply embedded in social and ecological systems and the blossoming fields of sociohydrology and ecohydrology continue to reveal a wide array of social, economic, ecological, and Earth system functions that groundwater resources provide. While several global classification systems exist for groundwater, all to date have focused exclusively on the physiographic attributes of groundwater and do not consider the patters of groundwater interactions with these social and biophysical systems. While physiographic attributes remain important and valuable from a hydrogeological perspective, a more comprehensive classification system is needed for sustainability-oriented research that requires consideration of the social-ecological context of groundwater systems. Here, we provide a global mapping of groundwater system archetypes based on biophysical and social system interactions with groundwater. We define seven archetypes, each with unique patterns of interactions with the climate, surface water, wetlands, social inequality, irrigation, economic value, and integrated water resources management. We find that economic, climate, irrigation, and wetland interactions are the most important variables in driving archetype membership, a finding that underscores the social-ecological nature of these archetypes. We provide an outlook for each archetype by evaluating the distribution of a set of pressures which include cropland expansion potential, likelihood for hydropolitical interaction, groundwater depletion rates, and population growth factors. The combination of our archetypes with this post-hoc analysis of pressures provides a unique tool to develop data driven narratives of global groundwater futures and can help facilitate insights on causal mechanisms of global groundwater dynamics and generate theories of change to promote groundwater sustainability. This project remains a work in progress and a list of remaining modifications and analyses are provided.

About the author

Xander Huggins is a PhD candidate in the Department of Civil Engineering at the University of Victoria (Canada) and at the Global Institute for Water Security (University of Saskatchewan, Canada).

Acknowledgments

I would like to thank my IIASA mentors Taher Kahil and Amanda Palazzo for useful conversations and feedback on this work. I would also like to thank Dor Fridman for discussions on archetyping approaches, as well as summer conversations with Viviana Re (Università di Pisa) on scale and human-groundwater interactions that supported development of this work. I'd also like to thank my co-supervisors Tom Gleeson and Jay Famiglietti for their continued support and mentorship.

Disclaimer

This report documents work performed during the 2022 Young Scientists Summer Program at the International Institute for Applied Systems Analysis. The contents of this report will continue to be modified and refined for submission in a peer-reviewed journal. The purpose of this document is not to serve as the eventual publication's preprint and is best understood as simply a 'status update' of the project's progress. A subsequent pre-print will be released on this work once it is submitted for peer-review.

Introduction

We solve problems according to how we understand them. This report wrestles with the related ideas of problem framing and problem definition in the context of the global groundwater crisis. We provide a novel systems-oriented framing and spatial template to address the global groundwater sustainability problem space through a data driven geospatial archetyping analysis.

Groundwater is a globally distributed and massive natural resource. It is the world's largest store of liquid freshwater[1] and the water table lies anywhere from hundreds of meters below the ground surface to at the ground surface itself[2]. Across all of these environments, but in different ways, groundwater provides critical functions for social[3,4], ecological[5,6], and Earth systems[7]. Over two billion people rely on groundwater as their source of drinking water, and roughly half (43%) of the water used to irrigate the world's crops comes from groundwater[8]. This dependence on groundwater will continue to increase under climate change as precipitation and streamflow patterns become increasingly variable, rendering groundwater a more reliable source of freshwater[9]. Groundwater also sustains streamflow during dry, low-flow months in the form of baseflow. This baseflow is critical to sustain aquatic and riparian ecosystems[10]. Shallow water tables and groundwater discharge at the land surface support terrestrial groundwater-dependent ecosystems (GDEs)[11], and these GDEs in turn provide myriad ecosystem services[12]. Surface discharge in the form of springs are not only hotspots for biodiversity but are also culturally significant sites for people[13,14]. In sum, groundwater is a resource that is connected to, but also serves the *role of connecting* physical, ecological, and social processes and systems around the world.

Yet, groundwater is a resource in global crisis[15]. Over half of the major aquifers of the world are in states of depletion[16]. In some instances, policy changes have led to the beginning of recovery in areas long-experiencing groundwater depletion[17] however groundwater governance and management remain absent or in initial development stages in most regions of the world[18]. This mismanagement, coupled with continued climate change and the world's increasing population and shifting diet, all indicate that global groundwater resources are likely face increasing pressures in the decades ahead.

The persistence of groundwater depletion at the global scale poses significant risk to all interactions and functions we described above. Decreased water security[19] and greater food insecurity[20] are often discussed impacts of groundwater depletion. Yet, what can be lost in these overarching, traditional narratives are the deeply systematic impacts groundwater depletion can set in motion. For instance, groundwater depletion can amplify both economic and water security inequalities as only wealthy well owners and large corporations are able to afford to drill deeper wells to keep pace with the rate of falling water tables[21]. This can create scenarios where the wealthy retain their access to groundwater and the economic benefits provided by groundwater access. Simultaneously, the deepening of the water table reduces the number of users sharing groundwater from the aquifer as shallower wells are no longer able to access the resource. Conversely, shallow wells that have run dry not only lose access to the groundwater system but also the ability for their owners to use this water for domestic, industrial, agricultural, or cultural purposes. These inequalities are further aggravated when non-economic values of groundwater are neglected. Falling water tables alter groundwater-surface water interactions[22] that not only lead to transgressed environmental streamflow needs with devastating impacts on aquatic ecosystems [10] but also equally devastating impacts on communities and cultures that have deep connections to these surface water bodies[23]. While no existing global estimates exist for the potential for economic inequalities to be aggravated by groundwater depletion, or on the pervasiveness of compromised cultural ecosystem services of groundwater, it is estimated that over half of all streamflows will be in states where their environmental flow needs are transgressed by the year 2050[24].

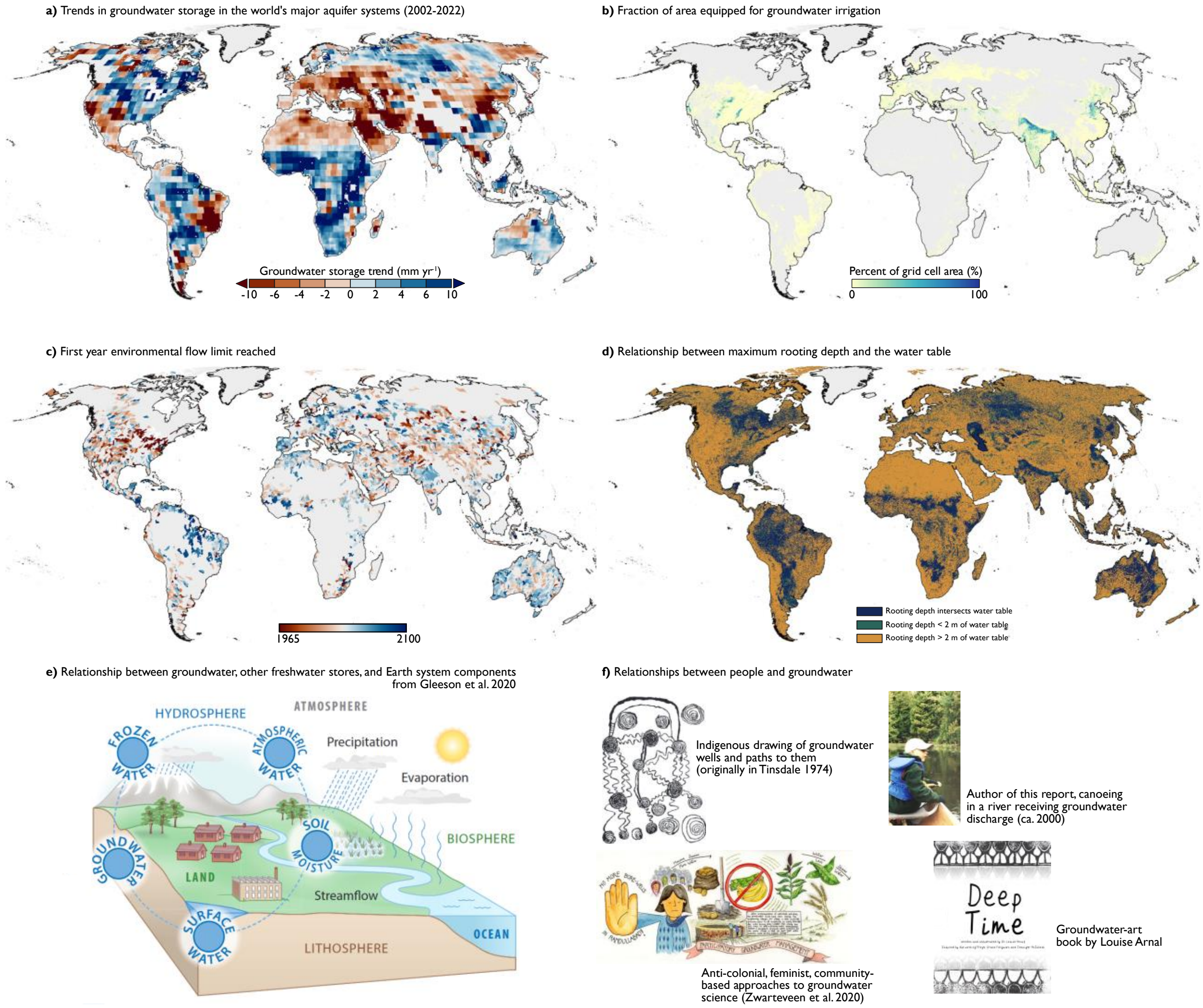
Global action is evidently needed to mitigate the current extent and rate of groundwater depletion and its impacts. Groundwater is a uniquely slow moving natural resource and inaction on the groundwater crisis means that impacts of groundwater depletion will persist over decades, centuries, and longer if they continue to be unaddressed. Thus, there is urgency to act on the global groundwater crisis, as is reflected in recent review articles and calls to action by the global groundwater research community[25–27].

It remains unclear, however, what this global action should entail to address the global groundwater crisis. In recent calls to action, recommendations include forming regional centers to generate solutions, improving the representation of groundwater processes and functions in global sustainability initiatives such as the United Nations sustainable development goals (SDGs), and convening global summits to continue elevating the profile of groundwater in policy dialogues. Yet, these calls to action all situate at an overarching planning level, and do not focus on the implementation-side of actions needed to set groundwater systems on sustainable pathways. What is clear, however, is that there are no one-size-fits-all solutions to addressing the groundwater crisis. As echoed throughout the sustainability science literature, locally-attuned, contextually-appropriate strategies are necessary for successful sustainability transformations[28]. Refining how groundwater resources and systems should be approached from this perspective is an important area for development in the groundwater sustainability dialogue.

One conceptual approach to facilitate this growth and refinement is considering the groundwater sustainability problem space through the lens of a social-ecological system. Social-ecological systems are coupled social and biophysical systems that are intertwined such that understanding system behaviour(s) requires understanding of both biophysical and social system processes and interactions[29]. The social-ecological system perspective and its associated frameworks are widely used across sustainability science[30], particularly for the study of common pool resources. There is increasing recognition of groundwater as a resource embedded in social-ecological systems[31] yet such conceptual foundations have yet to be applied in the form of data driven global analyses.

Figure 1 (shown on next page): An overview of the breadth of the global groundwater problem space.

(a) Trends in groundwater storage for the world's major aquifers [32]*. (b) Fraction of area equipped for irrigation from groundwater* [33]*; a representation of agricultural dependence on groundwater. (c) The estimated year in which environmental flow transgressions will occur due to groundwater pumping.*[24] *(d) The relationship between the water table and plat rooting depth* [2,6]*. (c) and (d) combine to represent ecological dependencies on groundwater. (e) A summary of groundwater functions within the Earth system, from Gleeson et al.* [7] *. (f) A number of examples demonstrating socio-cultural relationships and functions of groundwater* [4,34] *– as annotated in the panel.*



One approach to both (1) apply this developing paradigm of groundwater in social-ecological systems and (2) develop contextually appropriate problem and solution formulations for groundwater sustainability topics is found in systems archetyping. Systems archetyping is a broad group of methodologies used to categorize complex human-environmental systems and their interactions into a finite set of reoccurring interactions and system behavior types. Systems archetyping has been performed to date, for example, for individual river basins in western Africa[35], for vulnerability patterns in global drylands[36], in food system regimes[37], and in global land systems[38]. There are special issues on social-ecological archetyping[39] and research groups dedicated to system archetyping (https://glp.earth/how-we-work/working-groups/archetype-analysis-sustainability-and-land-governance-research). All together these indicate that the field is an emerging discipline within social-ecological systems research. However, such archetyping has yet to be applied to global groundwater systems.

There are many existing typologies of global groundwater systems. These include the World Hydrogeological Maps[40] (WHYMAP), Groundwater provinces[41], and other study-specific approaches to classify global groundwater systems based on multidimensional similarity[42]. However, all of these approaches focus exclusively on the physiographic aspects of groundwater. Similarly, there are myriad studies that focus on suitability analyses of various approaches and technologies that can help facilitate more sustainable groundwater futures[43] but these, too, integrate only physiographic constraints and criteria. While these studies are no doubt useful, there is dissonance between the interwoven human and environmental systems that typify the groundwater crises (as described above) and the strict disciplinarity of these classifications and evaluations. We argue that creating a typology of groundwater systems as they are embedded within social-ecological systems could prove very useful in facilitating and implementing the described calls to action listed above.

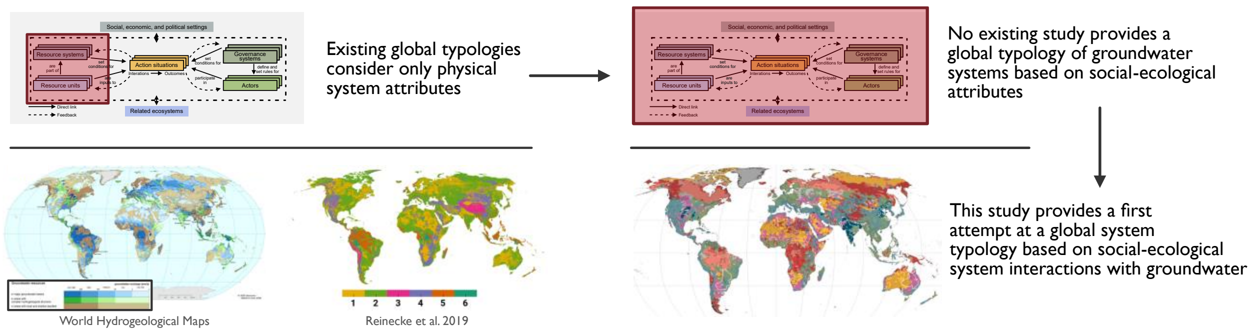


Figure 2: Existing typologies of global groundwater systems, their limited coverage of social-ecological dimensions of groundwater, and the more comprehensive social-ecological scope of the archetypes derived in this work.

Examples of existing global groundwater typologies shown are the World Hydrogeological Map and a clustering output from Reinecke et al.[42] *The social-ecological system diagram used to compare scopes of previous work and this work is redrawn from McGinnis and Ostrom* [44]*.*

In this study, we perform social-ecological systems archetyping for global groundwater systems for the first time. We describe our archetyping approach and philosophy briefly (in '*Archetyping to identify patterns of system interactions*'), present results of this approach (in '*Global patterns in social-ecological interactions with groundwater*') and assess the distribution of these system archetypes in the world's major aquifer systems (in '*Multiple archetypes found within world's major aquifers*'). We conclude by exploring patterns in current and future pressures facing archetypes (in '*Pressures facing archetypes*'). As this remains a work in progress, we list limitations of the presented work and also list remaining modifications and next steps for the project (in the Supplementary Information).

Archetyping to identify patterns of system interactions

The purpose of this archetyping analysis is to explore, group, and visualize the heterogeneity of global groundwater systems in the context that behave as social-ecological systems. To do so requires a conceptual foundation that expands beyond the physical attributes and processes of groundwater that groundwater hydrologists are familiar working with, and into interdisciplinary spaces and data that reflect how groundwater interacts with social, ecological, and Earth systems. Returning to our leading paragraph of this report, how a problem is conceptualized a central determinant in how it is addressed. This archetyping work aims to present a first attempt at redefining the global groundwater problem space through a social-ecological systems perspective.

We base our archetypes on a perceptual model of a groundwater-connected landscape at a regional scale (base of Figure 3). This perceptual model reflects a general understanding of the complex system interactions that occur between groundwater and other social-ecological system elements as represented in the schematic. However, this complex perception of the landscape requires simplification and specification in order to represent specific interactions with data and develop a data-driven archetyping analysis. From the complex, perceived landscape, we identify seven core interactions between groundwater and social-ecological system processes or attributes. These are listed below and are shown in the middle row of Figure 3. The social (S; n = 4) or biophysical (B; n = 3) nature of these interactions is also indicated below, which demonstrates the range of these indicators across the social-ecological data space.

1. Groundwater-climate interactions (B),
2. Groundwater-surface water interactions (B),
3. Groundwater-wetland interactions (B),
4. Equity of human access to groundwater (S),
5. Agricultural dependence on groundwater (S)
6. Economic value of groundwater (S), and
7. Level of integrated water resources management practices (S).

The social-ecological system perspective argues that this enumerated conceptualization of the groundwater sustainability problem space is more robust and comprehensive than approaches that consider groundwater as an isolated resource. While this list is far from comprehensive, it provides a robust overview of the dominant connections with groundwater in social-ecological systems. The selection of these interactions was also unavoidably driven by current data availability. In our archetyping analysis, each interaction listed above is represented by an existing global data set (shown in the top row in Figure 3). These seven data sources form the informational foundation that drive our ultimate archetype results. The data sets selected and used to represent these interactions are described in the Supplementary Information and in Supplementary Table 1.

As one moves upwards in Figure 3, the complexity of the system is reduced and the nuanced, realistic perceptual model is increasingly abstracted. While this is true and one outcome of this system simplification is that it reduces the ability to reflect local, nuanced social-ecological complexities, it simultaneously is necessary to guide insights and action at regional to global scales (represented on the right side of Figure 3).

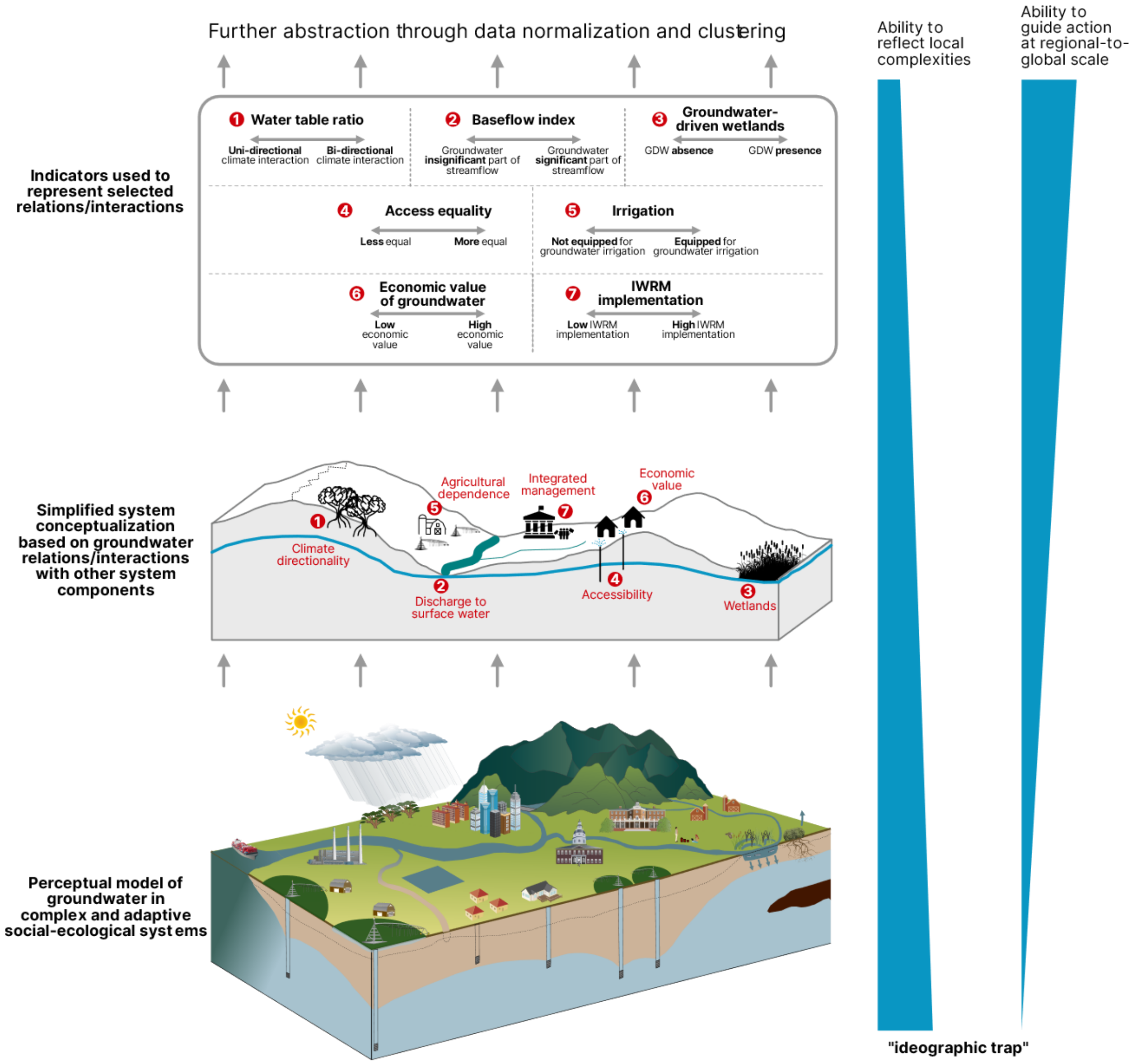
A great benefit of social-ecological systems archetyping is that it aims to avoid the two opposing traps of overspecification and overgeneralization in problem definition[45]. The ideographic trap (overspecification) is a planning trap that occurs when one assumes the normative position that every system is sufficiently unique to warrant individualized treatment, planning, and management. This, of course, is impracticable at the global scale as human and technical resources are insufficient to meet with such specific needs. Alternatively, there is the nomothetic trap (overgeneralization) that can be understood to represent simplification in problem definition and solution generation processes. Silver-bullet solutions, or one-size-fits-all paradigms are emblematic of the nomothetic trap, which are frequently associated with fixes that backfire and other unintended consequences of solutions and other implemented actions due to a lack of consideration for local context. One could understand framings of the 'global groundwater crisis' that omit discussions on regional variation in drivers and impacts of groundwater depletion to fall into this nomothetic trap. The goal of archetyping approaches is to find the 'right' level of aggregation and system abstraction that balances these two extremes and avoids both traps. In doing so, the archetyping paradigm is based on the hypothesis that regional-scale system types can enable resources and planning to be used and implemented more efficiently and effectively.

Once all data sources are identified, the procedure to derive recurring patterns in the data space becomes a clustering exercise. How data were accessed and preprocessed, the clustering methods that were used, and how final archetypes were assigned are all described in the Methods.

This archetyping work deviates from conventional social-ecological archetyping methods in that it explicitly focuses on representing social-ecological system interactions rather than attributes. Conventional archetyping studies use more isolated forms of data (e.g., such as economic indicators like GDP, or governance indicators such as governance effectiveness) and use the archetyping analysis to derive recurring relationships between these attributes. Conversely, our approach intentionally seeks to use data that represent system interactions, and to use this relationally embedded data to derive system archetypes. In this way, we do not use the archetyping analysis as a tool to understand system interactions but rather to as a tool to look for patterns in and among these interactions. This approach necessarily is predicated on, and benefits from a rich and emerging literature that has documented these interactions with groundwater. In this regard, the approach taken in this study can be understood as an integration of sorts of social-ecological system archetyping and ecosystem service bundling [e.g. 46] methods. Thus, this work can be alternatively interpreted as an effort to identify recurring 'bundles' of groundwater interactions with social-ecological systems through conventional methods used to derive social-ecological system archetypes.

Figure 3 (shown on next page): Abstracting complex groundwater-connected environments into a set of interaction indicators to derive archetypes.

We begin at the bottom of this figure with a perceptual understanding and mental model of groundwater interactions in complex social-ecological systems. From this perceptual model, seven social-ecological system interactions with groundwater are identified which are then represented by available global data sets.



Global patterns in social-ecological interactions with groundwater

We identify seven global archetypes of groundwater interactions in social-ecological systems (Figure 4). These archetypes are diverse in composition and visualize the heterogeneity of system interactions of groundwater systems at the global scale for the first time. The archetypes are derived using the spatial template of basins (see Methods). Though all archetypes are not found on every continent, each archetype has a wide geographic range. We observe spatial 'clusters' of archetypes, such as groupings of archetype 6 throughout the western USA, archetype 7 across northern India, archetype 2 in central Africa, archetype 5 across northern Europe, and archetype 3 across the Amazon and northern Canada. This global patchwork of seven archetypes both affirms and contradicts Tobler's First Law of Geography, that "near things are more related than distant things". The regional "patchiness" of archetypes reflects that groundwater interactions with social-ecological systems are generally consistent across sub-continental regions but dissipate and transition into other archetypes as distance increases. However, the same archetypes are also found in distant locations (such as archetype 2 across central Africa, northwestern Canada, and eastern Russia; archetype 4 across central Australia, the Sahara, and the Namib desert; and so on). See the Discussion for an elaboration on what this can imply for groundwater science, governance, and management approaches.

As the derivation of archetypes is methodologically a data clustering exercise, the selection of the number of clusters to derive is a core methodological step. However there is no consensus on the best approach or metric to use to guide selection of cluster numbers. Following other social-ecological archetyping studies, we were guided based on the consensus recommendation of clusters across a set of metrics used for cluster number selection. Regardless, the selection of the number of clusters is central to the results of this work and the global map of archetypes would necessarily appear different if six or eight clusters, for instance, were derived rather than seven. To address this, we compared archetype results for alternative numbers of archetypes and qualitatively compared these maps (results not shown). In our judgement, the variation between archetypes remained useful and differences between archetype compositions (i.e., the distributions of the underlying data) were easily interpretable as clusters increased to seven. Yet, once the number of archetypes increased beyond seven, we had a more difficulty explaining differences and developing unique and characteristic narratives of the archetypes. As this qualitative outcome matched the outcome of the data-driven consensus of metrics suggesting to use seven cluster centers, we proceeded using this value.

To improve confidence in archetype results, we used three clustering techniques rather than only a single algorithm. Though there is no consensus on the best clustering algorithm, the common approach in the archetyping literature remains to use only a single method per study. We perceive this approach to have pitfalls that include rendering clustering results to appear as the product of a series of relatively subjective methodological choices if not outright arbitrary. We sought to improve the robustness of such data-driven social-ecological archetyping by using three clustering algorithms and assigning final cluster membership based on the level of agreement across the three algorithms (see Methods). This approach enables more robust cluster assignment (with less algorithmically biased outcomes) and creates the opportunity to provide a heuristic measure of cluster membership confidence. We believe this benefit adds transparency to a data-driven methodology which can have subsequent benefits of greater trust and application of the results by end-users. We were able to compare cluster outputs across different methods by adapting the methodology of Sietz et al.[36] to create "comparable maps" (see Methods).

Following calls for greater transparency and embracing of uncertainty in hydrology research [47], we decided to *not* assign archetype membership to basins that had no agreement of cluster assignment across the three clustering algorithms. In doing this, we were able to represent the basins that occupy the liminal, boundary spaces between archetypes. By engaging more seriously with the uncertainties of our methodology, we hope the effect will be to instill greater confidence in using the archetype results that did receive archetype memberships.

We provide a brief description of each archetype in Table 1. These descriptions are summaries of the characteristic archetype radar plots in Figure 4b. For narrative building purposes, characteristic archetype interactions are compared to the global median interaction value across basins with available data. However, these descriptions need to be caveated carefully. The archetypes do not identify where interactions do or do not exist, or where interactions do or do not matter. Rather, the archetypes differ from one another based on recurring patterns in their interactions with social-ecological systems. This understanding of the archetypes hypothesis that similar system interactions (as represented by the multidimensional data space) are due to similar system structures at the basin scale. This archetyping approach performed here is thus not intended to facilitate trade-off analysis between functions in archetypes, but rather is to be used to generate regional networks for solution generation based on general system similarity.

As this work remains in progress, we opt to not provide names for each archetype. However, note that archetype names will be provided prior to this work being submitted for review, and indeed archetype naming will be an important component of framing and sharing of this work to broader audiences. See the Supplementary Material for a list of remaining improvements to make on this project.

Table 1: Description of archetypes, underlined description components represent distinctive interactions of each archetype.

Archetype	**Description** (summarized from Figure 4b)
1	Moderate-high inequality, moderate management, bidirectional climate interactions, low baseflow and wetland density, little agricultural dependence on groundwater, and high net present value of groundwater.
2	Moderate-low inequality, moderate management, unidirectional climate interaction, moderate-high baseflow, moderate wetland density, low agricultural dependence, and low net present value. Uniformly low interactions levels.
3	Moderate-high inequality, moderate management, presence of bi-directional climate interactions, high baseflow, high wetland density, low agricultural dependence, and low economic value.
4	Low inequality, moderate management, high climate interactions, low baseflow, low wetland density, low agricultural dependence, and low economic value. Low interaction levels across all indicators except for climate.
5	Moderate-high inequality, moderate-high management, moderate-low climate interactions, moderate-high baseflow, moderate-high wetland density, low agricultural dependence, high economic value.
6	High inequality, moderate management, unidirectional climate interaction, moderate baseflow, moderate-low wetland density, low agricultural dependence, high economic value.
7	Moderate inequality, moderate management, low climate interactions, moderate baseflow, moderate-high wetland density, high agricultural dependence, and high economic value.

From these summaries of characteristic archetype interactions, we can further group the seven archetypes into four more general parent categories: (1) generally low interaction levels with social-ecological systems and processes, (2) high economic value but low current agricultural dependence, (3) a biophysical outlier, and (4) an agricultural outlier. These four parent categories are visually summarized in Figure 7a. The minimal influence of inequality and management levels in driving archetype membership is evident in the archetype summaries above. To this end, we have performed a variable importance analysis to compare the role of the seven interactions in driving archetype results (see Supplementary Information).

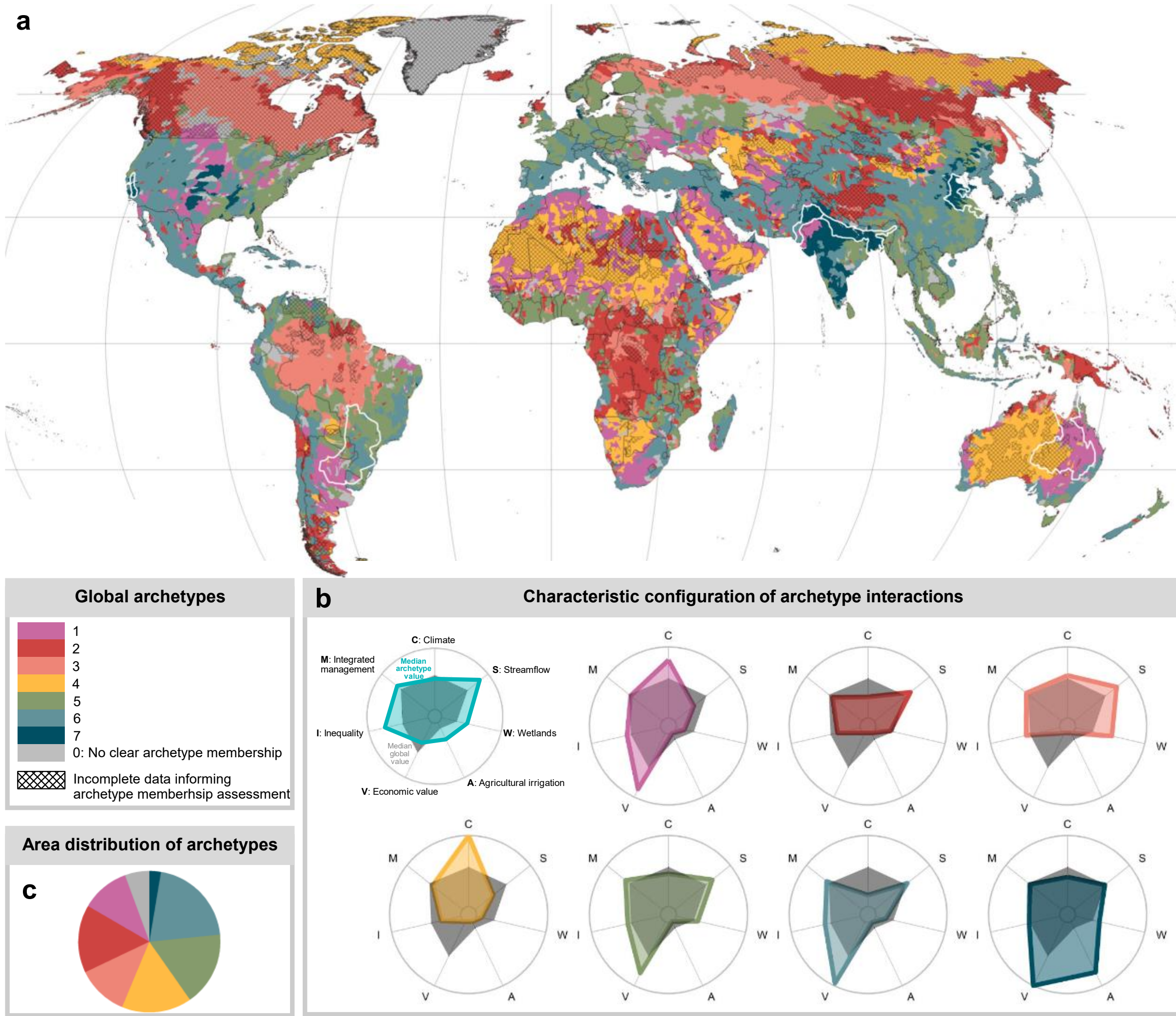


Figure 4: Global groundwater-connected system archetypes.

(a) Global map of archetypes. (b) Radial plots showing characteristic functional patters of each archetype. In each plot, the global median value for each function is shown in grey. (c) Area distribution of archetypes.

Multiple archetypes found within world's major aquifer systems

As groundwater systems are often globally approached and analyzed using the spatial 'framework' of the WHYMAP major aquifer systems, it is instructive to compare our archetyping results to these aquifer systems. Importantly, these two classification systems are derived entirely independently with no overlap in input data or methodology. Indeed, the two systems represent two alternative and parallel perspectives on global groundwater resources: the WHYMAP aquifers from a hydrogeologically-centric perspective, and the archetypes presented here from a social-ecological systems perspective.

When we compare the two maps (Figure 5), we find significant heterogeneity in archetypes across WHYMAP aquifers. For instance, archetypes 1, 2, 3, 4, 5, and 6 are all found in the Guarani Aquifer System (no.1 in Figure 5). For the Nubian Aquifer System (no.8), we find archetypes 2, 3, 4, and 6. Some aquifers are more homogeneous, such as California Central Valley (no.5), with only archetypes 6 and 7, or the Paris Basin (no.30) where only archetypes 5 and 6 are found. These comparisons show that while each WHYMAP aquifer is a contiguous groundwater body, these groundwater bodies are interacting with their connected social-ecological systems in complex, heterogeneous ways. Thus, this suggests that the scale of WHYMAP aquifers is an insufficient level for solution generation if approaches are to be attuned to social-ecological system dynamics. Thus, this comparison suggests that each of these major aquifers require multiple, locally attuned strategies implemented across their domains.

This mapping provides a needed visualization of the system heterogeneity of groundwater interactions which is yet unexplored or synthesized in the literature. We believe this form of analysis and visualization can offer a compelling basis for changing mental models in the groundwater research community to embrace complex social-ecological system problem conceptualizations by providing a visual impression of how solutions need to be derived and contextualized differently from one another (e.g., approaches should differ between the northern and southern Indus Basin on the basis that these regions are characterized by two different social-ecological archetypes). This mapping simultaneously shows the potential for regional and global solution networks based on archetype membership (e.g., solution sharing could occur between the northern Ogallala Aquifer and the southern Guarani Aquifer System on the basis they share a social-ecological archetype).

Pressures facing archetypes

We conclude by performing post-hoc analysis on the distribution of a set of pressures facing global groundwater archetypes. We do so to provide a more comprehensive system assessment that (1) describes our understanding of groundwater systems (i.e., the archetypes), and (2) provides an outlook for pressures facing each archetype. Together, this system assessment and appraisal of challenges can offer a great resource to build data driven narratives around groundwater systems and guide action at the global scale.

The four pressures we considered were: groundwater storage trends, potential for agricultural expansion, likelihood for hydropolitical interactions, and population growth factor projected to the year 2050. We plan to include additional pressures in the final version of this work. We include these four pressures as they span physical, food security, political, and water security aspects of groundwater sustainability. Thus, these pressures themselves are social-ecological in composition to match the social-ecological approach we took to archetyping.

Summarizing these pressures across archetypes (Figure 6) allows for considerations that anticipate the functional impacts of the pressures based on the functions identifiable through the archetyping analysis. For instance, the wetting (increasing groundwater storage) trends in archetype 2 (shown in Figure 6a) has the potential to raise the water table and lead to increasing climate interactions, increase the density of groundwater-driven wetlands, and increase baseflow. Under these circumstances, we could anticipate how some basins in archetype 2 could move into the domain of archetype 3 and become more ecologically significant basins requiring more integrated management approaches. Alternatively, we can observe the elevated potential for cropland expansion in archetypes 5 and 6 (shown in Figure 6c) and compare to the currently low levels of agricultural dependence on groundwater but high potential economic value of groundwater in these archetypes. In these archetypes, this archetyping and potential pressures analysis suggests we can expect the greatest expansion of irrigated agriculture in these basins. Yet, for example, when we compare to the moderate-high density of groundwater-driven wetlands in archetype 5, this analysis also brings to our attention the need for expanded groundwater irrigated agriculture in these basins to not come at the cost of degraded wetlands.

This form of narrative building can facilitate improving causal understanding in these complex systems. In this study, we are limited to the static nature of the data and thus of the archetypes. However, when temporally dynamic global data become available to track groundwater interactions, future work that considers the role of archetype configurations in both driving and responding to pressures can become possible. For the meantime, this analysis is useful in constructing causal hypotheses in complex systems that can be tested in future work.

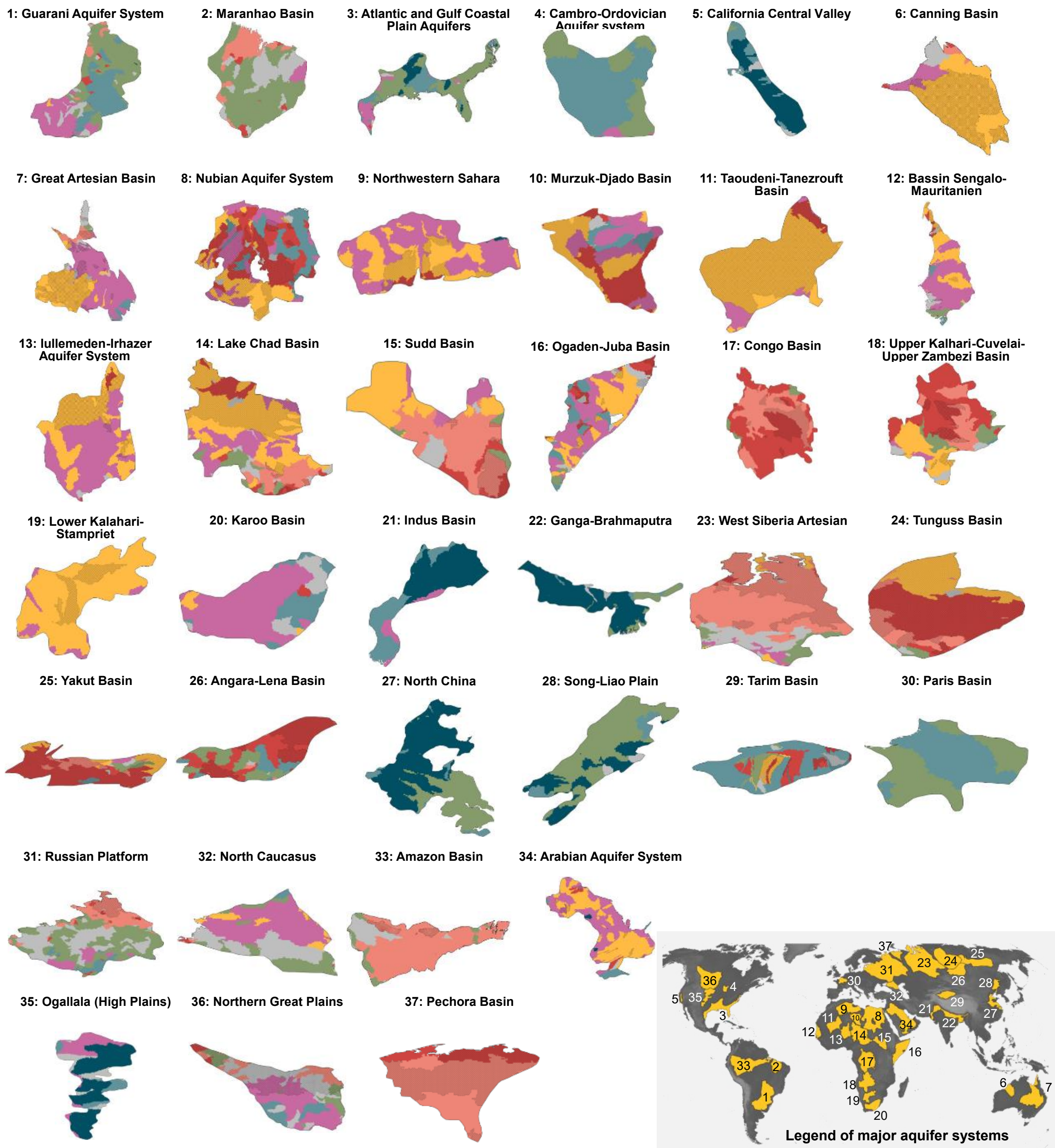
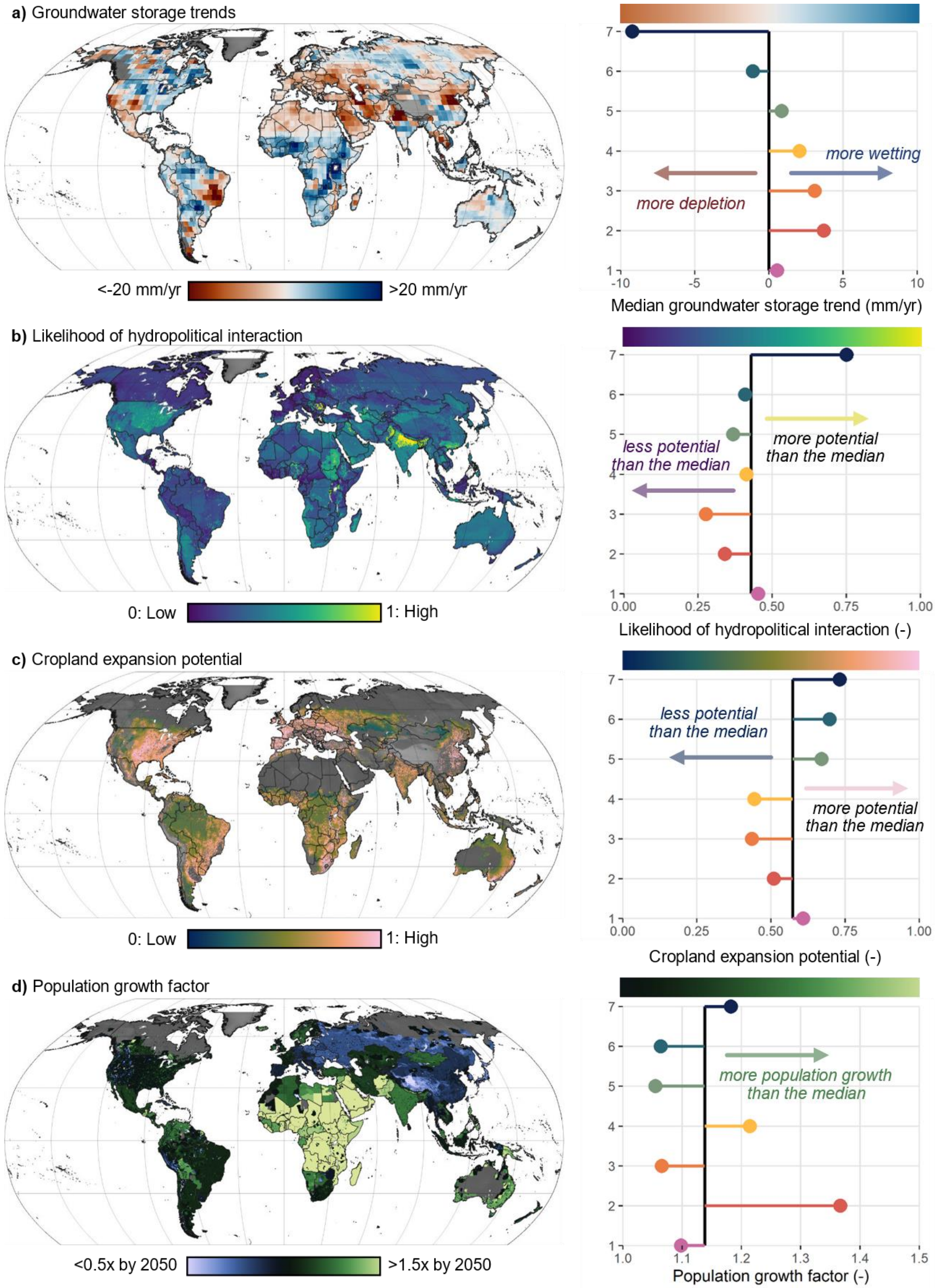


Figure 5: The distribution of groundwater system archetypes across the 37 WHYMAP major aquifers systems of the world.

Consult with the legend in Figure 4 for interpretation of archetype symbology.

Figure 6 (shown on next page): Challenges facing archetypes.

(a) Groundwater storage trends [32]*. (b) Likelihood of hydropolitical interaction*[48]*. (c) Cropland expansion potential* [49]*. (d) Population growth factor by the year 2050* [50]*. For each pressure, the median value per archetype is shown on the right. The vertical bar in each plot represents either a physically meaningful threshold (such as the difference between drying and wetting storage trends), or the median value across all archetypes.*


a) Groundwater storage trends
<-20 mm/yr
>20 mm/yr
more wetting
more depletion
Median groundwater storage trend (mm/yr)
b) Likelihood of hydropolitical interaction
0: Low
1: High
less potential than the median
more potential than the median
Likelihood of hydropolitical interaction (-)
c) Cropland expansion potential
0: Low
1: High
less potential than the median
more potential than the median
Cropland expansion potential (-)
d) Population growth factor
<0.5x by 2050
>1.5x by 2050
more population growth than the median
Population growth factor (-)

This post-hoc analysis is also useful in supporting prioritization schemes for archetypes at the global scale. When we compare the rank-order of archetypes facing each of the four pressures (Figure 7), we can build composite narratives across social-ecological pressures and develop an understanding of overall system states. For instance, we see that the agricultural outlier archetype 7 (Figure 7a) ranks among archetypes as the most depleting, with the greatest likelihood for hydropolitical interaction, the greatest potential for cropland expansion, and the third greatest rate of population growth (Figure 7b). This composite picture of the archetype allows us to build a causal hypothesis that the agricultural dependence of the archetype drives not only groundwater depletion, but a cascade of impacts, yet are anticipated to continue in the absence of mitigative policy action due to high cropland expansion potential. Further, this composite picture reflects the urgency that groundwater sustainability needs to be promoted within the archetype to limit these impacts and serves as an analysis that underscores the need to prioritize the archetype across global contexts.

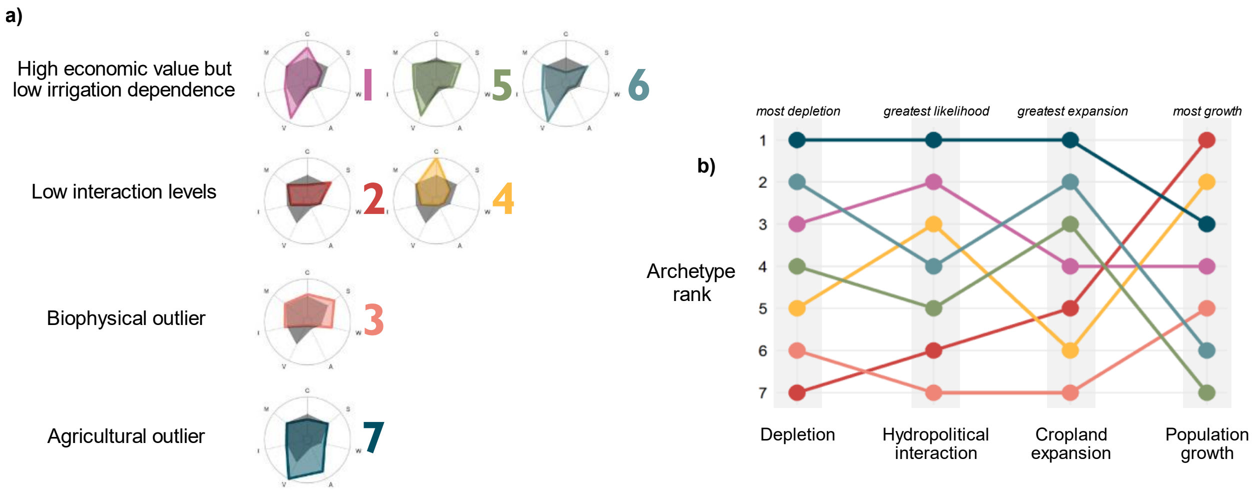


Figure 7: Rank order of challenges and opportunities facing archetypes.
(a) Parent groupings of archetypes, and (b) archetype rank order for each of the four pressures evaluated. The rank-order plot is colour-coordinated to the archetype which are likewise colour-coordinated and labelled in (a).

Discussion

The archetypes presented here offer a new, systems view of global groundwater systems. As the archetypes build on existing data sets that each represent individual interactions of groundwater with larger social, ecological, and Earth systems, this archetyping is largely an exercise in identifying recurring patterns in these multiple interactions when they are overlaid with each other. These co-occurrences of interactions become critical to understand when addressing global challenges facing groundwater resources. Though we have mentioned benefits of this approach throughout the report, we summarize three key benefits below.

1. **Promote mental models that consider groundwater connections with social-ecological systems when addressing global groundwater challenges.** This study's approach, that seeks to present a holistic set of groundwater interactions, can promote more system-wide conceptualizations

and approaches to address groundwater issues. These archetypes can introduce, reinforce, and promote thinking about groundwater in complex system among researchers and policy makers. As perceptions of the groundwater sustainability problem space are central to what science is conducted and approaches to governance and management, these groundwater system archetypes have the potential to shift underlying assumptions in groundwater researchers, practitioners, and policy makers and can generate myriad opportunities for growth and innovation solely through changes to how the challenges are perceived and a greater awareness around the degree of interconnectedness between groundwater system interactions.

2. **Support regionalized solution networks based on contextual fit rather than geographic proximity and provide a framework to quantify system differences.** Though suitability analyses are common [e.g. 43], they are dominated by physiographic constraints and criteria. The archetypes presented here offer a systems-oriented alternative to such approaches. Given that solutions to complex groundwater management issues are dictated not only by physical processes but the interactions between these processes with social and ecological systems, it follows that the suitability of an approach will vary based on this more expansive, system-wide context. As the archetypes present functionally similar groundwater systems, the potential exists for the archetypes to be used to guide solution sharing within archetypes and to contextualize critical functional differences when developing solutions across archetypes. Furthermore, as the archetypes present a heterogeneous global view of the groundwater sustainability problem space, this work acts to combat one-size-fits-all paradigms and approaches in addressing global scale groundwater challenges in favor of locally attuned, contextually appropriate strategies.

3. **Generate and test hypotheses of mechanistic causality in the complex system behaviour of groundwater interactions in social-ecological systems.** The narratives we constructed through the combination of archetyping with post-hoc analysis of future pressures (in *Pressures facing archetypes*) illustrate how this work can facilitate mechanistic theories of causality in these complex social-ecological systems. As more time series data become available on global groundwater interactions with social, ecological, and Earth system processes, these causal hypotheses can be confronted and tested with observations to better understand theories of social-ecological change in global groundwater systems.

There are several more potential benefits of the archetypes. For instance, in the global hydrological modelling community, there is discussion on parameterizing human decision making and improving ecological process representation in the next generation of global models[51]. In parallel, there are discussions about "patchwork" model approaches to global model development[52]. The archetypes or the concept of archetyping applied to be specifically fit for this purpose, could serve as a spatial template to parameterize similar social and ecological processes in these models.

Lastly, there is the potential for this archetyping work to be applied to other components of the hydrosphere such as green water or surface water. Together, an ensemble of such archetypes can provide a more complete picture of social-ecological interactions with freshwater that better aligns with conjunctive water management and "One Water" paradigms[53].

Though much work remains in refining the archetypes presented here, we perceive them to be an exciting and expansive conceptual advance in the global groundwater literature. The archetypes help us to see that there is not one single global groundwater crisis, but rather several regional, co-occurring, unique but interacting groundwater crises.

Methods

All input data used in this study are obtained from published data sets or methods and are described in Supplementary Table 1. Input data used to drive archetype analysis are plotted in Supplementary Figure 2. The scripts generated to conduct analysis are made available online at https://github.com/XanderHuggins/gcs-archetypes. All scripts are written in the R coding language and R packages necessary for this study are all found in the GitHub repository in the code `setup/packages.R`. The core packages used in this analysis are: terra[54], sf[55], NbClust[56], kohonen[57], ggplot2[58] and tmap[59].

Our archetypes use level 6 HydroBASINS as the base spatial template. While it may appear counterintuitive for a groundwater archetyping analysis to use surface water basins as the spatial template, there are several considerations that contributed to this core methodological decision. First, surface basins are increasingly used as the spatial template for freshwater-focused social-ecological systems analysis [35,60,61] and thus there is existing literature to support this methodological decision. Second, as there is no global mapping of groundwater system divides or groundwater basins, it would require additional geoprocessing and analysis to derive a groundwater equivalent. Third, many of the ecological functions and interactions of groundwater occur in the phreatic aquifer and the water table is often a subdued replica of the topography[62]. Thus, if we were to use a water table-derived groundwater-basin derivation, it would not deviate significantly from the already-available surface watershed template we used. We selected the level 6 HydroBASINS set (that are offered across levels 1 – 12) as it is a commonly used HydroBASINS level in other integrated global hydrological assessments, such as the World Resources Institute Aqueduct platform.

Using this basin template, we first pre-harmonized all input data to a 5 arc-minute grid and subsequently summarized all input data to the HydroBASIN level 6 scale. For all intensive data sets (e.g., baseflow), we took an area-weighted average across each basin; while for all extensive data sets (e.g., net present value of groundwater), we calculated within-basin sums. As clustering requires all data be normalized to a common range, we subsequently normalized basin data following steps outlined in Supplementary Table 2. The data we selected as input data for clustering were focused on system interactions and not the outcomes of these interactions. That is, we sought to generate archetypes of system functions and interactions and not archetypes of system outcomes. This conceptual approach explains why outcomes such as groundwater storage trends and hydropolitical interaction are used in post-hoc analysis of the archetypes rather than as input data for archetyping.

We applied three alternative clustering algorithms to our input data. These are: k-means clustering, a self-organizing map (SOM), and partitioning around medoids (PAM) clustering. We apply three alternative clustering methods to reduce bias of outcomes to the strengths and limitations of any single clustering algorithm and in doing so improve the robustness of the cluster analysis. Though there are myriad other clustering algorithms not applied here, we select these three algorithms as they are frequently used in the archetyping literature. Though we will not detail all methodological aspects of the clustering analysis here, the code to perform clustering is visible on the GitHub repository at `analysis/a2-archetyping.R`.

As all three clustering algorithms were executed to generate the same number of clusters, we were able to compare the similarity of cluster outputs to gain insight on cluster consistency between methods. As the allocation of cluster ID is arbitrary, we reclassified cluster outputs between methods based on Euclidean distance cluster centroid proximity. To do this, we calculated the centroid of each cluster in the multi-dimensional input data space for each clustering method. We then calculated the Euclidean distance between cluster centroids between clustering method 1 (k-means) and clustering method 2 (SOM). We then reclassified the cluster IDs from method 2 to cluster IDs from method 1 based on the nearest cluster in Euclidean space. We then repeated this between clustering method 1 and clustering method 3 (PAM).

This created a set of three comparable cluster maps all aligned to the arbitrary cluster ID numbering system assigned to clustering method 1. Subsequently, for each basin, we were able to assess the degree of agreement across clustering methods. For each basin, we declared a basin as a member of a specific cluster if at least two of the three methods assigned the basin to the same cluster. However, if all three methods assigned the basin to a different cluster, we did not assign the basin to a cluster in our final archetype result map. These steps we describe here are visible in Supplementary Figure 2, and the associated function written to perform this is visible on the GitHub repository at **`setup/cluster_reconcile.R`**.

The full set of input data (7/7) are available for approximately 70% of basins globally (Supplementary Figure 1). However, for the remaining 30% of basins, ≥ 1/7 of the input data sets are not available. This presents a challenge to the clustering methods, we the algorithms we used to not accommodate **`NA`** data entries. To resolve this, we conducted our clustering analysis for only the set of basins globally with complete data coverage. For the remaining basins, we assigned cluster membership based on the location of the basin in the *available* multi-dimensional data space in reference to the basins with complete data coverage that received an archetype membership based through clustering analysis. For each basin with incomplete input data availability, we computed the Euclidean distance of its input data to all basins with complete input data for all attributes that are mutually available. We then select the three nearest basins in Euclidean space to this basin with incomplete data and pull the archetype they were assigned. We then evaluate the consistency of these nearest-neighbor basin archetypes. Similarly to how we assigned archetypes across three clustering algorithms, if ≥ 2/3 of the nearest neighbor basins were assigned to the same archetype we then assigned the basin with missing data to this archetype. However, if all three of the nearest neighbor basins were assigned to different archetypes, we had no subsequent way of assignment membership to one archetype over the other and thus the basin was assigned to the '0: archetype not-assigned' class of basins. We then repeated this process for all basins with incomplete data coverage. This assignment process is visualized in Supplementary Figure 4.

As groundwater system archetypes are not derived at any scale or for any jurisdiction in the literature, we did not have a data set with which to perform validation of the archetype results. In place of a 'ground-truth' to compare our archetype results to, we developed a set of 'over-specified' basin classes and compared the distribution of these over-specific basins to our archetype results. This process is described in the Supplementary Text in the section '*Simplification of system heterogeneity*'.

Lastly, we performed a variable importance analysis to gain insight into the role of individual input data sets in driving archetype membership. In effect, the archetyping we performed generated a set of delineated zones in multi-dimensional space that characterized each archetype. That is, for each input indicator, there is a range of values that are found within each archetype and similarly a range of values that is not represented in the archetype. To generate an estimate of variable importance, we randomly imputed each input data set individually for each basin and calculated the frequency at which the randomly imputed data would move the basin outside of the data space of its assigned archetype. If the imputed data did not have a large effect on the archetype membership, we infer from this that the input variable does not play a significant role in driving archetype membership (as randomly introduced data does not affect the clustering outcome). Conversely, if the imputed input data frequently alters archetype membership, we infer that this input variable was important in the clustering process. The results of this variable importance analysis are shown in Supplementary Figure 5.

Supplementary information

Supplementary text

Simplification of system heterogeneity

We prefaced our archetypes with a discussion of the ideographic trap and attempting to find a balance in system aggregation and abstraction. We thus wanted to evaluate the success of this study in aggregating and simplifying the diverse multi-dimensional data space characterizing these groundwater systems. To do this, we set about archetyping if we *were* intending on producing a set of over-specified archetypes (i.e., the ideographic trap approach). This entailed classifying each of the seven interactions individually and simply creating overall system archetypes by assessing unique combinations of classes across all seven interactions. When doing so for the basins with complete interaction attribute data, we found 485 unique system types. Thus, the archetyping analysis performed here offers a 70x reduction in system types (485:7). Developing a set of seven solution networks and strategy portfolios is significantly more practicable and feasible than doing so for 485 system types.

To assess the effectiveness of this system reduction, we made the assumption that the ideographic (over-specified) set of basins is a more reflective, truer representation of system uniqueness and individuality. We then compared the distribution of these unique system types across the seven archetypes. For all unique system types that were found for >2 basins, we assumed that the modal system archetype was the 'correct' archetype for the unique system. Doing this enabled a comparison of consistency of archetype assignment across unique system types. We found that 85% of all unique system types were classified consistently with the modal unique system type – system archetype pairing. This heuristic serves to indicate that the archetyping analysis we performed retains a significant degree of the dominant variability in the underlying multi-dimensional data space but simultaneously provides critical and needed simplifications in the problem space. As these archetypes are without precedent, it is challenging to perform a validation exercise when there is no 'ground truth'. The analyses we described above serves as our place-holder validation exercise for the archetype analysis.

Preliminary sub-clustering at a nested scale

As we observed one archetype (archetype 7) to be facing the most acute challenges, we were compelled to explore the idea of generating sub-archetypes. The idea of sub-archetyping archetype 7 is further supported due to the dominant role of agricultural irrigation in separating the archetype from all others in the data space. Though this is useful to distinguish agriculturally intensive systems from non-intensive systems, we perceived this dominance of the agricultural signal to lead to a diminished ability to understand spatial patterns within this archetype in the other interactions we considered. Thus, we recursively applied our archetype analysis to exclusively archetype 7, yet omitted the agricultural data, and performed this analysis at the grid scale (5 arc-minute). This preliminary analysis is shown in Supplementary Figure 6 and represents an interesting potential further development of the social-ecological archetyping concept at multiple scales.

List of modifications and next steps

As mentioned throughout this report, this work remains an active work in progress. To illustrate the degree of modifications we intend to make, and for transparency regarding how we perceive the work will evolve, we list our intended modifications and next steps for this work below.

- Refinement of input data
  - Given the sensitivity of the water table ratio to low recharge rates, consider masking arid regions or setting these basins to the minimal water table ratio
  - Replace general baseflow data with groundwater discharge data when it becomes available
- Consider adding additional clustering methods
- Add additional pressures data in post-hoc analysis
  - E.g., agricultural intensification, biodiversity data, etc.
- Provide names to each archetype rather than the numbering system used in this report
- Provide policy recommendations per archetype
- Continue exploring sub-clustering idea for archetype 7

Supplementary Table 1: Data sources, description, justification, and summary of any data preprocessing applied.

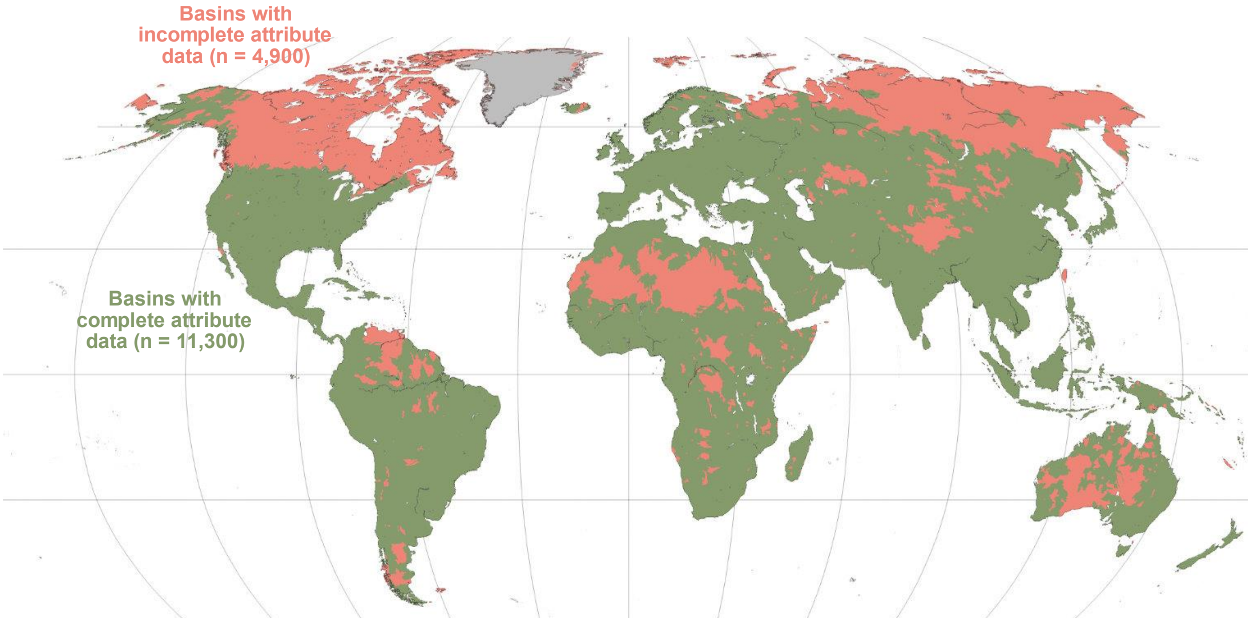
Dataset	Data source. Persistent web link. Temporal range. Spatial resolution. Resolution harmonization method. Description and justification.
Basins	**Data source:** HydroBASINS, Lehner and Grill[63] **Persistent web link:** https://www.hydrosheds.org/page/hydrobasins **Temporal range:** N/a **Spatial resolution:** Vector **Resolution harmonization method:** Rasterization at 5 arcminutes. **Description and justification:** HydroBASINS are the standard global basin discretization scheme. In this study, we used level 6 HydroBASINS.
Groundwater-climate interactions (B)	**Data source:** Water Table Ratio, Cuthbert et al.[64] **Persistent web link:** https://figshare.com/articles/dataset/Global_water_table_ratio_and_groundwater_response_time_raster_data/7393304 **Temporal range:** N/a **Spatial resolution:** 1 km **Resolution harmonization method:** Bilinear interpolation to 5 arcminutes. **Description and justification:** The water table ratio (WTR) is a measure of the relative fullness of the subsurface and that represents whether groundwater-climate interactions are unidirectional (where recharge occurs) or bidirectional (where both recharge and root water uptake and evapotranspiration occur).
Groundwater-streamflow interactions (B)	**Data source:** Baseflow from the Global Streamflow Characteristics Dataset, Beck et al.[65] **Persistent web link:** http://www.gloh2o.org/gscd/ **Temporal range:** N/a (reference) **Spatial resolution:** 0.05 decimal degrees **Resolution harmonization method:** Bilinear interpolation to 5 arcminutes. **Description and justification:** Global reference data on streamflow characteristics derived from neural network ensembles trained with global network of observed streamflow records. We use the BFI1 (Baseflow Index 1) data set from the GSCD records.
Groundwater-wetland interactions (B)	**Data source:** Groundwater-driven wetlands, Tootchi et al. [11] **Persistent web link:** https://doi.org/10.1594/PANGAEA.892657 **Temporal range:** N/a **Spatial resolution:** 15 arc-seconds (~500 m) **Resolution harmonization method:** Binary aggregation to 5 arcminutes. **Description and justification:** Global composite wetland maps that specify both routinely flooded wetlands (RFW) and groundwater-driven wetlands (GDW). To our knowledge, this is the only global dataset that explicitly maps groundwater-driven wetlands.

Groundwater access inequalities (S)	**Data source:** Night lights Gini Index, Usman Mirza et al. [66] **Persistent web link:** https://zenodo.org/record/4635734#.YxCMl7RBzdU **Temporal range:** 2010 **Spatial resolution:** 1 decimal degree **Resolution harmonization method:** Resampled to 5 arcminutes. **Description and justification:** Gini inequality coefficient estimated using spatial variation in nighttime light emitted per capita.
Agricultural irrigation dependence on groundwater (S)	**Data source:** Global maps of irrigated areas [33] **Persistent web link:** https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version/ **Temporal range:** 2005 **Spatial resolution:** 5 arcminutes. **Resolution harmonization method:** N/a **Description and justification:** Amount of area equipped for irrigation from groundwater per grid cell.
Economic value of groundwater (S)	**Data source:** Bierkens et al. [67] **Persistent web link:** https://zenodo.org/record/5576446 **Temporal range:** 2015 **Spatial resolution:** 5 arc-minute **Resolution harmonization method:** N/a **Description and justification:** Net present value of economically recoverable groundwater.
Status of management (S)	**Data source:** IWRM Data Portal [68] **Persistent web link:** http://iwrmdataportal.unepdhi.org/countrydatabase **Temporal range:** 2020 **Spatial resolution:** National jurisdictions **Resolution harmonization method:** Rasterized at 5 arcminutes. **Description and justification:** A composite indicator of "the global status and progress on SDG 6.5.1 … [the] degree of integrated water resources management implementation". The indicator is derived using 33 sub-indicators spanning four IWRM components: an enabling environment, institutions and participation, management instruments, and financing.
Groundwater storage trends (P)	**Data source:** Chandanpurkar et al. [32] **Persistent web link:** N/a – provided by author. **Temporal range:** 2002-2022 **Spatial resolution:** 0.5 decimal degrees **Resolution harmonization method:** Resampled at 5 arcminutes. **Description and justification:** Global gridded trends in groundwater storage derived from the GRACE and GRACE-FO satellite missions. GRACE data provide monthly storage anomalies of terrestrial (total) water storage. Groundwater storage trends are derived by removing estimated monthly storages of soil moisture from the land surface models: VIC and NOAH. Surface water storage anomalies are assumed to be negligible, and glaciated areas are masked from the data.

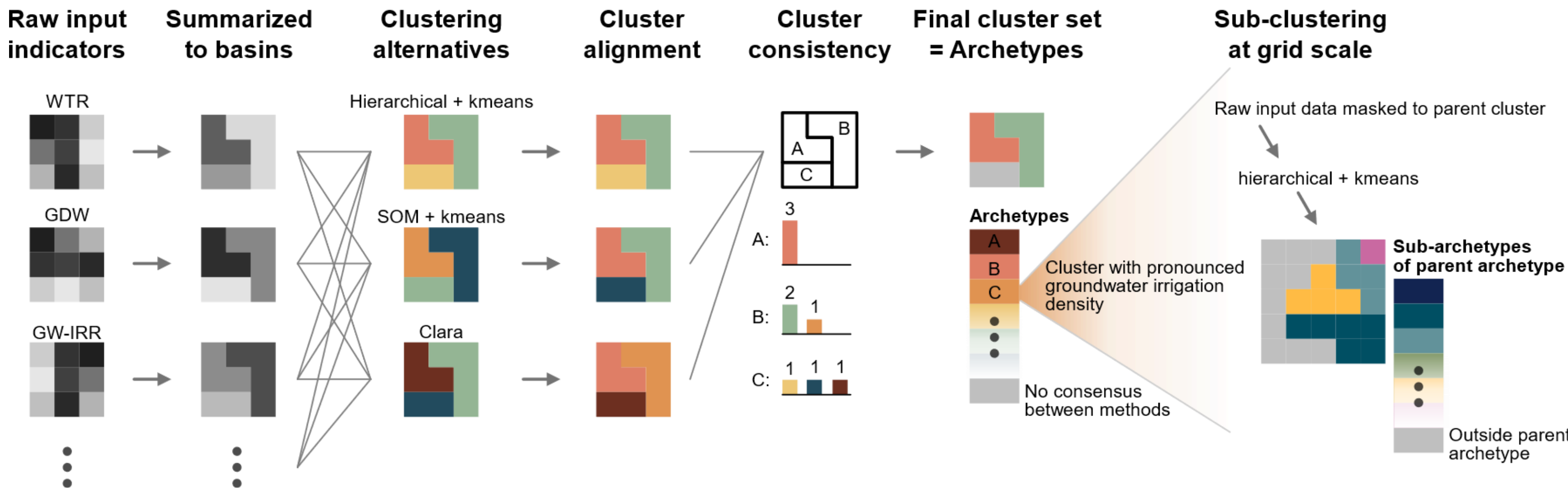
Likelihood of hydropolitical interactions (P)	**Data source:** Farinosi et al. [48] **Persistent web link:** N/a – provided by author, but visible at: https://waterscarcityatlas.org/hydro-political-interactions/ **Temporal range:** 2050 under scenario RCP 4.5 **Spatial resolution:** 0.25 decimal degree **Resolution harmonization method:** Resampled to 5 arcminutes. **Description and justification:** A random-forest derived index of potential for future transboundary hydro-political interactions.
Cropland expansion potential (P)	**Data source:** Oakleaf et al. [49] **Persistent web link:** aaa **Temporal range:** 2015 **Spatial resolution:** 30 arc-second **Resolution harmonization method:** Maximum aggregation to 5 arcminutes. **Description and justification:** Indicator of the suitability for cropland expansion.
Population growth factor (P)	**Data source:** Jones et al. [50] **Persistent web link:** https://sedac.ciesin.columbia.edu/data/set/popdynamics-1-8th-pop-base-year-projection-ssp-2000-2100-rev01 **Temporal range:** 2010-2100 **Spatial resolution:** 0.125 degree **Resolution harmonization method:** Ratio of 2050 population (SSP2) to 2010 population is bilinearly interpolated to 5 arcminutes. **Description and justification:** Gridded estimates of population for each decade under various SSP scenarios.

Supplementary Table 2: Data normalization for clustering input data.
This information is also attainable by viewing the associated code on the GitHub repository at **`setup/custom-normalizations.R`**.

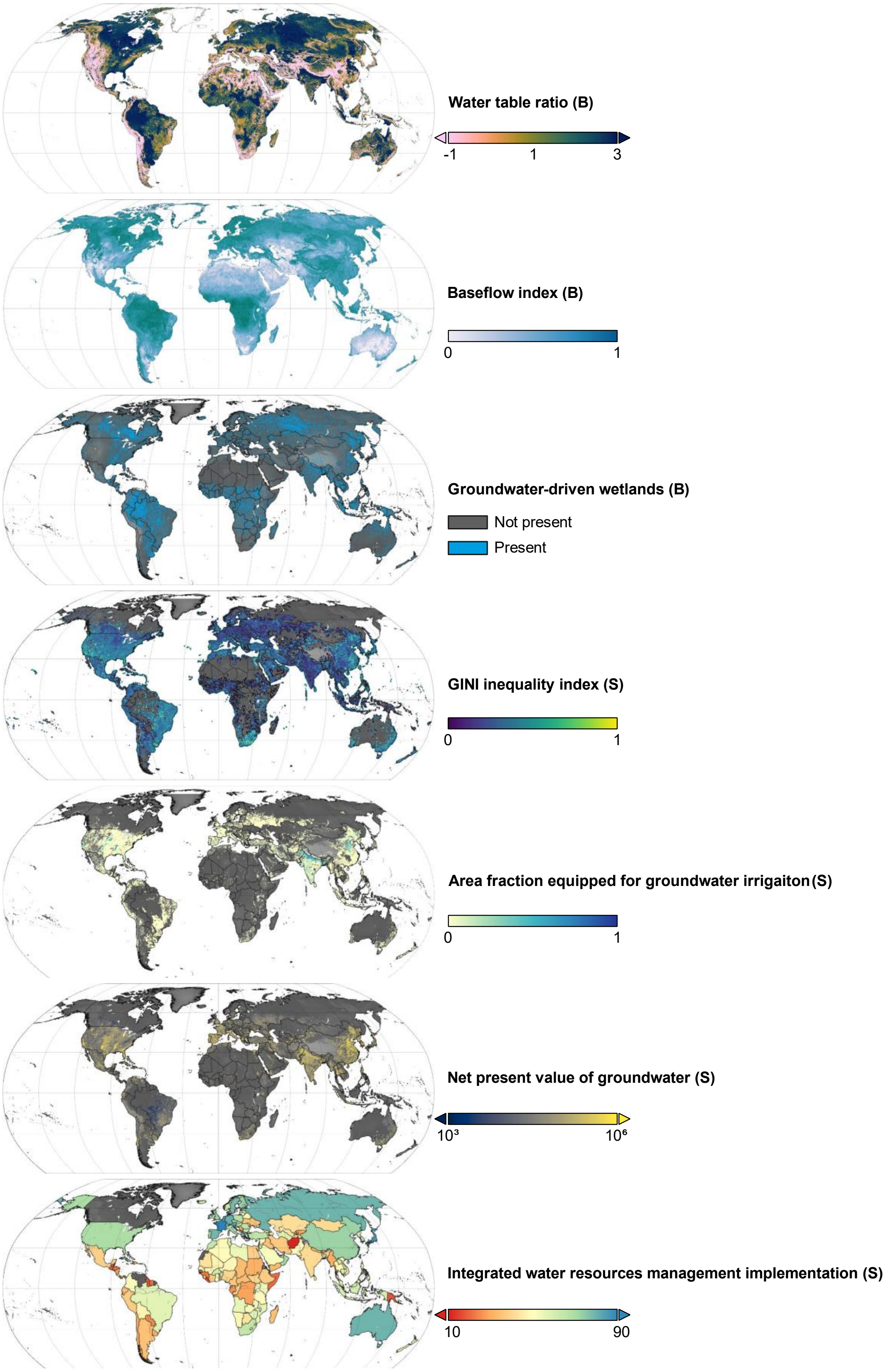
Data set	Normalization
Water table ratio Groundwater-climate interactions (B)	Log10(WTR) of ≤-1 is set to normalized scale of 0. Log10(WTR) of ≥3 is set to normalized scale of 1. Log10(WTR) is linearly scaled between these values.
Baseflow Groundwater-streamflow interactions (B)	Already ranges from [0,1], and thus does not require normalization.
Groundwater-driven wetlands Groundwater-wetland interactions (B)	We calculate the density of wetlands per basin, which ranges from [0,1] an thus does not require normalization.
Gridded Gini Index Groundwater access inequalities (S)	Already ranges from [0,1] and thus does not require normalization.
Area equipped for groundwater irrigation Agricultural irrigation dependence on groundwater (S)	Represents the area fraction of land equipped for groundwater irrigation. As this data's distribution is heavy-tailed, we set a maximum value of 20% to 1, and scale all values between 0% and 20% linearly.
Net present value of groundwater for irrigation Economic value of groundwater (S)	Log10(NPV) of ≤3 is set to normalized scale of 0. Log10(NPV) of ≥9 is set to normalized scale of 1. Log10(NPV) is linearly scaled between these values.
Level of IWRM implementation Status of management (S)	Ranges from [0,100]. After dividing by 100, the indicator ranges from [0,1] and thus does not require normalization.



Supplementary Figure 1: Data coverage across basins.

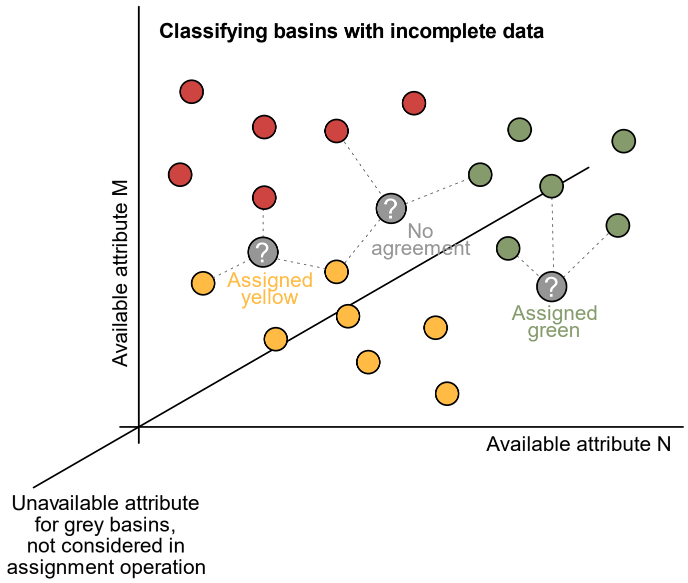


Supplementary Figure 2: Methodology overview.

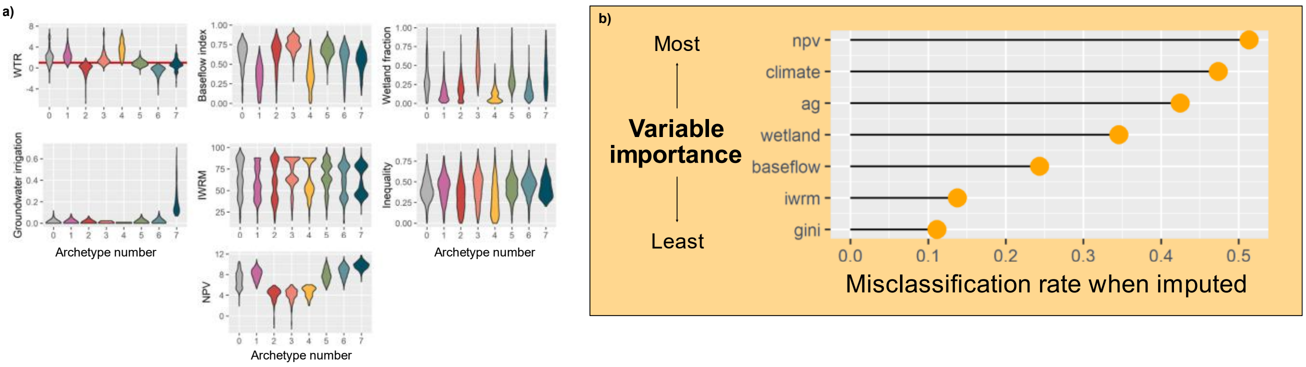


Supplementary Figure 3: Maps of input data.

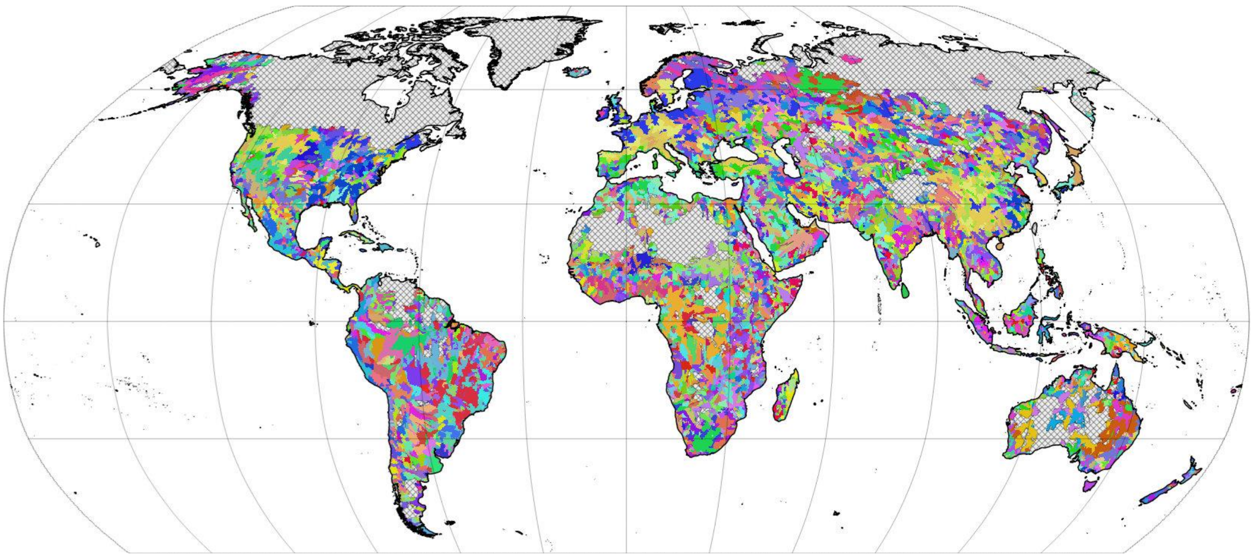
Consult Supplementary Table 1 for data sources and descriptions.



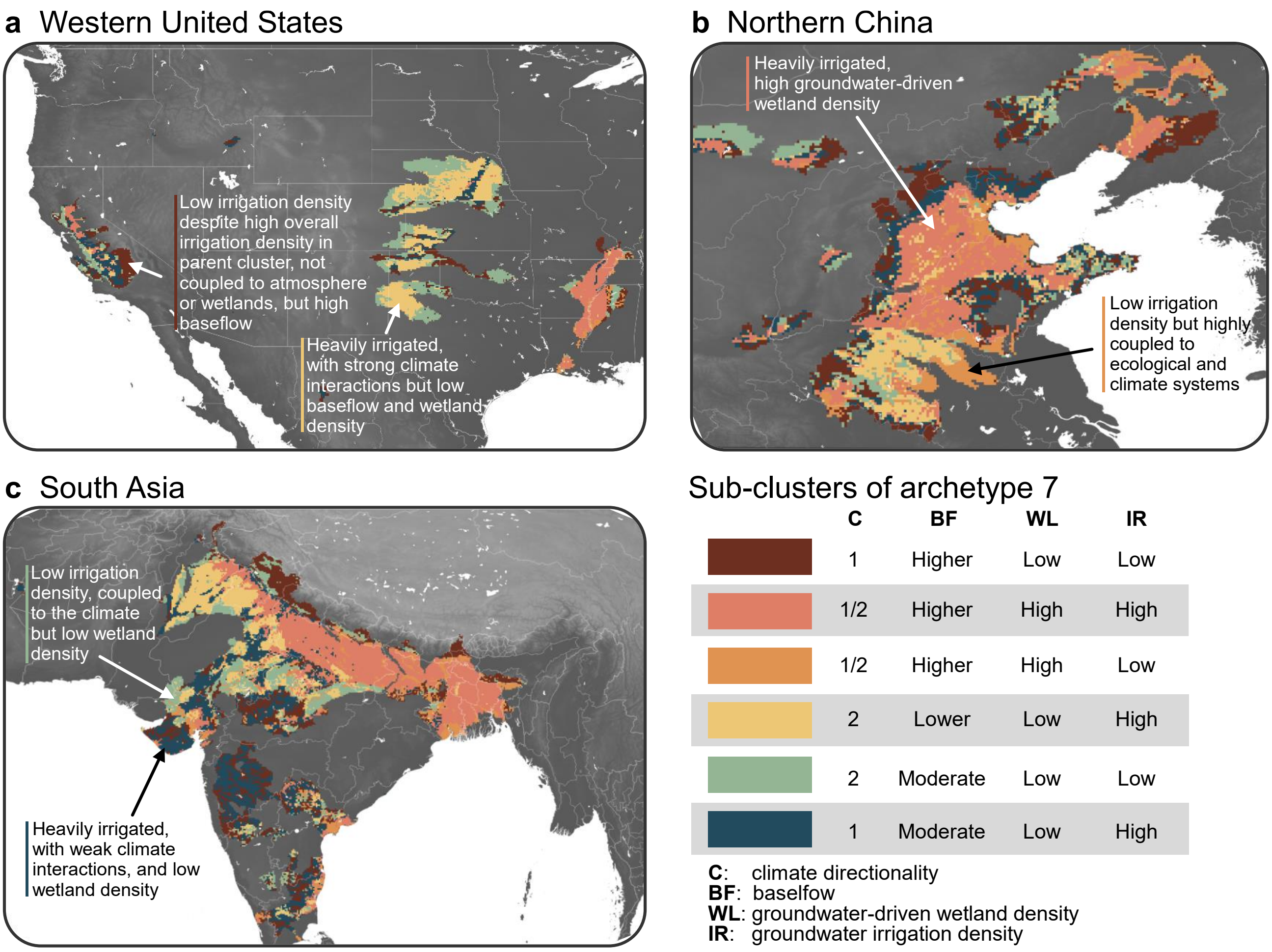
Supplementary Figure 4: Archetype assignment for basins with incomplete interaction data



Supplementary Figure 5: Variable importance of interactions in driving archetype membership



Supplementary Figure 6: Validation of groundwater archetypes through comparison to unique system types.



Supplementary Figure 7: Preliminary sub-clusters for archetype 7.

References

1. Aeschbach-Hertig, W. & Gleeson, T. Regional strategies for the accelerating global problem of groundwater depletion. *Nat. Geosci.* **5**, 853–861 (2012).
2. Fan, Y., Li, H. & Miguez-Macho, G. Global Patterns of Groundwater Table Depth. *Science* **339**, 940–943 (2013).
3. Re, V. Incorporating the social dimension into hydrogeochemical investigations for rural development: the Bir Al-Nas approach for socio-hydrogeology. *Hydrogeol. J.* **23**, 1293–1304 (2015).
4. Zwarteveen, M. *et al.* Transformations to groundwater sustainability: from individuals and pumps to communities and aquifers. *Curr. Opin. Environ. Sustain.* **49**, 88–97 (2021).
5. Cantonati, M. *et al.* Ecohydrogeology: The interdisciplinary convergence needed to improve the study and stewardship of springs and other groundwater-dependent habitats, biota, and ecosystems. *Ecol. Indic.* **110**, 105803 (2020).
6. Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B. & Otero-Casal, C. Hydrologic regulation of plant rooting depth. *Proc. Natl. Acad. Sci.* **114**, 10572–10577 (2017).
7. Gleeson, T. *et al.* The Water Planetary Boundary: Interrogation and Revision. *One Earth* **2**, 223–234 (2020).
8. Siebert, S. *et al.* Groundwater use for irrigation – a global inventory. *Hydrol. Earth Syst. Sci.* **14**, 1863–1880 (2010).
9. Taylor, R. G. *et al.* Ground water and climate change. *Nat. Clim. Change* **3**, 322–329 (2013).
10. Gleeson, T. & Richter, B. How much groundwater can we pump and protect environmental flows through time? Presumptive standards for conjunctive management of aquifers and rivers. *River Res. Appl.* **34**, 83–92 (2018).
11. Tootchi, A., Jost, A. & Ducharne, A. Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth Syst. Sci. Data* **11**, 189–220 (2019).
12. Gleeson, T., Huggins, X., Connor, R., Arrojo-Agudo, P. & Vázquez Suñé, E. Groundwater and ecosystems. In *The United Nations World Water Development Report 2022: Groundwater: making the invisible visible*. 89-100 (2022).
13. Eamus, D., Fu, B., Springer, A. E. & Stevens, L. E. Groundwater Dependent Ecosystems: Classification, Identification Techniques and Threats. in *Integrated Groundwater Management: Concepts, Approaches and Challenges* (eds. Jakeman, A. J., Barreteau, O., Hunt, R. J., Rinaudo, J.-D. & Ross, A.) 313–346 (Springer International Publishing, 2016). doi:10.1007/978-3-319-23576-9_13.
14. Moggridge, B.J., & Thompson, R.M. Cultural value of water and western water management: an Australian Indigenous perspective. *Australasian Journal of Water Resources* **25**, 4-14 (2021).
15. Famiglietti, J. S. The global groundwater crisis. *Nat. Clim. Change* **4**, 945–948 (2014).
16. Richey. A.S. *et al.* Quantifying renewable groundwater stress with GRACE. *Water Resour. Res.* **51**, 5217-5238. (2015).
17. Rodell, M. *et al.* Emerging trends in global freshwater availability. *Nature* **557**, 651–659 (2018).
18. Villholth, K.G., & Conti, K.I. Groundwater governance: rationale, definition, current state and heuristic framework. In *Advances in Groundwater Governance*. (1st ed) 3-32 (2018).
19. Gain, A. K., Giupponi, C. & Wada, Y. Measuring global water security towards sustainable development goals. *Environ. Res. Lett.* **11**, 124015 (2016).
20. Gleeson, T., Wada, Y., Bierkens, M. F. P. & van Beek, L. P. H. Water balance of global aquifers revealed by groundwater footprint. *Nature* **488**, 197–200 (2012).
21. Jasechko, S. & Perrone, D. Global groundwater wells at risk of running dry. *Science* **372**, 418–421 (2021).
22. Winter, T. C., Harvey, J. W., Franke, O. L. & Alley, W. M. *Ground water and surface water: A single resource*. *Ground water and surface water: A single resource* vol. 1139 http://pubs.er.usgs.gov/publication/cir1139 (1998).
23. Anderson, E.P. *et al*. Understanding rivers and their social relations: A critical step to advance environmental water management. *WIREs Water*, **6**, e1381 (2019).
24. de Graaf, I. E. M., Gleeson, T., (Rens) van Beek, L. P. H., Sutanudjaja, E. H. & Bierkens, M. F. P. Environmental flow limits to global groundwater pumping. *Nature* **574**, 90–94 (2019).
25. Gleeson, T., Cuthbert, M., Ferguson, G. & Perrone, D. Global Groundwater Sustainability, Resources, and Systems in the Anthropocene. *Annu. Rev. Earth Planet. Sci.* **48**, 431–463 (2020).

26. Lall, U., Josset, L. & Russo, T. A Snapshot of the World's Groundwater Challenges. *Annu. Rev. Environ. Resour.* **45**, 171–194 (2020).
27. Gleeson, T., Villholth, K., Taylor, R., Perrone, D. & Hyndman, D. Groundwater: a call to action. *Nature* **576**, 213–213 (2019).
28. Loring, P.A. Threshold concepts and sustainability: features of a contested paradigm. *FACETS,* **5**, 182-199 (2020).
29. Ostrom, E. A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science* **325**, 419–422 (2009).
30. Partelow, S. A review of the social-ecological systems framework: applications, methods, modifications, and challenges. *Ecol. Soc.* **23**, (2018).
31. Huggins, X. *et al.* Groundwater in complex adaptive social-ecological systems. Preprint available at: https://eartharxiv.org/repository/view/3248/ (2022).
32. Chandanpurkar, H. et al. *Manuscript in preparation* (2022).
33. Siebert, S. *et al.* Development and validation of the global map of irrigation areas. *Hydrol. Earth Syst. Sci.* **9**, 535–547 (2005).
34. Tindale, N. B. *Aboriginal tribes of Australia : their terrain, environmental controls, distribution, limits, and proper names*. (Australian National University Press, 1974).
35. Rocha, J., Malmborg, K., Gordon, L., Brauman, K. & DeClerck, F. Mapping social-ecological systems archetypes. *Environ. Res. Lett.* **15**, 034017 (2020).
36. Sietz, D., Lüdeke, M. K. B. & Walther, C. Categorisation of typical vulnerability patterns in global drylands. *Glob. Environ. Change* **21**, 431–440 (2011).
37. Fridman, D., Koellner, T. & Kissinger, M. Exploring global interregional food system's sustainability using the functional regions typology. *Glob. Environ. Change* **68**, 102276 (2021).
38. Václavík, T., Lautenbach, S., Kuemmerle, T. & Seppelt, R. Mapping global land system archetypes. *Glob. Environ. Change* **23**, 1637–1647 (2013).
39. Eisenack, K., Oberlack, C. & Sietz, D. Avenues of archetype analysis: roots, achievements, and next steps in sustainability research. *Ecol. Soc.* **26**, (2021).
40. Richts, A., Struckmeier, W. F. & Zaepke, M. WHYMAP and the Groundwater Resources Map of the World 1:25,000,000. in *Sustaining Groundwater Resources: A Critical Element in the Global Water Crisis* (ed. Jones, J. A. A.) 159–173 (Springer Netherlands, 2011). doi:10.1007/978-90-481-3426-7_10.
41. Margat, J., & van der Gun, J. Groundwater around the World: A Geographic Synopsis. *Routledge & CRC Press* (2013).
42. Reinecke, R. *et al.* Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body parameterization. *Hydrol. Earth Syst. Sci.* **23**, 4561–4582 (2019).
43. Dupont, F., Sterckx, A. & Dijksma, R. Managed Aquifer Recharge (MAR) Suitability maps and standardized suitability index, the case study of the Occitanie region (South France). 60.
44. McGinnis, M. & Ostrom, E. Social-ecological system framework: initial changes and continuing challenges. *Ecol. Soc.* **19**, (2014).
45. Eisenack, K. *et al.* Design and quality criteria for archetype analysis. *Ecol. Soc.* **24**, (2019).
46. Raudsepp-Hearne, C., Peterson, G. D. & Bennett, E. M. Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. *Proc. Natl. Acad. Sci.* **107**, 5242–5247 (2010).
47. Wagener, T. *et al.* On doing hydrology with dragons: Realizing the value of perceptual models and knowledge accumulation. *WIREs Water* **8**, e1550 (2021).
48. Farinosi, F. *et al.* An innovative approach to the assessment of hydro-political risk: A spatially explicit, data driven indicator of hydro-political issues. *Glob. Environ. Change* **52**, 286–313 (2018).
49. Oakleaf, J. R. *et al.* Mapping global development potential for renewable energy, fossil fuels, mining and agriculture sectors. *Sci. Data* **6**, 101 (2019).
50. Jones, B. & O'Neill, B. C. Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environ. Res. Lett.* **11**, 084003 (2016).
51. Bierkens, M. F. P. Global hydrology 2015: State, trends, and directions. *Water Resour. Res.* **51**, 4923–4947 (2015).
52. Gleeson, T. *et al.* GMD perspective: The quest to improve the evaluation of groundwater representation in continental- to global-scale models. *Geosci. Model Dev.* **14**, 7545–7571 (2021).
53. Villholth, K. G. One water: Expanding boundaries for a new deal and a safe planet for all. *One Earth* **4**, 474–477 (2021).

54. Hijmans, R. J. *et al.* terra: Spatial Data Analysis. R package version 1.6-17. https://cran.r-project.org/package=terra (2022).
55. Pebesma, E. et al. sf: simple features for R. R package version 1.0–8. https://cran.r-project.org/package=sf (2022).
56. Charrad, M. *et al.* NbClust: Determining the Best Number of Clusters in a Data Set. https://cran.r-project.org/package=NbClust (2022).
57. Wehrens, R. & Kruisselbrink, J. kohonen: Supervised and Unsupervised Self-Organising Maps. R package version 3.0.11. https://cran.r-project.org/package=kohonen (2022).
58. Wickham, H. *et al.* ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. R package version 3.3.6. https://cran.r-project.org/package=ggplot2 (2022).
59. Tennekes, M. *et al.* tmap: Thematic Maps. R package version 3.3-3. https://cran.r-project.org/package=tmap (2022).
60. Varis, O., Taka, M. & Kummu, M. The Planet's Stressed River Basins: Too Much Pressure or Too Little Adaptive Capacity? *Earths Future* **7**, 1118–1135 (2019).
61. Huggins, X. *et al.* Hotspots for social and ecological impacts from freshwater stress and storage loss. *Nat. Commun.* **13**, 439 (2022).
62. Haitjema, H. M. & Mitchell-Bruker, S. Are Water Tables a Subdued Replica of the Topography? *Groundwater* **43**, 781–786 (2005).
63. Lehner, B. & Grill, G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186 (2013).
64. Cuthbert, M. O. *et al.* Global patterns and dynamics of climate–groundwater interactions. *Nat. Clim. Change* **9**, 137–141 (2019).
65. Beck, H. E. *et al.* Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resour. Res.* **49**, 7843–7863 (2013).
66. Mirza, M. U., Xu, C., Bavel, B. van, van Nes, E. H. & Scheffer, M. Global inequality remotely sensed. *Proc. Natl. Acad. Sci.* **118**, e1919913118 (2021).
67. Bierkens, M. *et al. Global Economic Limits of Groundwater When Used as a Last Resort for Irrigation*. Prepreint available at: https://www.researchsquare.com/article/rs-1874539/v1 (2022).
68. UNEP & UNEP-DHI. IWRM Data Portal. http://iwrmdataportal.unepdhi.org/ (2021).