Water Resources Research

Supporting Information for

Technical Report – Methods: Automated Discovery of Functional Relationships in Earth System Data

R. Reinecke[1,2], F. Pianosi[3,4], and T. Wagener[1]

[1]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

[2]Institute of Geography Johannes Gutenberg-University Mainz, Mainz, Germany

[3]Department of Civil Engineering, University of Bristol, Bristol, UK

[4]Cabot Institute, University of Bristol, Bristol, UK

Introduction

This supplemental material includes additional material that explains the process of groundwater recharge, different tests of data binning, in-depth explanations of variables used in the main manuscript, in-depth descriptions of the recharge model implementation and multiple experiments to test the validity of SONAR.

Table of Contents

S1 Groundwater recharge

Groundwater recharge can be defined as downward flow of water towards the water table adding water to an aquifer. This could be through downward percolation of soil water excess or through seepage from surface water bodies. The definition however can vary largely in its details between research communities and models (see Table S1).

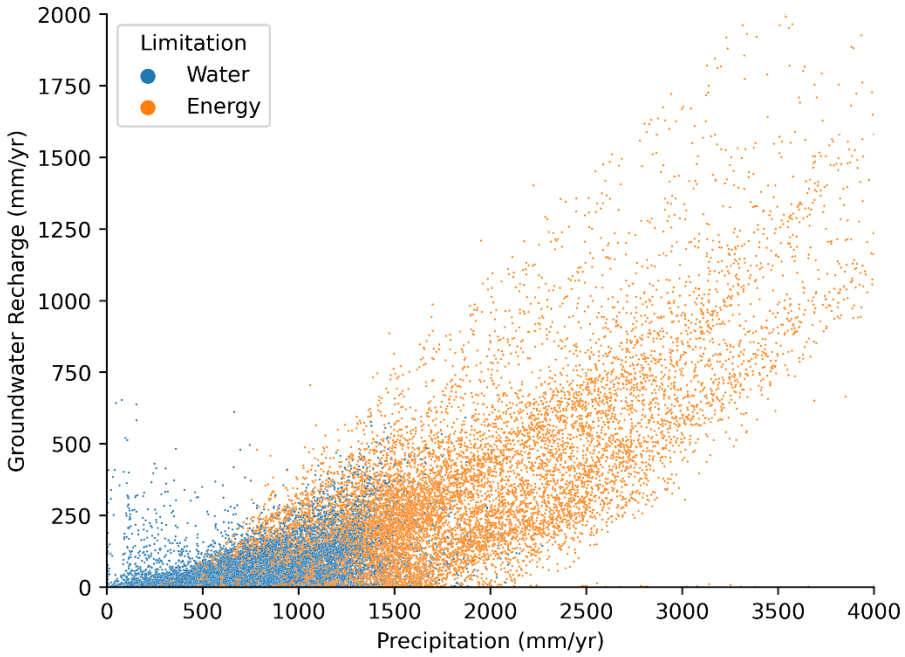

Figure S1 Global groundwater recharge simulated by a global hydrological model on a 0.5° spatial resolution. Scatter points are colored by two different climatic areas: areas where there is more water than potential evapotranspiration (energy-limited) and areas where there is more evapotranspiration than water (water-limited). The x and y-axis are limited to the majority of points for better readability (precipitation may reach over 8000 mm/yr and recharge over 4000 mm/yr).

S2 Strategies for testing subsets of binned data

The following figures assess how the correlation metric works in determining the first split decision. First, for equally-sized bins (as used in the main manuscript) and then for equally-spaced bins. The split decision is reached by taking one bin at a time and putting it into a virtual “bucket” that is then used to calculate the correlation of the data inside the bucket. In the “from=left” approach we start adding bins starting with small values on the x-axis and with “from=right” we start adding bin starting from high values. Thus, in the end we test different subsets of data. In SONAR both methods are implemented since the correlation is calculated both on the data inside the bucket and outside the bucket.

To show differences the figures are shown for two example models: CLM4.5 and PCR-GlobWB. The black line indicates how much of the data was used at a particular moment to calculate the correlation. The dotted line indicates a possible first split (when the correlation was highest.

Equally-sized bins

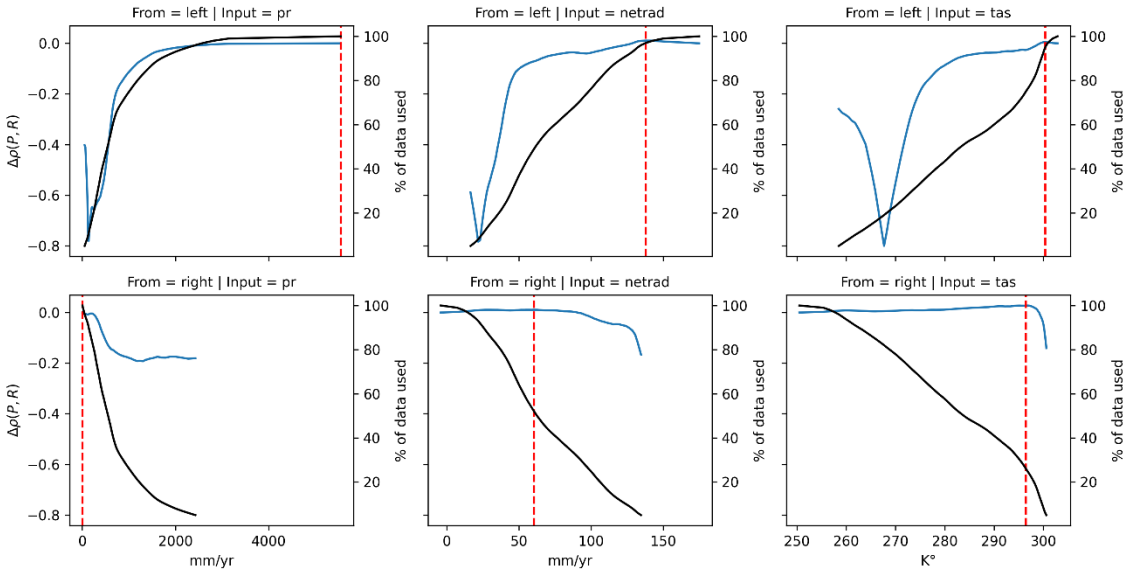

Figure S2 Change of correlation between precipitation and recharge calculated by the model CLM4.5. by selecting different data subsets for three different variables: precipitation, net radiation, and temperature and two different strategies in selecting subsets of data. The black line indicates the %of data points used to calculate the correlation at a given point. The red line indicates a possible split (point with highest correlation.

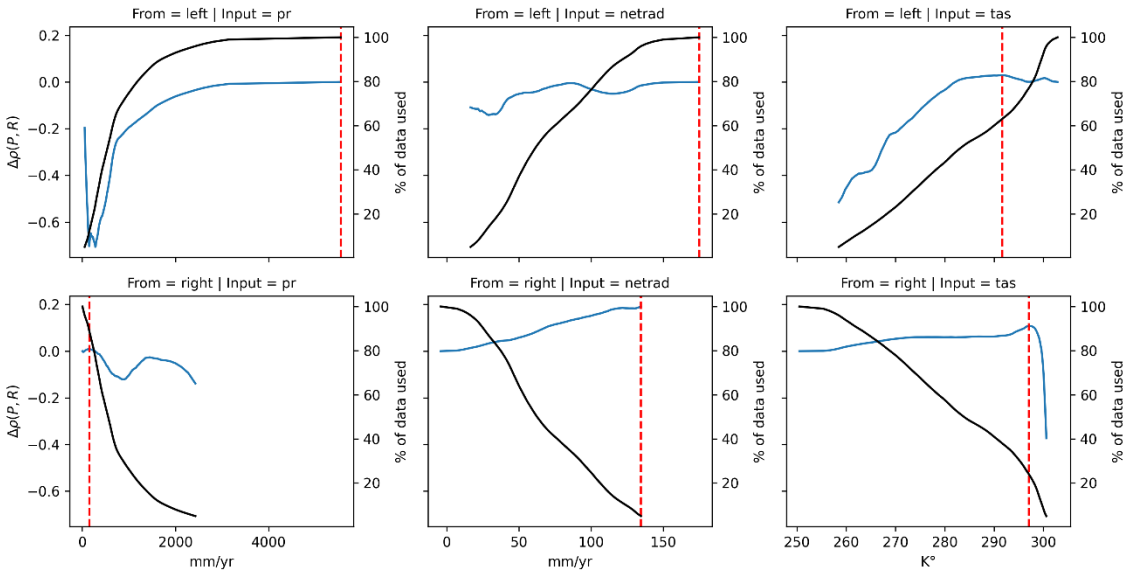

Figure S3 Change of correlation between precipitation and recharge calculated by the model PCR-GlobWB. by selecting different data subsets for three different variables: precipitation, net radiation, and temperature and two different strategies in selecting subsets of data. The black line indicates the %of data points used to calculate the correlation at a given point. The red line indicates a possible split (point with highest correlation.

Equally-spaced bins

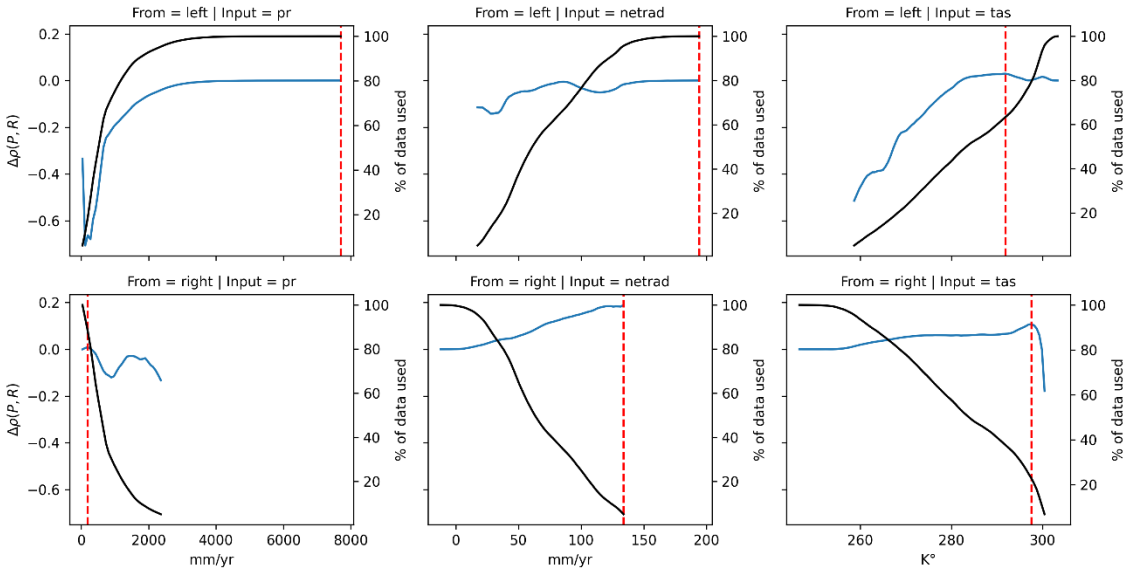

Figure S4 Change of correlation between precipitation and recharge calculated by the model CLM4.5. by selecting different data subsets for three different variables: precipitation, net radiation, and temperature and two different strategies in selecting subsets

of data. The black line indicates the %of data points used to calculate the correlation at a given point. The red line indicates a possible split (point with highest correlation.

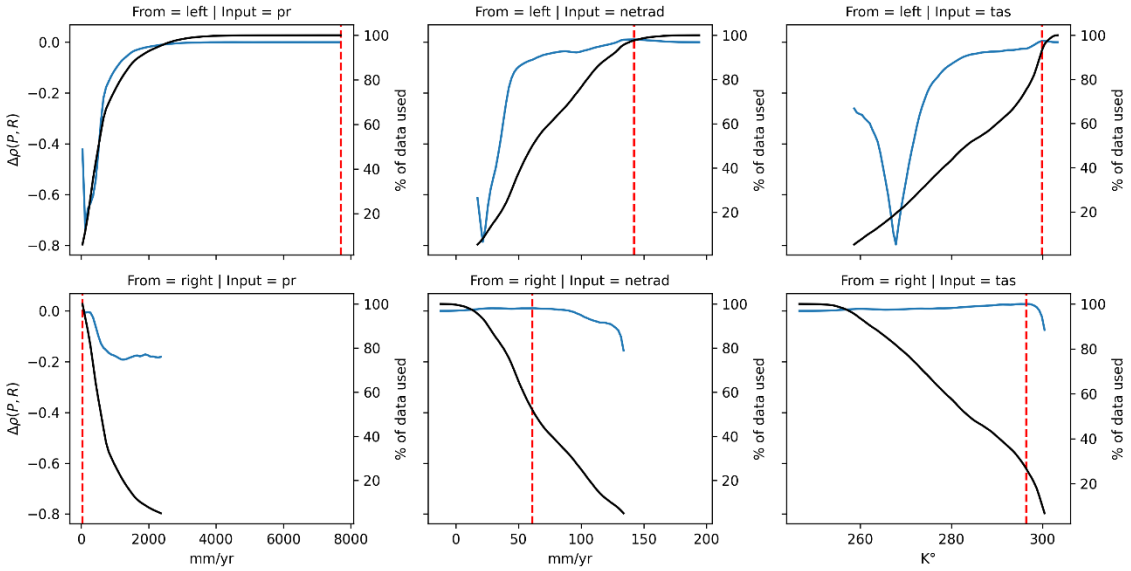

Figure S5 Change of correlation between precipitation and recharge calculated by the model PCR-GlobWB. by selecting different data subsets for three different variables: precipitation, net radiation, and temperature and two different strategies in selecting subsets of data. The black line indicates the %of data points used to calculate the correlation at a given point. The red line indicates a possible split (point with highest correlation.

S3 Explanatory variables used for the evaluation

The model outputs are based on the ISIMIP (Warszawski et al., 2014) framework and were aggregate into yearly means. The data used here is equal to the data available in Gnann et al. (2023) and the models used equal to Reinecke et al. (2021).

Table S1 List of global hydrological models (GHMs) used in the example analysis. The groundwater recharge (GWR) implementation description is adapted from Reinecke et al. (2021).

Model	Groundwater Recharge implementation	Model Reference
WaterGAP	GWR in WaterGAP2 is calculated as being a fraction of runoff from land based on soil texture, relief, aquifer type, and the existence of permafrost or glaciers, taking into account a soil-texture dependent maximum daily groundwater recharge rate (P. Döll & Fiedler, 2008). If a grid cell is defined as semiarid or arid and has a medium or coarse soil texture, GWR will only occur if daily precipitation exceeds a critical value (P. Döll & Fiedler, 2008); otherwise, the water runs off. Runoff from land that does not contribute to GWR is transferred to surface water bodies as fast surface runoff. WaterGAP further computes focused recharge beneath surface water bodies in semiarid and arid grid cells, which is not considered in this study.	(Müller Schmied et al., 2021)
PCR-GlobWB	PCR-GLOBWB (PCRaster Global Water Balance; Sutanudjaja et al., 2018); simulates the water storage in two vertically stacked soil layers and an underlying groundwater layer. Water exchanges are simulated between the layers (infiltration, percolation, and capillary rise) and the interaction of the top layer with the atmosphere (rainfall, evapotranspiration, and snowmelt). PCR-GLOBWB also calculates canopy interception and snow storage. Natural groundwater recharge is fed by net precipitation, and additional recharge from irrigation occurs as the net flux from the lowest soil layer to the groundwater layer, i.e., deep percolation minus capillary rise. The ARNO (a semi-distributed conceptual rainfall–runoff model; (Todini, 1996)) scheme is used to separate direct runoff, interflow, and GWR. Groundwater recharge can be balanced by capillary rise if the top of the groundwater level is within 5 m of the topographical surface (calculated as the height of the groundwater	(Sutanudjaja et al., 2018)

	storage over the storage coefficient on top of the streambed elevation and the sub-grid distribution of elevation).	
MATSIRO	The Minimal Advanced Treatments of Surface Interaction and RunOff (MATSIRO; Takata et al., 2003) is a global land surface model initially developed for an atmospheric–ocean general circulation model, the Model For Interdisciplinary Research On Climate (https://ccsr.aori.u-tokyo.ac.jp/~hasumi/miroc_description.pdf). This process-based model calculates water and energy flux and storage at and below the land surface, also considering the stomatal response to CO2 increase in the photosynthesis process. The offline version of MATSIRO used for the ISIMIP2b simulation explicitly takes vertical groundwater dynamics into account, including groundwater pumping (Y. Pokhrel et al., 2012; Y. N. Pokhrel et al., 2015). Soil moisture flux between the 15 soil layers is expressed as a function of the vertical gradient of the hydraulic potential, which is the sum of the matric potential and the gravitational head, and the soil moisture movement is calculated by Richards equation. MATSIRO calculates net groundwater recharge as a budget of gravitational drainage into and capillary rise from the layer where the groundwater table exists. A simplified TOPMODEL (TOPography-based MODEL; (Beven & Kirkby, 1979)) is used to represent surface runoff processes, and groundwater discharge is simulated by using an unconfined aquifer model (Koirala et al., 2014).	(Takata et al., 2003)
LPJML	Lund Potsdam Jena managed Land (LPJmL) is a dynamic global vegetation model that simulates the growth and productivity of both natural and agricultural vegetation as being coherently linked through their water, carbon, and energy fluxes (Schaphoff et al., 2018). The soil column is divided into six active hydrological layers, with a total thickness of 13 m depth. Percolation of infiltrated water through the soil column is	(Schaphoff et al., 2018)

	calculated according to a storage routine technique that simulates free water in the soil bucket. Excess water over the saturation levels produces lateral runoff in each layer (subsurface runoff). GWR is considered to be percolation (seepage) from the bottom soil layer. As there is no groundwater storage in LPJmL, for the ISIMIP2b protocol, seepage from the base soil layer is reported as both GWR and groundwater runoff, which is routed directly (with no time delay) back into the river system.	
JULES-W1	The Joint UK Land Environment Simulator (JULES; Best et al., 2011; W1 stands for water-related simulations in the ISIMIP framework) is a land surface model initially developed by the Met Office as the land surface component of the Met Office Unified Model. JULES is a process-based model that simulates the carbon, water, energy, and momentum fluxes between land and atmosphere, including plant–carbon interactions (Clark et al., 2011). The rainfall that reaches the ground is partitioned into Hortonian surface runoff and an infiltration component. A total of four soil layers represent the soil column, with a total thickness of 3 m, with a unit hydraulic head gradient lower boundary condition and no groundwater component. The water that infiltrates the soil moves down the soil layers that are updated using a finite difference form of the Richards equation (Best et al., 2011). The saturation excess water from the bottom soil layer becomes subsurface runoff that can be considered to be GWR (Le Vine et al., 2016).	(Best et al., 2011)
H08	H08 (Hanasaki et al., 2018) is a GHM that includes various components for water use and management. It consists of five major components, namely a simple bucket-type land surface model, a river routing model, a crop growth model, which is mainly used to estimate the timing of planting, harvesting, and irrigation in cropland, a reservoir operation model, and a water abstraction model. The abstraction model supplies water to meet the daily water demand	(Hanasaki et al., 2018)

	of three sectors (irrigation, industry, and municipality) from six available and accessible sources (river, local reservoir, aqueduct, seawater desalination, renewable groundwater, and non-renewable groundwater) and one hypothetical one termed unspecified surface water. It has two soil layers; one is to represent the unsaturated rootzone and the other the saturated zone (groundwater). The scheme of GWR computation is identical to Döll and Fiedler (2008).	
CWATM	The Community Water Model (CWatM) is a large-scale integrated hydrological model which encompasses general surface and groundwater hydrological processes, including human hydrological activities such as water use and reservoir regulation (Burek et al., 2020). CWatM takes six land cover classes into account and applies the tile approach. This hydrological model has three soil layers and one groundwater storage. The depth of the first soil layer is 5 cm, and the depth of second and third layers vary over grids, depending on the rootzone depth of each land cover class, resulting in total soil depth of up to 1.5 m. Groundwater storage is designed being as a linear reservoir. CWatM includes preferential bypass flow directly into groundwater storage and capillary rise from groundwater storage and percolation from the third soil layer to groundwater storage. Hence, the groundwater recharge reported by CWatM in ISIMIP2b is the net recharge calculated from these three terms.	(Burek et al., 2020)
CLM4.5	The Community Land Model version 4.5 (CLM4.5; Swenson and Lawrence, 2015) is the land component of the Community Earth System Model (CESM), a fully coupled, state-of-the-art Earth system model. CLM is a land surface model representing the physical, chemical, and biological processes through which terrestrial ecosystems influence and are influenced by climate, including CO2, across a variety of spatial and temporal scales (Lawrence et al. 2015). Individual land grid points can be	(S. C. Swenson & Lawrence, 2015)

	composed of multiple land units due to the nested tile approach, which enables the implementation of multiple soil columns and represents biomes as a combination of different plant functional types. Groundwater processes, including sub-surface runoff, recharge, and water table depth variations, are simulated based on the SIMTOP scheme (SImple groundwater Model TOPgraphy based; (Oleson et al., 2013).	

Table S2 Explanatory variables used. Except for landcover all explanatory variables are based on ISIMIP (Warszawski et al., 2014) data aggregated

Feature	**Temporal aggregation**	**Source**
Precipitation	Long-term mean (30-years; bias-corrected GCMs)	ISIMIP, (Gnann et al., 2023)
PET	Long-term mean (model ensemble)	ISIMIP, (Gnann et al., 2023)
Aridity (PET/P)	See PET and P	-
Temperature	Long-term mean (30-years, bias-corrected GCM)	ISIMIP (Gnann et al., 2023)
Temperature (cold day indicator)	Days below 1°C (30-years, bias-corrected GCM	ISIMIP (Gnann et al., 2023)
Land cover (Forest, Shrubland, Grassland, Sparsely Veg., Bare areas, Wetlands, Cropland, Waterbodies, Snow/ice, Artificial)	GlobCover (aggregated to 0.5° with area-weighted Mode)	(ESA, 2010)

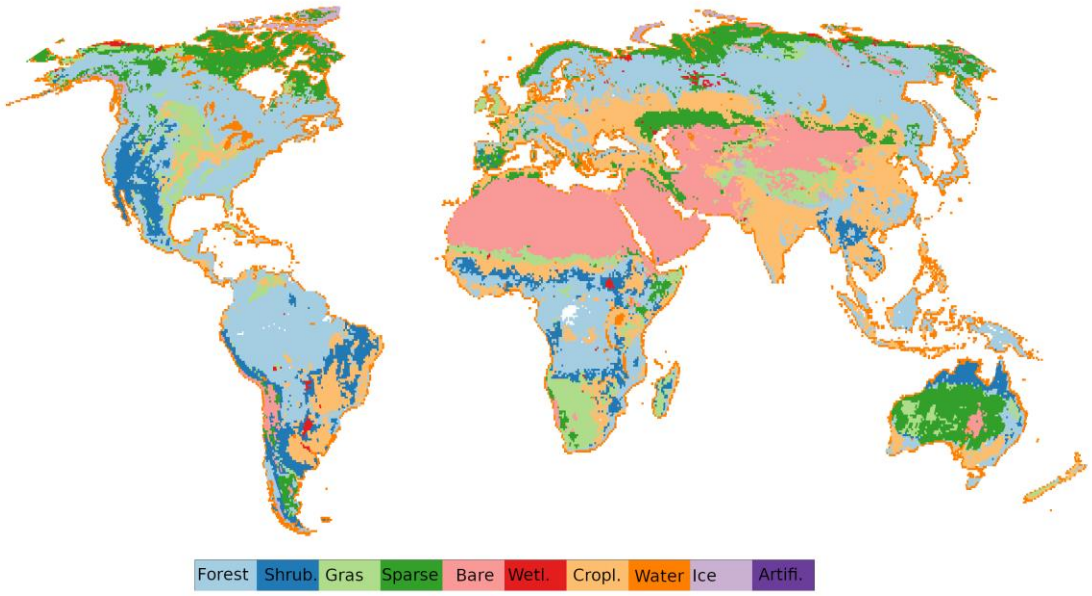

Figure S6 Landcover classes.

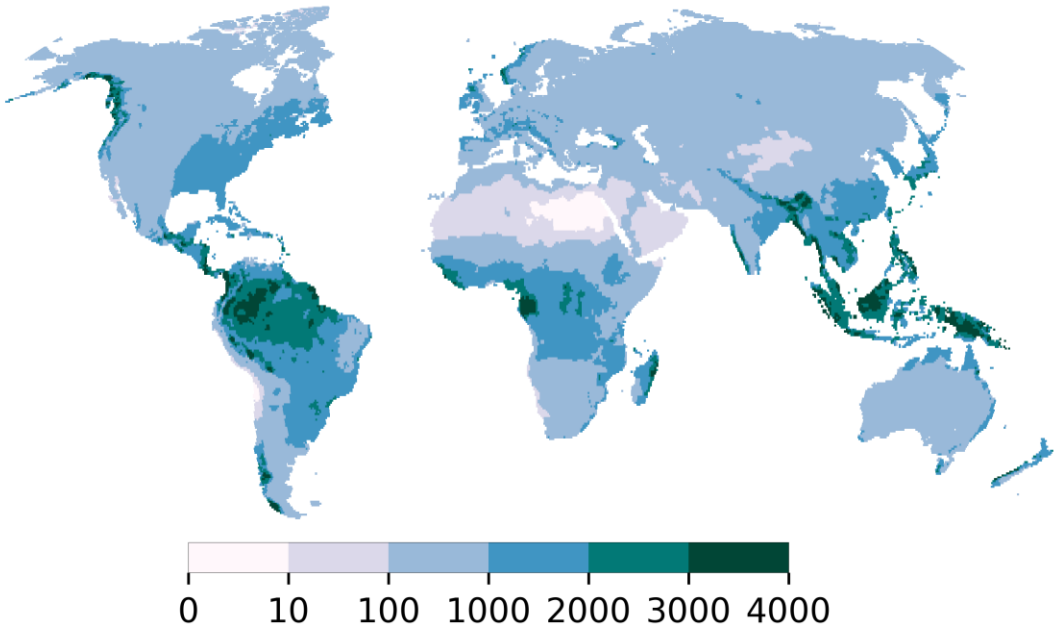

Figure S7 Precipitation in mm/yr.

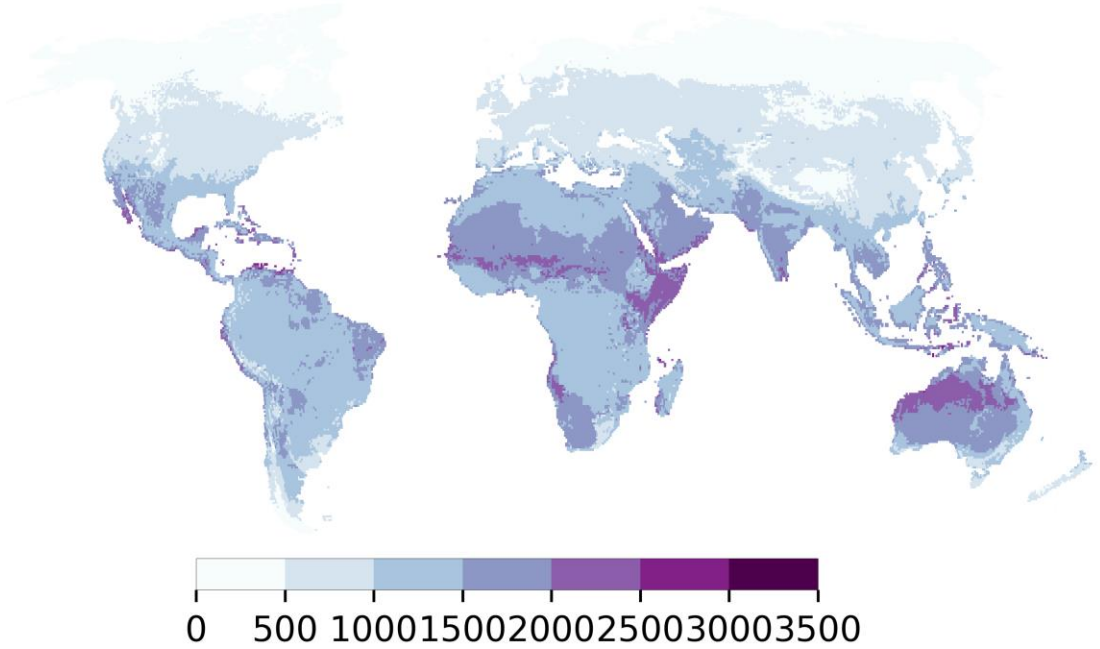

Figure S8 PET in mm/yr.

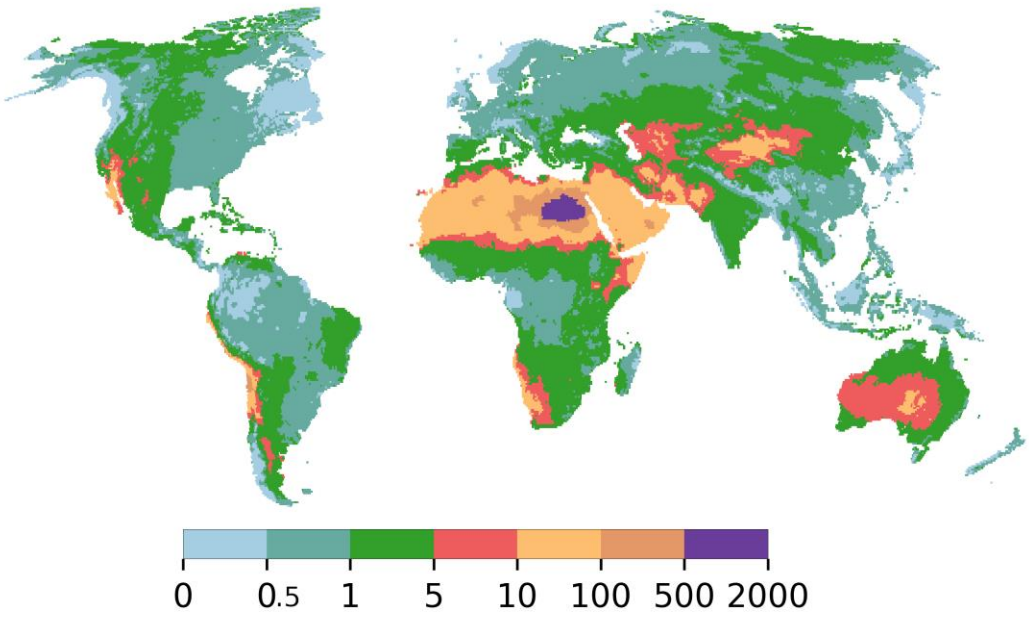

Figure S9 Aridity index as PET/P.

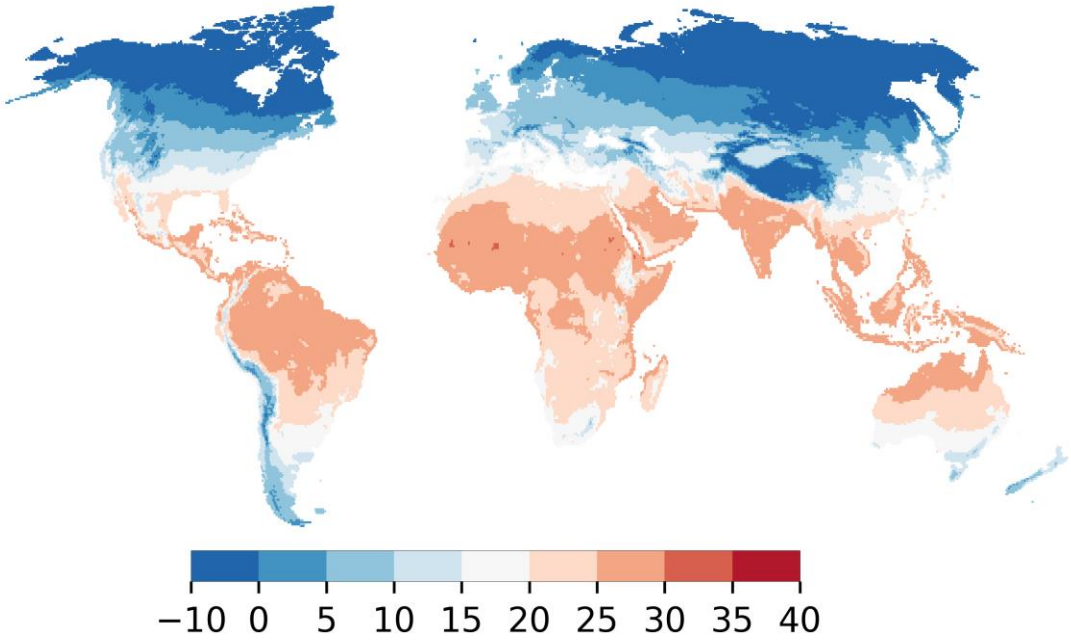

Fig S10 Daily mean temperature in °C.

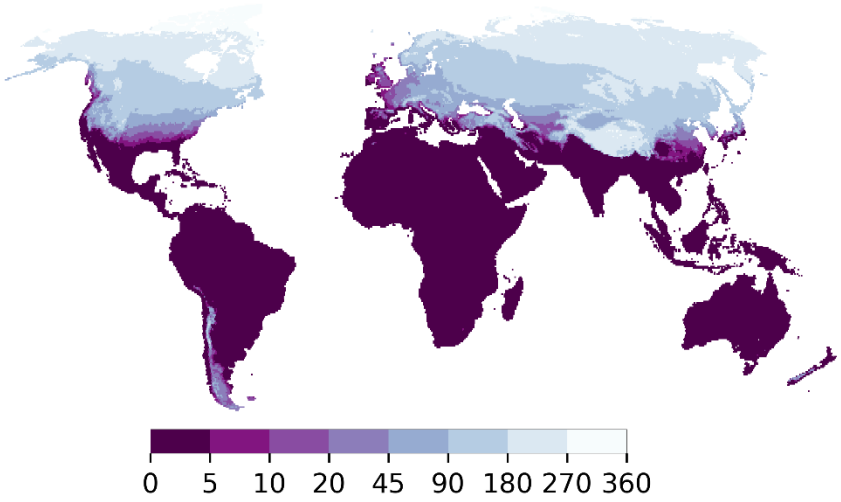

Fig S11 Days below 0°C.

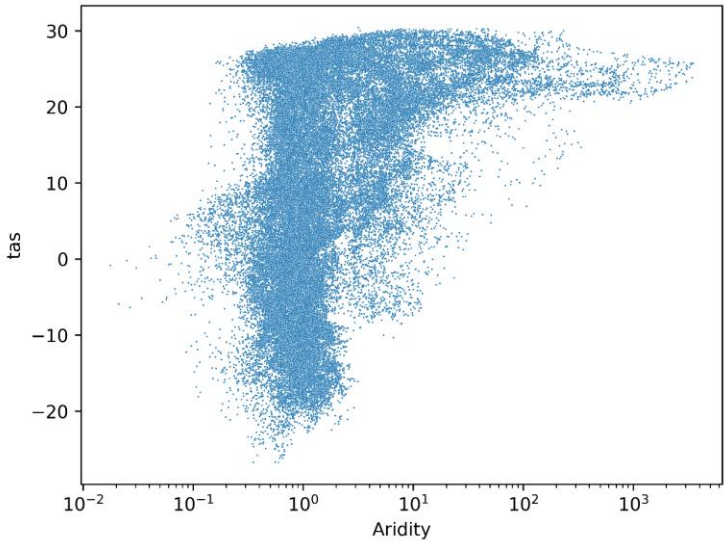

Fig S12 Scatterplot of the aridity index and the mean daily temperature in °C.

S3 How do different categorizations of recharge effect the results?

A difference in recharge classes only affects the results of the classification algorithms of CART (Fig. S13). SONAR does not make any prior assumptions about classes.

a)

```
|--- pr <= 74.65
|   |--- pr <= 39.91
|   |   |--- PET <= 1719.07
|   |   |   |--- pr <= 36.61
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  36.61
|   |   |   |   |--- class: 0
|   |   |--- PET >  1719.07
|   |   |   |--- pr <= 28.64
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  28.64
|   |   |   |   |--- class: 0
|   |--- pr >  39.91
|   |   |--- tas <= 24.08
|   |   |   |--- pr <= 58.11
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  58.11
|   |   |   |   |--- class: 0
|   |   |--- tas >  24.08
|   |   |   |--- PET <= 1823.81
|   |   |   |   |--- class: 1
|   |   |   |--- PET >  1823.81
|   |   |   |   |--- class: 0
|--- pr >  74.65
|   |--- pr <= 2217.79
|   |   |--- Aridity <= 2.46
|   |   |   |--- pr <= 1828.68
|   |   |   |   |--- class: 1
|   |   |   |--- pr >  1828.68
|   |   |   |   |--- class: 1
|   |   |--- Aridity >  2.46
|   |   |   |--- 1d <= 29.95
|   |   |   |   |--- class: 1
|   |   |   |--- 1d >  29.95
|   |   |   |   |--- class: 0
|   |--- pr >  2217.79
|   |   |--- pr <= 2847.89
|   |   |   |--- tas <= 25.76
|   |   |   |   |--- class: 1
|   |   |   |--- tas >  25.76
|   |   |   |   |--- class: 1
|   |   |--- pr >  2847.89
|   |   |   |--- tas <= 23.33
|   |   |   |   |--- class: 1
|   |   |   |--- tas >  23.33
|   |   |   |   |--- class: 2
```

b)

```
|--- pr <= 424.96
|   |--- pr <= 302.94
|   |   |--- pr <= 140.30
|   |   |   |--- pr <= 119.57
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  119.57
|   |   |   |   |--- class: 0
|   |   |--- pr >  140.30
|   |   |   |--- tas <= -8.46
|   |   |   |   |--- class: 0
|   |   |   |--- tas >  -8.46
|   |   |   |   |--- class: 0
|   |--- pr >  302.94
|   |   |--- tas <= -7.91
|   |   |   |--- tas <= -8.85
|   |   |   |   |--- class: 0
|   |   |   |--- tas >  -8.85
|   |   |   |   |--- class: 0
|   |   |--- tas >  -7.91
|   |   |   |--- PET <= 733.52
|   |   |   |   |--- class: 1
|   |   |   |--- PET >  733.52
|   |   |   |   |--- class: 0
|--- pr >  424.96
|   |--- pr <= 2217.79
|   |   |--- tas <= -6.89
|   |   |   |--- pr <= 659.61
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  659.61
|   |   |   |   |--- class: 1
|   |   |--- tas >  -6.89
|   |   |   |--- Aridity <= 2.14
|   |   |   |   |--- class: 1
|   |   |   |--- Aridity >  2.14
|   |   |   |   |--- class: 1
|   |--- pr >  2217.79
|   |   |--- pr <= 2847.89
|   |   |   |--- tas <= 25.76
|   |   |   |   |--- class: 1
|   |   |   |--- tas >  25.76
|   |   |   |   |--- class: 1
|   |   |--- pr >  2847.89
|   |   |   |--- tas <= 23.33
|   |   |   |   |--- class: 1
|   |   |   |--- tas >  23.33
|   |   |   |   |--- class: 2
```

c)

```
|--- pr <= 917.09
|   |--- pr <= 589.98
|   |   |--- pr <= 500.93
|   |   |   |--- pr <= 445.92
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  445.92
|   |   |   |   |--- class: 0
|   |   |--- pr >  500.93
|   |   |   |--- Aridity <= 1.01
|   |   |   |   |--- class: 0
|   |   |   |--- Aridity >  1.01
|   |   |   |   |--- class: 0
|   |--- pr >  589.98
|   |   |--- Aridity <= 1.17
|   |   |   |--- tas <= -5.88
|   |   |   |   |--- class: 0
|   |   |   |--- tas >  -5.88
|   |   |   |   |--- class: 1
|   |   |--- Aridity >  1.17
|   |   |   |--- Aridity <= 1.73
|   |   |   |   |--- class: 0
|   |   |   |--- Aridity >  1.73
|   |   |   |   |--- class: 0
|--- pr >  917.09
|   |--- pr <= 2217.79
|   |   |--- Aridity <= 1.09
|   |   |   |--- pr <= 1592.24
|   |   |   |   |--- class: 1
|   |   |   |--- pr >  1592.24
|   |   |   |   |--- class: 1
|   |   |--- Aridity >  1.09
|   |   |   |--- pr <= 1098.23
|   |   |   |   |--- class: 0
|   |   |   |--- pr >  1098.23
|   |   |   |   |--- class: 1
|   |--- pr >  2217.79
|   |   |--- pr <= 2681.47
|   |   |   |--- tas <= 25.93
|   |   |   |   |--- class: 1
|   |   |   |--- tas >  25.93
|   |   |   |   |--- class: 1
|   |   |--- pr >  2681.47
|   |   |   |--- tas <= 23.81
|   |   |   |   |--- class: 1
|   |   |   |--- tas >  23.81
|   |   |   |   |--- class: 2
```

Fig S13 CART classification for the same global model and three different choices of what constitutes low recharge (class 0 in this text representation)): a) less than 1mm, b) less than 10mm and c) less than 100mm (as use in the main text). Text representation need to be read from left (values on the far left represent the first split, values on the far right the leaf nodes with the different recharge classifications) to right.

S4 Path comparison between trees

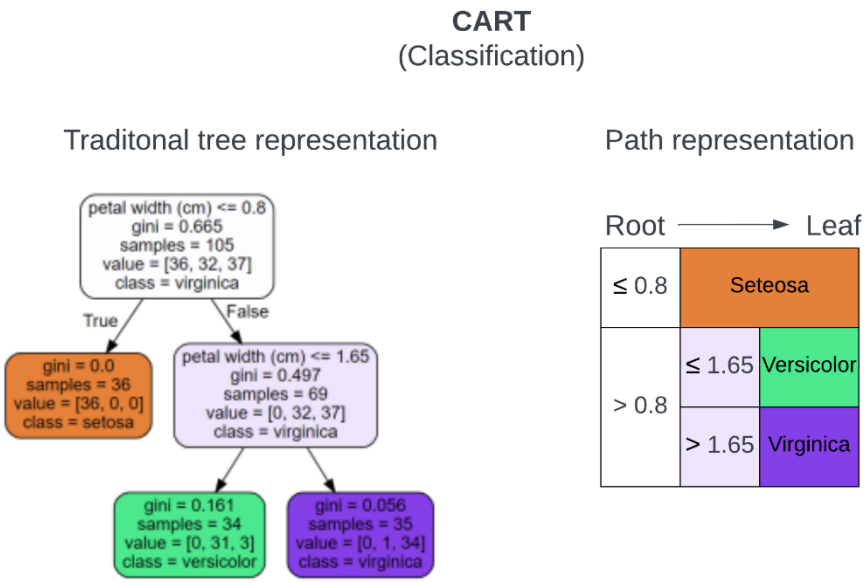

Fig S14 A simple example of the path visualization used in this manuscript with the established flower classification problem (Unwin & Kleinman, 2021). Left the CART tree and right the path representation in the same colors for the explanatory variables.

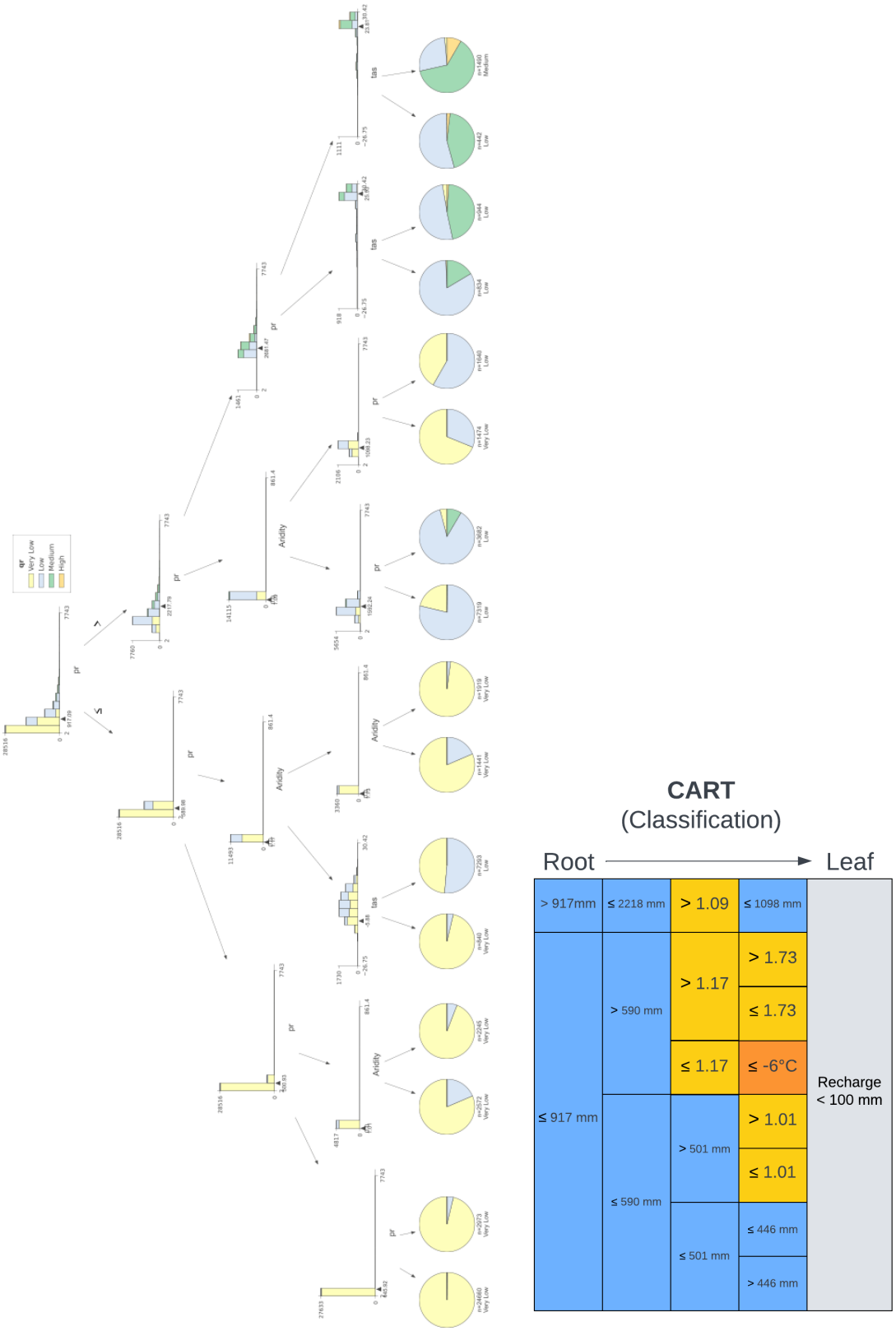

Fig S15 Full CART tree of the three depicted in Fig. 3 of the main manuscript next to the corresponding path visualization.

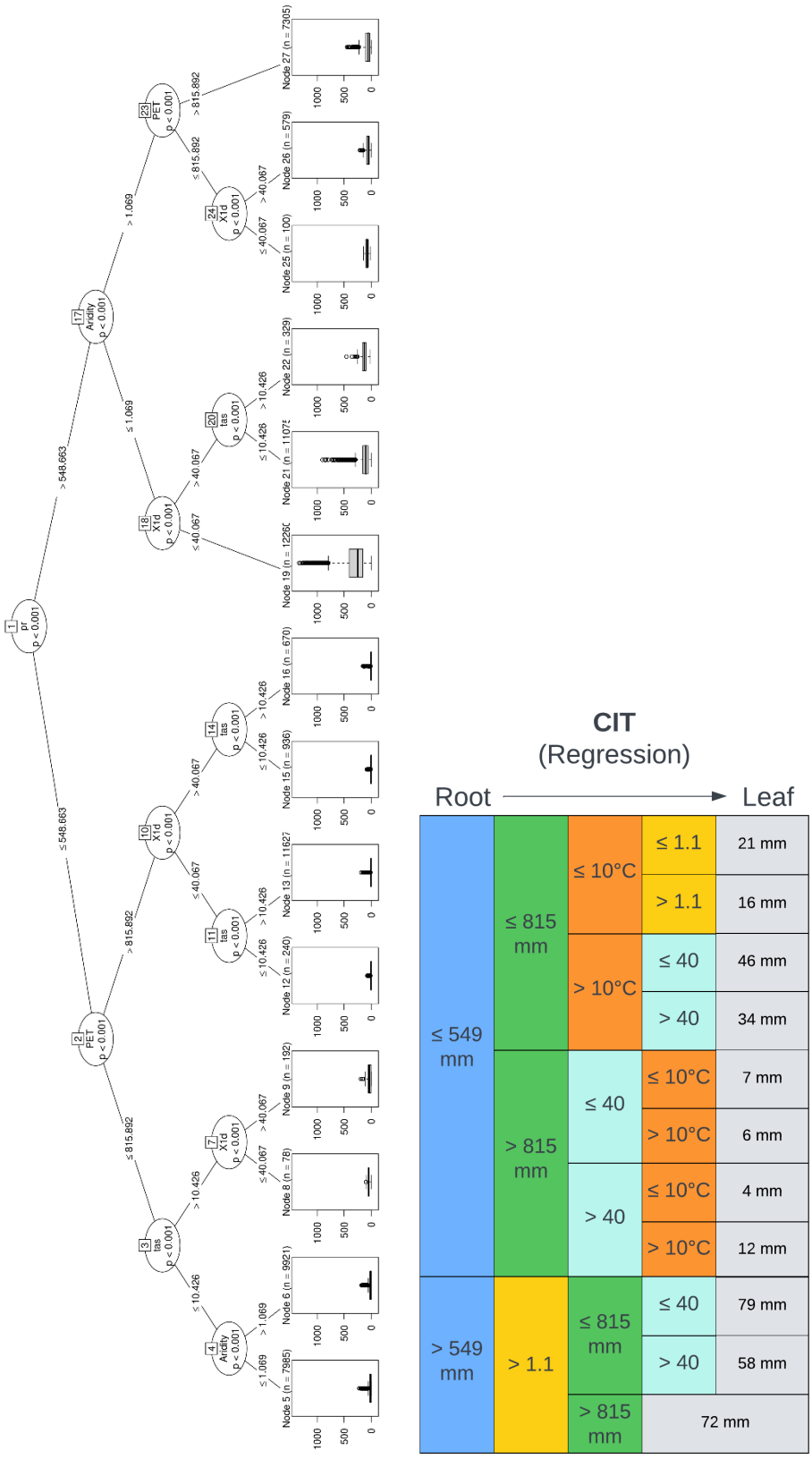

Fig S16 Full CIT tree of the three depicted in Fig. 3 of the main manuscript next to the corresponding path visualization.

S5 Additional SONAR trees of other all models

The following shows all SONAR trees of the 8 investigated models. The trees of the two models shown in the main manuscript are equal to the ones shown here in text representation. Text representation need to be read from left (values on the far left represent the first split, values on the far right the leaf nodes with the different recharge classifications) to right.

```
Checking pcr-globwb
Max initial correlation is 0.74 to variable pr
Root node
|-  1: Aridity <= 0.46 with 4332 points; p = 0.87; dri. = pr
|-  2: Aridity > 0.46 with 57550 points; p = 0.70; dri. = pr
```

Fig S17 SONAR tree for the model PCR-GlobWB.

```
Checking watergap2
Max initial correlation is 0.89 to variable pr
Root node
|-  1: tas <= 25.52 with 51755 points; p = 0.64; dri. = Aridity
|-  2: tas > 25.52 with 9858 points; p = 0.95; dri. = Aridity
```

Fig S18 SONAR tree for the model WaterGAP2.

```
Checking clm45
Max initial correlation is 0.89 to variable pr
Root node
|-  1: Aridity <= 0.38 with 1874 points; p = 0.97; dri. = pr
|-  2: Aridity > 0.38 with 35597 points; p = 0.88; dri. = pr
|---  3: Aridity <= 0.51 with 2623 points; p = 0.97; dri. = pr
|---  4: Aridity > 0.51 with 32974 points; p = 0.87; dri. = pr
```

Fig S19 SONAR tree for the model CLM4.5.

```
Checking cwatm
Max initial correlation is 0.91 to variable pr
Root node
|-  1: PET <= 1297.08 with 27178 points; p = 0.88; dri. = pr
|-  2: PET > 1297.08 with 18118 points; p = 0.94; dri. = pr
|---  3: tas <= 26.53 with 12701 points; p = 0.95; dri. = pr
|---  4: tas > 26.53 with 5417 points; p = 0.94; dri. = pr
```

Fig S20 SONAR tree for the model CWATM.

```
Checking h08
Max initial correlation is 0.91 to variable pr
Root node
|-  1: tas <= 27.05 with 59108 points; p = 0.91; dri. = pr
|-  2: tas > 27.05 with 3773 points; p = 0.98; dri. = pr
```

Fig S21 SONAR tree for the model H08.

```
Checking jules-w1
Max initial correlation is 0.67 to variable pr
Root node
|-  1: LC == Sparsely Veg. with 8633 points; p = 0.93; dri. = Aridity
|---  2: PET <= 375.26 with 4918 points; p = 0.95; dri. = Aridity
|---  3: PET > 375.26 with 3715 points; p = 0.53; dri. = Aridity
|-  4: LC != Sparsely Veg. with 54025 points; p = 0.87; dri. = Aridity
```

Fig S22 SONAR tree for the model Jules-W1.

```
Checking lpjml
Max initial correlation is 0.77 to variable pr
Root node
|-  1: Aridity <= 0.46 with 3528 points; p = 0.90; dri. = pr
|---  2: Aridity <= 0.36 with 1764 points; p = 0.90; dri. = pr
|---  3: Aridity > 0.36 with 1764 points; p = 0.94; dri. = pr
|-  4: Aridity > 0.46 with 31749 points; p = 0.75; dri. = pr
```

Fig S23 SONAR tree for the model LPJML.

```
Checking matsiro
Max initial correlation is 0.79 to variable pr
Root node
|-  1: tas <= 20.18 with 37998 points; p = 0.72; dri. = Aridity
|-  2: tas > 20.18 with 17881 points; p = 0.91; dri. = Aridity
|---  3: PET <= 1432.46 with 7830 points; p = 0.95; dri. = pr
|-----  4: PET <= 1326.65 with 3818 points; p = 0.93; dri. = pr
|-----  5: PET > 1326.65 with 4012 points; p = 0.96; dri. = pr
|---  6: PET > 1432.46 with 10051 points; p = 0.84; dri. = pr
```

Fig S24 SONAR tree for the model MATSIRO.

S6 Tree growth without minimum number of point requirement

If the number of point requirement is set to a very low value (in the following: > 0.1% of points of parent node and at least 10) even SONAR trees grow bigger. However, the number of points in splits is likely to low to allow a meaningful calculation of a correlation.

```
Checking lpjml
Max initial correlation is 0.77 to variable pr
Root node
|-  1: Aridity <= 0.46 with 3528 points; p = 0.90; dri. = pr
|---  2: Aridity <= 0.29 with 1059 points; p = 0.88; dri. = pr
|-----  3: Aridity <= 0.23 with 706 points; p = 0.86; dri. = pr
|-----  4: Aridity > 0.23 with 353 points; p = 0.95; dri. = pr
|-------  5: LC == Waterbodies with 88 points; p = 0.85; dri. = PET
|-------  6: LC != Waterbodies with 265 points; p = 0.96; dri. = PET
|---------  7: LC == Forest with 166 points; p = 0.96; dri. = pr
|-----------  8: tas <= 23.40 with 130 points; p = 0.95; dri. = tas
|-----------  9: tas > 23.40 with 36 points; p = 0.00; dri. = tas
|---------  10: LC != Forest with 99 points; p = 0.91; dri. = pr
|-----------  11: tas <= 1.80 with 57 points; p = 0.72; dri. = pr
|-----------  12: tas > 1.80 with 42 points; p = 0.92; dri. = pr
|---  13: Aridity > 0.29 with 2469 points; p = 0.95; dri. = pr
|-----  14: Aridity <= 0.33 with 353 points; p = 0.97; dri. = tas
|-----  15: Aridity > 0.33 with 2116 points; p = 0.93; dri. = tas
|-  16: Aridity > 0.46 with 31749 points; p = 0.75; dri. = pr
|---  17: Aridity <= 0.49 with 706 points; p = 0.91; dri. = pr
|-----  18: LC == Forest with 424 points; p = 0.87; dri. = PET
|-----  19: LC != Forest with 282 points; p = 0.94; dri. = PET
|-------  20: LC == Cropland with 57 points; p = 0.86; dri. = PET
|-------  21: LC != Cropland with 225 points; p = 0.94; dri. = PET
|---------  22: LC == Waterbodies with 126 points; p = 0.90; dri. = tas
|---------  23: LC != Waterbodies with 99 points; p = 0.95; dri. = tas
|-----------  24: Aridity <= 0.48 with 53 points; p = 0.96; dri. = tas
|-----------  25: Aridity > 0.48 with 46 points; p = 0.95; dri. = tas
|---  26: Aridity > 0.49 with 31043 points; p = 0.74; dri. = pr
```

Fig S25 SONAR tree of the model LPJML with almost no requirements on the minimum number of points per split.

S7 Artificial generation of recharge test data

Experiment 1: Completely random groundwater recharge

In this experiment the groundwater recharge data is substituted by randomly generated data. The data lies within the same ranges as the original but does not follow its distribution or any spatial patterns. Fig. S26 shows the resulting values plotted as a global map. With the chosen explanatory variables SONAR does not find any splits for the randomly generated data.

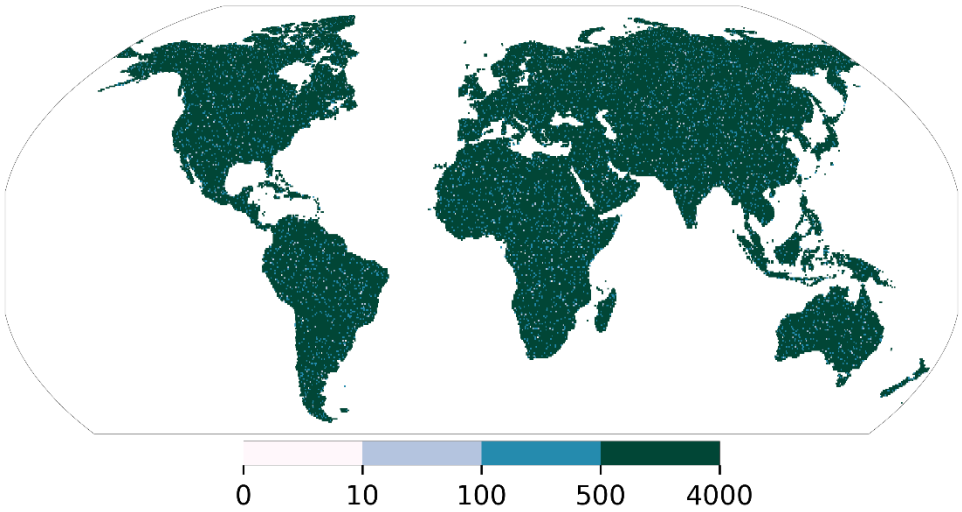

Fig S26 Randomly generated recharge data plotted as a global map.

Experiment 2: Precipitation as dominant control

In this experiment we also generate groundwater recharge data based only on precipitation. To create a perfect correlation, we simply turn precipitation into recharge based on the following rules (Fig. S27).

Multiplier k for the four climatic regions:

Wet cold regions: 0.2
Dry cold regions: 0.4
Dry warm regions: 0.6
Wet warm regions: 0.8

Groundwater recharge = Precipitation * k

This tests whether SONAR is correctly picking up this introduced signal. The resulting tree is shown in the text representation in Fig. S28. The dominant driver is always precipitation confirming that SONAR correctly picks up the artificially introduced relationship. Splits are based on PET which is likely because PET is a good proxy for separating water and energy limited regions (Fig. S8).

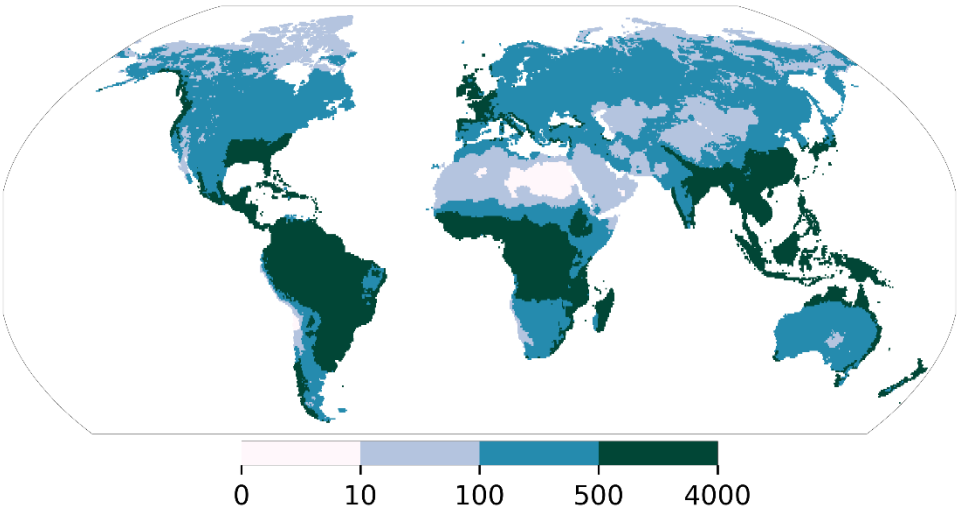

Fig S27 Groundwater recharge based on precipitation.

```
Checking artificial
Max initial correlation is 0.89 to variable pr
Root node
|-  1: PET <= 1860.28 with 60132 points; p = 0.89; dri. = pr
|---  2: PET <= 1686.52 with 56967 points; p = 0.90; dri. = pr
|---  3: PET > 1686.52 with 3165 points; p = 1.00; dri. = pr
|-  4: PET > 1860.28 with 3165 points; p = 1.00; dri. = pr
```

Fig S28 SONAR tree of the generated groundwater recharge.

Experiment 3: PET as dominant control

This experiment works equally to experiment 2, but with PET instead of precipitation. PET is here turned directly into groundwater recharge: 10% of PET = recharge. The resulting SONAR tree is shown in text from in Fig. S30. Even if the tree grows relatively large PET is always identified as the dominant control (as the introduced correlation is 1).

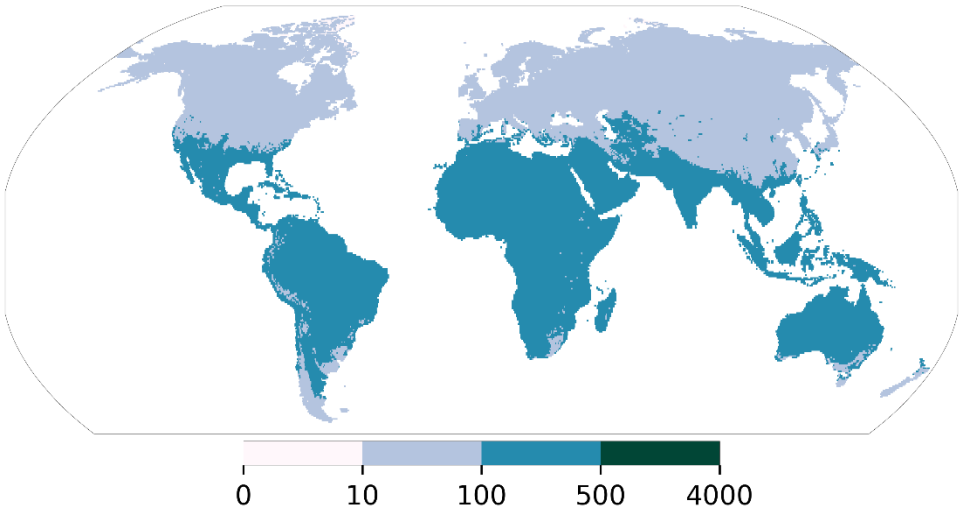

Fig S29 Groundwater recharge generated based on PET.

```
Checking artificial
Max initial correlation is 1.00 to variable PET
Root node
|-  1: LC == Forest with 20650 points; p = 1.00; dri. = PET
|---  2: pr <= 394.58 with 3215 points; p = 1.00; dri. = PET
|---  3: pr > 394.58 with 17435 points; p = 1.00; dri. = PET
|-----  4: pr <= 552.77 with 3199 points; p = 1.00; dri. = PET
|-----  5: pr > 552.77 with 14236 points; p = 1.00; dri. = PET
|-------  6: pr <= 690.29 with 3428 points; p = 1.00; dri. = PET
|-------  7: pr > 690.29 with 10808 points; p = 1.00; dri. = PET
|---------  8: pr <= 1034.77 with 3170 points; p = 1.00; dri. = PET
|---------  9: pr > 1034.77 with 7638 points; p = 1.00; dri. = PET
|-----------  10: pr <= 1605.25 with 3494 points; p = 1.00; dri. = PET
|-----------  11: pr > 1605.25 with 4144 points; p = 1.00; dri. = PET
|-  12: LC != Forest with 42491 points; p = 1.00; dri. = PET
|---  13: LC == Shrubland with 3816 points; p = 1.00; dri. = PET
|---  14: LC != Shrubland with 38675 points; p = 1.00; dri. = PET
|-----  15: LC == Grasland with 4469 points; p = 1.00; dri. = PET
|-----  16: LC != Grasland with 34206 points; p = 1.00; dri. = PET
|-------  17: LC == Sparsely Veg. with 8685 points; p = 1.00; dri. = PET
|---------  18: pr <= 287.40 with 3215 points; p = 1.00; dri. = PET
|---------  19: pr > 287.40 with 5470 points; p = 1.00; dri. = PET
|-------  20: LC != Sparsely Veg. with 25521 points; p = 1.00; dri. = PET
|---------  21: LC == Bare areas with 8314 points; p = 1.00; dri. = PET
|-----------  22: pr <= 100.86 with 4570 points; p = 1.00; dri. = PET
|-----------  23: pr > 100.86 with 3744 points; p = 1.00; dri. = PET
|---------  24: LC != Bare areas with 17207 points; p = 1.00; dri. = PET
|-----------  25: LC == Cropland with 9289 points; p = 1.00; dri. = PET
|-------------  26: pr <= 575.81 with 3211 points; p = 1.00; dri. = PET
|-------------  27: pr > 575.81 with 6078 points; p = 1.00; dri. = PET
|-----------  28: LC != Cropland with 7918 points; p = 1.00; dri. = PET
|-------------  29: pr <= 517.45 with 3240 points; p = 1.00; dri. = PET
|-------------  30: pr > 517.45 with 4678 points; p = 1.00; dri. = PET
```

Fig S30 SONAR tree for recharge that is only based on PET.

References

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R.. H., & Ménard, C. B., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011

Beven, K. J., & Kirkby, M. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, *24*(1), 43–69. https://doi.org/10.1080/02626667909491834

Burek, P., Satoh, Y., Kahil, T., Tang, T., Greve, P., & Smilovic, M., et al. (2020). Development of the Community Water Model (CWatM v1.04) – a high-resolution hydrological model for global and regional assessment of integrated water resources management. *Geoscientific Model Development*, *13*(7), 3267–3298. https://doi.org/10.5194/gmd-13-3267-2020

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., & Best, M. J., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, *4*(3), 701–722. https://doi.org/10.5194/gmd-4-701-2011

Döll, P., & Fiedler, K. (2008). Global-scale modeling of groundwater recharge. *Hydrology and Earth System Sciences*, *12*(3), 863–885. https://doi.org/10.5194/hess-12-863-2008

ESA. (2010). Global land cover map. Retrieved from http://due.esrin.esa.int/page_globcover.php

Gnann, S., Reinecke, R., Stein, L., Wada, Y., Thiery, W., & Müller Schmied, H., et al. (2023). Functional relationships reveal differences in the water cycle representation of global water models. Preprint. https://doi.org/10.31223/X50S9R

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., & Kanae, S. (2018). A global hydrological simulation to specify the sources of water used by humans. *Hydrology and Earth System Sciences*, *22*(1), 789–817. https://doi.org/10.5194/hess-22-789-2018

Koirala, S., Yeh, P. J.-F., Hirabayashi, Y., Kanae, S., & Oki, T. (2014). Global-scale land surface hydrologic modeling with the representation of water table dynamics. *Journal of Geophysical Research: Atmospheres*, *119*(1), 75–89. https://doi.org/10.1002/2013JD020398

Le Vine, N., Butler, A., McIntyre, N., & Jackson, C. (2016). Diagnosing hydrological limitations of a land surface model: application of JULES to a deep-groundwater chalk basin. *Hydrology and Earth System Sciences*, *20*(1), 143–159. https://doi.org/10.5194/hess-20-143-2016

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., & Niemann, C., et al. (2021). The global water resources and use model WaterGAP v2.2d: model description and evaluation. *Geoscientific Model Development*, *14*(2), 1037–1079. https://doi.org/10.5194/gmd-14-1037-2021

Oleson, K., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., & Koven, C., et al. (2013). *Technical description of version 4.5 of the Community Land Model (CLM).*

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P. J.-F., & Kim, H., et al. (2012). Incorporating Anthropogenic Water Regulation Modules into a Land Surface Model. *Journal of Hydrometeorology*, *13*(1), 255–269. https://doi.org/10.1175/JHM-D-11-013.1

Pokhrel, Y. N., Koirala, S., Yeh, P. J.-F., Hanasaki, N., Longuevergne, L., Kanae, S., & Oki, T. (2015). Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. *Water Resources Research*, *51*(1), 78–96. https://doi.org/10.1002/2014WR015602

Reinecke, R., Müller Schmied, H., Trautmann, T., Andersen, L. S., Burek, P., & Flörke, M., et al. (2021). Uncertainty of simulated groundwater recharge at different global warming levels: a global-scale multi-model ensemble study. *Hydrology and Earth System Sciences*, *25*(2), 787–810. https://doi.org/10.5194/hess-25-787-2021

Schaphoff, S., Bloh, W. von, Rammig, A., Thonicke, K., Biemans, H., & Forkel, M., et al. (2018). LPJmL4 – a dynamic global vegetation model with managed land – Part 1: Model description. *Geoscientific Model Development*, *11*(4), 1343–1375. https://doi.org/10.5194/gmd-11-1343-2018

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., & Drost, N., et al. (2018). PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. *Geoscientific Model Development*, *11*(6), 2429–2453. https://doi.org/10.5194/gmd-11-2429-2018

Swenson, S. C., & Lawrence, D. M. (2015). A GRACE -based assessment of interannual groundwater dynamics in the C ommunity L and M odel. *Water Resources Research*, *51*(11), 8817–8833. https://doi.org/10.1002/2015WR017582

Takata, K., Emori, S., & Watanabe, T. (2003). Development of the minimal advanced treatments of surface interaction and runoff. *Global and Planetary Change*, *38*(1-2), 209–222. https://doi.org/10.1016/S0921-8181(03)00030-4

Todini, E. (1996). The ARNO rainfall—runoff model. *Journal of Hydrology*, *175*(1-4), 339–382. https://doi.org/10.1016/S0022-1694(96)80016-3

Unwin, A., & Kleinman, K. (2021). The Iris Data Set: In Search of the Source of Virginica. *Significance*, *18*(6), 26–29. https://doi.org/10.1111/1740-9713.01589

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2014). The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): project framework. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3228–3232. https://doi.org/10.1073/pnas.1312330110