

Environmental flow limits to global groundwater pumping

Inge E. M. de Graaf^{1,2,3*}, Tom Gleeson⁴, L. P. H. (Rens) van Beek², Edwin H. Sutanudjaja² & Marc F. P. Bierkens^{2,5}

Groundwater is the world's largest freshwater resource and is critically important for irrigation, and hence for global food security^{1–3}. Already, unsustainable groundwater pumping exceeds recharge from precipitation and rivers⁴, leading to substantial drops in the levels of groundwater and losses of groundwater from its storage, especially in intensively irrigated regions^{5–7}. When groundwater levels drop, discharges from groundwater to streams decline, reverse in direction or even stop completely, thereby decreasing streamflow, with potentially devastating effects on aquatic ecosystems. Here we link declines in the levels of groundwater that result from groundwater pumping to decreases in streamflow globally, and estimate where and when environmentally critical streamflows—which are required to maintain healthy ecosystems—will no longer be sustained. We estimate that, by 2050, environmental flow limits will be reached for approximately 42 to 79 per cent of the watersheds in which there is groundwater pumping worldwide, and that this will generally occur before substantial losses in groundwater storage are experienced. Only a small decline in groundwater level is needed to affect streamflow, making our estimates uncertain for streams near a transition to reversed groundwater discharge. However, for many areas, groundwater pumping rates are high and environmental flow limits are known to be severely exceeded. Compared to surface-water use, the effects of groundwater pumping are markedly delayed. Our results thus reveal the current and future environmental legacy of groundwater use.

During dry seasons or times of drought when surface water is insufficient to meet human water demands, groundwater often sustains people and ecosystems¹. With a growing world population and continuing economic development, our freshwater resources are under simultaneous threat from increasing human water consumption and human-induced climate change, the latter of which will lead to more frequent and severe droughts⁸. Already, groundwater is widely exploited, increasingly at rates that exceed recharge from rain and rivers over longer time periods and larger areas, leading to substantial and persistent drops in levels of groundwater and losses of groundwater from its storage (called groundwater depletion)^{5–7}. Groundwater storage can recover only when pumping decreases or more groundwater is recharged from precipitation, rivers or engineered managed aquifer recharge (MAR) systems. About 70% of the pumped groundwater worldwide is used to sustain irrigation and is important for food security, as groundwater is used to maintain agricultural production during short- and long-term droughts^{2,3}. When the groundwater level drops, pumping costs increase, potentially resulting in a rise in food prices. When wells run dry, local and possibly larger-scale food security can be threatened⁹. It is expected that over the coming decades global food demands will rise further and food production will compete for space and water resources with crops used as biofuels, increasing the dependency on groundwater globally and highlighting the central place of groundwater resources within the food–water–energy nexus^{10,11}. In addition, declining groundwater levels induce land subsidence,

affecting infrastructure and increasing flood risks in coastal cities¹. Last, declining groundwater levels reduce the groundwater discharge that is essential for sustaining river flows, lake levels, springs, groundwater-fed wetlands and related ecosystems, especially during droughts^{12,13}. The net result is a slow desiccation of the landscape, with the progression largely hidden by precipitation variability. The effect of groundwater pumping on groundwater discharge is our primary focus. The expected increase in groundwater dependence and the negative effects related to over-abstraction means we must urgently identify the limits to global groundwater pumping and determine where and when these limits will be reached.

In this study, we estimate where and when environmentally critical streamflow will be reached because of groundwater pumping. Unlike previous large-scale impact assessments of groundwater depletion^{5,6}, we pay particular attention to the intricate effects of groundwater pumping on interactions between groundwater and surface water through groundwater discharge and river infiltration. For this assessment we used a physically based global-scale surface water–groundwater model^{12–15} (GSGM; see Methods). We simulated groundwater heads and head-dependent groundwater fluxes globally at a high resolution. We have used the best currently available datasets for the model parameterization of, for example, aquifer transmissivity and surface-water depth, and tested the model sensitivity for different values of these parameters (see Methods).

More specifically, we consider that streamflow reaches its environmental flow limit for the first time as a result of groundwater pumping when the threshold of environmentally critical streamflow is crossed for at least three consecutive months for two consecutive years. Environmentally critical streamflow is defined as the 90th percentile over five years (10% exceedance) of groundwater discharge⁴. This approach focuses on the dependence of ecosystem functions and services on streamflow under low flow conditions, when the contribution of groundwater discharge to streamflow is largest. By comparing the 'natural' and human-influenced results, we can exclude environmental flow limits that are reached as a result of climate-driven drought events¹⁶ alone (see Methods).

The GSGM simulations cover the period 1960–2100, using previously published climate reanalysis products¹⁷ and gridded human water demands⁷ as model input between 1960 and 2010. For 2011–2100 we assume a 'business-as-usual' scenario in which industrial and domestic water demands, as well as irrigated area, remain constant after 2010. Irrigation water demands change only as a result of climate change. For climate change, we assume the RCP 8.5 emission scenario and use the driest, wettest, and average climate projections in terms of future global precipitation change¹⁸ to represent climate change uncertainty. The maps in the main Figures show the results of the average scenario (that is, HadGEM2-ES) (for results of other climate scenarios see Extended Data Fig. 1).

Our results show that environmental flow limits caused by groundwater pumping have already been reached for a substantial number of

¹Chair of Environmental Hydrological Systems, Faculty of Environmental and Natural Resources, University of Freiburg, Freiburg, Germany. ²Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands. ³Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO, USA. ⁴Department of Civil Engineering and School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada. ⁵Unit Soil and Groundwater Systems, Deltares, Utrecht, The Netherlands. *e-mail: inge.de.graaf@hydrology.uni-freiburg.de

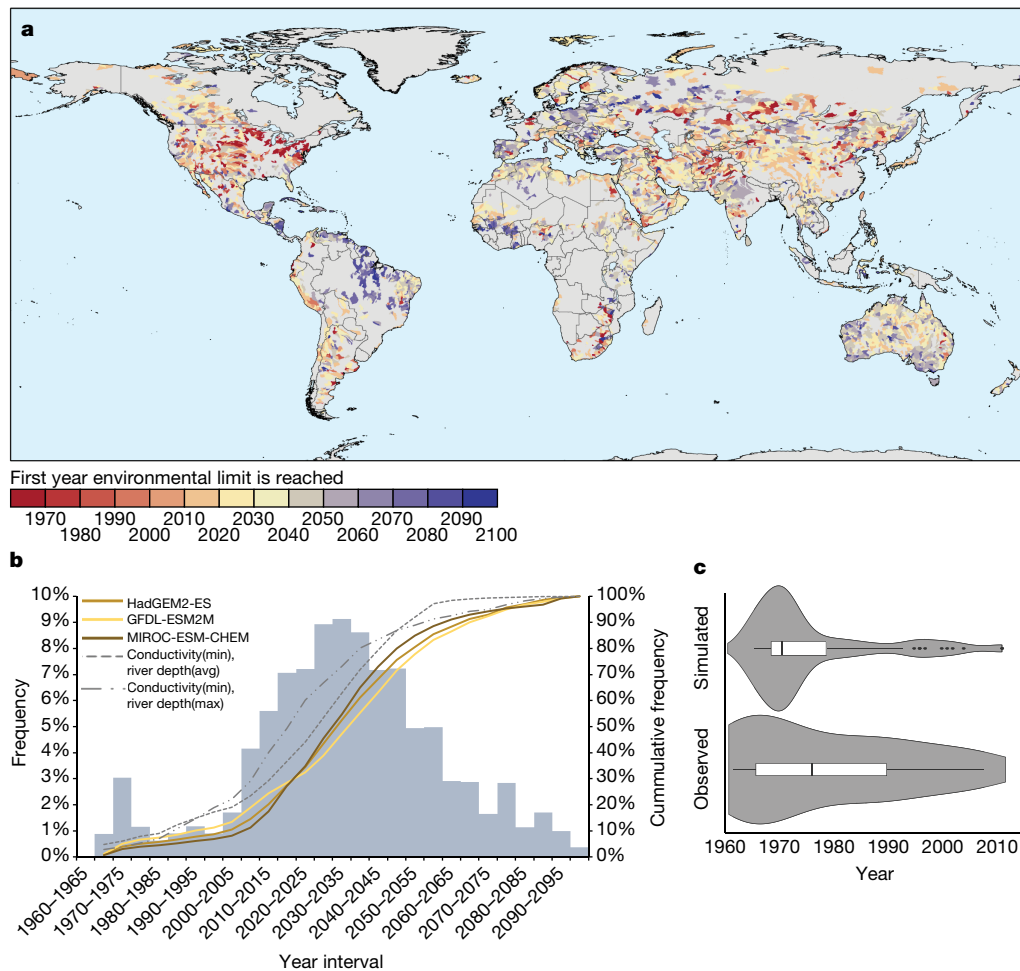


Fig. 1 | Estimated first time environmental flows have been or will be reached globally, and evaluation of results and model sensitivity. **a**, The first time at which environmental flow limits have been, or will be, reached, by year, averaged per sub-watershed (using the sub-watershed level of the HydroBASINS dataset²³). **b**, Global distribution of estimated first times at which environmental flow limits have been, or will be, reached. The histogram shows the estimates using the average climate input (HadGEM2-ES). Cumulative frequency is plotted for all three climate inputs used (coloured lines), and for the two runs with different parameter values resulting in the minimum and maximum number

of watersheds reaching their environmental flow limit (see Methods; grey lines, with minimum conductivity values and showing average and maximum river depth). **c**, Evaluation of this study's estimations for the first time at which environmental flow limits have been reached (simulated) and 'observed' environmental flow limits estimated from streamflow observations for 42 stations in Kansas, USA. The violin plot shows the distribution of the data, the bottom and top of each box represent the 25% and 75% quantiles, and the line inside each box represents the median.

watersheds (currently estimated as approximately 15%, 17% and 21% for the wettest, average and driest climate projections, respectively) and are likely to be reached for more than half of the watersheds before the end of 2050 (Fig. 1a, b; approximately 42%, 58% and 79%). Globally, the estimated first times at which environmental flow limits will be reached peak at around 2030 (Fig. 1b). Regions that have already reached their environmental flow limit are mainly found in the drier climates of the world, where discharge is small and irrigation depends more on groundwater^{3,13}. In general, environmental flow limits have not yet been reached for regions with lower groundwater pumping and/or higher streamflow, where streamflow is less dependent on groundwater discharge, or where additional recharge from irrigation helps to maintain groundwater discharge to streams (Fig. 1a). Hotspots of 'early limits', which were reached before 2010, are found for groundwater depletion hotspots (Extended Data Figs. 2, 3) such as the High Plains aquifer, part of the Central Valley aquifer, parts of Mexico, and the Upper Ganges and Indus basins. However, a considerable number of watersheds where environmental flow limits have been reached are found outside the estimated depletion hotspots, such as in the north-east USA and parts of Argentina. In the near future, before 2050, new regions that are reaching their environmental flow limit will develop;

in particular where the pressure on groundwater resources will increase owing to projected drier climate conditions that will cause an increase in irrigation water demand, such as in southern and central Europe and part of Africa.

A one-at-a-time sensitivity analysis showed that varying model parameters (hydraulic conductivity and surface-water depths) or using different climate models changed the number of watersheds and timing of environmental limits reached (see Methods). However, both the timing of the peak (around 2030) and the spatial distribution of watersheds that reached their environmental flow limit do not differ much between the different runs. This shows that the results we present are robust to uncertainties in parameters and climate models. Currently, environmental flow limits are reached for 10% to 23% of the watersheds under different parameter settings (by the end of 2050 this is 16% to 53%; Fig. 1b, Extended Data Fig. 1). The cumulative frequency distributions of the estimated first times at which environmental flow limits are, or will be, reached for the runs with different sensitivity are shown in Fig. 1b.

We compared our global-scale results to watershed-scale estimates of the first time environmental flow limits are reached for watersheds that undergo substantial pumping and for which observed streamflow

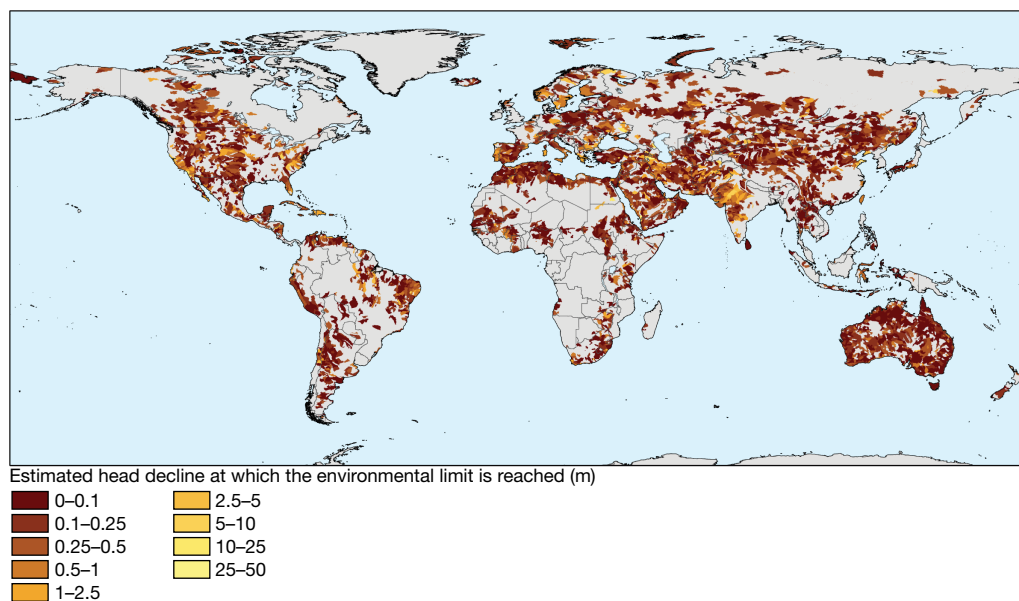


Fig. 2 | Estimated head decline, caused by groundwater pumping, associated with reaching the environmental limit. Estimated head decline is shown in metres.

records exist for periods long enough to estimate environmentally required streamflow and the impact of groundwater pumping (see Methods). Although rare, we found these data for Kansas, USA (Fig. 1c) and a few other watersheds in several countries in different climate zones (Extended Data Fig. 4). Global-scale estimates show similar frequencies and distribution of timing to catchment-scale estimates.

We find that the declines in groundwater level (estimated over 1960–2100 and for average climatology) that are associated with reaching the environmental flow limit for the first time are unexpectedly small (Fig. 2). Only a very small decline in groundwater level is needed to alter streamflow. This makes our estimates uncertain for streams that are near the transition from gaining to losing flow. However, for many areas, groundwater pumping for irrigation has been extreme and environmental flow limits are very likely to have been exceeded. The small head declines also indicate that, for many regions, environmental flow limits are reached before substantial groundwater depletion occurs.

This difference in timing (spatially and temporally) between reaching environmental flow limits and groundwater depletion can be explained by the complex dynamics of the effects of groundwater pumping on groundwater–surface water interactions (see Methods; Fig. 3). The effects of pumping on groundwater levels and streamflow vary widely depending on the groundwater–surface water regime. When the groundwater body and stream are still connected, river infiltration constrains the drop in groundwater levels and thus groundwater depletion. However, groundwater discharge may have already decreased markedly at this point (Fig. 3b, c). When the groundwater body is disconnected from the stream, a further decrease in groundwater levels occurs when pumping continues and substantial groundwater depletion is experienced (Fig. 3d). Unlike surface-water use, which immediately affects streamflow, the effect of groundwater pumping on streams can be substantially delayed (of the order of months to decades; for example, Fig. 3b), turning unsustainable groundwater withdrawals into a ‘ticking time bomb’ for streamflow.

From Extended Data Fig. 2 we also note that there are regions that are notable depletion hotspots in terms of head decline, but where environmental flow limits are not (yet) reached, such as the north China plain and the southeastern part of the Ganges basin. This difference in pattern might be partly explained by long delays between the start of pumping and the maximum effect on streamflow (Fig. 3b, c), but are more likely to be due to a disconnection between the riverbed and

groundwater levels (Fig. 3d), when the influence of head declines on streamflow is much smaller.

Although groundwater depletion estimates change under different climate projections, using a ‘business-as-usual’ scenario (Extended Data Fig. 1a), the spatial patterns and global distribution of the estimated first time at which the environmental flow limit is reached are similar (Fig. 1b, Extended Data Fig. 1b, c). However, a higher or lower number of watersheds exceed environmental flow limits for dryer or wetter climate input, respectively.

Our estimated depletion rates and the estimated first time that environmental flow limits are reached are likely to be optimistic, as they do not take into account projected increases in groundwater demand due to population growth or economic development in the emerging economies of the world¹⁹. In addition, our analysis did not include surface-water withdrawals, which also have an effect on river low flows²⁰. We estimate that for 60% of the watersheds with substantial groundwater pumping, surface-water withdrawal is also substantial (exceeding 0.01 m³ per m² per year) and environmental flow limits are likely to not be reached by groundwater pumping only.

We have focused on the environmental flow limit, but the successive depletion of groundwater resources will eventually hit physical and economic limits as well. The economic limit, which is assumed to be reached when pumping groundwater is no longer profitable, depends on many factors, such as labour and material costs, type of aquifer, energy costs, required pumping capacity and crop prices. It is safe to say that the economic limits to global groundwater pumping have not yet been reached, considering that in the Indian Punjab province farmers are currently abstracting groundwater from depths²¹ of over 40 m and some agricultural wells in the USA are close to 300 m deep²². Apparently, revenues still exceed pumping costs in these cases. Also, as indicated in this study, only a limited water level drop is needed to reach environmental flow limits, suggesting that, in general, it is very likely that environmental flow limits will be reached before any economic limit is encountered.

Given that healthy streamflow regimes are essential for aquatic ecosystems and provide an invaluable ecosystem service, environmental groundwater requirements should be part of any water-resource assessment at the global scale. The insights provided by this study could be used as a starting point for more detailed continental- to regional-scale studies to ensure sustainable and efficient groundwater use as much as possible, especially for regions in which detailed regional-to-continental studies have not been carried out.

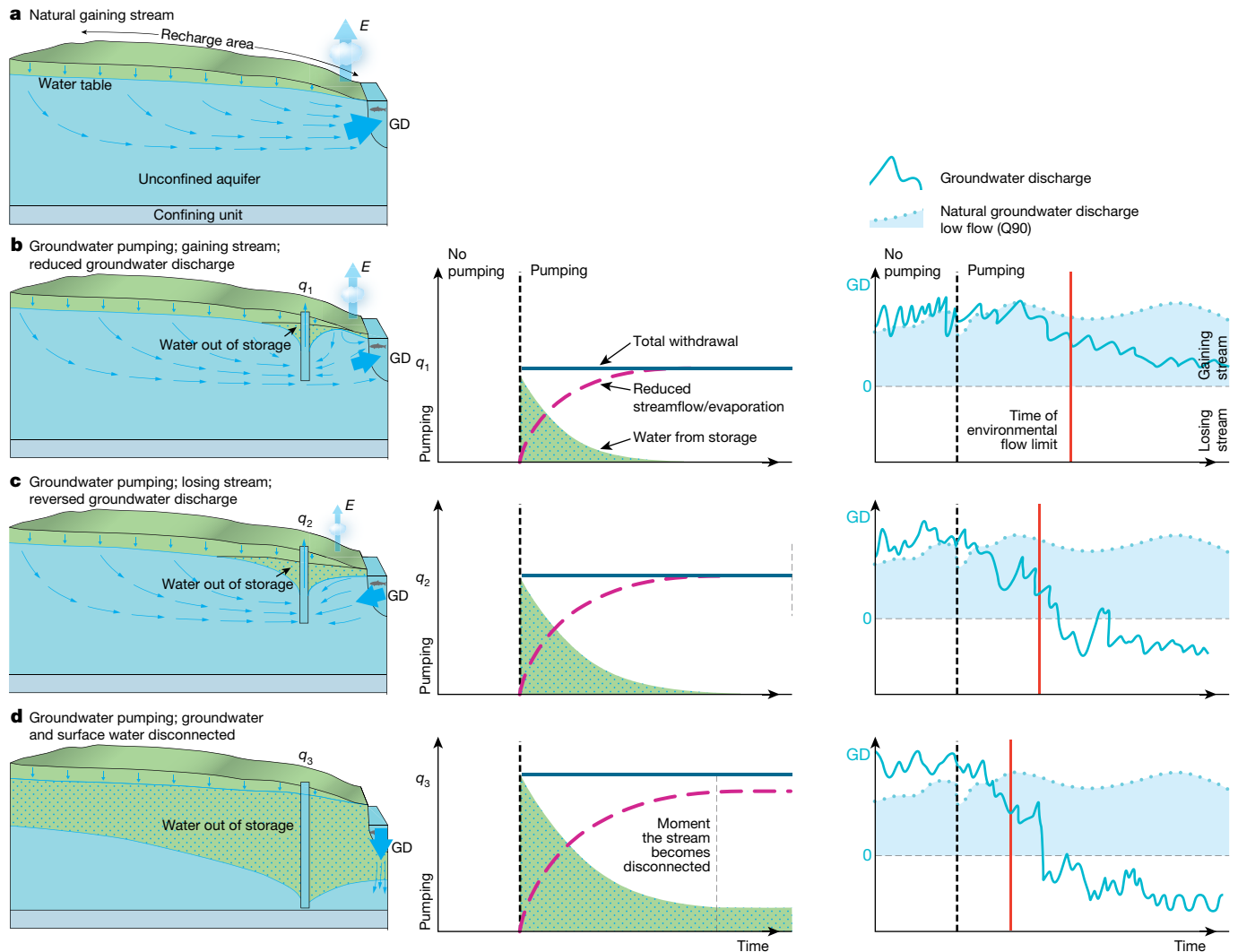


Fig. 3 | Effect of groundwater pumping on groundwater discharge, groundwater storage change, reaching environmental critical streamflow and possibly regime changes. **a**, A natural gaining stream. **b**, Left, limited pumping rate (q_1), reducing groundwater discharge (GD). At first, groundwater is taken out of storage. Middle, eventually a new equilibrium is reached where all pumped water comes from reduced groundwater discharge and evaporation. Right, streamflow is reduced and environmentally critical streamflow can be exceeded, but the stream is still gaining. **c**, Left, higher pumping rates (q_2), reversing groundwater discharge. Middle, more groundwater is taken out of storage, but again a new equilibrium is reached. Right, pumping instead results in surface

water infiltration. **d**, Left, even more intense pumping rates (q_3), leading to a disconnection of the groundwater and surface water systems. Surface water infiltration reaches a maximum, independent of groundwater depth. Middle panel, groundwater is persistently taken out of storage leading to a continuous lowering of the water table at a faster rate if pumping rates are higher than surface water infiltration and diffuse recharge over the depression cone. Right panel, further declines in groundwater level will not affect streamflow further. Left and middle panels of **a–d** are modified from United States Geological Survey (USGS) publications^{24,25}. *E*, rate of evapotranspiration.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1594-4>.

Received: 19 June 2018; Accepted: 16 July 2019;
Published online 2 October 2019.

- Giordano, M. Global groundwater? Issues and solutions. *Annu. Rev. Environ. Resour.* **34**, 153–178 (2009).
- Döll, P. & Siebert, S. Global modelling of irrigation water requirements. *Wat. Resour. Res.* **38**, 8-1–8-10 (2002).
- Siebert, S. & Döll, P. Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. *J. Hydrol.* **384**, 198–217 (2010).
- Gleeson, T. & Richter, B. How much groundwater can we pump and protect environmental flow through time? Presumptive standards for conjunctive management of aquifers and rivers. *River Res. Appl.* **34**, 83–92 (2018).

- Rodell, M., Velicogna, I. & Famigletti, J. S. Satellite-based estimates of groundwater depletion in India. *Nature* **460**, 999–1002 (2009).
- Aeschbach-Hertig, W. & Gleeson, T. Regional strategies for the accelerating global problem of groundwater depletion. *Nat. Geosci.* **5**, 853–861 (2012).
- Wada, Y., Wisser, D. & Bierkens, M. F. P. Global modelling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth Syst. Dynam.* **5**, 15–40 (2014).
- Taylor, R. G. et al. Groundwater and climate change. *Nat. Clim. Chang.* **3**, 322–329 (2013).
- Konikow, L. F. & Kendy, E. Groundwater depletion: a global problem. *Hydrogeol. J.* **13**, 317–320 (2005).
- Lawford, R. et al. Basin perspectives on the water–energy–food security nexus. *Curr. Opin. Environ. Sustain.* **5**, 607–616 (2013).
- Postel, S. L. Entering an era of water scarcity: the challenges ahead. *Ecol. Appl.* **10**, 941–948 (2000).
- de Graaf, I. E. M., van Beek, L. P. H., Wada, Y. & Bierkens, M. F. P. Dynamic attribution of global water demand to surface water and groundwater resources: effects of abstractions and return flows on river discharges. *Adv. Water Resour.* **64**, 21–33 (2014).
- van Beek, L. P. H., Wada, Y. & Bierkens, M. F. P. Global monthly water stress: 1. Water balance and water availability. *Wat. Resour. Res.* **47**, W07517 (2011).

14. de Graaf, I. E. M. et al. A global-scale two-layer transient groundwater model: development and application to groundwater depletion. *Adv. Water Resour.* **102**, 53–67 (2017).
15. Sutanudjaja, E. H., et al. PCR-GLOBWB 2: a 5 arc-minute global hydrological and water resources model. *Geosci. Model Dev.* **11**, 2429–2453 (2018).
16. Ficklin, D. L., Robeson, S. M. & Knouft, J. H. Impacts of recent climate change on trends in baseflow and stormflow in United States watersheds. *Geophys. Res. Lett.* **43**, 5079–5088 (2016).
17. Weedon, G. P. et al. Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J. Hydrometeorol.* **12**, 823–848 (2011).
18. Hempel, S., Frieler, K., Warszawski, L., Schewe, J. & Piontek, F. A trend-reserving bias correction—the ISI-MIP approach. *Earth Syst. Dynam.* **4**, 219–236 (2013).
19. World Water Assessment Programme *The United Nations World Water Development Report 4: Managing Water under Uncertainty and Risk*. Report No. 978-92-3-104235-5, 407 (UNESCO, 2012).
20. Döll, P., Fiedler, K. & Zhang, J. Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrol. Earth Syst. Sci.* **13**, 2413–2432 (2009).
21. India Central Ground Water Board *Ground Water Yearbook 2013–14* <http://cgwb.gov.in/Documents/Ground%20Water%20Year%20Book%202013-14.pdf> (2014).
22. Perrone, D. & Jasechko, S. Dry groundwater wells in the western United States. *Environ. Res. Lett.* **12**, 104002 (2017).
23. Lehner, B., Verdin, K. & Jarvis, A. New global hydrography derived from spaceborne elevation data. *Eos* **89**, 93–94 (2008).
24. Winter, T. C., Harvey, J. W., Franke, O. L. & Alley, W. M. *Ground Water and Surface Water: A Single Resource*. Circular 1139 (United States Geological Survey, 1998).
25. Alley, W. M., Reilly, T. E. & Franke, O. L. *Sustainability of Groundwater Resources*. Circular 1186 (United States Geological Survey, 1999).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

Groundwater–streamflow interaction under groundwater pumping. Figure 3 provides a schematic overview of groundwater–streamflow dynamics affected by groundwater pumping. Figure 3a shows the natural situation of a gaining stream, in which groundwater discharge contributes substantially to streamflow. Figure 3b–d schematically shows the effect of wells with increasing pumping rates, assuming the same hydrogeological environment. The left and middle columns are typical groundwater impact model conceptualizations^{24,25}, while the right-hand images are the ‘next step’ we have made in this study, linking groundwater and surface water systems.

Figure 3b shows a pumping regime in which the pumping rate (q_1) is limited. Just after pumping starts, the principle source of water to a well comes out of groundwater storage of the porous medium. As time passes, the groundwater table starts to develop a gradient towards the pumping well and part of the groundwater recharge—that otherwise would have supplied water to the stream—now contributes to the pumped water (Fig. 3b, left). Consequently, groundwater discharge to the stream is reduced. Also, evapotranspiration (E) from groundwater-dependent vegetation will decrease owing to falling groundwater levels. The deeper the groundwater table falls, the larger the contribution of reduced groundwater discharge and reduced evaporation (which, together with increased recharge, is ‘increased capture’) to pumped water becomes, until a new equilibrium is reached where all pumped water comes out of capture (Fig. 3b, middle) and groundwater levels stabilize again. Figure 3b, right shows the impact of limited pumping (q_1) on groundwater discharge to the stream (solid line). When pumping starts, the groundwater discharge is not affected directly, because most of the pumped groundwater comes out of storage (Fig. 3b, middle). However, as time progresses, groundwater discharge starts to decrease. Here, in the case of limited pumping rates, groundwater discharge will remain positive and the stream will remain a gaining stream, albeit with lower discharge rates.

If the pumping rate is higher (q_2 , Fig. 3c), groundwater levels drop further, evaporation is further reduced, and groundwater discharge may shift to surface water infiltration, making the stream a losing stream (Fig. 3c, left). The contribution of recharge from infiltrating river water increases with falling groundwater levels, as long as the groundwater level and rivers are connected. Also, in this case with higher pumping rates, a new equilibrium may be reached, in which all pumped water comes from streamflow infiltration and decreased evapotranspiration (Fig. 3c, middle, right). The scenario illustrated in Fig. 3c is a typical example in which groundwater abstraction already markedly affects streamflow but does not yet lead to major losses in groundwater storage, and hence groundwater depletion.

If the pumping rate is even higher (q_3 , Fig. 3d) and groundwater levels drop below the river bed, groundwater becomes disconnected from the stream. This situation can be seen as a critical threshold, as the stream recharge rate remains almost constant when groundwater levels drop further, which induces an acceleration of groundwater decline and successive reduction of evaporation. This can be readily seen by calculating sensitivities of the infiltration flux to surface water and groundwater levels using Darcy’s law in Fig. 3c, d. Denoting I the infiltration flux from the stream, s the surface water level and h the groundwater level, it follows that in the case of a connected stream (Fig. 3c) the infiltration flux is equally sensitive to s and h : $\partial I/\partial s = \partial I/\partial h = \text{constant}$, whereas in the case of a disconnected stream (Fig. 3d) we have $\partial I/\partial s \propto 1/h$ and $\partial I/\partial h \propto s/h^2$.

Moreover, if the pumping rate becomes higher than the maximum stream infiltration rate and higher than the recharge over the depression cone (as often the case in areas with predominantly groundwater dependent irrigation), the excess rate of pumping will come out of storage. As a result, groundwater levels will continue to decline and groundwater storage will be persistently depleted (Fig. 3d, middle), while further lowering of groundwater levels has only a limited effect on infiltration from the stream (Fig. 3d, right).

The actual response time for the groundwater levels to reach a new equilibrium in the scenarios shown in Fig. 3b and Fig. 3c depends on hydrogeological properties and the dimensions of the aquifer and boundary conditions (that is, the well location and pumping rate, river levels and precipitation surplus) and can be years to decades²⁶. Given that groundwater withdrawal for irrigation often occurs in semi-arid areas with little to no precipitation surplus during the growing season²⁷, the regime shown in Fig. 3d is prominent for these regions and progressive groundwater depletion is the rule. In many regions of the world, groundwater is pumped from deeper (semi-)confined aquifers¹⁴. Under confined conditions, groundwater–streamflow interaction occurs only for larger rivers that are deep enough to penetrate the confining layer or move slowly through the confining layer. Moreover, because the storage coefficients of (semi-)confined aquifers are much smaller than the specific yield of phreatic aquifers, head declines in (semi-)confined aquifers are much larger than the declines in phreatic aquifers. The complex dynamics shown in Fig. 3 and the added complexities that occur in the case of confined aquifers illustrate that, in order to correctly simulate the effects of groundwater pumping on streamflow in general and environmentally critical

flows in particular, dynamic coupling between surface water and groundwater in the model is imperative.

Extended Data Fig. 5 shows maps of the occurrence of the different types of groundwater–surface water interactions of the USA for 1964 and 2004 under natural conditions and under conditions including human water withdrawal (Extended Data Fig. 5a). The spatial patterns, as expected, show predominantly gaining streams in the wetter east and in valleys close to mountain areas, disconnected streams in the dryer west and south, and losing and a small area of intermittently disconnected streams in the transition zones. It is difficult from the spatial pattern to detect any differences in time, or between the runs with water withdrawal and without withdrawal (the ‘natural’ runs). Therefore, Extended Data Fig. 5b shows time series of changes in total area covered by the various interaction types from 1960 to 2005 (as five-year averages). These time series show a slightly rising trend in total area with gaining streams, which is likely to be caused by increased rainfall. Otherwise, the introduction of water withdrawal increases the total area containing losing streams and intermittently disconnected streams at the expense of the total area containing gaining streams. Differences between the natural run and the run including water withdrawal do not increase much over time because irrigation water withdrawal was already substantial before the 1960s in this part of the world, as is also evident from the timing of reaching environmental flow limits (Fig. 1). Differences between the natural run and the one with water withdrawal are very small for the continuously disconnected streams because they do not permit surface water withdrawal and groundwater withdrawal does not change the classification of these streams.

Global-scale surface water–groundwater model and model runs. We used a physically based GSGM that simulates hydrological processes at 5 arc-minute resolution (approximately 10 km × 10 km at the equator). The model consists of the global hydrology and water-resources model PCR-GLOBWB^{14,15} that is dynamically (two-way) coupled via groundwater recharge and capillary rise, and by groundwater discharge and river infiltration to a two-layer global groundwater flow model based on MODFLOW¹⁴ that simulates lateral groundwater flow; the hydrological model runs at a daily time step, the groundwater model runs at a monthly time step. The hydrological model includes a water-use module that dynamically allocates sectoral water demand from irrigated agriculture, industries, households or livestock, to withdrawal of desalinated water, groundwater or surface water, based on the availability of these resources¹². Return flows of unconsumed withdrawn water, flowing back to surface water or groundwater resources, are included in the estimate of water availability. The model parameterization is based on the best available global-scale datasets, including hydrogeological information essential for the aquifer parameterization^{28,29}. We refer to ref. ¹⁵ for an extensive description of PCR-GLOBWB and its parameterization and to refs ^{14,30} for a detailed description of the global groundwater model and its hydrogeological schematization. Details about dynamic coupling between PCR-GLOBWB and the groundwater model can be found in a previous work³¹. Our model is, to our knowledge, the first to simulate groundwater–surface water interactions dynamically at the global scale, which is a prerequisite for analysing the effects of groundwater withdrawal on streamflow discharge. Two important parameters for the calculation of groundwater–surface water interactions are the river bottom elevation and the river bed conductance³¹ (Fig. 3). The river bottom elevation is estimated by subtracting the estimated channel depth from surface elevation, and the river bed conductance from the wetted perimeter. The channel wetted perimeter and depth are estimated by combining Lacey’s and Manning’s formulae, assuming a rectangular channel, and bankfull discharge calculated from the simulated river discharge^{30,31}. For surface elevation we used data from HydroSHEDS²³ and Hydro1k³². Model outcomes have been extensively validated against observed river discharges and heads, and additionally, in this study, to water table depths, fluctuations and declines, and timing of environmental flow limits, showing good results^{12–15,30,31} (see Methods section ‘Evaluation of simulated groundwater heads, head declines and trends’).

We ran the GSGM with past and future climate forcing (over 1960–2010 and 2011–2100, respectively); once with groundwater and surface water withdrawal and once without (a natural run). For the period 1960–2010 we used the WATCH forcing dataset¹⁷ for the meteorological input of the model and gridded sectoral demand data⁷ (from irrigation, industry, domestic and livestock) to estimated water withdrawals. We projected future water withdrawals by assuming a ‘business-as-usual’ scenario, in which industrial and domestic demands, as well as the extent of irrigated areas, stay unchanged after 2010, and where irrigation demands increase or decrease as a result of climate change only. We used CMIP5 RCP8.5 as an emission scenario (the worst-case scenario) and tested sensitivity to climate input by running the GSGM using the results of three global climate models (GCMs) as provided by the Inter-Sectoral Impact Model Intercomparison Project ISIMIP (<https://www.isimip.org/>) and bias-corrected using the WATCH forcing dataset¹⁸. We selected the wettest (GFDL-ESM2M), average (HadGEM2-ES) and driest (MIROC-ESM-CHEM) model outcomes in terms of projected future global precipitation change. The ‘business-as-usual’ scenario reduces uncertainty

associated with projections of future population growth, socio-economic development and technological development that opts for more sustainable water use. Similarly, the ‘business-as-usual’ scenario provides a more positive view of the future world than what may be expected, as we omit future increases in water demand and withdrawal resulting from population growth and economic development.

Calculation of environmentally critical flow and environmental flow limits. To calculate the first time that environmental flow limits have been, or will be, reached necessitates estimating the environmentally critical streamflow threshold as well as frequency criterion of how often streamflow can cross the threshold before the environmental flow limit is reached.

We ran the GSGM with past and future climate forcing; once with groundwater and surface water withdrawal and once without (a natural run). For the natural run, the environmentally critical streamflow threshold was estimated for every grid cell and year (1965–2099) as the Q90 of monthly groundwater discharge applying a five-year window over the past five years. The Q90 is used as a low-flow index and indicates the groundwater discharge needed to sustain a minimal flow required for aquatic habitats⁴; it means that for 90% of the months (that is, 54 of the 60 months in the five-year window) groundwater discharge is above low-flow conditions. Next, we estimate the timing of environmental flow limits reached by evaluating simulated monthly groundwater discharges against the estimated Q90s. To separate out limits reached by changes in climate only (and not by pumping) we estimated the timing of environmental flow limits reached both for the natural run and human-impacted run and removed the limits that were reached in both the natural and human-impacted runs to exclude the limits that could not be associated with pumping. We defined the timing of reaching the environmental flow limit as the moment in time at which monthly groundwater discharge falls 10% below the natural Q90 for at least three months a year (in general the summer months/dry months of the year), for at least two consecutive years. The value of 10% was based on previous environmentally required streamflow criteria⁴. The three months criterion was based on a simple probability calculation (three months in a row represents a probability of about 0.01); a statistically significant alteration that is likely to suggest a regime change. The two consecutive years criterion is motivated by the assumption that at least two dry years, with increased groundwater pumping, are needed before water management strategies are changed. We also refer to Fig. 3b, right for an explanation of the time the environmental flow limit is reached. The width of the moving year window used to estimate the Q90s and the required number of consecutive years of exceedance, were subjected to sensitivity analysis (see Methods section ‘Model sensitivity to the definition of environmental flow requirements’).

For visualization of the results at the global scale, we averaged the timing of the reached environmental flow limits per grid cell over the sub-watershed levels defined by the HydroBASINS database (level 7)²³. For the estimation of the percentage of watersheds that have exceeded the environmental flow limit we included all watersheds with an average groundwater withdrawal (over 1960–2099) exceeding 0.01 m³ per m² per year.

Evaluation of simulated groundwater heads, head declines and trends. We first evaluated observed versus simulated hydraulic heads at the sub-watershed scale (Extended Data Fig. 6). The observation dataset³³ provides time-averaged head observations of 1,603,871 stations worldwide. These observations were averaged over the sub-watersheds and compared to the simulated averages. It should be noted that, in general, the observed data are biased towards river valleys, coastal regions and regions where productive aquifers occur. At the sub-watershed level this means that, in general, observations are made for regions of lower elevation (and thus lower hydraulic heads) and higher-elevated regions are under-represented in the observations. Additionally, we estimated groundwater depletion from the simulated groundwater declines and aquifer specific yields and/or storage coefficients (similar to ref. ¹⁴; Extended Data Fig. 2).

Second, we evaluated interpolated reported head declines over three well monitored, heavily exploited aquifers; the Central Valley and High Plains aquifers in the USA³⁴ and the Upper Ganges and Indus basin, India³⁵ (Extended Data Fig. 3).

Third, we compared water table observations for the High Plains and Central Valley aquifer systems (<https://www.usgs.gov/mission-areas/water-resources>) to the model results (Extended Data Fig. 7). Observations were included in the analysis if at least two observations per year were available covering at least five years within the period 1960–2010. For the Central Valley, 105 wells (of 120) were included, covering 17 sub-watersheds (of 20); for the High Plains aquifer 276 wells (of 910) were included, covering 28 sub-watersheds (of 44). Well data were clustered over the sub-watersheds and, per well, average heads, standard deviations and head decline were calculated, and also averaged over the sub-watershed level.

The scatter in Extended Data Fig. 6a shows that simulated hydraulic heads match the observed values reasonably well ($R^2 = 0.86$). The residuals (the differences between the observed and simulated heads) are large for some catchments but are small in many other catchments, and are proportional to the magnitude of the simulated head. The largest residuals are found in the catchments where

the simulated head values are large. These are predominantly mountain catchments where the simulated groundwater level is deep, groundwater–surface water interaction is limited and no groundwater withdrawal occurs. The smallest residuals are obtained for catchments where the simulated head values are small. These represent low-lying catchments with alluvial aquifers, where groundwater tables are shallow and groundwater pumping will affect groundwater discharge. Extended Data Fig. 6b shows a histogram of the relative residuals and Extended Data Fig. 6c the global spatial distribution of relative residuals at the catchments scale. Approximately 65% of the watersheds have a relative error between –20% and 20%, and 83% between –50% and 50%.

The estimated cumulative depletion (Extended Data Fig. 3) shows some of the well known groundwater depletion hotspots^{27,36,37}, such as the Upper Ganges and Indus basin, the north China plain, the Central Valley of California, USA, and the High Plains aquifer. For the period 1960–2010 our depletion estimates are similar to estimates for specific aquifer regions where observations are available^{9,34,35,38} and to indirect estimates following from gravimetric anomalies^{5,37}. The global model is able to capture the spatial pattern of groundwater level decline and recovery well (Extended Data Fig. 3), although locally head declines can be over- or underestimated. However, simulated head declines are within the same order of magnitude as the observations. The exceptions are parts of the northern High Plains, where recovery is underestimated, and parts of the upper Ganges, where head declines are overestimated. This local mismatch is likely to be driven by the uncertainty in hydrogeological parameter sets (the subject of previous uncertainty analyses^{14,30}) and/or the allocation of water demand to groundwater and surface-water resources (also the subject of previous uncertainty analyses^{12,15}). These uncertainties in hydrogeological parameterization and demand allocation are not easy to overcome yet, as the currently used hydrogeological information comes from the best available global datasets, but with their own uncertainties. Results are affected by uncertainties in aquifer transmissivities resulting in uncertainty in estimated lateral flows and specific yields and storage of aquifer systems. This affects estimated groundwater head locally. Water allocation is, for the most part, driven by surface water availability, which is mainly driven by climate inputs coming from the climate model used, with its own uncertainties in estimated precipitation and temperature.

Most important for the timing of the environmental limit is the head decline. For the Central Valley aquifer, simulated head declines matches to observed declines well (Extended Data Fig. 7). For the south part of this aquifer, a slight overestimation of head declines is found, which could lead to an environmental limit being reached too soon. Water table depths and standard deviations are captured well by the model, resulting in a R^2 for water table depths of 0.80 (when watersheds with only one observation are excluded; 0.60 if all wells are included). For the High Plains, larger deviations in head declines are present (with a R^2 for water table depths of 0.4; Extended Data Fig. 7). For the south, head declines are, in general, underestimated, possibly leading to an estimated environmental limit being reached too late. For the middle and north of the aquifer, head declines are better matched, although the average water table depths are overestimated in the north by the model. However, as we estimate the environmental flow limit by comparing the natural simulation run to the human-impacted simulation run in terms of decline only (the absolute level is not very important), the uncertainty in water table depth will probably not affect the estimation of environmental limits to any appreciable extent (see Methods section ‘Evaluation of groundwater discharge and environmental flow limits’, below). The statistics for the individual wells illustrate the wide spread in observations, which shows that directly comparing global-scale models with well data is challenging.

Evaluation of groundwater discharge and environmental flow limits. First, as observations of groundwater discharge (used for the estimation of environmentally critical flow as defined in this study) are unavailable, we benchmark our results against results from a detailed (grid size, 1 square mile \approx 2.59 km²) and calibrated regional groundwater model for the Republican River Basin in central USA³⁹ (Extended Data Fig. 8). Both models—the global and the regional model—use MODFLOW and have similar model structures and setup. The regional model is assumed to be more accurate locally than the global-scale model and the objective of this evaluation is to show that, globally, we are able to represent the relevant processes well enough and able to capture the same trends and anomalies as more detailed regional-scale calibrated models. As the calibrated model used here did not provide groundwater discharge at the grid cell resolution nor a natural run, model results could not be used to estimate environmentally critical streamflow and the time the environmental flow limits were reached.

Second, we indirectly compared our estimates of the first time environmental flow limits were reached to streamflow observations. We used estimates of groundwater discharge by baseflow separation performed on streamflow data for 50 stations in Kansas, USA, where groundwater pumping is the dominant water use. The streamflow records of these stations covered a period before pumping, from which we could calculate the Q90 of natural groundwater discharge, after baseflow separation from the streamflow record. This Q90 was used to evaluate

the groundwater discharge after pumping, and from this estimate the timing of reaching the environmental limit was evaluated. Results are presented in the main text in the violin plots of Fig. 1c, comparing the observed and simulated limits.

Last, as a proxy to the long-term observed time series, we used modelled groundwater discharges for a limited number of calibrated catchment scale models⁴⁰—similar to the sub-watershed size we used to average our results—spread over the different continents of the world. The calibrated models were run once with groundwater and surface water abstractions and once without. Thus we could perform the same analysis as described in this paper. The results are shown in the scatter in Extended Data Fig. 4a.

Comparing the head declines estimated by a region-specific calibrated model to our global-scale estimate (Extended Data Fig. 8a) shows that, although there are some spatial differences, the global-scale model is able to capture the general spatial patterns and magnitudes well compared to the calibrated model results. Groundwater discharge to the stream is evaluated at different spatial scales (Extended Data Fig. 8b). For the full basin, the trend in groundwater reduction is captured well, and the same holds for the large sub-basins. In the global model we see more variability owing to the parametrization of the subsurface, the coarser resolved river network and the uncertainty in the pumping. Such differences may lead to a different timing for reaching environmental limits (which tends to be sooner). However, this impact may not be large, because the five-year temporal window used to estimate the environmental flow limits accounts for the outliers and/or fluctuation. When we move to the large sub-basins (shown as Regional basin levels 2 and 3 in Extended Data Fig. 8c), the global model trend in groundwater discharge reduction is slightly larger than the simulated trend, possibly leading to the limits being estimated too soon. The groundwater discharges at the smallest sub-basin level (shown as Regional basin level 1 in Extended Data Fig. 8c) are small and did not show significant differences, therefore trends at this level are not shown.

Our timing estimates compare reasonably well with observed timings (Extended Data Fig. 4), except for one location in Kansas, USA, a UK catchment, and the Czech catchment. Note, however, that the latter two estimates are only proxies of real observations and may overestimate the effects of pumping (and thus underestimate the time the limit is reached) because they do not include the delay in the impact of groundwater pumping on surface water discharge that occurs in real aquifer systems (for example, Fig. 3).

Model sensitivity to parameter settings and boundary conditions. In order to assess the robustness of our estimate of the first time environmental flow limits are reached as a result of pumping, we performed a one-at-a-time sensitivity analysis. Because of the large computation times involved with running the GSGM, a full Monte Carlo analysis, varying the parameters randomly, is not feasible. Therefore, we had to limit ourselves to a model sensitivity analysis in which we separately changed the settings of a few key parameters. As the estimated groundwater discharge flux is highly dependent on conductivity of the sub-surface and river drainage level, and additionally on riverbed conductance, we tested how sensitive our estimates are to these different parameter settings. In order to capture the full parameter uncertainty, we chose to vary the parameters across a large range. Eleven runs with different parameter settings were completed (Extended Data Fig. 9a). Hydraulic conductivities were increased or decreased by one order of magnitude, drainage levels were increased and decreased by 50% (based on observed variations in river depth over the USA), and river bed conductance adjusted by factors of 10 and 0.5. The eleven parameter settings were run under natural conditions and with water withdrawal for 1960–2010, and the first times at which the environmental limits were reached was estimated and averaged over the watersheds. The baseline run, as well as the two sets of parameter values that resulted in the minimum and maximum number of watersheds with the limit reached (by 2010), were subsequently run for 2010–2100 using the HadGEM2-ES GCM and ‘business-as-usual’ water demands. The first time at which environmental limits have been, or will be, reached was again estimated and the cumulative frequency shown in Fig. 1b.

Extended Data Fig. 9 shows that, in general, the estimated temporal patterns of the first time at which the environmental flow limits are reached is similar for all runs, meaning all runs react to an increase or decrease in groundwater pumping in the same way. In general, lower conductivities lead to larger head drops, with smaller cones of depression, and higher conductivities lead to smaller head drops but with larger cones of depression. Shallower rivers lead to smaller head gradients and earlier disconnection of the groundwater system and the surface water system, and deeper rivers lead to higher gradients and a longer connection between the groundwater system and the surface water system. Decreasing the river conductance did not substantially change the estimated time of limit reached, increasing the river conductance lead to smaller head gradients and an earlier disconnection of the groundwater system and surface-water system. The topology of the surface-water system, in combination with the spatial heterogeneity of hydraulic conductivity and the patterns of water use, result in nonlinear responses to changes

in parameters when calculating the fraction of watersheds affected, as can be seen by the lines crossing each other in time.

Model sensitivity to the definition of environmental flow requirements. Whether or not environmental flow limits are reached is also dependent on the definition of the environmentally critical streamflow and criteria of how often groundwater discharge is allowed to fall below this threshold (see Methods section ‘Calculation of environmentally critical flow and environmental flow limits’).

We tested the sensitivity of the estimates of the time the environmental flow limit is reached to the statistical criteria used. We varied both the estimate of environmentally critical streamflow (that is, the Q90 of groundwater discharge as a five-year or ten-year running average), and the number of consecutive years needed to reach the environmental flow limit (Extended Data Fig. 10). These results show that estimates of the first time that the environmental flow limits are reached are quite insensitive to its statistical definition. When choosing between a five- and ten-year window to calculate the Q90 we chose a period of five years, as it is more sensitive to changing pumping rates over time. On the basis of the assumption that water management tends to be altered when a water shortage is experienced for at least two years in a row, we use the two-consecutive-year criterion.

Model sensitivity to climate forcing. Estimates of environmentally critical flow and estimated groundwater discharge are dependent on climate input. Therefore, the last part of our sensitivity analysis focused on the sensitivity of the estimated time to reach environmental flow limits to climate forcing. We ran the GSGM using three GCMs (that is, HadGEM2-ES, GFDL-ESM2M and MIROC-ESM-CHEM). The first time environmental limits will be reached was again estimated and the cumulative frequencies of each model run are shown in Fig. 1b.

Globally (see Extended Data Fig. 1a) we see an increase in the depletion trend starting around the 1970s, and rapidly increasing between 2000 and 2010. After 2010, the depletion trends first start to decrease for two GCMs (2010–2030), after which it increases again, but not as rapidly as before. This increase in depletion after 2030 is also reflected in the environmental flow limit peak discussed in the main text. Differences between the GCMs are large—showing that, even under fixed industrial and domestic water demands and fixed irrigated area—climate change, through increased irrigation demand, may have a major impact on future groundwater depletion volumes. Overall, depletion is estimated to increase from 4,190 km³ in 2010 to 7,490 km³ (5,360–10,042 km³) depending on the climate model used.

Despite the uncertainty in depletion volumes, the spatial patterns (Extended Data Fig. 1b, c) of the first time at which the environmental flow limits are reached between the different GCMs are similar (Fig. 1), although with a higher or lower number of catchments reaching the limits for the dryer or wetter GCMs, respectively.

Data availability

All data needed to evaluate the conclusions in the paper are presented in the paper (Figs. 1, 2) and are available through the University of Victoria, <https://doi.org/10.5683/SP2/D7I7CC>. Additional model outputs (as part of the sensitivity analysis and model evaluation presented in the Extended Data) are prohibitively large to store in a repository but are available from the corresponding author on reasonable request.

Code availability

The model code used to run the global-scale surface water—groundwater model is provided through a GitHub repository, https://github.com/UU-Hydro/PCR-GLOBWB_model/tree/develop/modflow.

- Konikow, L. F. & Leake, S. A. Depletion and capture: revisiting “the source of water derived from wells”. *Ground Water* **52**, 100–111 (2014).
- Wada, Y. et al. Global depletion of groundwater resources. *Geophys. Res. Lett.* **37**, L20402 (2010).
- Hartmann, J. & Moosdorf, N. The new global lithological map database glim: a representation of rock properties at the earth surface. *Geochem. Geophys. Geosy.* **13**, Q12004 (2012).
- Gleeson, T., Moosdorf, N., Hartmann, J. & van Beek, L. P. H. A glimpse beneath earth’s surface: Global Hydrogeological MaPS (GLHYMPS) of permeability and porosity. *Geophys. Res. Lett.* **41**, 3891–3898 (2014).
- de Graaf, I. E. M., van Beek, L. P. H., Sutanudjaja, E. H. & Bierkens, M. F. P. A high-resolution global-scale groundwater model. *Hydrol. Earth Syst. Sci.* **19**, 823–837 (2015).
- Sutanudjaja, E. H., van Beek, L. P. H., De Jong, S. M., Van Geer, F. C. & Bierkens, M. F. P. Calibrating a large-extent high-resolution coupled groundwater–land surface model using soil moisture and discharge data. *Wat. Resour. Res.* **50**, 687–705 (2014).
- Verdin, K. L. & Greenlee, S. K. Development of continental scale digital elevation models and extraction of hydrographic features. In *Proc. Third Intl Conf., Workshop on Integrating GIS and Environmental Modeling, 1996* (National Center for Geographic Information and Analysis, 1996).
- Fan, Y., Li, H. & Miguez-Macho, G. Global patterns of groundwater table depth. *Science* **339**, 940–943 (2013).

34. McGuire, V. L. *Water Level Changes in the High Plain Aquifers, Predevelopment to 2007, 2005–06, and 2006–2007*. Scientific Investigations report 2009-5019 (United States Geological Survey, 2009).
35. Williamson, A. K., Prudic, D. E. & Swain, L. A. *Ground-Water Flow in the Central Valley, California*. Professional paper 1401-D (United States Geological Survey, 1989).
36. Konikow, L. F. Contribution of global groundwater depletion since 1900 to sea-level rise. *Geophys. Res. Lett.* **38**, L17401 (2011).
37. Richey, A. S. et al. Uncertainty in global groundwater storage estimates in a total groundwater stress framework. *Wat. Resour. Res.* **51**, 5198–5216 (2015).
38. MacDonald, A. M. et al. Groundwater quality and depletion in the Indo–Gangetic basin mapped from in situ observations. *Nat. Geosci.* **9**, 762–766 (2016).
39. Republican River Compact Administration, Groundwater Model Update 2001–2003 <http://www.republicanrivercompact.org/> (2005).
40. Van Loon, A. F. & Van Lanen, H. A. J. Making the distinction between water scarcity and drought using an observation-modelling framework. *Wat. Resour. Res.* **49**, 1483–1502 (2013).
41. Ahner, B. W. Towards a Representative Scale of the Groundwater Footprint Based on the Ecohydrological Impacts of Groundwater Abstraction, MSc thesis, Georg-August-Universität Göttingen, Göttingen, Germany. (2018)

Acknowledgements We thank S. Rangecroft, Y. Wada, E. Luijendijk, B. Ahner and S. de Kock for generously sharing data that supported the analysis done in this study. This study's calculations were computed on the Dutch national supercomputer Cartesius with the support of SURFsara. This research was funded by the Netherlands Organization for Scientific Research (NWO) under the programme 'Planetary boundaries of the global fresh water cycle'.

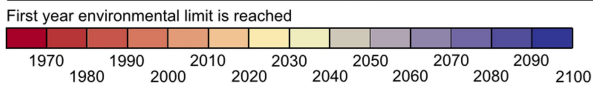
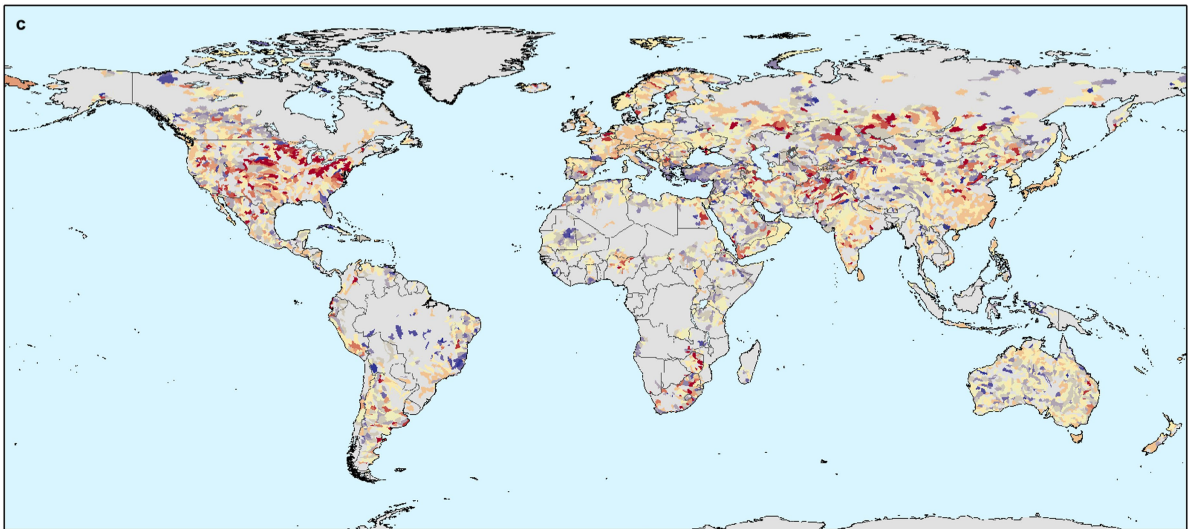
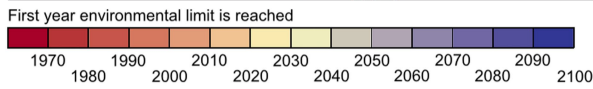
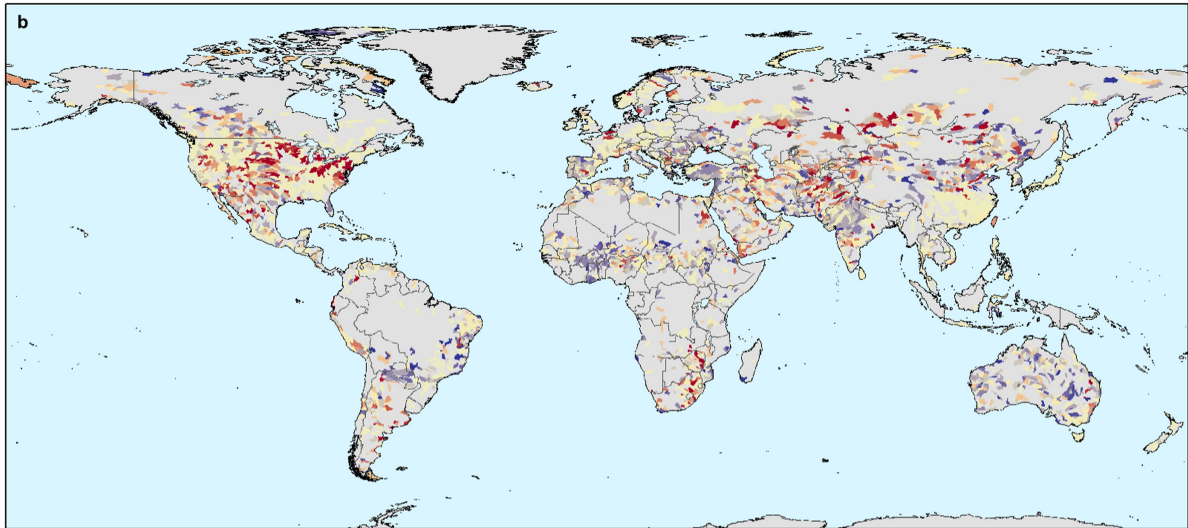
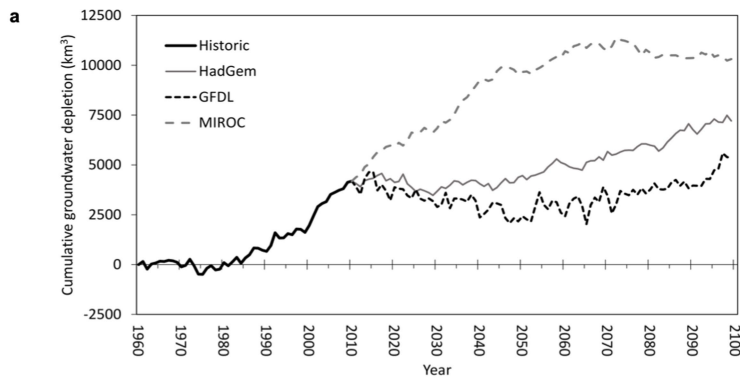
Author contributions I.E.M.d.G., T.G., M.F.P.B. and L.P.H.(R.)v.B. designed the study and wrote the manuscript. I.E.M.d.G. performed the model calculations and data analysis. E.H.S. assisted in the modelling and data analysis and helped improve the manuscript.

Competing interests The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to I.E.M.d.G.
Peer review information *Nature* thanks Laura Foglia, Mary C. Hill and Timothy R. Green for their contribution to the peer review of this work.

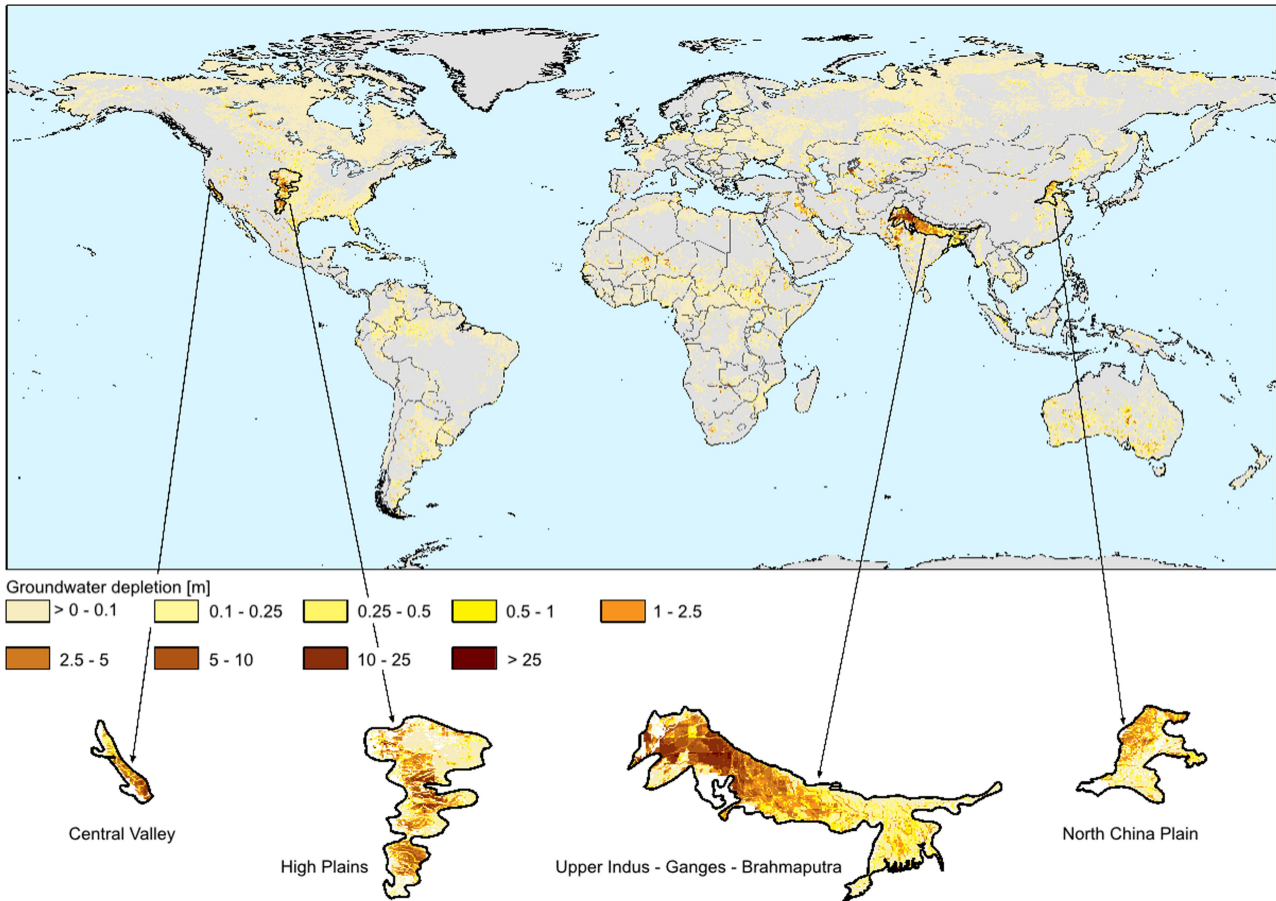
Reprints and permissions information is available at <http://www.nature.com/reprints>.



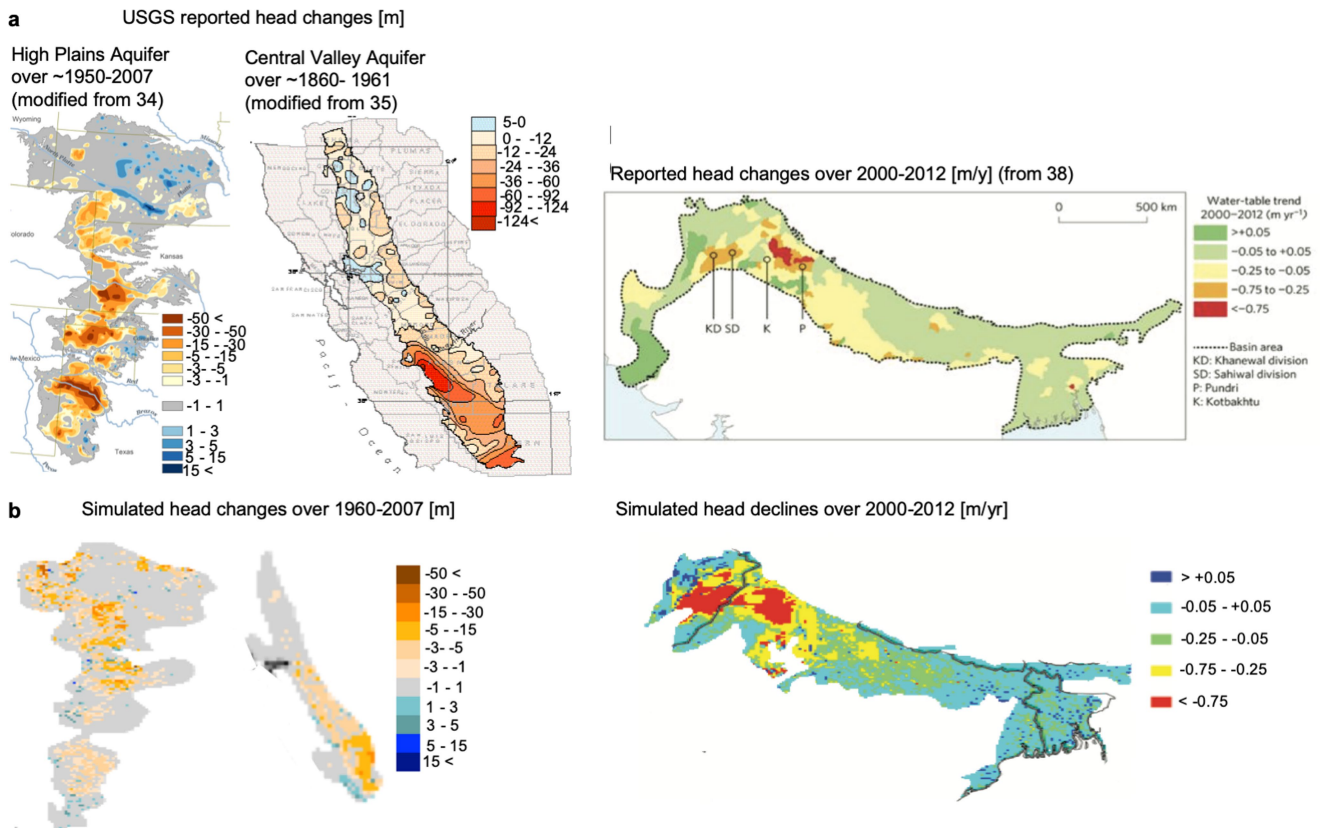
Extended Data Fig. 1 | Model sensitivity to climate forcing.

a, Cumulative groundwater depletion trend (in km³) since 1960 for different climate scenarios. After 2010, the estimate assumes a ‘business-as-usual’ scenario for water demands and uses three GCMs—HadGEM2-HS, GFDL-ESM2M and MIROC-ESM-CHEM—using

RCP 8.5. **b, c**, The first time the environmental flow limit is or will be reached (by year) averaged over watersheds using the sub-watershed level of hydroBASINS²⁸ for the climate scenarios GFDL-ESM2M (**a**) and MIROC-ESM-CHEM (**b**) using RCP 8.5.

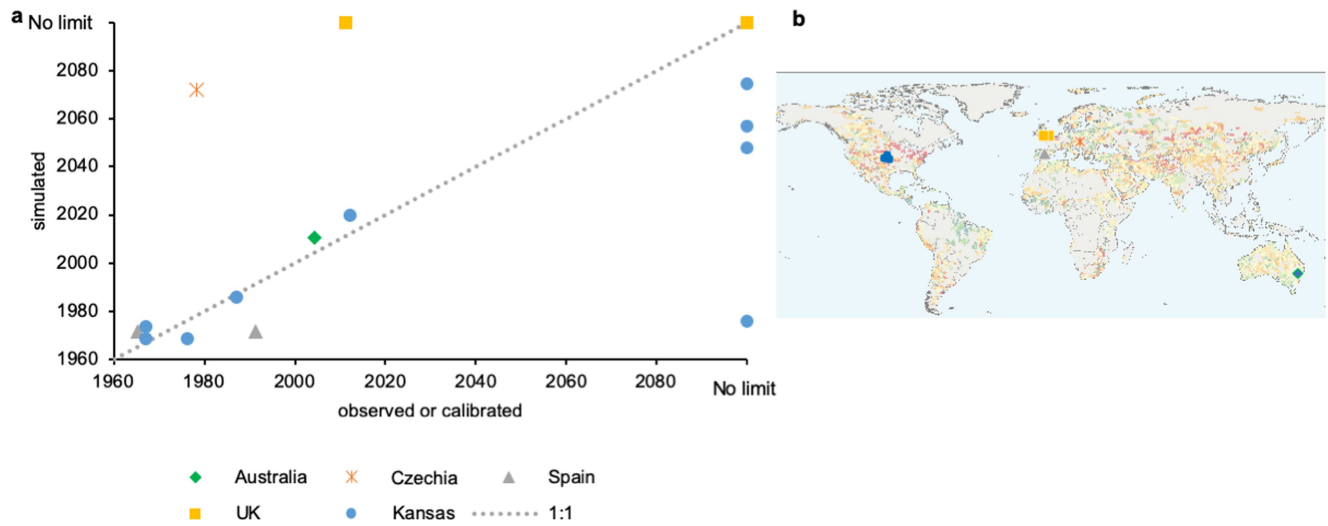


Extended Data Fig. 2 | Gridded estimates of cumulative groundwater water depletion (in m^3 per m^2) for 1960–2099. Four major heavily pumped aquifers are magnified. Aquifer magnifications are from WHYMAP, BGS/UNESCO.



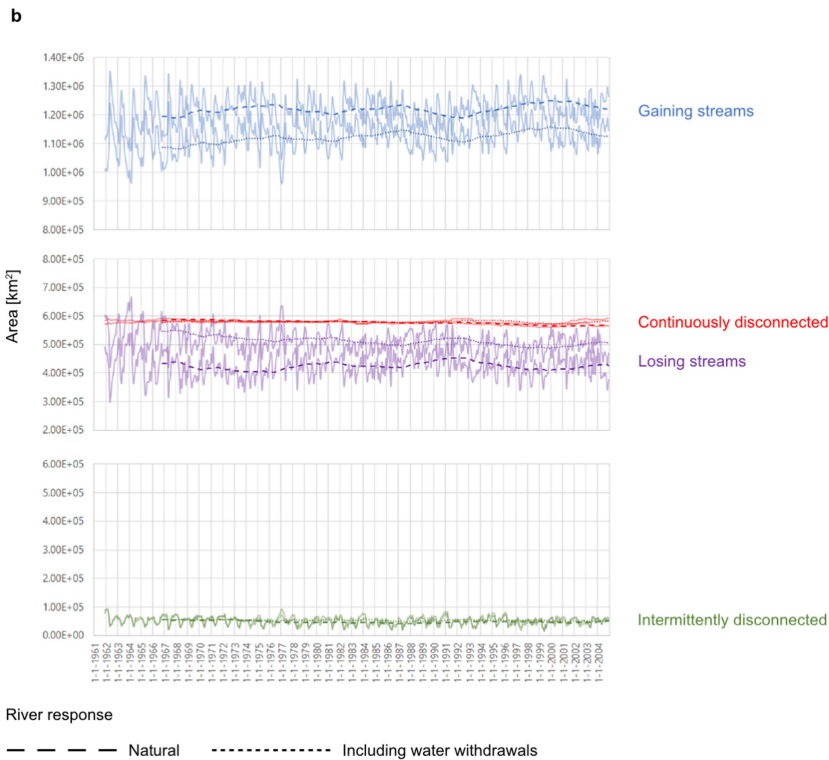
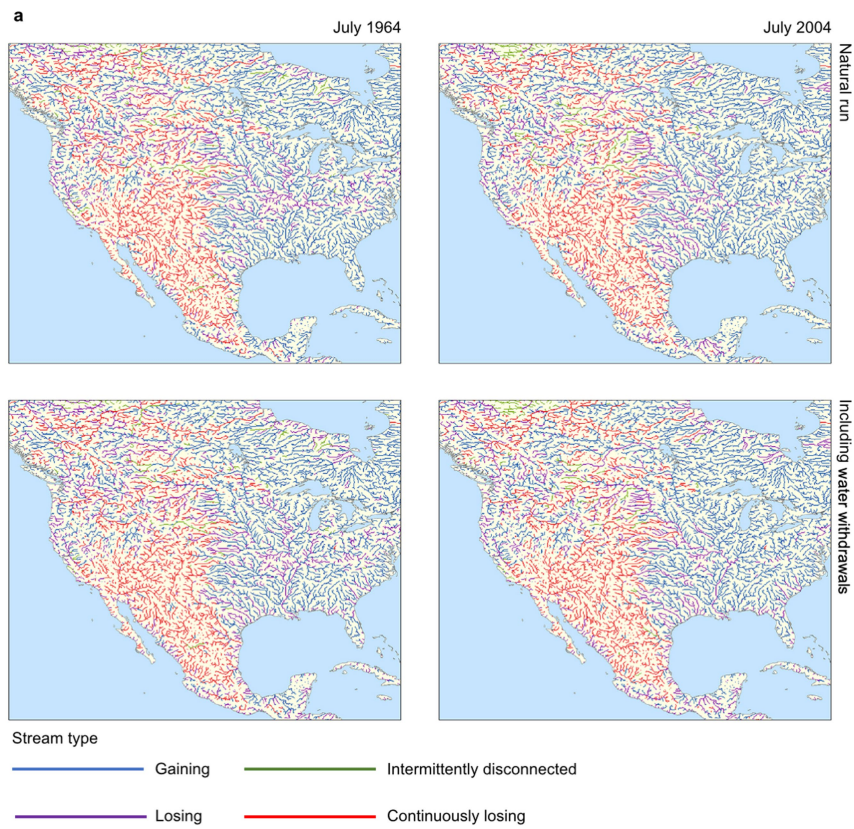
Extended Data Fig. 3 | Model evaluation of simulated groundwater head changes owing to pumping compared to observations. Three of the largest, intensively pumped, and best monitored alluvial aquifer systems of the world are shown: the High Plains aquifer (left) and Central Valley aquifer, USA, (middle) and the Upper Ganges and Indus basin, India

(right). **a**, Observed data and published maps. **b**, This study's estimates. A comparison between **a** and **b** shows that the model results matches the observations well. Nonlinear colour scales are used. The Ganges basin map is from a previous work³⁸.



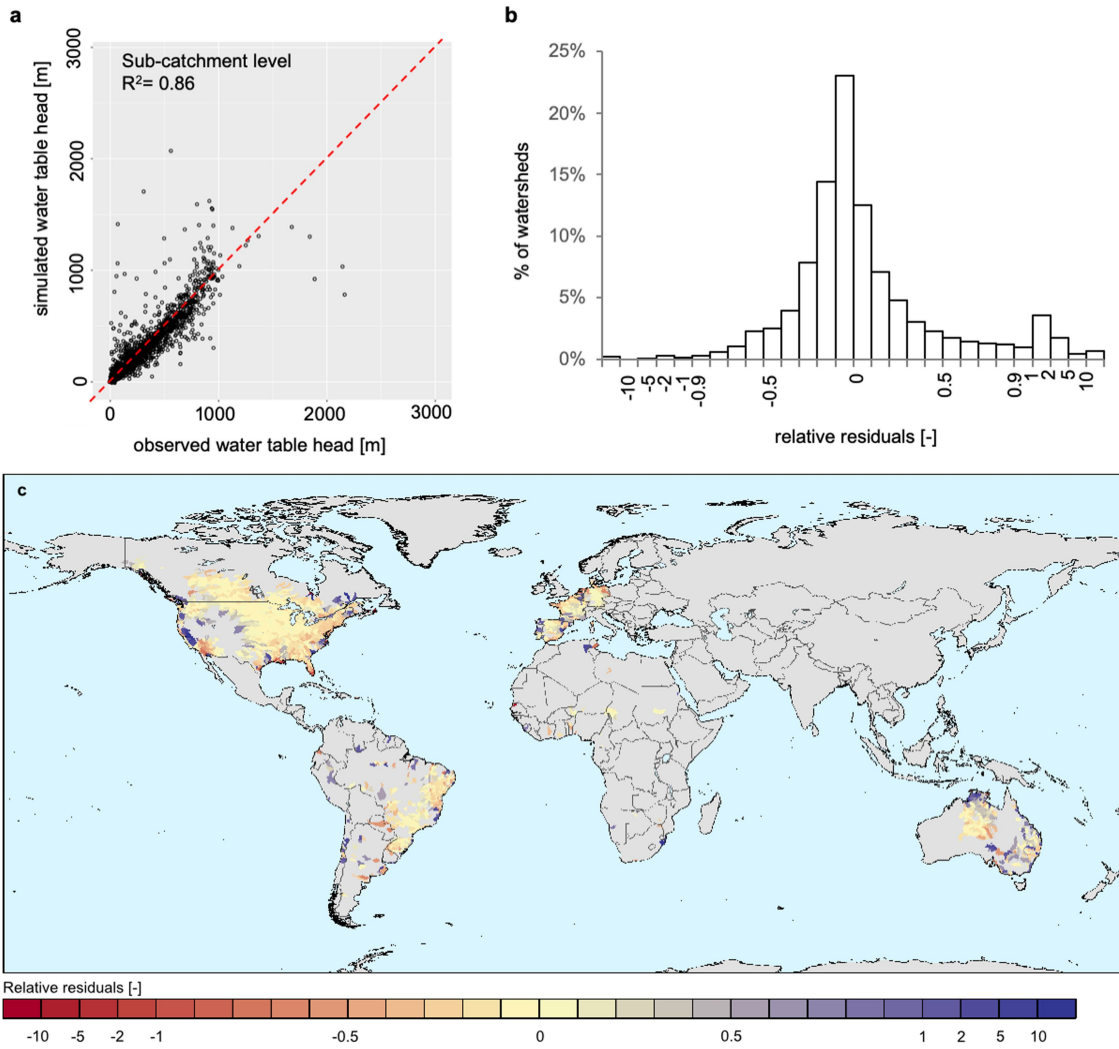
Extended Data Fig. 4 | Model evaluation of the estimated first time that the environmental flow limits are reached. **a**, Observed versus simulated first time that the environmental flow limits are reached (x and

y axes in years) for several groundwater-pumping-impacted catchments. **b**, Locations of the studied catchments are indicated on the map; the three dots in Kansas represent nine sub-catchments used in the analysis.



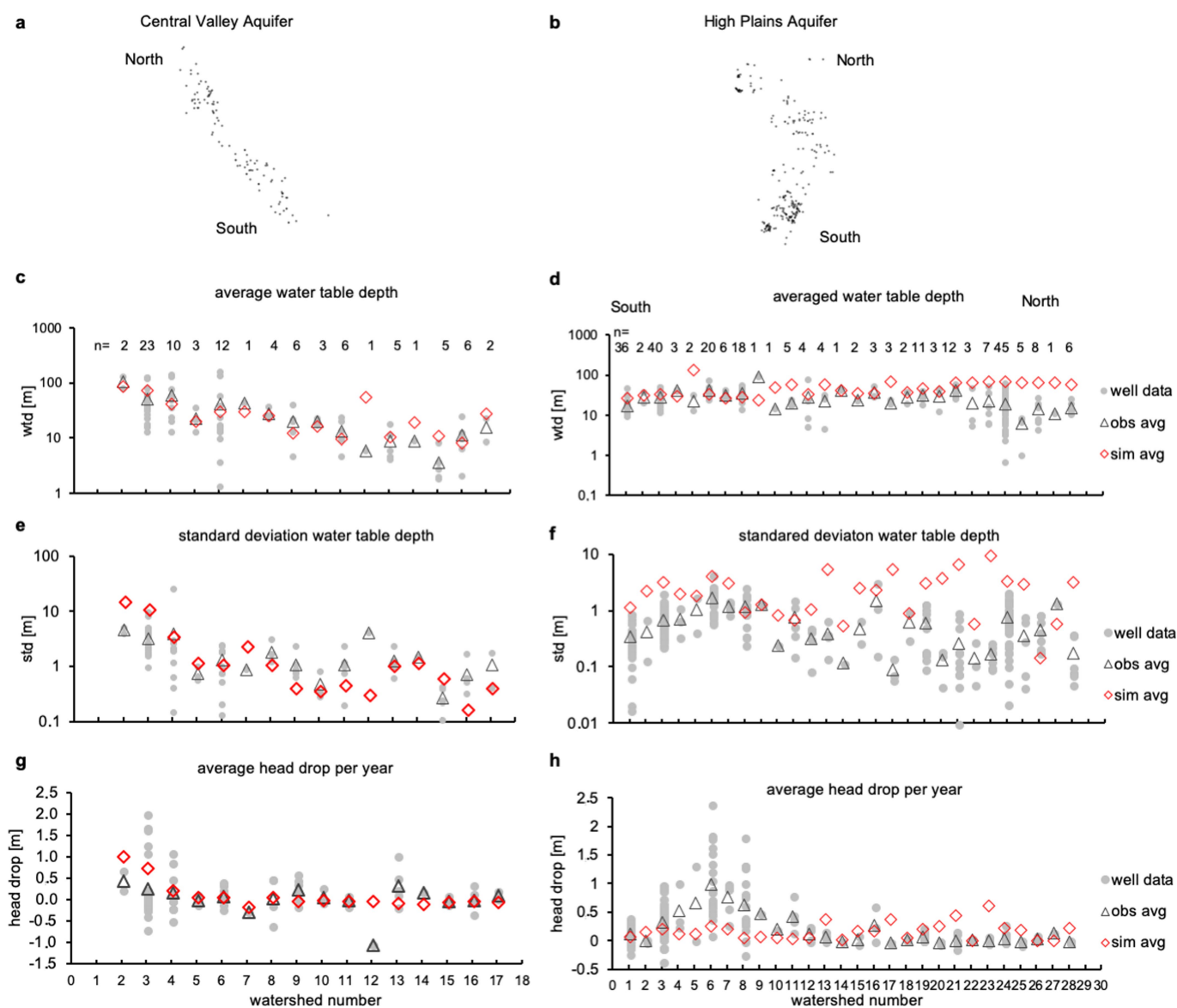
Extended Data Fig. 5 | Distribution of surface water–groundwater interaction classes for north America as simulated using the physically based global-scale GSGM. The figure distinguishes four classes: gaining streams, losing streams, intermittently disconnected streams and continuously disconnected streams. Results are shown for the month of July (generally the driest month of the year, on the basis of monthly discharge) and represent five-year moving averages of river drainage and groundwater levels. A stream is classified as a continuously disconnected

stream if the stream is disconnected for at least two years in a row over the moving average window of five years. **a**, Spatial maps of surface water–groundwater interaction for July 1964 and 2004 under natural conditions and including human water withdrawal. **b**, Temporal variation of the total area covered by each surface water–groundwater interaction class. The thinner continuous lines are yearly values, thicker dashed and dotted lines are the five-year averages.



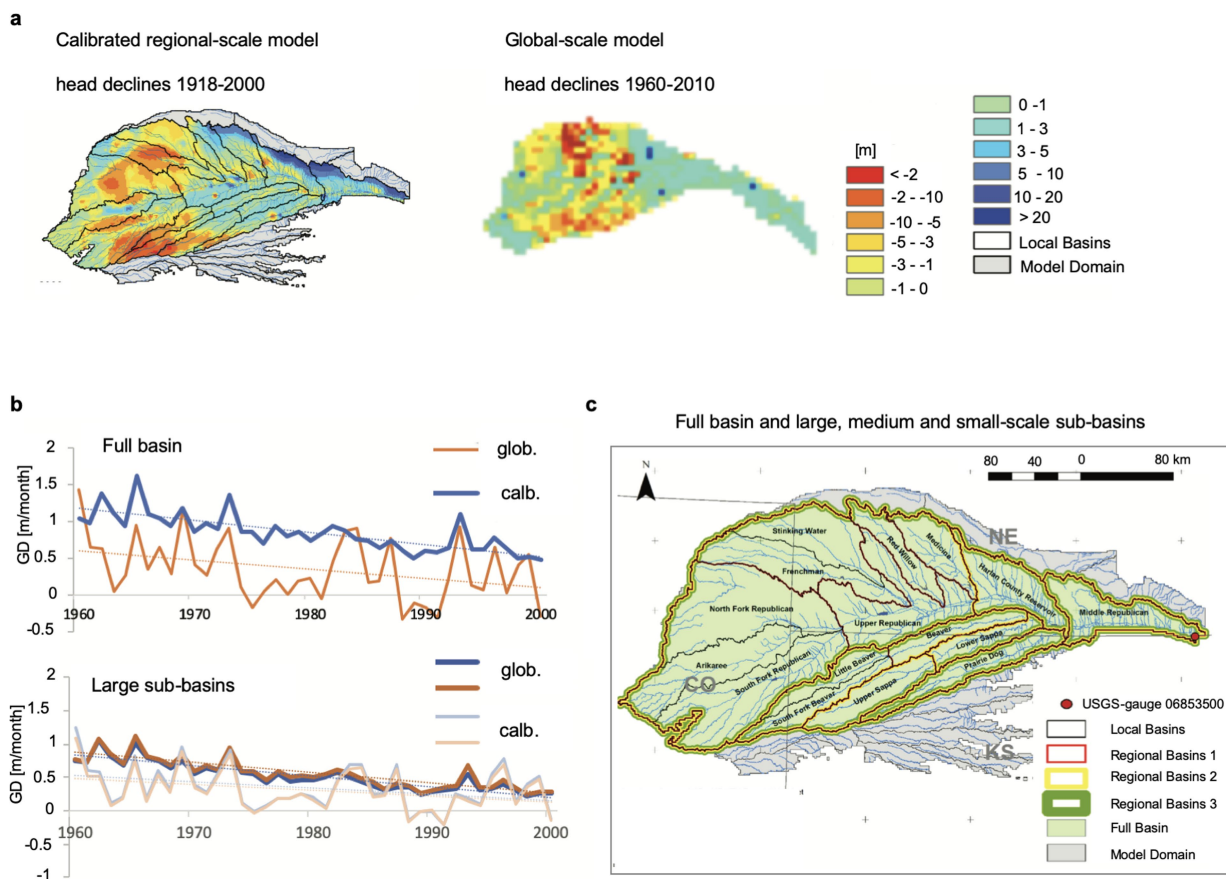
Extended Data Fig. 6 | Evaluation of observed versus simulated water table averaged for sub-watersheds. a, Scatter plot; the red line shows the 1:1 slope. **b,** Histogram of relative residuals, calculated as

$(\text{observed} - \text{simulated})/\text{observed}$. **c,** Global map of relative residuals. All watersheds with no available data are mapped in grey.



Extended Data Fig. 7 | Model evaluation of simulated groundwater table depth to well observations for two well monitored aquifer systems in the USA. a, b, Location of the wells used for this analysis within the Central Valley aquifer system (a) the High Plains aquifer system (b). **c–h,** Averaged water table depths ('wtd') (c, d); standard deviation ('std') of monthly wtd (e, f) and head drops (g, h) were estimated ('obs avg') and

compared to simulated results ('sim avg'). In each plot, the values per well in the watershed are given in grey (well data) and the statistics show the wide spread in observations. The value n indicates the number of wells within the watershed. The watersheds are numbered from north to south over both aquifers, indicated by the watershed number (the exact location of the watersheds is not relevant).



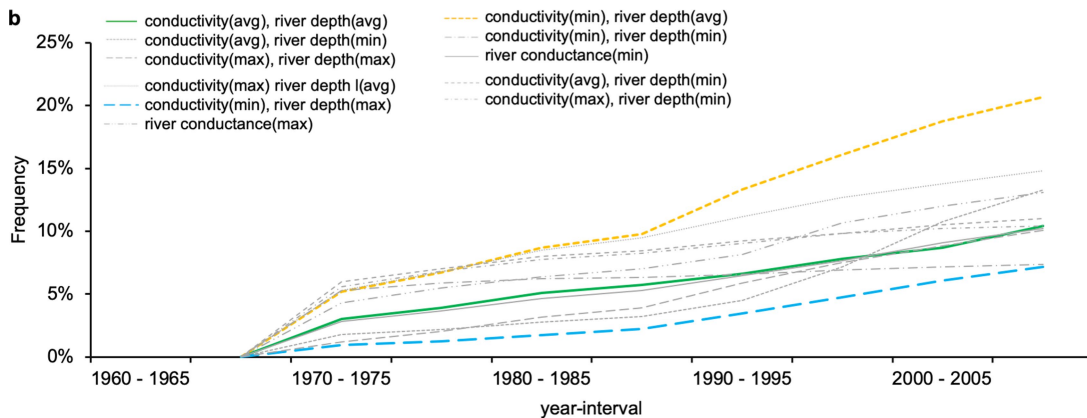
Extended Data Fig. 8 | Inter-scale model comparison. **a, b**, Comparison of groundwater head declines (**a**) and groundwater discharge (GD; **b**) at different spatial levels, simulated by a calibrated regional-scale model ('calb.'): results modified from previous work⁴¹ using the Republican River Groundwater Model³⁹ and the global-scale surface water-groundwater model ('glob.'). **c**, The full basin covers the entire Republican River basin,

which is situated in the central-north of the High Plains aquifer, USA. The larger sub-basins are the level 2 and 3 regional basins that consist of more level 1 basins. In **b**, the solid lines present the simulated groundwater discharge, the dashed lines present the trends of the groundwater discharge. Groundwater discharge trends simulated by both models are comparable.

a

Model Run Name	Parameter value conductivity	Parameter value drainage level	Fraction limits reached by 2010 compared to Baseline
conductivity(min), river depth(min)	bl *0.1	bl *0.5	0.70
conductivity(min), river depth(avg)	bl *0.1	bl	1.98
conductivity(min), river depth(max)	bl *0.1	bl *1.5	0.69
conductivity(avg), river depth(min)	bl	bl *0.5	1.06
conductivity(avg), river depth(avg)	bl	bl	1
conductivity(avg), river depth(max)	bl	bl *1.5	1.28
conductivity(max), river depth(min)	bl *10	bl *0.5	1.00
conductivity(max), river depth(avg)	bl *10	bl	1.42
conductivity(max), river depth(max)	bl *10	bl *1.5	0.97

	Parameter value river conductance	
river conductance (min)	bl *0.5	0.99
river conductance (max)	bl *10	1.2

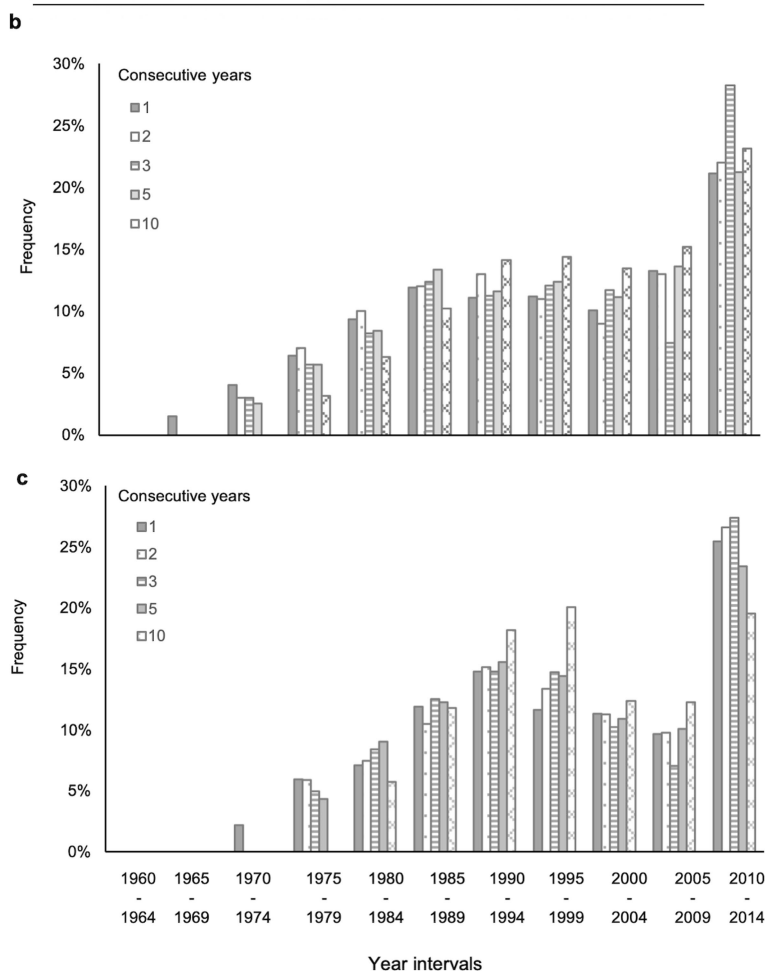


Extended Data Fig. 9 | Model sensitivity to parameter settings and boundary conditions. **a**, Table showing the different parameter settings used; varying the sub-surface conductivity, the river depth and the river conductance by decreasing or increasing the settings of the baseline ('bl') run. The baseline run uses the average parameter settings. **b**, Frequency plot of the first time the environmental flow limits are reached under

different parameter settings for 1960–2010. The fractional increase or decrease of estimated environmental limits compared to the baseline run is given in the fourth column of the table in **a**. The runs with the smallest and largest limits are indicated in bold in the table and presented in blue and yellow, respectively, in the graph.

a

Environmental flow requirements	Consecutive years	Fraction limits reached by 2010 compared to Baseline
Q90 over 5 years	1 year	1.04
Q90 over 5 years	2 years	1
Q90 over 5 years	3 years	1.03
Q90 over 5 years	5 years	1.02
Q90 over 5 years	10 years	1.01
Q90 over 10 years	1 year	0.97
Q90 over 10 years	2 years	0.96
Q90 over 10 years	3 years	0.93
Q90 over 10 years	5 years	0.92
Q90 over 10 years	10 years	0.86



Extended Data Fig. 10 | Model sensitivity to the definition of environmental flow requirements. **a**, Table giving the different criteria of the Q90 windows and consecutive years used to estimate the environmental flow limits. **b**, **c**, Histograms of the limits reached, estimated using the Q90 over five years (**b**) and the Q90 over ten years (**c**).

The fractional increase or decrease of the estimated environmental limits compared to the baseline run (Q90 over five years, for two consecutive years) is given in the third column of the table in **a**. Difference in the estimated environmental flow limits are only limited when using different criteria.