nature geoscience

ARTICLES

PUBLISHED ONLINE: 16 NOVEMBER 2015 | DOI: 10.1038/NGEO2590

The global volume and distribution of modern groundwater

Tom Gleeson[1,2*], Kevin M. Befus[3], Scott Jasechko[4], Elco Luijendijk[2,5] and M. Bayani Cardenas[3]

Groundwater is important for energy and food security, human health and ecosystems. The time since groundwater was recharged—or groundwater age—can be important for diverse geologic processes, such as chemical weathering, ocean eutrophication and climate change. However, measured groundwater ages range from months to millions of years. The global volume and distribution of groundwater less than 50 years old—modern groundwater that is the most recently recharged and also the most vulnerable to global change—are unknown. Here we combine geochemical, geologic, hydrologic and geospatial data sets with numerical simulations of groundwater and analyse tritium ages to show that less than 6% of the groundwater in the uppermost portion of Earth's landmass is modern. We find that the total groundwater volume in the upper 2 km of continental crust is approximately 22.6 million km^3, of which 0.1–5.0 million km^3 is less than 50 years old. Although modern groundwater represents a small percentage of the total groundwater on Earth, the volume of modern groundwater is equivalent to a body of water with a depth of about 3 m spread over the continents. This water resource dwarfs all other components of the active hydrologic cycle.

The inventory of groundwater ages on Earth is largely unknown. Groundwater ages in excess of a million years have been documented in desert regions[1], whereas groundwater ages of months to decades are common in more humid regions. Groundwater of various ages also commonly mix in aquifers, which results in a distribution of groundwater ages for even a small groundwater sample and complicates measurement interpretation[2–6]. Our objective is to quantify the global volume and distribution of young groundwater and provide an updated estimate of the total volume of crustal groundwater to 2 km depth. We use tritium (^{3}H), a radioactive isotope of hydrogen whose concentration spiked in precipitation approximately 50 years ago as a result of above-ground thermonuclear testing[6], to calculate the storage of modern groundwater recharged after the onset of thermonuclear testing. We then use simulations of groundwater age to quantify modern groundwater as well as a broader range of young groundwater up to 100 years old. We define young groundwater as that which is younger than 100 years old, whereas groundwater recharged in the past 50 years, since the above-ground thermonuclear testing[6], is considered modern[4,5].

Quantifying the volume and distribution of young groundwater is crucial as it may be: an important and more renewable groundwater resource rather than older 'fossil' groundwater; more vulnerable to industrial or agricultural contamination, as well as land-use changes at and near the surface of the Earth[7]; more strongly coupled and actively interacting with the broader hydrologic cycle, climate[8] and oceans[9]; and an important driver of global biogeochemical cycles as part of chemical weathering in the critical zone[10]. We focus on the upper 2 km of the continental crust, where fresh, young groundwater is located and the best geochemical and geologic data constrain our analysis. We exclude high-latitude North America and Asia[11] where permafrost exists, but this may not significantly impact the calculation of young groundwater, as permafrost blocks modern groundwater recharge over much of this region[12].

Global total groundwater storage

Although the storage of fresh groundwater and its temporal distribution are critical components of hydrologic and climatic processes, no new and rigorous estimates of the volume of total groundwater have been made for the past 40 years (Supplementary Table 1). Current models that address the water budget[13] rely on groundwater estimation studies from the mid-1970s. Although more recent calculations are widely available[14,15], these ultimately draw from the same 40-year-old estimates[13,16,17]. Here we report the total (young and old) global volume of groundwater in the upper 2 km of continental crust at ~22.6 million km^3 (16–30 million km^3 accounting for uncertainty in our porosity estimate), which is equivalent to 180 m of groundwater if extracted and pooled evenly across the global land surface like a flood. Much of the older groundwater, and some of the young groundwater, is probably brackish, saline, or of low quality[18]. The total volume of stored groundwater was calculated from porosity–depth relationships of four broad rock types (carbonate sediments, siliciclastic sediments, volcanic rocks and crystalline rocks) derived from >40,000 porosity measurements (Fig. 1 and Supplementary Fig. 3) and water table depths[19]. Our estimate is consistent with some previous estimates conducted decades ago[20] of the groundwater volume stored in continents, but our analysis is more robustly data-driven and does not assume that topography entirely controls the porosity and volume of groundwater (see Supplementary Information for detailed methodology and comparison).

[1]Civil Engineering, University of Victoria, Victoria, British Columbia V8P 5C2, Canada. [2]Department of Civil Engineering, McGill University, Montreal, Quebec H3A 0C3, Canada. [3]Department of Geological Sciences, The University of Texas at Austin, Austin, Texas 78712, USA. [4]Department of Geography, University of Calgary, Calgary, Alberta T2N IN4, Canada. [5]Geoscience Centre, Georg-August-Universität Göttingen, Göttingen 37077, Germany.
*e-mail: tgleeson@uvic.ca

© 2016 Macmillan Publishers Limited. All rights reserved

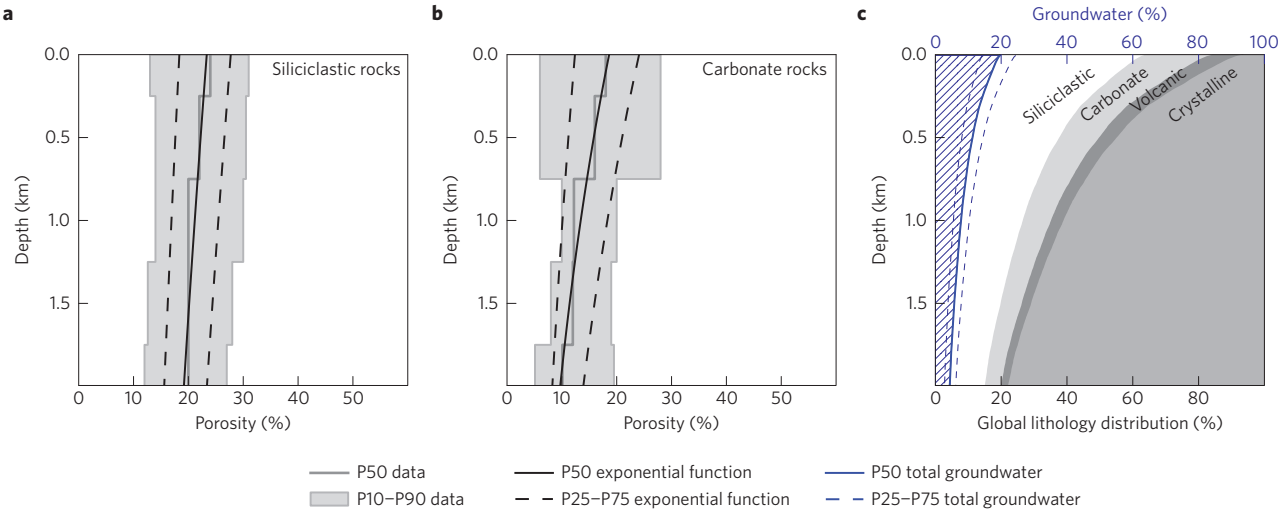

Figure 1 | The global relationship between porosity, lithology and total groundwater volume with depth. a,**b**, Porosity–depth data of 30,122 siliciclastic (**a**) and 10,481 carbonate (**b**) reservoirs[32] and best-fit exponential porosity–depth curves. **c**, The global distribution of lithology with depth and the global average porosity with depth that is filled with groundwater (in blue). P is the percentile distribution.

Modern groundwater storage from tritium data

The International Atomic Energy Agency's TWIN groundwater database reports 9,737 groundwater tritium measurements. However, associated well screen depth data is unavailable in the TWIN database, making it impossible to calculate the volume and distribution of modern groundwater from these existing databases. We circumvented this limitation by developing a new global groundwater tritium database of 3,769 tritium measurements of groundwater samples from 55 countries (Fig. 2). We have quantified the distribution of modern groundwater in depth, showing that modern groundwater is predominantly in the first few hundred metres below ground and that the percentage of modern groundwater generally decreases with depth, although mixtures of modern and old groundwater can be present at any depth (Fig. 2a). Less groundwater is modern below 250 m depth. We use tritium data to calculate the proportion of modern groundwater, $R_{\mathrm{modern,3H}}$, at different depths, rather than calculating groundwater ages, as groundwater ages calculated using tritium can be non-unique and problematic[4,5]. The compiled tritium samples are predominantly located in sedimentary basins with higher permeability than the global average[21], which biases the volume of modern groundwater we calculate using tritium (Fig. 2b). Another issue is that samples are often collected from long well screens, potentially with multiple flow paths. Unfortunately, most aquifers have not been sampled for tritium, which motivates a separate analysis of modern groundwater storage using a combination of geospatial synthesis and numerical flow and age modelling.

Modern groundwater storage from groundwater modelling

A second estimate of young groundwater storage was derived using generic, cross-sectional steady-state models of groundwater flow and age transport parameterized by binned geospatial data from 933,639 watersheds[11]. We use two-dimensional (2D) rather than 3D models because of the computational demands of simulating groundwater age, and we use surface watersheds to delineate shallow groundwater systems, where most young groundwater would be circulating[22]. Groundwater flow models can use recharge or the water table as the input, with the other being an output. We primarily use the water table as an input to our models (Fig. 2c), as water tables are much more frequently measured ($n = 1{,}603{,}781$ globally[19]), vary more smoothly in space than recharge, and have been calculated globally with a groundwater flow model[19]. In contrast, groundwater recharge is difficult to directly and accurately measure even at local scales, and is impossible to directly measure at regional scales, while also being more heterogeneous than water table gradients[23–25]. These 'water-table-driven flow models' avoid limitations with previous 'topography-driven flow models', such as assuming the water table follows topography, and ignore vadose processes, which is consistent with simulating groundwater age as the ageing of groundwater begins at the water table. The median values of the model input parameters are 0.013 for the water table gradient, 5.2 km for the watershed half length, 1.5×10^{-14} m^2 for the near-surface permeability, and 0.19 for the near-surface porosity. Modelled groundwater age fields were most sensitive to permeability and water table gradient, but the decrease of permeability and porosity with depth also influenced the results (see Supplementary Information). We ignore areas with deep water tables (>100 m) from this modelling analysis (~6% by area of the continents and accounts for <11% of young groundwater volumes) owing to a limitation in the water table data[19] (see Supplementary Information), but it is unlikely that significant young groundwater is found in regions with deep water tables, where less recharge is likely to occur. We tested the impact of the choice of water table gradient rather than recharge as an upper boundary condition with a local sensitivity analysis of deep, flat water table conditions where results from these two upper boundary conditions will be most different (see Supplementary Information). Results indicate that, in deep water table conditions, recharge-based models have modern groundwater distributed across more of the domain, with up to approximately double the volume of modern groundwater, whereas with water tables closer to the surface, the groundwater age fields will become more similar to water-table-driven flow models. The simulations that are recharge-based are much more computationally intensive, making it impossible for simulating the large, multi-dimensional parameter space of global hydrogeologic conditions.

To test the consistency between the tritium estimates and models, we calculate aquifer-specific volumes of modern groundwater for 30 aquifers with the largest number of tritium samples (see Supplementary Information). For these 30 aquifers, the storage of modern groundwater, V_{storage}, from the numerical models compares reasonably well to tritium-based values, $V_{\mathrm{storage,3H}}$, given the diversity of hydrogeologic conditions, the simplicity and assumptions of the models, and the difficulty of comparing environmental tracers and groundwater age simulations[2] (Fig. 3 and Supplementary Fig. 3 and Supplementary Table 4). The numerical models generally result in

© 2016 Macmillan Publishers Limited. All rights reserved

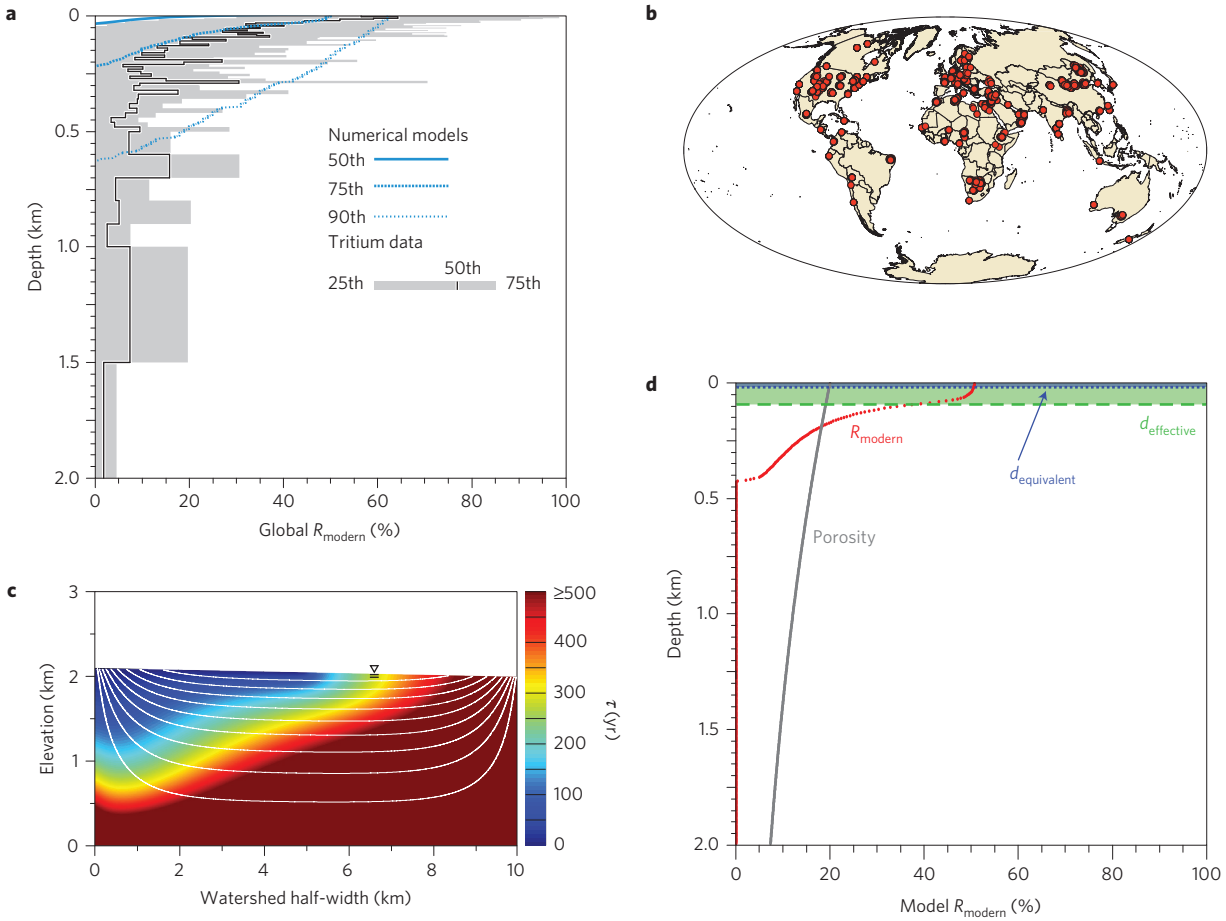

Figure 2 | Quantifying modern groundwater from tritium data and numerical groundwater models of groundwater age. a, The global relationship between the proportion of modern groundwater, R_{modern}, based on tritium, and the numerical simulations at different depths. **b**, The locations of 3,769 measurements of ^{3}H in groundwater were compiled from 160 globally distributed published data sets. **c**,**d**, An example of a groundwater age τ field (**c**) and the calculated R_{modern} (**d**) for a flow system with a surface porosity $n_0 = 0.2$, permeability decay $\beta\alpha = 0.001\,m^{-1}$, surface permeability $k_o = 10^{-12}\,m^2$ and a water table gradient of 0.01.

smaller values for $V_{storage}$. Both the model and tritium approaches are conceptually different, and each have their own uncertainties, biases and simplifications (see Supplementary Information for details). Therefore, we do not attempt to calibrate numerical models to the tritium data. The comparison to tritium estimates and the simulations with recharge-based models described above both indicate the modern groundwater volumes in the water-table-driven flow models are conservative underestimates.

Global modern groundwater storage

From the tritium data and modelled groundwater age fields, we calculated two new metrics of young groundwater storage: the effective depth of young groundwater ($d_{effective}$), which is the depth to which young groundwater is present underground below the water table, and the groundwater equivalent ($d_{equivalent}$), the height of young groundwater if extracted and pooled evenly at the land surface like a flood (see Supplementary Information for detailed methodology). The spatiotemporal storage of modern groundwater is integrated over the global scale for the tritium analysis or over individual watersheds for the numerical simulations (Fig. 2d). Thus, both approaches can estimate the global volume of modern groundwater, although the simulations can estimate storage for other timescales and map its spatial distribution.

Using the proportions of modern groundwater derived from tritium concentrations (see Supplementary Information) and the global average porosity distribution with depth (Fig. 1a and Supplementary Fig. 3), the global volume of modern groundwater was estimated to be 1.3 million km^3 (0.1–5.0 million km^3 accounting for uncertainty in the mixing and recharge models, porosity and tritium, as well as the uncertainty in total groundwater volume). This calculation indicated that 5.6% of groundwater is modern (1–17% with uncertainties), although this is a spatially aggregated global average that does not differentiate rock type or hydrogeologic conditions and is likely to overestimate the volume of modern groundwater as the samples are predominantly from sedimentary basins with relatively more porous and permeable aquifers.

To calculate the global volume of modern groundwater from simulations, geomatic data from each watershed were paired with a numerical model with similar input values. The global groundwater volume was calculated by multiplying the $d_{equivalent}$ by their area for the 933,639 watersheds in non-permafrost regions. We used three different strategies to pairing models and geomatic data as each strategy has limitations. In the first pairing strategy, we paired watersheds using all geomatic data used in the 'water-table-driven' modelling (permeability, porosity and water table gradient). Using this strategy, with mean permeability values for each lithology[26] as an input, we significantly overestimate global recharge (48.0×10^3 km^3 yr^{-1}) compared to global hydrologic models[26]. However, changing the permeability by an order of magnitude results in modelled recharge of $5–497 \times 10^3$ km^3 yr^{-1}, which brackets published estimates of global recharge ($12.0–24.8 \times 10^3$ km^3 yr^{-1}). Culling our calculated recharge

© 2016 Macmillan Publishers Limited. All rights reserved

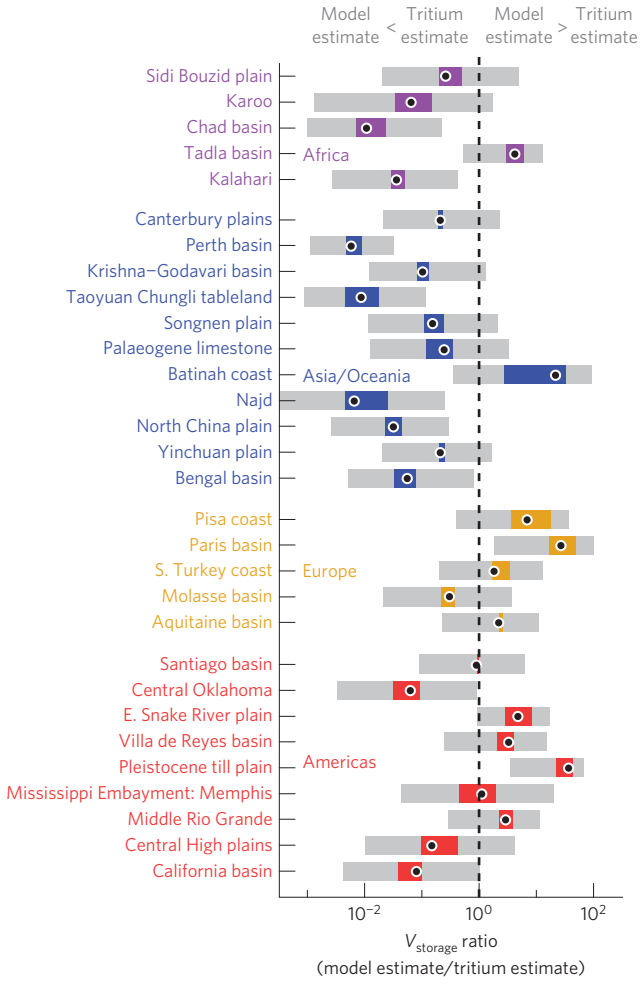

Figure 3 | Comparing estimates of the volume of modern groundwater derived from models and tritium analysis for 30 aquifers with the greatest number of ^{3}H samples. $V_{storage}$ is the volume of modern groundwater stored. The coloured bar shows the uncertainty range considering only the ^{3}H analysis, and the grey bars show the combined uncertainty from the ^{3}H calculation and permeability in the numerical simulations. Black dots show the ratio of storage volumes calculated from the average permeability models and the median $d_{equivalent,3H}$.

values in areas where modelled groundwater recharge is greater than precipitation (10% of land surface), reveals that the majority of overestimated recharge values (83% of modelled global recharge with average permeability values) are located in mountainous and arid regions (Supplementary Fig. 12). High recharge rates in our simulations could be due to an overestimated water table gradient or underestimated permeability in the simulations. Removing watersheds with recharge in excess of precipitation reduces the estimate of modern groundwater volumes to 0.35 million km^3 (1.5% of groundwater to 2 km depth) for the average permeability cases and to 0.14–0.54 million km^3 (0.1–2.4% of groundwater to 2 km depth) for permeability changed by one order of magnitude.

In the second pairing strategy, we paired recharge and water table gradient, but in this case the permeability and porosity may not match the current geomatic data[21] (Table 1). The second strategy resulted in a higher estimate of modern groundwater of 0.67 million km^3 (3.0% of groundwater to 2 km depth). In the third pairing strategy, we paired recharge and porosity, which essentially keeps the lithology the same but locally uses models with different water table gradients and permeabilities than in the geomatic data. The third strategy resulted in a slightly higher estimate of modern groundwater of 0.72 million km^3 (3.2% of groundwater to 2 km depth). In sum, these three strategies result in similar estimates of modern groundwater, with modern groundwater being 1.5–3.2% of the total groundwater globally. These estimates are less than the 5.6% estimate derived from tritium, potentially owing to the approximately two times underestimate of modern groundwater in some regions due to the boundary conditions (Supplementary Fig. 14) and/or the overestimate of tritium samples due to the sample bias. Importantly, all estimates of modern groundwater are in a relatively small range, especially compared to the difference in volume between modern groundwater and total groundwater or all other stores of the global hydrologic cycle. We consider the first pairing strategy the best estimate as it directly uses all geomatic data used as inputs for the 'water-table-driven' modelling.

The model-derived spatial distribution of modern groundwater is extremely heterogeneous, with $d_{equivalent}$ ranging from <0.1 m to >50 m (Fig. 4). The smallest $d_{equivalent}$ values are often found in more arid regions, such as the Sahara and Gobi deserts, central North America, and Australia. Although the three pairing strategies result in similar estimates of the volume of modern groundwater, the patterns of $d_{equivalent}$ are slightly different between the different strategies. The first strategy results in deeper $d_{equivalent}$ in some mountainous regions such as the Andes or the North American Cordillera, which is probably due to high water table gradients in these regions. The second and third strategies result in a more continuous function of $d_{equivalent}$ and deeper $d_{equivalent}$ in humid regions such as the Amazon and Congo, as these pairings explicitly consider recharge estimates.

Figure 5 compares our new estimates of total and modern groundwater to other stores in the global hydrologic cycle using the results of the first pairing strategy with a global volume of modern groundwater of 0.35 million km^3 (0.24 to 3.8 million km^3 accounting for permeability uncertainty, the largest source of uncertainty in groundwater flow models, and the uncertainty in total groundwater volume). Modern groundwater dwarfs all the other components of the active hydrologic cycle, being three times larger than surface water, the next largest component. Yet, the total groundwater volume is vastly larger than the volume of modern groundwater.

Table 1 | Young groundwater (with different age cutoffs of 25, 50, 75 and 100 years) as a percentage of total groundwater with the three different approaches to calculating young groundwater globally.

	25 years (%)	50 years (%)	75 years (%)	100 years (%)
1. Pairing all geomatic data used in modelling (permeability, porosity and water table gradient) and removing watersheds where recharge is unreasonable	0.8	1.5	2.2	2.8
2. Pairing water table gradient and recharge (permeability and porosity may not be reasonable)	1.5	3.0	4.4	5.8
3. Pairing recharge and porosity (water table gradient and permeability may not be reasonable)	1.6	3.2	4.7	6.3
Tritium analysis	–	5.6	–	–

© 2016 Macmillan Publishers Limited. All rights reserved

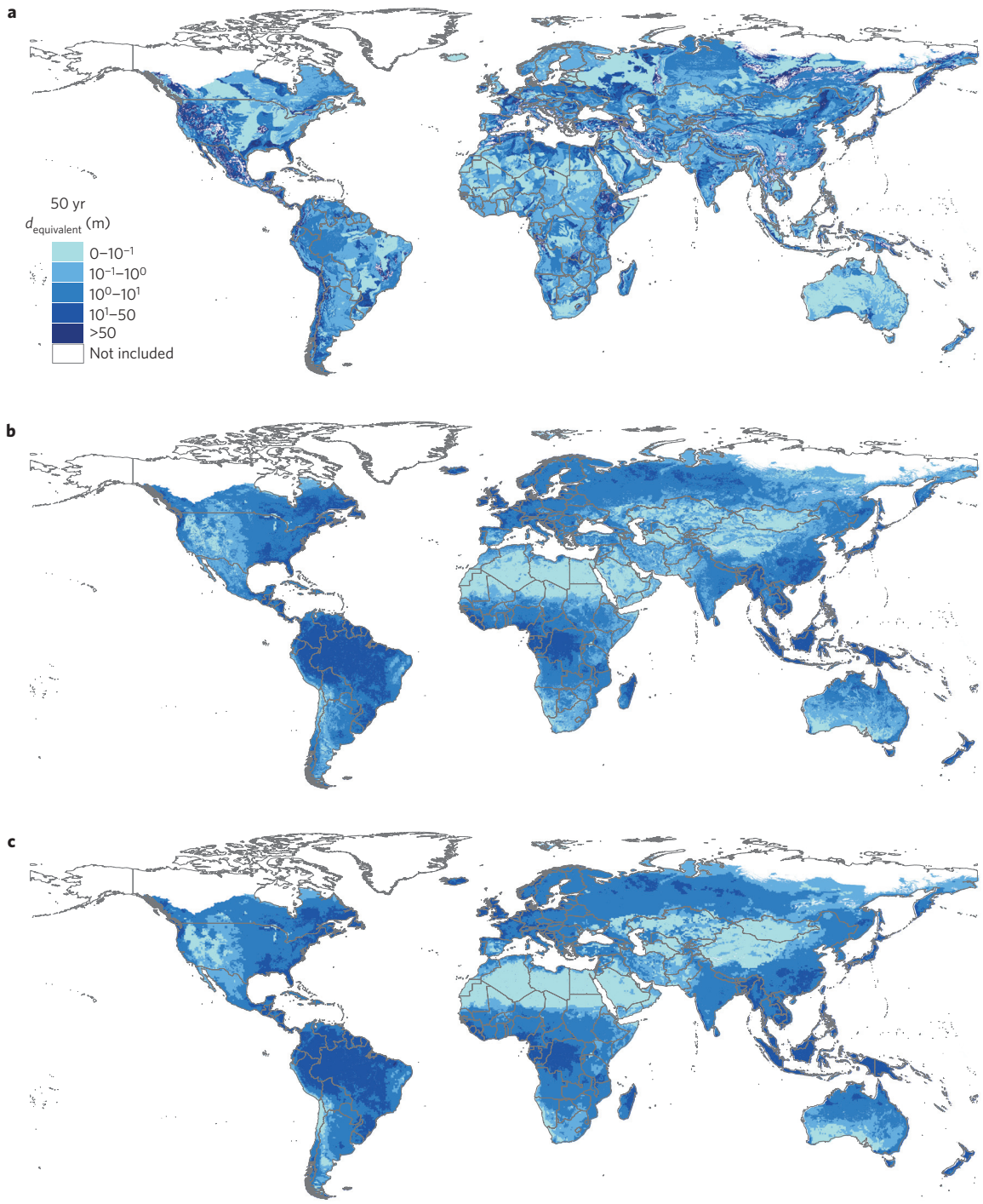

Figure 4 | The global distribution of modern groundwater as a depth if it was extracted and pooled at the land surface like a flood. a–c, $d_{equivalent}$ for the analyses is calculated using the geomatic data as model input parameters (**a**), using groundwater recharge[26] and water table gradient[19] (**b**), and using groundwater recharge[26] and porosity[21] (**c**).

Another way to visualize the amount of modern groundwater is the globally averaged modern groundwater equivalent, which is ~3 m (2.7–5.4 m). Therefore, modern groundwater if extracted and pooled evenly across the global land surface, would be a ~3 m deep body of water.

Groundwater age fields from the simulations are also used to investigate the storage of young groundwater over other timespans. Over 25 years, the volume of global groundwater storage is 0.8 million km^3 (0.1–2.9 million km^3 with permeability uncertainty). Over 75 years, the groundwater storage is 1.6 million km^3 (0.3–4.4 million km^3 with permeability uncertainty). With a timespan of 100 years, the groundwater storage volume is 1.9 million km^3 (0.4–4.9 million km^3 with permeability uncertainty). These global volumes of young groundwater show a linear increase with age for the relatively short groundwater timescales we consider, but this linearity is not expected to hold across longer timespans[22]. Regardless of what timespan is chosen to delineate young and old groundwater, these results indicate that young groundwater remains a limited resource which comprises a minority of the groundwater in the upper 2 km of the Earths' crust.

© 2016 Macmillan Publishers Limited. All rights reserved

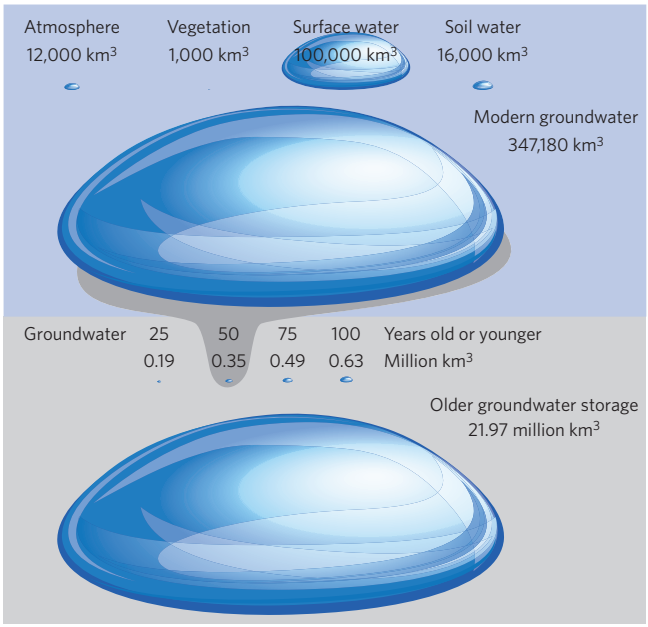

Figure 5 | The different volumes of water stored in the global water cycle. Global volumes of young (<100 years old), modern (<50 years old) and total groundwater to 2 km depth compared with other fresh water volumes stored in the atmosphere[33], in surface waters (that is, wetlands, rivers and lakes)[33], within plants or in soils[33].

Global and local groundwater budgets affect energy and food security, ecosystem services, and drinking water availability[27–30]. Sustainable groundwater use requires systematic quantification not only of groundwater fluxes[26,31] but also of groundwater storage[20]. Here, we provide four consistent and independent estimates of the stored volume of modern groundwater and an estimate of the total volume of groundwater in the upper 2 km of the crust. We find that groundwater replenished over a human lifetime of 25–100 years is a finite, limited resource with a spatially heterogeneous distribution dependent on geographic, geologic and hydrologic conditions. Groundwater comprises a vast water resource that vastly exceeds all other water sources on our planet. However, the younger and more readily available portions of the subterranean water cycle are much smaller drops and are at most 6% of the total volume of groundwater in the upper crust (Fig. 5). A key application of our new estimate is quantifying groundwater storage and availability. In addition, the distribution of young groundwater storage can be used to delineate groundwater vulnerability to contamination, as recently recharged groundwater is more vulnerable to contamination by industrial or agricultural activities. Our estimates of spatiotemporally constrained groundwater storage may also indicate hotspots for chemical weathering, landscape evolution[10] and coastal eutrophication[9], helping us discern the role of groundwater in the Earth system.

Methods

Methods and any associated references are available in the online version of the paper.

Received 18 May 2015; accepted 14 October 2015; published online 16 November 2015

References

1. Sturchio, N. C. *et al.* One million year old groundwater in the Sahara revealed by krypton-81 and chlorine-36. *Geophys. Res. Lett.* **31,** L05503 (2004).
2. McCallum, J. L., Cook, P. G. & Simmons, C. T. Limitations of the use of environmental tracers to infer groundwater age. *Groundwater* **53,** 56–70 (2014).
3. Weissmann, G. S., Zhang, Y., LaBolle, E. M. & Fogg, G. E. Dispersion of groundwater age in an alluvial aquifer system. *Wat. Resour. Res.* **38,** 1198 (2002).
4. Bethke, C. M. & Johnson, T. M. Groundwater age and groundwater age dating. *Annu. Rev. Earth Planet. Sci.* **36,** 121–152 (2008).
5. Kazemi, G., Lehr, J. & Perrochet, P. *Groundwater Age* (Wiley-Interscience, 2006).
6. Alley, W. M., Healy, R. W., LaBaugh, J. W. & Reilly, T. E. Flow and storage in groundwater systems. *Science* **296,** 1985–1990 (2002).
7. Foster, S. S. D. & Chilton, P. J. Groundwater: The processes and global significance of aquifer degradation. *Phil. Trans. R. Soc. Lond. B* **358,** 1957–1972 (2003).
8. Taylor, R. G. *et al.* Ground water and climate change. *Nature Clim. Change* **3,** 322–329 (2013).
9. Moore, W. S. Large groundwater inputs to coastal waters revealed by ^{226}Ra enrichments. *Nature* **380,** 612–614 (1996).
10. Maher, K. & Chamberlain, C. P. Hydrologic regulation of chemical weathering and the geologic carbon cycle. *Science* **343,** 1502–1504 (2014).
11. Lehner, B., Verdin, K. & Jarvis, A. New global hydrography derived from spaceborne elevation data. *Eos* **89,** 93–94 (2008).
12. Gruber, S. Derivation and analysis of a high-resolution estimate of global permafrost zonation. *Cryosphere* **6,** 221–233 (2012).
13. Garmonov, I. V., Konoplyantsev, A. A. & Lushnikova, N. P. in *The World Water Balance and Water Resources of the Earth* (ed. Korzun, K. I.) 48–50 (Hydrometeoizdat, 1974).
14. Chahine, M. T. The hydrological cycle and its influence on climate. *Nature* **359,** 373–380 (1992).
15. Schneider, U. *et al.* GPCC's new land surface precipitation climatology based on quality-controlled *in situ* data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.* **115,** 15–40 (2014).
16. L' Vovich, M. I. in *World Water Resources and their Future* (ed. Nace, R. L.) 13–23 (American Geophysical Union, 1979); http://onlinelibrary.wiley.com/book/10.1029/SP013
17. Nace, R. L. in *Water, Earth, and Man: A Synthesis of Hydrology, Geomorphology, and Socio-Economic Geography* (ed. Chorley, R. J.) 31–42 (Methuen and Co., 1969).
18. Holland, H. D. & Turekian, K. K. (eds) in *Treatise on Geochemistry* 2nd edn (Pergamon, 2003); http://www.sciencedirect.com/science/referenceworks/9780080983004
19. Fan, Y., Li, H. & Miguez-Macho, G. Global patterns of groundwater table depth. *Science* **339,** 940–943 (2013).
20. USSR Committee for the International Hydrologic Decade *World Water Balance and Water Resources of the Earth* (UNESCO, 1978).
21. Gleeson, T., Moosdorf, N., Hartmann, J. & vanBeek, L. P. H. A glimpse beneath Earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophys. Res. Lett.* **41,** 3891–3898 (2014).
22. Cardenas, M. B. Potential contribution of topography-driven regional groundwater flow to fractal stream chemistry: Residence time distribution analysis of Tóth flow. *Geophys. Res. Lett.* **34,** L05403 (2007).
23. Lerner, D. N. in *Geochemical Processes, Weathering and Groundwater Recharge in Catchments* (eds Saether, O. M. & de Caritat, P.) 109–150 (Balkema, 1997).
24. Scanlon, B., Healy, R. & Cook, P. Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeol. J.* **10,** 18–39 (2002).
25. Scanlon, B. R. *et al.* Global synthesis of groundwater recharge in semiarid and arid regions. *Hydrol. Process.* **20,** 3335–3370 (2006).
26. Döll, P. & Fiedler, K. Global-scale modeling of groundwater recharge. *Hydrol. Earth Syst. Sci.* **12,** 863–885 (2008).
27. Giordano, M. Global groundwater? Issues and solutions. *Annu Rev. Environ. Resour.* **34,** 153–178 (2009).
28. Morris, B. L. *et al. Groundwater and its Susceptibility to Degradation: A Global Assessment of the Problem and Options for Management* (UNEP Early Warning and Assessment Report Series RS 03-3, 2003).
29. Fleckenstein, J. H., Krause, S., Hannah, D. M. & Boano, F. Groundwater-surface water interactions: New methods and models to improve understanding of processes and dynamics. *Adv. Water Resour.* **33,** 1291–1295 (2010).
30. Nicot, J.-P., Scanlon, B. R., Reedy, R. C. & Costley, R. A. Source and fate of hydraulic fracturing water in the Barnett Shale: A historical perspective. *Environ. Sci. Technol.* **48,** 2464–2471 (2014).
31. Wada, Y., van Beek, L. P. H. & Bierkens, M. F. P. Nonsustainable groundwater sustaining irrigation: A global assessment. *Wat. Resour. Res.* **48,** W00L06 (2012).
32. Athy, L. F. Density, porosity, and compaction of sedimentary rocks. *AAPG Bull.* **14,** 1–24 (1930).
33. Fowler, A. & Yang, X. Fast and slow compaction in sedimentary basins. *SIAM J. Appl. Math.* **59,** 365–385 (1998).

© 2016 Macmillan Publishers Limited. All rights reserved

Acknowledgements

T.G. and E.L. were supported by the NSERC and a CIFAR Junior Fellowship. M.B.C. and K.M.B. were supported by the NSF (EAR-0955750) and the Geology Foundation at the University of Texas at Austin. K.M.B. and S.J. were supported by American Geophysical Union Horton Research Grants.

Author contributions

T.G. conceived and led the project and the writing of the paper. K.M.B. led and conducted the modelling, geomatic analysis and model-related calculations as well as developed the mathematical methods for calculating the metrics. S.J. conducted the tritium data collection and analysis. E.L. derived the original geomatic data and a method for coupling geomatic data to models, as well as conducted the data analysis of total groundwater storage. M.B.C. brainstormed ideas and analysed results. All authors co-developed the methods, wrote text for their respective sections, and heavily discussed and edited all drafts of the manuscript.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.G.

Competing financial interests

The authors declare no competing financial interests.

© 2016 Macmillan Publishers Limited. All rights reserved

Methods

Our methods are illustrated in Supplementary Fig. 1 and include: global data synthesis and calculation of the total groundwater volume; synthesis of geochemical data; geomatic analysis and numerical groundwater age simulations; analysis of thirty specific aquifers using both tritium data and numerical simulations; and the calculation of the distribution and global volume of modern and young groundwater.

Global data synthesis and calculation of the total groundwater volume. *Porosity–depth models.* Porosity–depth profiles were used in calculating the total global volume of groundwater, calculating the global volume of modern groundwater based on tritium data at different depths, and guiding the parameterization of porosity and permeability decay in numerical groundwater age transport modelling. Porosity–depth relationships for four different rock types (carbonate sediments, siliciclastic sediments, volcanic and crystalline rocks) were compiled. Depending on lithology, porosity (n) was considered constant with depth or modelled as exponentially decaying with depth[32,33]:

$$n = n_0 e^{-\beta z'} \quad (1)$$

where n_0 is the surface porosity, β is the matrix compressibility (m^{-1}) and z' the is depth (m) below the ground surface. For carbonate and siliciclastic rocks we fit equation (1) with global porosity–depth compilations of siliciclastic (30,122 values) and carbonate (10,481 values) petroleum reservoirs[34] (Supplementary Fig. 3). As the porosity of volcanic rocks is highly variable both at the surface and at depth, and no discernible depth–porosity trends are evident in the literature, we applied a depth-invariant and uncertain porosity of 0.09 ± 0.09 for volcanic rocks (Supplementary Table 2). We used a depth-invariant and uncertain porosity of 0.01 ± 0.01 for crystalline rocks (Supplementary Table 2). See Supplementary Information for more information on how porosity functions were chosen.

Calculation of the total groundwater volume using porosity–depth models. The area of each lithology was calculated[35]. The distribution of rock types with depth incorporated global data on the thickness of sedimentary basins[36]. If sedimentary basins were thinner than 2 km, materials below the sedimentary basin were assigned the non-sedimentary porosity (crystalline or volcanic depending on the fraction of each lithology at the surface). Results of the compilation and analysis are presented in the Supplementary Information.

For each lithology, a groundwater equivalent (the height of water if removed from the ground and pooled over the land surface) was calculated by integrating the lithology-specific porosity decay and weighting each lithology by its volumetric fraction in the upper crust (Supplementary Table 3). The global volume of groundwater is then the product of the global sum of the groundwater equivalent and the total land area.

Global tritium (^3H) data synthesis and calculation of modern groundwater. A global data base of ^3H concentrations in groundwater was compiled for 3,769 globally distributed samples from 160 publications (Fig. 2b). Samples came from confined (12% of data set), partially confined (23% of data set) and unconfined aquifers (65% of data set).

To estimate the mass (m) proportion of modern groundwater at a given location, we developed a ^3H-based mixing model for each sample. The proportion of modern groundwater (m_{modern}) within a given water sample (m_{sample}) taken from depth z was quantified using a mixing model:

$$R_{modern,3H}(z) = \frac{m_{modern}}{m_{sample}} = \frac{{}^3H_{sample} - {}^3H_{old}}{{}^3H_{modern} - {}^3H_{old}} \quad (2)$$

where $^3H_{sample}$ represents the ^3H concentration of a given sample, and $^3H_{old}$ or $^3H_{modern}$ represent ^3H values for groundwater that recharged within ($^3H_{modern}$) or before 50 years ($^3H_{old}$) of the sampling date, and m_{modern}/m_{sample} is the mass fraction of young groundwater within a given water sample. $R_{modern,3H}$ is the mass proportion of young groundwater within any given water sample. This model necessarily accounts for changes to atmospheric ^3H levels due to thermonuclear testing, radioactive decay of tritium, and the mixing of groundwaters of different ages within aquifers.

As $^3H_{sample}$ is the measured quantity, implementing the mixing model requires the calculation of $^3H_{modern}$ and $^3H_{old}$. However, both values have varied through time—from 'infinitely' long ago to 50 years before sampling for $^3H_{old}$, and the 50 years preceding the time of sampling for $^3H_{modern}$. For $^3H_{modern}$, ^3H in precipitation was determined at an annual time step for individual sample locations using a global model of tritium in precipitation[37] developed for 1960–2005 (Supplementary Fig. 4). For $^3H_{old}$, ^3H in precipitation before 1960 was set to a range of 1–10 T.U. (before accounting for radioactive decay), as supported by pre-bomb ^3H records from ice cores, wine and lakes[38–40] (Supplementary Fig. 5). And finally, the radioactive decay of all ^3H pools was accounted for before calculation with the mixing model. These steps are explained in detail in the Supplementary Information.

Global spatial data analysis and assignment of properties for watersheds. Spatial data were synthesized for 933,639 watersheds from the HydroSHEDS watershed data[11] by extracting other input parameters for each watershed. This study focused on six continents (total area of 135 million km^2), excluding Antarctica (14 million km^2). For each watershed, the water table gradient, near-surface permeability and porosity, and distance between streams were analysed.

The HydroSHEDS watershed data[11] are not available for Antarctica, Greenland, and high-latitude regions in North America, and our analysis excludes these regions. Subglacial or sub-permafrost groundwater is unlikely to be modern ($\leq$50 yr old), but excluding these regions increases the uncertainty of our global estimates of young groundwater. The total land area used in our analysis was 126.3 million km^2.

The average water table gradient was calculated at 500 × 500 m scale from a recent groundwater flow model calculating the depth to the water table globally[19]. Spatially averaged mean near-surface permeability and porosity for each watershed were derived ref. 21. The distance between streams was calculated at 1 km resolution globally by summing the distance to two nearest streams[11] for all pixels. The maximum distance between streams was set at 200 km to reduce the computational burden of the geospatial analysis. The watershed half-width used for creating the numerical model domains is half the distance between these streams. A maximum half-width of 100 km does not affect the calculation significantly because watersheds with a half-width of greater than 100 km will have minimal young groundwater and mainly occur in deserts. Most watersheds have a <10 km half-width.

Near-surface permeability, porosity, and the decay of each of these parameters were assigned based on surface lithology, which was classified into nine groups[21]. When a watershed overlapped lithologic contacts, the most areally extensive rock or sediment type within the watershed was used—this resulted in no more than a quarter order of magnitude difference in permeability, and was much less than the uncertainty in the data[21].

Numerical modelling of groundwater age transport. *Numerical simulations.* Groundwater age fields for the watersheds were modelled with 2D steady-state groundwater flow and age transport simulations. The models were solved numerically with a finite-element model implemented in COMSOL Multiphysics for a range of parameter combinations based on the global data set analysis in the previous section. (permeability[21], water table gradient[19], porosity[21], compressibility[34], and watershed half-widths[11]). The domain geometry was designed to describe groundwater flow from a watershed boundary (for example, a ridge) to the corresponding drainage feature (for example, a river)[41]. We constructed the model domains with a length based on the average distance from the watershed boundary to the corresponding river, calculated from the mean distance between streams[11]. The depth of the aquifer was assigned to be one fifth of the domain length. These spatial approximations do not include larger regional flow below and across the watershed boundaries used for the current analysis. However, the focus of these models is to elucidate the groundwater systems with water younger than 50 years old.

As modelling the groundwater flow and age transport field for all 933,639 watersheds is prohibitive even with 2D models, we reduced the number of models by binning watershed properties. The frequency distributions of water table gradients and watershed half-widths guided the discretization. The roughly log-normally distributed water table gradients were divided into 33 bins based on the logarithms of water table gradients increasing in width away from the median value (Supplementary Fig. 6). The watershed half-widths, which exhibited a long-tailed normal distribution, were also split into 33 bins (Supplementary Fig. 6). To better represent the larger (wider) watersheds, 16 evenly spaced bins for every 5 km were added to the original 33 bins, starting with 10 km up to 100 km.

In our numerical simulations, we modelled the saturated subsurface (that is, below the water table). Previous models of groundwater age and regional groundwater flow systems have often assumed the water table is near or at the surface as a subdued replica of topography[41–44], which can lead to artificially deeper groundwater circulation and/or greater than measured recharge rates.

The permeability used at the surface of the flow models was derived from global permeability maps[21]. These permeability values decayed as a function of depth based on the lithologic porosity decay models in equation (1):

$$\mathbf{k}(x,z) = k_0 e^{-\beta\alpha[z_s(x)-z]} \quad (3)$$

with k_0 the permeability of near-surface materials, β the compressibility from equation (1), $\alpha = 2$ is a theoretical factor scaling porosity to permeability decay[45], and z_s the elevation of the ground surface. No anisotropy was assigned to $\mathbf{k}$. Porosity was similarly assigned to the models with:

$$\mathbf{n}(x,z) = n_0 e^{-\beta[z_s(x)-z]} \quad (4)$$

© 2016 Macmillan Publishers Limited. All rights reserved

The steady-state groundwater flow equation:

$$\nabla \cdot (\mathbf{k} \nabla h(x,z)) = 0 \quad (5)$$

was solved for the hydraulic head (h) distribution and used to set the advection and dispersion terms in the age-as-mass transport equation, with $\mathbf{k}$ the two-dimensional, spatially heterogeneous but locally isotropic permeability tensor calculated from the permeability distribution in equation (3). The top boundary condition for groundwater flow was a prescribed head based on the head gradient calculated from ref. 19, and all other boundaries were impermeable (that is, no-flow).

The groundwater age transport equation treats groundwater age (τ) as mass in a steady-state advection–dispersion–diffusion equation (after equation (10) of ref. 46):

$$\nabla \cdot (\mathbf{n}\mathbf{D} \nabla \tau(x,z)) - \nabla (\mathbf{u}\mathbf{n}\tau(x,z)) + \mathbf{n} = 0 \quad (6)$$

with porosity, $\mathbf{n}$, as an age-as-mass source term, $\mathbf{u} = [u_x, u_z]$ the average linear groundwater velocity, and $\mathbf{D}$ the diffusion–dispersion coefficient tensor:

$$\mathbf{D} = \alpha_T |\mathbf{u}| \delta_{ij} + \frac{(\alpha_L - \alpha_T) u_i u_j}{|\mathbf{u}|} + D_m \quad (7)$$

In $\mathbf{D}$, the molecular diffusion coefficient, D_m, was assigned a constant value of 10^{-8} m^2 s^{-1}, the longitudinal dispersivity (α_L) was scaled as a tenth of the square root of the domain length and thickness with the transverse dispersivity (α_T) an order of magnitude lower, and δ_{ij} is the Kronecker delta function. No-flux boundary conditions were imposed on all but the upper boundary, which was split into $\tau = 0$ years across the recharge zone and only advection of τ (that is, no diffusion or dispersion) across the discharge zone. Using the parameter assignment scheme described above, 43,659 coupled groundwater flow and age transport models were run to solve for each of the parameter combinations in the global watershed data.

With the binning procedure explained above, the groundwater age distributions for Earth's watersheds could have been modelled using 14,553 combinations of the hydrologic input parameters, but we ran an additional 29,106 models, changing the original combinations by one order of magnitude in k_0 to provide an estimate of the uncertainty in each modelled groundwater age distribution. Details on solving equations (5) and (6) can be found in the Supplementary Information.

Calculating the effective depth to young groundwater and the young groundwater equivalent. The age distributions from the numerical simulations were used to calculate both the effective depth and the young groundwater equivalent. We calculated the relative frequency of given groundwater ages, $f(\tau(z'))$, across 500 evenly spaced horizontal cross-sections to the maximum domain depth and using 500 logarithmically spaced age bins from 10^{-1} to 10^4 years (Supplementary Fig. 6). With the large number of samples (that is, grid cells) in this analysis, the relative frequency distribution approximates the probability distribution. This depth-specific age probability distribution was integrated to chosen threshold ages for every depth, giving a depth profile for the cumulative probability of finding groundwater T years old or younger (Supplementary Fig. 7). We consider this basin-wide, truncated cumulative probability the ratio of young groundwater, $R_{young}(z')$, as:

$$R_{young}(z') = \int_0^T f(\tau(z'))\mathrm{d}\tau \quad (8)$$

where $f(\tau(z'))$ is the number of samples per age bin divided by the total number of bins with units of yr^{-1}, and $R_{young}(z')$ is unitless. R_{young} above is conceptually similar to $R_{modern,3H}$ in equation (2) in that they both represent the proportion of groundwater that is younger than $T = 50$ years old. However, it is important to note their fundamental differences. R_{young} results from integration across an entire aquifer's width and thus effectively represents all of the groundwater at some depth. On the other hand, $R_{modern,3H}$ is the proportion of a groundwater sample that is younger than 50 years. These two will be perfectly equivalent only when the groundwater sample analysed for tritium is a mixture of all groundwater in an aquifer at the sampling depth. Note that this analysis was performed with $T = 25$, 50, 75 and 100 years. These groundwater ages are all young (<100 years old), but we specifically denote groundwater <50 years old as 'modern'.

Integrating $R_{young}(z')$ in depth gives either the effective depth of young groundwater ($d_{effective}$):

$$d_{effective} = \int_0^{z'_{max}} R_{young}(z')\mathrm{d}z' \quad (9)$$

or the young groundwater equivalent ($d_{equivalent}$) if porosity is included in the integrand:

$$d_{equivalent} = \int_0^{z'_{max}} n(x,z') R_{young}(z')\mathrm{d}z' \quad (10)$$

with z'_{max} the maximum depth of the models from the top boundary. $d_{effective}$ represents the depth to which $R_{young}(z') = 1$ if all $R_{young}(z') > 0$ contributions were summed starting from the surface. Thus, $d_{effective}$ represents a characteristic depth scale for encountering a specific age T of groundwater in a hydrogeologic system, chosen as 25, 50, 75 and 100 years for this study. Similarly, $d_{equivalent}$ provides a porosity-corrected characteristic depth scale that more readily accounts for the total volume of groundwater $\leq T$ yr. Equations (8)–(10) are also applicable for both the discrete ^{3}H-derived ratio of modern groundwater (defined in equation (2)) and numerical model results. The integration was implemented numerically via a three-point Newton–Cotes quadrature for equations (9) and (10) for both the age-transport-model-derived and ^{3}H-derived ratio of young groundwater and a summation for equation (8).

Modern groundwater storage for 30 specific aquifers. Tritium concentrations are available only for a minority of aquifers globally. Therefore, we tested the agreement between the simulation and ^{3}H-derived storage volumes of modern groundwater storage for 30 aquifers with the greatest number of tritium samples. ^{3}H estimates of young groundwater storage volumes were calculated using well samples from each aquifer with aquifer-average porosity values. To calculate the modelling estimate of the modern groundwater storage volume for an aquifer, all of the watersheds within the areal extent of that aquifer were summed, using the well sample locations to define the aquifer extent. We also compared model and aquifer $d_{effective}$ values and the alternate model–watershed pairing approaches (see Supplementary Information).

Calculation of the global volume of modern groundwater. *Global estimate of modern groundwater based only on tritium.* We calculated a ^{3}H-based value for the global modern groundwater by integrating the proportion of young groundwater (equation (2), $R_{modern,3H}$; Fig. 2a) in depth. This is conceptually similar to the approach applied to the modelled age distributions (equation (10)):

$$d_{equivalent,3H} = \int_0^{z'_{max}} n(z') R_{modern,3H}(z')\mathrm{d}z' \quad (11)$$

where $R_{modern,3H}(z')$ represents an average of all ^{3}H-based estimates of modern groundwater ($R_{modern,3H}$) at given depth intervals from Earth's surface down to 2 km (the black line in Fig. 2a). The porosity profile $n(z')$ is the global porosity function calculated from Methods 1.2. Dividing $d_{modern,3H}$ by the global total groundwater equivalent (179 m, Supplementary Table 3), we obtained the globally integrated ^{3}H-based estimate of the proportion of modern groundwater within the uppermost 2 km of Earth's crust.

Assignment of groundwater age transport model results to global watershed distribution. In the geomatic-controlled assignment, all model results, including $d_{effective}$ and $d_{equivalent}$, were assigned to the 933,639 HydroSHEDS watersheds by mapping each watershed to the binned input hydrologic parameters driving the 43,659 modelled generic groundwater systems. The model results with one order of magnitude change in the k_0 above and below the published data were also assigned to each watershed to account for uncertainty in the permeability data[21]. Global maps of young groundwater storage were created for $T = 25$, 50, 75 and 100 years (Supplementary Fig. 5). Watersheds located where water table depths were >100 m were removed from the analysis owing to the original analysis[19] constraining a maximum water table depth of 100 m. We also removed watersheds with $d_{equivalent}$ values that were greater than precipitation (see Supplementary Information). To calculate the global volume of young groundwater, we summed the volume of young groundwater from each watershed, calculated by multiplying the watershed area by $d_{equivalent}$.

For the recharge-based mapping of models to watersheds, groundwater recharge estimates[26] were assigned to each watershed and used as the master variable. Recharge from the models was calculated by integrating the flux of groundwater across the upper model boundary, where recharge occurs, and then dividing this total domain recharge by the model half-width to give the average recharge per unit length of the model. This model recharge could then be multiplied by the area of the watershed to calculate the total volumetric recharge occurring in a particular watershed and is equivalent to the recharge modelled[26].

As the model recharge can be changed by varying permeability or the water table gradient, we assigned model results to watersheds based on recharge using two different approaches. We first paired models with the same half-width and water table gradient with groundwater recharge differences <5% and let permeability and porosity (that is, lithology) change freely. The second method matched recharge and kept watershed half-width and porosity constant, allowing permeability to range by one order of magnitude around the lithologic median value, while letting the water table gradient change freely. Where models could not be found with recharge within 5% of the previous values[26], the model with the most similar recharge for a given length scale was used and accounted for no more than 5% of the total number of ice-free watersheds per continent.

© 2016 Macmillan Publishers Limited. All rights reserved

Data. The tritium data are available at S.J.'s research website (http://www.isohydro.ca/global-water-datasets.html). All the geomatic input data and results are available at http://figshare.com/articles/Global_modern_groundwater_Gleesonetal/1560081.

Code availability. The program used to generate all the results is COMSOL Multiphysics (https://www.comsol.com/comsol-multiphysics); a representative model input file for COMSOL Multiphysics used to derive the results in this study can also be accessed at http://figshare.com/articles/Global_modern_groundwater_Gleesonetal/1560081.

References

34. Ehrenberg, S. N. & Nadeau, P. H. Sandstone vs. carbonate petroleum reservoirs: A global perspective on porosity-depth and porosity-permeability relationships. *Am. Assoc. Petrol. Geol. Bull.* **89,** 435–445 (2005).
35. Hartmann, J. & Moosdorf, N. The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochem. Geophys. Geosyst.* **13,** Q12004 (2012).
36. Laske, G. & Masters, G. A global digital map of sediment thickness. *Eos* **78,** F483 (1997).
37. Zhang, Y., Ye, S. & Wu, J. A modified global model for predicting the tritium distribution in precipitation, 1960–2005. *Hydrol. Process.* **25,** 2379–2392 (2011).
38. Begemann, F. & Libby, W. F. Continental water balance, ground water inventory and storage times, surface ocean mixing rates and world-wide water circulation patterns from cosmic-ray and bomb tritium. *Geochim. Cosmochim. Acta* **12,** 277–296 (1957).
39. Clark, I. & Fritz, P. *Environmental Isotopes in Hydrogeology* (Lewis, 1997).
40. Kotzer, T. G., Kudo, A., Zheng, J. & Workman, W. Natural and anthropogenic levels of tritium in a Canadian Arctic ice core, Agassiz Ice Cap, Ellesmere Island, and comparison with other radionuclides. *J. Glaciol.* **46,** 35–40 (2000).
41. Tóth, J. A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.* **68,** 4795–4812 (1963).
42. Jiang, X.-W., Wan, L., Cardenas, M. B., Ge, S. & Wang, X.-S. Simultaneous rejuvenation and aging of groundwater in basins due to depth-decaying hydraulic conductivity and porosity. *Geophys. Res. Lett.* **37,** L05403 (2010).
43. Jiang, X.-W., Wan, L., Wang, X.-S., Ge, S. & Liu, J. Effect of exponential decay in hydraulic conductivity with depth on regional groundwater flow. *Geophys. Res. Lett.* **36,** L24402 (2009).
44. Cardenas, M. B. & Jiang, X. W. Groundwater flow, transport, and residence times through topography-driven basins with exponentially decreasing permeability and porosity. *Wat. Resour. Res.* **46,** W11538 (2010).
45. Bernabé, Y., Mok, U. & Evans, B. Permeability-porosity relationships in rocks subjected to various evolution processes. *Pure Appl. Geophys.* **160,** 937–960 (2003).
46. Goode, D. J. Direct simulation of groundwater age. *Wat. Resour. Res.* **32,** 289–296 (1996).

© 2016 Macmillan Publishers Limited. All rights reserved