Hydrology and
Earth System
Sciences

Predicting groundwater recharge for varying land cover and climate conditions – a global meta-study

Chinchu Mohan[1], **Andrew W. Western**[1], **Yongping Wei**[2], and **Margarita Saft**[1]

[1]Department of Infrastructure Engineering, University of Melbourne, Melbourne, Victoria, Australia
[2]School of Geography, Planning and Environmental Management, The University of Queensland, Brisbane, Australia

Correspondence: Chinchu Mohan (mohanc@student.unimelb.edu.au)

Abstract. Groundwater recharge is one of the important factors determining the groundwater development potential of an area. Even though recharge plays a key role in controlling groundwater system dynamics, much uncertainty remains regarding the relationships between groundwater recharge and its governing factors at a large scale. Therefore, this study aims to identify the most influential factors of groundwater recharge, and to develop an empirical model to estimate diffuse rainfall recharge at a global scale. Recharge estimates reported in the literature from various parts of the world (715 sites) were compiled and used in model building and testing exercises. Unlike conventional recharge estimates from water balance, this study used a multimodel inference approach and information theory to explain the relationship between groundwater recharge and influential factors, and to predict groundwater recharge at 0.5° resolution. The results show that meteorological factors (precipitation and potential evapotranspiration) and vegetation factors (land use and land cover) had the most predictive power for recharge. According to the model, long-term global average annual recharge (1981–2014) was $134\,\mathrm{mm\,yr^{-1}}$ with a prediction error ranging from -8 to $10\,\mathrm{mm\,yr^{-1}}$ for 97.2 % of cases. The recharge estimates presented in this study are unique and more reliable than the existing global groundwater recharge estimates because of the extensive validation carried out using both independent local estimates collated from the literature and national statistics from the Food and Agriculture Organization (FAO). In a water-scarce future driven by increased anthropogenic development, the results from this study will aid in making informed decisions about groundwater potential at a large scale.

1 Introduction

Human intervention has dramatically transformed the planet's surface by altering land use and land cover and consequently the hydrology associated with it. In the last 100 years the world population has quadrupled, from 1.7 billion (in 1900) to more than 7.3 billion (in 2014), and is expected to continue to grow significantly in the future (Gerland et al., 2014). During the last century, rapid population growth and the associated shift to a greater proportion of irrigated food production led to an increase in water extraction by a factor of ~ 6. This eventually resulted in the overexploitation of both surface and groundwater resources, including the depletion of 21 of the world's 37 major aquifers (Richey et al., 2015). This depletion threatened human lives in many ways, ranging from critical reductions in water availability to natural disasters such as land subsidence (Chaussard et al., 2014; Ortiz-Zamora and Ortega-Guerrero, 2010; Phien-Wej et al., 2006; Sreng et al., 2009). Therefore, there is a need to closely examine approaches for sustainably managing this resource by controlling withdrawal from the system.

Groundwater recharge is one of the most important limiting factors for groundwater withdrawal and determines the groundwater development potential of an area (Döll and Flörke, 2005) Groundwater recharge connects atmospheric, surface and subsurface components of the water balance and is sensitive to both climatic and anthropogenic factors (Gurdak, 2008; Herrera-Pantoja and Hiscock, 2008; Holman et al., 2009; Jyrkama and Sykes, 2007). Various studies have employed different methods to estimate groundwater recharge including tracer methods, water table fluctuation methods, lysimeter methods and simple water balance techniques. Some of these studies input recharge to numeri-

cal groundwater models or dynamically link it to hydrological models to estimate variations under different climate and land cover conditions (Aguilera and Murillo, 2009; Ali et al., 2012; Herrera-Pantoja and Hiscock, 2008; Sanford, 2002).

In the last few decades, interest in global-scale recharge analysis has increased for various scientific and political reasons (Tögl, 2010). L'vovich (1979) made the first attempt at a global-scale analysis by creating a global recharge map using baseflow derived from river discharge hydrographs. The next large-scale groundwater recharge estimate was done by Döll (2002) who modelled global groundwater recharge at a spatial resolution of 0.5^0 using the WaterGAP Global Hydrological model (WGHM; Alcamo et al., 2003; Döll, 2002). In this study, the runoff was divided into fast surface runoff, slow subsurface runoff and recharge using a heuristic approach. This approach considered relief, soil texture, hydrogeology and occurrence of permafrost and glaciers for the runoff partitioning. However, WGHM failed to reliably estimate recharge in semi-arid regions (Döll, 2002). Importantly, in that study, there was no consideration of the influence of vegetation which has been reported to be the second most important determinant of recharge by many researchers (Jackson et al., 2001; Kim and Jackson, 2012; Scanlon et al., 2005). In subsequent years, several researchers have attempted to model global groundwater recharge using different global hydrological models and global-scale land surface models (Koirala et al., 2012; Scanlon et al., 2006; Wada et al., 2010).

Although a fair amount of research has been carried out to model groundwater recharge at a global scale, most studies compared results to country-level groundwater information from the FAO (FAO, 2005). FAO statistics were based on estimates compiled from national institutions. The data estimation and reporting capacities of national agencies vary significantly and raise concerns about the accuracy of the data (Kohli and Frenken, 2015). In addition, according to FAO AQUASTAT reports, most national institutions in developing countries prioritise subnational-level statistics over national-level statistics, and in most cases data are not available for all subnational entities. This decreases the accuracy of country-wide averages and raises concerns about the reliability of using them as standard comparison measures. Only a few studies have validated modelled estimates against small-scale recharge measurements. Döll and Fiedler (2008) used 51 recharge observations from arid and semi-arid regions to correct model outputs. This study develops a recharge model and undertakes a more extensive validation of it using 715 local recharge measurements. Moreover, previous research has mostly been restricted to studying meteorological influences on recharge, few studies have systematically explored global-scale factors governing recharge. Much uncertainty still exists about the relationship between groundwater recharge and topographical, lithological and vegetation factors. Without adequate knowledge of these controlling factors, our capacity to sustainably manage groundwater globally will be seriously compromised.

The major objectives of this study are to identify the most influential factors of groundwater recharge and to develop an empirical model to estimate diffuse rainfall recharge. Specifically, to quantify regional effects of meteorological, topographical, lithological and vegetation factors on groundwater recharge using data from 715 globally distributed sites. These relationships are used to build an empirical groundwater recharge model and then the global groundwater recharge is modelled at a spatial resolution of $0.5° \times 0.5°$ for the time period 1981–2014.

2 Methods

2.1 Dataset

This study is based on a compilation of recharge estimates reported in the literature from various parts of the world. This dataset is an expansion of previously collated sets of recharge studies along with the addition of new recharge estimates (Döll and Flörke, 2005; Edmunds et al., 1991; Scanlon et al., 2006; Tögl, 2010; Wang et al., 2010). The literature search was carried out using Google Scholar, Scopus and Web of Science with related keywords "groundwater recharge", "deep percolation", "diffuse recharge" and "vertical groundwater flux". Several criteria were considered in including each study. To ensure that the data reflects all seasons, recharge estimates for time periods less than 1 year were excluded. The sites with significant contribution to groundwater from streams or by any artificial means were also eliminated as the scope of this research was to model naturally occurring recharge. In order to maximise the realistic nature of the dataset, all studies using some kind of recharge modelling were removed from the dataset. After all exclusions, 715 data points spread across the globe remained (Fig. 1) and were used for further analysis. Of these studies, 345 were estimated using the tracer method, 123 using the water balance method and the remaining studies used baseflow method, lysimeter or water table fluctuation method. This diversity in recharge estimation has enabled us to evaluate systematic differences in various measurement techniques. The year of measurement or estimation of recharge estimates in the final dataset differed (provided as supplementary material), and ranged from 1981 to 2014 (Fig. 2a). This inconsistency in the data raised a challenge when choosing the time frame for factors in the modelling exercise, particularly those showing inter-annual variation. Moreover, the compiled dataset does not represent all climate zones well (Fig. 2c), as most of the studies used were done either in arid, semi-arid or temperate zones. Pasture and cropland were the dominant land uses in the dataset (Fig. 2b).

The next step was to identify potential explanatory factors that could influence recharge (henceforth referred to as
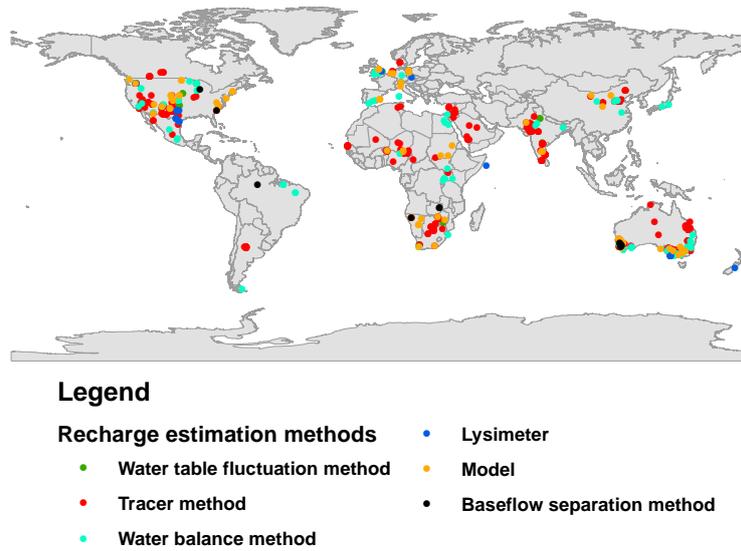
Figure 1. Locations of the 715 selected recharge estimation sites and the corresponding recharge estimation method, used for model building.
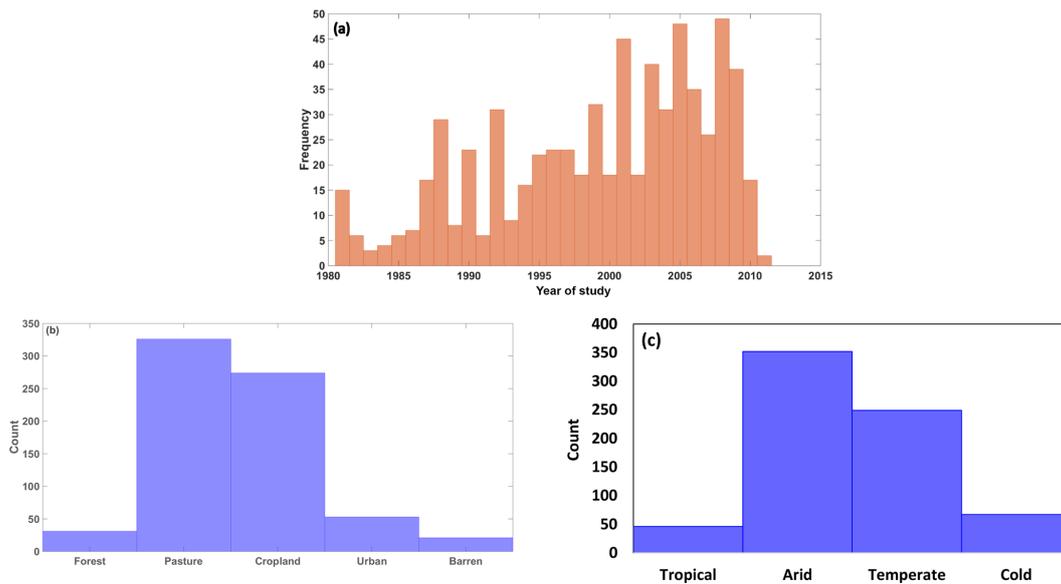
Figure 2. Histograms showing frequency of **(a)** study year, **(b)** land use and **(c)** Köppen–Geiger climate zones for the recharge estimates used.

predictors). Potential predictors that were reported in the literature as having some influence on recharge were identified (Athavale et al., 1980; Bredenkamp, 1988; Edmunds et al., 1991; Kurylyk et al., 2014; Nulsen and Baxter, 1987; O'Connell et al., 1995; Pangle et al., 2014). The choice of predictors was made based on the availability of global gridded datasets and their relative importance in a physical sense, as informed by the literature. According to the literature, the water availability on the surface for infiltration and the potential of the subsurface system to intake water are the two major controls on recharge. Different variables that can

potentially represent these two factors were chosen as predictors in this study. The water availability is represented mainly by using meteorological predictors including precipitation, potential evapotranspiration, aridity index and number of days with rainfall and vegetation characteristics (land use and land cover), whereas the intake potential is represented using various quantifiable characteristics of the vadose zone. We employed 12 predictors comprising meteorological factors, soil/vadose zone factors, vegetation factors and topographic factors. However, other factors which could have a sizable influence on recharge were not included in this

Table 1. Description of predictors used for recharge model building.

Predictors	Symbol	Unit	Resolution	Temporal span	Source	Description	Reference
Precipitation	P	mm yr^{-1}	$0.5° \times 0.5°$	1981–2014	Climatic Research Unit, University of East Anglia, United Kingdom	Mean annual precipitation	Harris et al. (2014)
Mean temperature	T	$°C$	$0.5° \times 0.5°$	1981–2014	Climatic Research Unit, University of East Anglia, United Kingdom	Mean annual temperature	Harris et al. (2014)
Potential evapotranspiration	PET	mm yr^{-1}	$0.5° \times 0.5°$	1981–2014	Climatic Research Unit, University of East Anglia, United Kingdom	Penman–Monteith reference crop evapotranspiration	Harris et al. (2014)
No. of rainy days	R_d		5 arcmin	1981–2014	AQUAMAPS, FAO	Average number of wet days per year defined as having ≥ 0.1 mm of precipitation	New et al. (2002)
Slope	S	fraction	$0.5° \times 0.5°$	–	Earthdata, NASA	Mean surface slope	Verdin (2011)
Saturated hydraulic conductivity	k_{sat}	cm day^{-1}	$1° \times 1°$	–	Earthdata, NASA	Saturated hydraulic conductivity at 0–150 cm depth	Webb et al. (2000)
Soil water storage capacity	SWSC	mm	$1° \times 1°$	–	Earthdata, NASA	Texture-derived soil water storage capacity in soil profile (up to 15 m depth)	Webb et al. (2000)
Excess water (without irrigation)	EW	mm	–	1981–2014	–	$\sum_{i=1}^{12}(P_i - \text{PET}_i)$ where $P_i > \text{PET}_i$	
Aridity index	AI	–	–	1981–2014	–	$\text{AI} = P/\text{PET}$	
Clay content	Clay	%	$1° \times 1°$	–	Earthdata, NASA	0–150 cm profile	DAAC (2016)
Bulk density	ρ_b	gm cm^{-3}	$1° \times 1°$	–	Earthdata, NASA	0–150 cm profile	DAAC (2016)
Land use land cover	LU	–	15 arcsec	–	USGS/Literature	Forest, pasture, cropland, urban/built up, barren	Kim and Jackson (2012), Broxton et al. (2014)

study because of insufficient data. Thus, we did not consider the effects of irrigation on recharge, limiting the scope of the study to rainfall-induced recharge. Subsurface lithology which could be another important recharge factor, was also eliminated from the study, due to a lack of suitable lithological and geological datasets at a larger scale. Better quality information about various predictors would have been desirable to enhance the accuracy of prediction. Details of predictors are given in Table 1.

Data for the chosen predictors corresponding to 715 recharge study sites were extracted from global datasets. Meteorological datasets (P, T and PET) were obtained from the Climatic Research Unit, University of East Anglia, United Kingdom. Even though daily data were available from 1901 to 2014 at a resolution of $0.5° \times 0.5°$, in this

study the mean annual average of the latest 34 years (1981 to 2014) was used to reduce the inconsistency in year of recharge measurements in the final dataset. Topographic and soil data were acquired from the NASA Earth observation dataset. Both datasets were of $0.5° \times 0.5°$ spatial resolution. A few of the predictors, including number of rainfall days (R_d) and land use/land cover (LU) data were obtained from AquaMaps (by FAO) and USGS (United States Geological Survey) at a spatial resolution of $0.5° \times 0.5°$ and 15 arcmin respectively. Thus-obtained LU data were compared with land cover reported in literature and corrected for any discrepancies. The spatial resolution of the different data used was diverse. This was dealt with by extracting the values for each recharge site from the original grids using the nearest neighbour interpolation method. As a result, predictor data

extracted for each recharge site will differ from the actual value due to scaling and interpolation errors. Out of the 12 predictors LU was not a quantitative predictor and was transformed into a categorical variable in the modelling exercise.

2.2 Recharge model development

With empirical studies, the science world is always sceptical about whether to use a single best-fit model or to infer results from several better-predicting and plausible models. The former option is feasible only if there exists a model which clearly surpasses other models, which is rare in the case of complex systems like groundwater. Usually cross correlation and multiple controlling influences on the system lead to more than one model having similarly good fits to the observations. Thus choosing explanatory variables and model structure is a significant challenge. In the past this challenge was often addressed using various step-wise model construction methods, with the final model being selected based on some model fit criteria that penalise model complexity (Fenicia et al., 2008; Gaganis and Smith, 2001; Jothityangkoon et al., 2001; Sivapalan et al., 2003). These approaches were pragmatic responses to the large computational load involved in trying all possible models. The disadvantage of this method is that the final model will be dependent on the step-wise selection process used (Sivapalan et al., 2003). An alternative approach for addressing this high level of uncertainty in model structure is to adopt a multi-model inference approach that compares many models (Duan et al., 2007; Poeter and Anderson, 2005). It typically results in multiple final models and an assessment of the importance of each explanatory variable. Therefore, this approach was used to develop an understanding of the role of different controlling factors on recharge in a data-limited condition.

Choosing predictors that are capable of representing the system and selecting the right models for prediction are the key steps in the multi-model inference approach. Here, models were chosen by ranking the fitted models based on performance, and comparing this to the best-performing model in the set (Anderson and Burnham, 2004). This model ranking also provided a basis for selecting individual predictors. The analysis progressed through three key stages: exploratory analysis, model building and model testing.

2.2.1 Multi-model analysis

A multi-model selection process aims to explore a wide range of model structures and to assess the predictive power of different models in comparison with others. Essentially, models with all possible combinations of selected predictors are developed and assessed via traditional model performance metrics (discussed later). By conducting such an exhaustive search, multi-model analysis avoids the problems associated with selection methods in step-wise regression

approaches (Burnham and Anderson, 2003). Importantly, it reduces the chance of missing combinations of predictors with good predictive performance. However, a disadvantage of this approach is that the number of predictor combinations grows rapidly with the number of factors considered. To make the analysis computationally efficient, we set an upper limit for the number of predictors used. Another problem with this approach is that it can result in overfitting. To address this issue we evaluated model performance with metrics that penalise complexity and tested the model robustness with a cross-validation analysis. The model development procedure using multi-model analysis is described in detail below.

(a) Exploratory analysis

Firstly, all the chosen predictors were individually regressed against the compiled recharge dataset. This was carried out with the main objective of finding the predictors with significant control over recharge and to gain an initial appreciation of how influential each predictor is compared to the others. This understanding will aid in eliminating the least influential predictors from further analysis. Then assumptions involved in regression analysis, such as linearity, low multicollinearity (important for later multivariate fitting) and independent identically distributed residuals were analysed using residual analysis. Following the residual analysis, various data transformations (square root, logarithmic and reciprocal) were carried out to reduce heteroscedasticity and improve linearity of the variables. The square root transformed recharge along with non-transformed predictors gave the most homoscedastic relationships (results not shown). Therefore, these transformed values were used in further model-building exercises. Predictors were selected and eliminated based on statistical indicators such as adjusted coefficient of determination (R^2_{adj}) value and root mean square error (RMSE).

(b) Model building

Multiple linear regression was employed for building the models as the transformed dataset did not exhibit any nonlinearity. Furthermore, the presence of both negative and positive values in the dataset restricted the applicability of other forms of regression analysis like log-linear and exponential (Saft et al., 2016). Linear regression is known for its simple and robust nature in comparison to higher-order analysis. The robustness of linear regression helped to maintain parsimony together with reasonable prediction accuracy. A rigorous model-building approach was adopted in order to capture the interplay between predictors with combined/interactive effects on groundwater recharge. This is an exhaustive search in which all candidate models are fitted and inter-compared using performance criteria. In a way, this modelling exercise used a top-down approach, starting with a simple model

which is expanded as shortcomings were identified (Fenicia et al., 2008).

(c) Model testing

The analysis above provided insight into the relative performance of the models. However, it is also important to assess the dependence of the results on the particular sample. Therefore, we conducted a subsample analysis in which the same method was re-applied to subsamples of the data. Finally, predictive uncertainty was estimated through leave-one-out cross validation. In the first case, the whole model development process was redone multiple times using subsamples of the data. To achieve this, the entire dataset was randomly divided into 80 and 20 % subsets and 80 % of the data were used for building the model. The predictive performance of the developed model was tested against the omitted 20 % of data. This was repeated 200 times in order to eliminate random sampling error. The leave-one-out cross validation was applied to the best few individual model structures and provided an estimate of predictive performance for those particular models. It also gave an indication of data quality at each point.

In summary the key steps in the multi-model analysis were

1. selecting predictors;

2. fitting all possible models consisting different combinations of predictors;

3. calculating model performance metrics for each model;

4. calculating the "weight of evidence" for each predictor based on the performance metric of all models containing that predictor;

5. testing the predictive performance of the models.

2.2.2 Ranking models and predictors

This part of the analysis has closely followed the approach developed in Saft et al. (2016). Model performance was evaluated using several information criteria. These information criteria include a goodness-of-fit term and an overfitting penalty based on the number of predictors in the model. In this study we used R^2_{adj}, the consistent Akaike information criterion (AIC_c) and the complete Akaike information criterion (CAIC) as the performance evaluation criteria. These criteria differ in terms of penalising overfitting. R^2_{adj} penalises overfitting the least, AIC_c moderately and CAIC heavily. However, when we are unsure of the true model and whether it overfits or not, there is some advantage in employing several criteria as it gives insight into how the results depend on the criteria used. Suitability of the information criteria also varies with the sample size. CAIC acts as an unbiased estimator for large sample size with relatively small candidate models, but produces large negative bias in

other cases. Conversely, AIC_c is well suited for small-sample applications (Cavanaugh and Shumway, 1997; Hurvich and Tsai, 1989). The formulas for the above criteria are as follows:

$$\text{AIC} = -2 \times \text{llf} + 2 \times k \text{ (Akaike, 1974)}, \quad (1)$$

$$\text{AIC}_\text{c} = \text{AIC} + (2 \times (k-1) \times \frac{k+2}{n-k-2}) \text{ (Hurvich and}$$
$$\text{Tsai, 1989)}, \quad (2)$$

$$\text{CAIC} = -2 \times \text{llf} + k \times (\ln(n)+1) \text{ (Bozdogan, 1987)}, \quad (3)$$

$$R^2 = 1 - \left[\frac{n-1}{n-k-1}\right] \times \left[1-R^2\right] \text{ (Ezekiel, 1929;}$$
$$\text{Wang and Thompson, 2007)}, \quad (4)$$

where "llf" is the log-likelihood function, k is the dimension of the model and n is the number of observations.

When assessing candidate models there are two questions which are of particular interest. (1) Which models are better? (2) How much evidence exists for each predictor in predicting recharge? Analysis of the AIC_c and CAIC was used to answer both these questions. Models were ranked using information criteria, with smaller values indicating better performance. Information criteria are more meaningful when they are used to evaluate the relative performance of the models (Poeter and Anderson, 2005). Models were ranked from best to worst by calculating model delta values (Δ) and model weights (W) as follows:

$$\Delta_i = \text{AIC}_i - \text{AIC}_{\text{min}}, \quad (5)$$
$$W_i = \exp(-0.5 \times \Delta_i) / \Sigma \exp(-0.5 \times \Delta_m,) \quad (6)$$

where AIC_{min} is the information criteria value of the best model. Δ_i and W_i represent the performance of ith model in comparison with the best-performing model in the set of M models.

Evidence ratios were then calculated as the ratio of the ith model weight to the best model weight. They can be used as a measure of the evidence for the ith model compared to the other models. They also provide means to estimate the importance of each predictor. This involves transformation of evidence ratios into a proportion of evidence (PoE) for each predictor. PoE for a predictor is defined as the sum of weights of all the models containing that particular predictor. PoE ranges from 0 to 1. The closer the PoE of a predictor is to 1, the more influential that predictor is.

2.3 Global groundwater recharge estimation

The best model (model 1, Table 3) from the above analysis was used to build a global recharge map at a spatial resolution of $0.5° \times 0.5°$. Recharge estimation was done annually for a study period of 34 years (1981–2014), and the estimated

Table 2. Summary statistics of potential predictors from the dataset used in this study.

| Parameters	Minimum	Maximum	Range	Mean	Standard deviation
P (mm yr^{-1})	1.30	2627.00	2625.70	572.82	305.65
T (°C)	1.60	30.62	29.02	17.73	6.04
PET (mm yr^{-1})	6.60	2600.00	2593.40	1356.17	401.77
R_d (day yr^{-1})	2.00	270.00	268.00	85.89	42.78
S	0.00	10.16	10.15	0.84	1.17
k_{sat} (cm day^{-1})	0.00	265.75	265.75	60.61	59.50
SWSC (mm)	2.00	1121.00	1119.00	517.38	240.81
AI	0.00	68.18	68.18	0.70	3.74
EW (mm yr^{-1})	0.01	1467.87	1467.86	125.41	188.07
ρ_b (gm cm^{-3})	0.15	1.67	1.51	1.44	0.20
Clay (%)	1.87	52.51	50.64	23.77	7.66
LU	1.00	5.00	4.00	2.58	0.81
Recharge (mm yr^{-1})	0.00	1375.00	1375.00	73.22	125.94

groundwater recharge was then averaged over the 34-year period to produce a global map. In addition to this, maps showing percentage of rainfall becoming recharge and the standard deviation of annual recharge over the 34 years were also generated. As recharge data from regions with frozen soil were scarce in the model-building dataset, the model predictions in those regions particularly for regions with Köppen–Geiger classification Dfc, Dfd, ET and EF were not highly reliable. EF regions of Greenland and Antarctica were excluded from the final recharge map due to lack of both recharge and predictor data. However, the modelled recharge for Dfc, Dfd and ET regions were included because of the availability of predictor data. In addition, the modelled recharge values were compared against country-level statistics from FAO (2005) for 153 countries.

3 Results

The results address three important questions. (1) What are the most influential predictors of groundwater recharge? (2) What are better models for predicting recharge? (3) How does groundwater recharge vary over space and time? The first question was answered by carrying out an exploratory data analysis and also by estimating the PoE for each predictor, the second using information criteria and the third by mapping recharge at 0.5° × 0.5° using the best model.

3.1 Exploratory data analysis

Table 2 gives the statistical summary of predictors and groundwater recharge at 715 data sites. It is apparent from the table that predictors varied considerably between sites, consistent with inter-site variability in regional physical characteristics. This variability provided an opportunity to explore recharge mechanisms in a range of different physical environments. As we used linear regression to study the one-
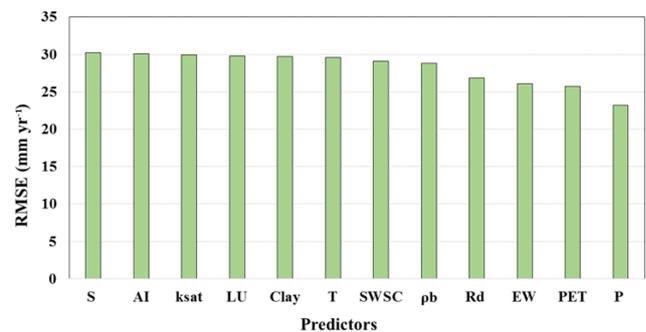
Figure 3. Model fit performance criteria for single predictor regressions.

to-one relationship of recharge with each of the predictors, RMSE and bias of fitting were used to identify the predictors with the most explanatory power. In this case, RMSE values ranged between 23.2 mm yr^{-1} for P and 30.21 mm yr^{-1} for S. Predictive potential of meteorological predictors was greater than for other classes of predictor (Fig. 3). P, AI, EW and ρ_b had a negative bias, whereas all other predictors had a positive bias.

3.2 Multi-model analysis

3.2.1 Proportion of evidence (PoE) for individual predictors

Figure 4 shows the PoE of the 12 predictors used in this study. According to this analysis, 3 of the 12 predictors stood out as having the greatest explanatory power (Fig. 4). Precipitation (P), potential evapotranspiration (PET) and land use and land cover (LU) had the highest proportions of evidence (~ 1). Subsurface percentage of clay (clay) and saturated hydraulic conductivity (k_{sat}) also had an important influence on recharge with PoE ~ 0.4. Aridity index (AI),

Table 3. Coefficient of predictors used in the top 10 models, ranked based on CAIC.

P	T	PET	R_{d}	S	k_{sat}	SWSC	AI	EW	ρ_{b}	Clay	LU	Constant	R^2_{adj}
0.0081		−0.0043									0.9567	5.3543	0.35
0.0086		−0.0044								−0.0606	1.0335	6.3781	0.25
0.0078		−0.0041							−1.9083		0.9667	7.8822	0.25
0.0076		−0.0055	−0.0247		0.0089			0.0040	−2.5857		1.0131	11.8652	0.34
0.0084		−0.0053	−0.0195					0.0036		−0.0758	1.0189	9.4112	0.33
0.0092		−0.0052	−0.0128							−0.0631	1.0409	8.2317	0.33
0.0075		−0.0050	−0.0194					0.0034	−2.3410		0.9370	11.2147	0.35
0.0084		−0.0049	−0.0130						−2.0104		0.9716	9.8549	0.35
0.0086		−0.0050	−0.0122								0.9607	7.0692	0.33
0.0086		−0.0053	−0.0166		0.0075				−2.1688		1.0402	10.2082	0.33
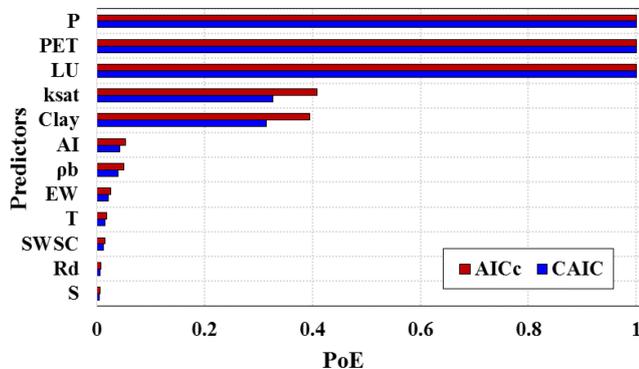
Figure 4. Proportion of evidence according to $\mathrm{AIC_c}$ and CAIC for 12 predictors (sorted in descending order of PoE).

rainfall days (R_{d}), mean temperature (T), bulk density (ρ_{d}), slope (S), excess water (EW) and soil water storage capacity at root zone (SWSC) were in the lower PoE range (< 0.1 according to both the criteria). There was some variation in the PoE value of the predictors depending on the performance metric, due to the diversity in overfitting penalty. However, ranking of the variables was identical irrespective of the performance metric used. The "best" and "worst" predictors ranked according to R^2_{adj} were also in agreement with the PoE analysis (not shown). In addition, results of the subsample analysis gave similar results (not shown).

3.2.2 Better-performing models

According to information criteria, the performance of models can only be evaluated relative to the best-performing model in the set. In this study, as per the model weights, no model exhibited apparent dominance. The evidence ratio (ratio between the weights of the best model and nth model) suggested that the best model according to CAIC was only 1.04 times better than the second best model. However, the evidence ratio increased exponentially with increase in model rank and there was a clear distinction between better models and worse models. Similar results were reported

by Saft et al. (2016) in her work for modelling rainfall–runoff relationship shift. The choice of better models was made by considering the PoE of individual predictors (refer Sect. 3.2.1) and the number of predictors in the model (V). Figure 5 shows the performance criteria for the top three models for different V values. The model performance increased with V up to 6 to 7 depending on the different criteria. After that, $\mathrm{AIC_c}$, CAIC, RMSE and R^2_{adj} values remained almost constant, indicating that further addition of predictors did not improve the model performance. In particular CAIC shows reaches a minimum at $V = 7$ and it penalises model complexity more rigorously. Table 3 illustrates the predictors in the top 10 models selected based on CAIC. All the top 10 models had $V <= 7$. P, PET and LU repeatedly appeared in the predictor list of the top 10 models substantiating their high predictive capacity, and the top ranked model includes these three predictors only. In this particular case, top-performing models according to both information criteria were the same; therefore results from only one criteria (CAIC) will be discussed.

3.2.3 Model testing

Models ranking from 1 to 10 according to CAIC (Table 3) were tested using both model testing techniques discussed in Sect. 2.2.1c. Figure 6 depicts model fit and model prediction RMSE values of 200 subsample tests. It is clear from the box plots that the difference between the RMSE of the 1st and the 10th model during both model fitting and prediction is less than $1\,\mathrm{mm\,yr^{-1}}$. In subsample tests, R^2_{adj} of the best model ranged from 0.42 to 0.56 implying 42 to 56 % of the variance was explained (please see Sect. 3.2.3 for details on sub-sample testing). The model errors at each data point ranged from −8 to $28\,\mathrm{mm\,yr^{-1}}$. However, 97.2 % of the points had errors between −8 and $10\,\mathrm{mm\,yr^{-1}}$. Figure 7 shows the relationship between precipitation and model errors and it is evident from this scatter plot that model predictions were not greatly influenced by low or high precipitation. In other words, the model was unbiased by precipitation trends. Similar checking was done for all other predictors
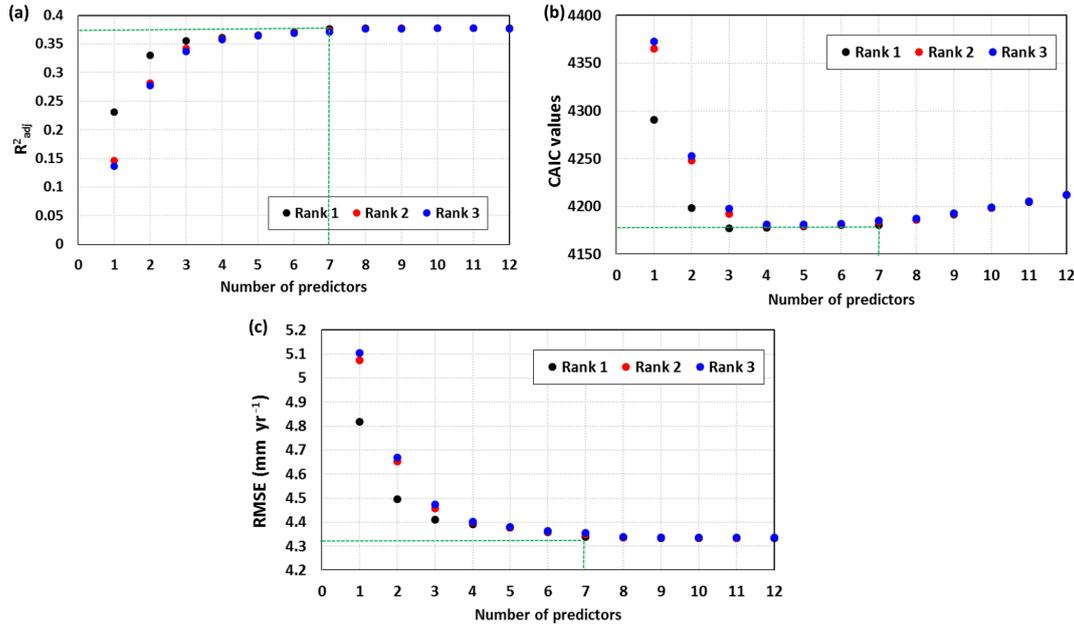
Figure 5. (a) R^2_{adj}, **(b)** CAIC and **(c)** RMSE for the top three models with 1 to 12 predictors and the green dotted lines representing the number of predictors for the best performance criteria value.
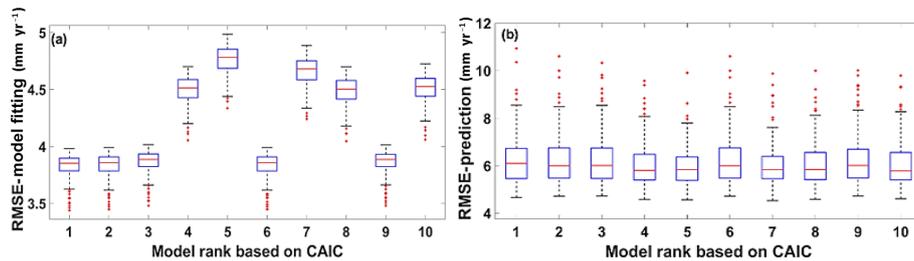
Figure 6. RMSE of sub-sample **(a)** model fitting and **(b)** model prediction of top 10 models according to CAIC.

(not shown) which all showed a similar pattern to precipitation. The dataset was classified based on recharge estimation techniques and model performance was tested with results showing no systematic difference (not shown).

3.3 Global groundwater recharge

The global long term (1981–2014) mean annual groundwater recharge map at a spatial resolution of 0.5° was made by the model developed in Sect. 3.2 (Fig. 8). In this study, the best model as defined by CAIC (model 1 in Table 3) was used to generate the recharge map. However, due to the similarity in structure of the top 10 models (Table 3), all models were equally good at predicting groundwater recharge and gave similar results (not shown). Grid-scale recharge ranged from 0.02 to 996.55 mm yr^{-1} with an average of 133.76 mm yr^{-1}. The highest recharge was associated with very high rainfall (> 4000 mm yr^{-1}). Humid regions such as Indonesia, Philippines, Malaysia, Papua New Guinea, Amazon, western Africa, Chile, Japan and Norway had very high recharge
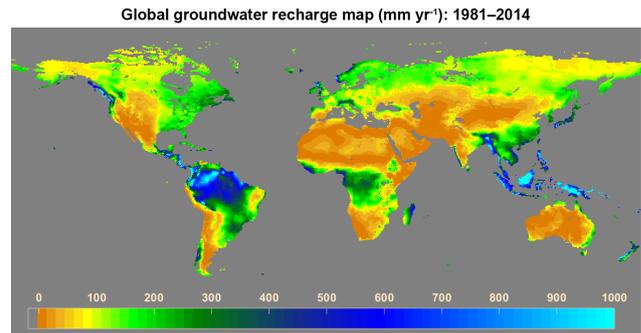
(> 450 mm yr^{-1}), whereas arid regions of Australia, the Middle East and Sahara had very low recharge (< 0.1 mm yr^{-1}). In humid areas, percentage of rainfall becoming groundwater recharge ($> 40\%$) was found to be very high in comparison to other parts of the world. However, the mean percentage of rainfall becoming recharge is only 22.06 % across the globe. Among all the continents, Australia had the lowest annual groundwater recharge rate.

Over the 34 years, global annual mean recharge followed the same pattern as that of global annual mean precipitation (Fig. 9). Least recharge was predicted in the year 1987 (groundwater recharge $= 95$ mm yr^{-1}), where the annual average rainfall was < 180 mm yr^{-1}. Variation in recharge over the years was maximal in arid regions of Australia and North Africa (Fig. 10a). However, the standard deviation of recharge was higher in humid areas than in arid regions (Fig. 10b). This indicates that standard deviation did not clearly represent year-to-year variations in recharge. Potentially, the advantage of using coefficient of variation over standard deviation is that it can capture variations even when
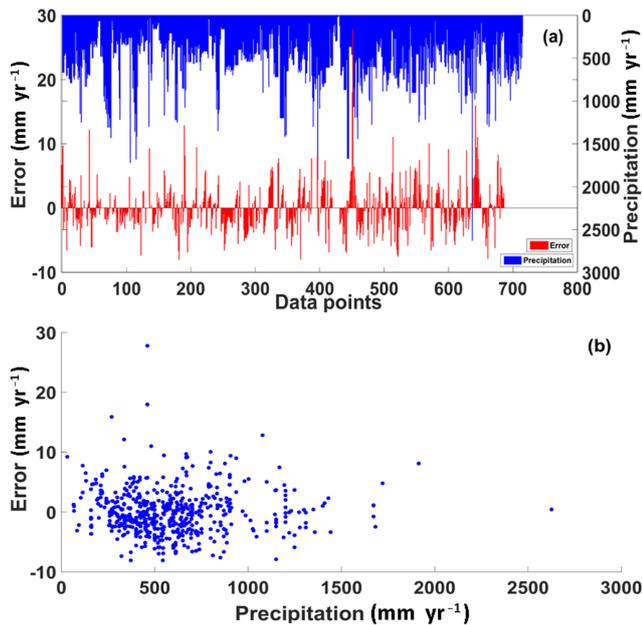
Figure 7. (a) Error at each data point along with the corresponding rainfall obtained using the leave-one-out model testing procedure and **(b)** Scatter plot between error at each data point and corresponding precipitation.

Figure 8. Long-term (1981–2014) average annual groundwater recharge estimated using the developed model.

mean values are very small. In this case precipitation and potential evapotranspiration were the two major predictors of recharge. Globally, variability in evapotranspiration is much less than variability in rainfall (Peel et al., 2001; Trenberth and Guillemot, 1995). Therefore, variability of groundwater recharge both temporally and spatially is due to variability in precipitation, which implies that arid regions are more susceptible to inter-annual variation in groundwater recharge. A comparison of predicted recharge against country-level recharge estimates from FAO (2005) shows that the model tends to overpredict recharge, particularly for low recharge areas. However, due to inaccuracies in the FAO estimates this cannot be considered as a reliable comparison (Fig. 11a). Recharge estimates from the best models in the present study were compared to recharge estimates from the complex hydrological model (WaterGAP; Fig. 11b). Even though the model in this study overestimates recharge for countries with fewer data points, the scatter shows a smaller spread compared to the FAO estimates. Figure 12 shows the country-wide distribution of errors in model prediction in comparison with FAO statistics. Very high errors were found in countries with fewer model-building data points. The model considerably overestimated recharge for Russia, Canada, Brazil, Indonesia, Malaysia and Madagascar.

4   Discussion

The aims of this study were to identify the factors with the most influence on groundwater recharge, and to develop a global model for predicting groundwater recharge under limited data conditions, without extensive water balancing. In this study, an empirical model-building exercise employing linear regression analysis, multimodel inference techniques and information criteria was used to identify the most influential predictors of groundwater recharge and use them to build predictive models. Finally, a global groundwater recharge map was created using the developed model. The key findings from this study and their implications for future research and practice with respect to global groundwater recharge are discussed below.

One of the findings to emerge is that, out of numerous models developed in this study there was no single best model for groundwater recharge. Instead, there were clear sets of better and worse models. However, there were predictors which stood out as having greater explanatory power. Of the 12 predictors chosen for the analysis, meteorological (P, PET) and vegetation predictors (LU) had the most explanatory information followed by saturated hydraulic conductivity and clay content. Thus models using these predictors ranked higher according to information criteria. It is reasonable that meteorological factors had the most explanatory information. In most cases, especially dry regions, groundwater recharge is controlled by the availability of water at the surface, which is mainly controlled by precipitation, evapotranspiration and geomorphic features (Scanlon et al., 2002). Numerous studies agree with this finding. For example, in south-western USA, 80 % of recharge variation is explained by mean annual precipitation (Keese et al., 2005). However, the influence of meteorological factors on groundwater recharge is highly site specific (Döll and Flörke, 2005). The effect of meteorological factors can also depend on whether the season or year is wet or dry, type of aquifer and irrigation intensity (Adegoke et al., 2003; Moore and Rojstaczer, 2002; Niu et al., 2007).
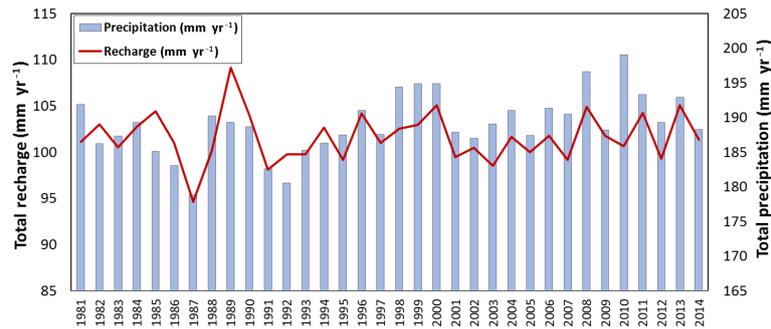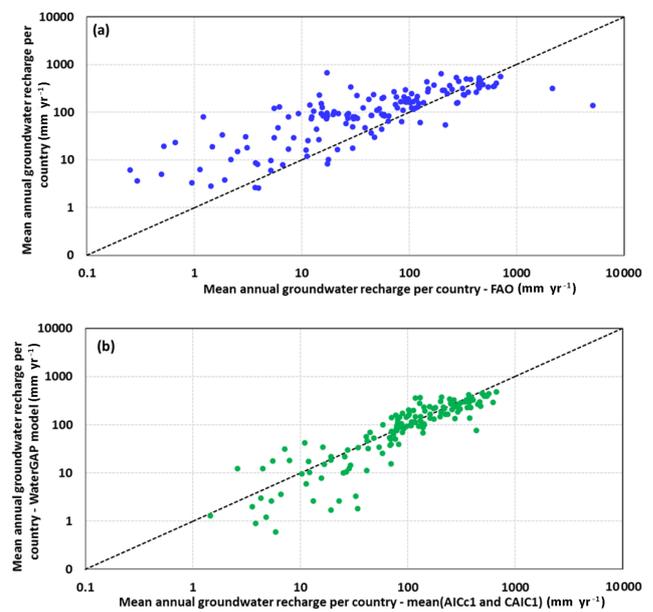
Figure 9. Temporal distribution of total global recharge along with total global precipitation of corresponding years for a period of 1981 to 2014.
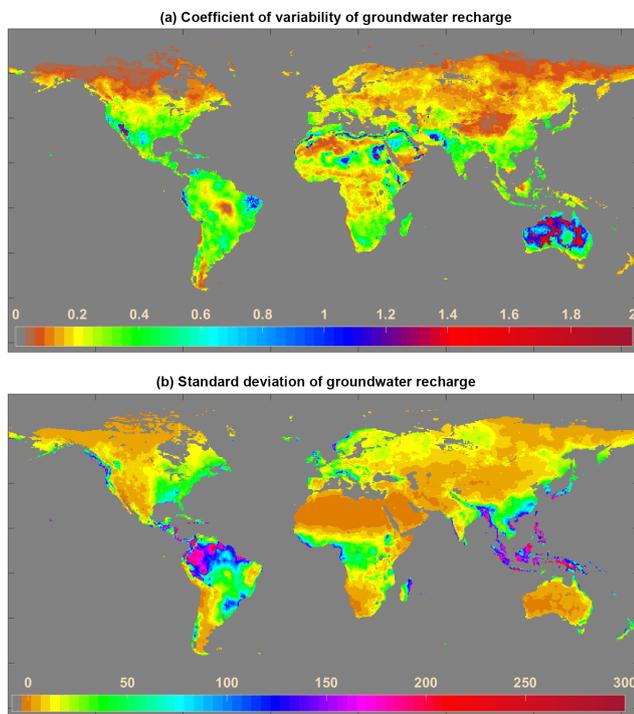
Figure 10. Map showing **(a)** coefficient of variability and **(b)** standard deviation of annual groundwater recharge from 1981 to 2014.

Figure 11. Comparison of predicted recharge against country-level estimates from **(a)** FAO and **(b)** WaterGAP model.

Many studies have reported vegetation-related parameters as the second most influential predictor of groundwater recharge. Vegetation has a high correlation with other physical variables such as soil moisture, runoff capacity and porosity, which adds to its recharge explanatory power (Kim and Jackson, 2012; Scanlon et al., 2005). In this study land use (LU) was used as a proxy for vegetation. According to the results, LU was found to be one of the predictors having the highest proportion of evidence (PoE; Fig. 4). In addition, all the better-performing models included LU as one of the predictors, which clearly indicates that vegetation is one of the most influential factors for groundwater recharge. Results indicates that recharge rates were high where runoff

water had greater retention time on the surface. This was mainly observed in areas with shallow-rooted vegetation like grasslands. In deep-rooted forest areas recharge was reduced because of increased evapotranspiration (Kim and Jackson, 2012). However, not all studies reported are in agreement with vegetation as an important predictor of recharge. For example, Tögl (2010) failed to find a correlation between vegetation/land cover and recharge. This may be the result of some peculiarity in the study dataset. Apart from the predictors discussed above, depth to groundwater and surface drainage density were also identified as potential predictors of recharge from literature (Döll and Flörke, 2005; Jankiewicz et al., 2005). Despite this they were excluded from this study because of the lack of appropriate resolution global datasets.

The total recharge estimated in this study is strongly consistent with results from complex global hydrological mod-

Table 4. Global estimates of groundwater recharge.

Model used	Spatial resolution	Temporal range	Total global recharge ($km^3\,yr^{-1}$)	Reference
Empirical model	0.5°	1981–2014	13 600	Current study
WaterGAP 2	0.5°	1961–1990	14 000	Döll (2002)
WaterGAP	0.5°	1961–1990	12 666	Döll and Flörke (2005)
PCR GlobWB	0.5°	1958–2001	15 200	Wada et al. (2010)
PCR GlobWB	0.5°	1960–2010	17 000	Wada et al. (2012)
MATSIRO	1°	1985–1999	29 900	Koirala et al. (2012)
FAO Statistics	Country	1982–2014	10 613	FAO (2016)
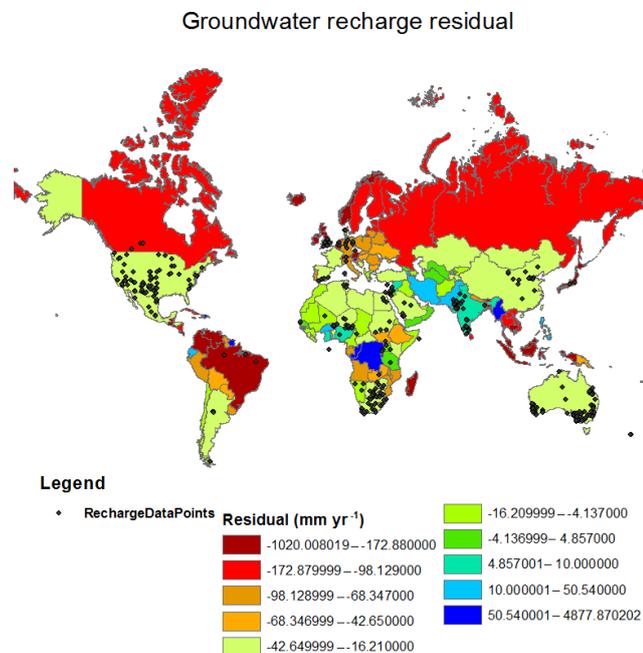
Figure 12. Spatial distribution of groundwater recharge residual (FAO estimates less model estimates) along with recharge sites selected for model building.

els. Long-term average annual recharge was found to be $134\,mm\,yr^{-1}$. The total recharge estimated in this study ($13\,600\,km^3\,yr^{-1}$) was very close to existing estimates of complex hydrological models except those using MATSIRO, which overestimates recharge in humid regions (Koirala et al., 2012). The results shown in Table 4 indicate that, compared to existing techniques, the model developed in this study can make recharge assessments with the same reliability but with fewer computational requirements. Moreover, the error in recharge prediction in this study was low, ranging from only -8 to $10\,mm\,yr^{-1}$ for 97.2 % of cases.

The global recharge map developed showed a similar pattern to recharge maps produced using complex global hydrological models. The results of this study indicate that

recharge across the globe varied considerably as a function of spatial region, and was analogous to global distribution of climate zones (Scanlon et al., 2002). Humid regions had very high recharge compared to arid (semi-arid) regions, which is obviously due to the higher availability of water for recharge. Recharge was also affected by climate variability and climate extremes at a regional level (Scanlon et al., 2006; Wada et al., 2012). However, an effect of climate variability on inter-annual recharge at a global scale was not pronounced in our results. The potential reason for this is that the El Niño Southern Oscillation (ENSO), the primary factor determining climate variability globally, has converse effects in different parts of the world. The effects of increased precipitation in some parts of the world would have been counteracted by reductions in precipitation in other areas resulting in relatively small effects on inter-annual variation in global recharge.

5 Conclusion

This study presents a new method for identifying the major factors influencing groundwater recharge and using them to model large-scale groundwater recharge. The model was developed using a dataset compiled from the literature and containing groundwater recharge data from 715 sites. In contrast to conventional water balance recharge estimation, a multi-model analysis technique was used to build the model. The model developed in this study is purely empirical and has fewer computational requirements than existing large-scale recharge modelling methods. The 0.5° global recharge estimates presented here are unique and more reliable because of the extensive validation done at different scales. Moreover, inclusion of a range of meteorological, topographical, lithological and vegetation factors adds to the predictive power of the model. The results of this investigation show that meteorological and vegetation factors had the most predictive power for recharge. The high dependency of recharge on meteorological predictors make it more vulnerable to climate change. Although it is a computationally efficient modelling method, the approach used in this study has some limita-

tions. Firstly it does not include direct anthropogenic effects on the groundwater system or recharge by natural or artificial means, suggesting scope for further future development. Secondly, the recharge dataset used in this study did not include data points from frozen regions. Therefore, Greenland and Antarctica were excluded from the final recharge map. However, the model developed in this study and the recharge maps produced will aid policy makers in predicting future scenarios with respect to global groundwater availability.

Data availability. Data used for this study are provided as a Supplement.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/hess-22-2689-2018-supplement.

Competing interests. The authors declare that they have no conflict of interest.

Edited by: Nandita Basu
Reviewed by: two anonymous referees

References

Adegoke, J. O., Pielke Sr., R. A., Eastman, J., Mahmood, R., and Hubbard, K. G.: Impact of irrigation on midsummer surface fluxes and temperature under dry synoptic conditions: A Fregional atmospheric model study of the US High Plains, Mon. Weather Rev., 131, 556–564, 2003.

Aguilera, H. and Murillo, J.: The effect of possible climate change on natural groundwater recharge based on a simple model: a study of four karstic aquifers in SE Spain, Environ. Geol., 57, 963–974, 2009.

Akaike, H.: A new look at the statistical model identification, IEEE T. Automat. Control, 19, 716–723, 1974.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global model of water use and availability, Hydrolog. Sci. J., 48, 317–337, 2003.

Ali, R., McFarlane, D., Varma, S., Dawes, W., Emelyanova, I., and Hodgson, G.: Potential climate change impacts on the water balance of regional unconfined aquifer systems in southwestern Australia, Hydrol. Earth Syst. Sci., 16, 4581–4601, https://doi.org/10.5194/hess-16-4581-2012, 2012.

Anderson, D. and Burnham, K.: Model selection and multi-model inference, in: 2nd Edn., Springer-Verlag, New York, 2004.

Athavale, R., Murti, C., and Chand, R.: Estimation of recharge to the phreatic aquifers of the Lower Maner Basin, India, by using the titrium injection method, J. Hydrol., 45, 185–202, 1980.

Bozdogan, H.: Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions, Psychometrika, 52, 345–370, 1987.

Bredenkamp, D.: Quantitative estimation of ground-water recharge in dolomite, in: Estimation of Natural Groundwater Recharge, Springer, Dordrecht, 449–460, 1988.

Broxton, P. D., Zeng, X., Sulla-Menashe, D., and Troch, P. A.: A Global Land Cover Climatology Using MODIS Data, J. Appl. Meteorol. Clim., 53, 1593–1605, https://doi.org/10.1175/jamc-d-13-0270.1, 2014.

Burnham, K. P. and Anderson, D. R.: Model selection and multimodel inference: a practical information-theoretic approach, Springer Science & Business Media, New York, 2003.

Cavanaugh, J. E. and Shumway, R. H.: A bootstrap variant of AIC for state-space model selection, Statist. Sin., 7, 473–496, 1997.

Chaussard, E., Wdowinski, S., Cabral-Cano, E., and Amelung, F.: Land subsidence in central Mexico detected by ALOS InSAR time-series, Remote Sens. Environ., 140, 94–106, 2014.

DAAC: Spatial Data Access Tool (SDAT), in: ORNL Distributed Active Archive Center, https://daac.ornl.gov/get_data/, last access: 23 September 2016.

Döll, P.: Impact of climate change and variability on irrigation requirements: a global perspective, Climatic Change, 54, 269–293, 2002.

Döll, P. and Fiedler, K.: Global-scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12, 863–885, https://doi.org/10.5194/hess-12-863-2008, 2008.

Döll, P. and Flörke, M.: Global-Scale estimation of diffuse groundwater recharge: model tuning to local data for semi-arid and arid regions and assessment of climate change impact, Frankfurt Hydrology Paper 03, Institute of Physical Geography, Frankfurt University, Frankfurt, Germany, http://www.geo.uni-frankfurtde/ipg/ag/dl/fpublikationen/2005/FHP03DoellFloerke2005.pdf (last access: May 2018), 2005.

Duan, Q., Ajami, N. K., Gao, X., and Sorooshian, S.: Multi-model ensemble hydrologic prediction using Bayesian model averaging, Adv. Water Resour., 30, 1371–1386, 2007.

Edmunds, W., Gaye, C., and Fontes, J.-C.: A record of climatic and environmental change contained in interstitial waters from the unsaturated zone of northern Senegal, in: International Symposium on Isotope Techniques in Water Resources Development, Vienna, 533–549, 1991.

Ezekiel, M.: The application of the theory of error to multiple and curvilinear correlation, J. Am. Stat. Assoc., 24, 99–104, 1929.

FAO – Food and Agriculture Organization of the United Nations: Aspects of FAO's policies, programmes, budget and activities aimed at contributing to sustainable development, Rome, Italy, 2005.

FAO: AQUASTAT Main Database, http://www.fao.org/nr/water/aquastat/sets/index.stm, last access: 20 March 2016.

Fenicia, F., Savenije, H. H., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, Water Resour. Res., 44, W01402, https://doi.org/10.1029/2006WR005563, 2008.

Gaganis, P. and Smith, L.: A Bayesian approach to the quantification of the effect of model error on the predictions of groundwater models, Water Resour. Res., 37, 2309–2322, 2001.

Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., and Lalic, N.: World population stabilization unlikely this century, Science, 346, 234–237, 2014.

Gurdak, J. J.: Ground-water vulnerability: nonpoint-source contamination, climate variability, and the high plains aquifer, VDM Verlag Publishing, Saarbrücken, Germany, p. 223, 2008.

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642, 2014.

Herrera-Pantoja, M. and Hiscock, K.: The effects of climate change on potential groundwater recharge in Great Britain, Hydrol. Process., 22, 73–86, 2008.

Holman, I., Tascone, D., and Hess, T.: A comparison of stochastic and deterministic downscaling methods for modelling potential groundwater recharge under climate change in East Anglia, UK: implications for groundwater resource management, Hydrogeol. J., 17, 1629–1641, 2009.

Hurvich, C. M. and Tsai, C.-L.: Regression and time series model selection in small samples, Biometrika, 76, 297–307, 1989.

Jackson, R. B., Carpenter, S. R., Dahm, C. N., McKnight, D. M., Naiman, R. J., Postel, S. L., and Running, S. W.: Water in a changing world, Ecol. Appl., 11, 1027–1045, 2001.

Jankiewicz, P., Neumann, J., Duijnisveld, W., Wessolek, G., Wycisk, P., and Hennings, V.: Abflusshöhe, Sickerwasserrate, Grundwasserneubildung – Drei Themen im Hydrologischen Atlas von Deutschland, Hydrol. Wasserbewirt., 49, 2–13, 2005.

Jothityangkoon, C., Sivapalan, M., and Farmer, D.: Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development, J. Hydrol., 254, 174–198, 2001.

Jyrkama, M. I. and Sykes, J. F.: The impact of climate change on spatially varying groundwater recharge in the grand river watershed (Ontario), J. Hydrol., 338, 237–250, 2007.

Keese, K., Scanlon, B., and Reedy, R.: Assessing controls on diffuse groundwater recharge using unsaturated flow modeling, Water Resour. Res., 41, W06010, https://doi.org/10.1029/2004WR003841, 2005.

Kim, J. H. and Jackson, R. B.: A global analysis of groundwater recharge for vegetation, climate, and soils, Vadose Zone J., 11, https://doi.org/10.2136/vzj2011.0021RA, 2012.

Kohli, A. and Frenken, K.: Renewable Water Resources Assessment – 2015 AQUASTAT methodology review, Food and Agricultural Organisation of the United Nations, Rome, Italy, 1–6, 2015.

Koirala, S., Yamada, H., Yeh, P., Oki, T., Hirabayashi, Y., and Kanae, S.: Global simulation of groundwater recharge, water table depth, and low flow using a land surface model with groundwater representation, J. Jpn. Soc. Civ. Eng., 68, I_211–I_216, 2012.

Kurylyk, B. L., MacQuarrie, K. T., and Voss, C. I.: Climate change impacts on the temperature and magnitude of groundwater discharge from shallow, unconfined aquifers, Water Resour. Res., 50, 3253–3274, 2014.

L'vovich, M. I.: World water resources and their future, American Geophysical Union, Washington, D.C., USA, 1979.

Moore, N. and Rojstaczer, S.: Irrigation's influence on precipitation: Texas High Plains, USA, Geophys. Res. Lett., 29, 1755, https://doi.org/10.1029/2002GL014940, 2002.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Clim. Res., 21, 1–25, 2002.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., Gulden, L. E., and Su, H.: Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res.-Atmos., 112, D07103, https://doi.org/:10.1029/2006JD007522, 2007.

Nulsen, R. and Baxter, I.: Water use by some crops and pastures in the southern agricultural areas of Western Australia, Report 32, Department of Agriculture and Food, Western Australia, 1987.

O'Connell, M. G., O'Leary, G. J., and Incerti, M.: Potential groundwater recharge from fallowing in north-west Victoria, Australia, Agr. Water Manage., 29, 37–52, 1995.

Ortiz-Zamora, D. and Ortega-Guerrero, A.: Evolution of long-term land subsidence near Mexico City: Review, field investigations, and predictive simulations, Water Resour. Res., 46, W01513, https://doi.org/10.1029/2008WR007398, 2010.

Pangle, L. A., Gregg, J. W., and McDonnell, J. J.: Rainfall seasonality and an ecohydrological feedback offset the potential impact of climate warming on evapotranspiration and groundwater recharge, Water Resour. Res., 50, 1308–1321, 2014.

Peel, M. C., McMahon, T. A., Finlayson, B. L., and Watson, F. G.: Identification and explanation of continental differences in the variability of annual runoff, J. Hydrol., 250, 224–240, 2001.

Phien-Wej, N., Giao, P., and Nutalaya, P.: Land subsidence in Bangkok, Thailand, Eng. Geol., 82, 187–201, 2006.

Poeter, E. and Anderson, D.: Multimodel ranking and inference in ground water modeling, Groundwater, 43, 597–605, 2005.

Richey, A. S., Thomas, B. F., Lo, M. H., Reager, J. T., Famiglietti, J. S., Voss, K., Swenson, S., and Rodell, M.: Quantifying renewable groundwater stress with GRACE, Water Resour. Res., 51, 5217–5238, 2015.

Saft, M., Peel, M. C., Western, A. W., and Zhang, L.: Predicting shifts in rainfall–runoff partitioning during multiyear drought: Roles of dry period and catchment characteristics, Water Resour. Res., 52, 9290–9305, 2016.

Sanford, W.: Recharge and groundwater models: an overview, Hydrogeol. J., 10, 110–120, 2002.

Scanlon, B. R., Healy, R. W., and Cook, P. G.: Choosing appropriate techniques for quantifying groundwater recharge, Hydrogeol. J., 10, 18–39, 2002.

Scanlon, B. R., Reedy, R. C., Stonestrom, D. A., Prudic, D. E., and Dennehy, K. F.: Impact of land use and land cover change on groundwater recharge and quality in the southwestern US, Global Change Biol., 11, 1577–1593, 2005.

Scanlon, B. R., Keese, K. E., Flint, A. L., Flint, L. E., Gaye, C. B., Edmunds, W. M., and Simmers, I.: Global synthesis of groundwater recharge in semiarid and arid regions, Hydrol. Process., 20, 3335–3370, 2006.

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrological prediction, Hydrol. Process., 17, 2101–2111, 2003.

Sreng, S., Li, L., Sugiyama, H., Kusaka, T., and Saitoh, M.: Upheaval phenomenon in clay ground induced by rising groundwa-

ter level, Proceedings of the fourth biot conference on poromechanics, Columbia University, New York, 106–203, 2009.

Tögl, A. C.: Global groundwater recharge: Evaluation of modeled results on the basis of independent estimates, Institut für Physische Geographie Fachbereich Geowissenschaften/ Geographie, Goethe-Universität, Frankfurt am Main, 2010.

Trenberth, K. E. and Guillemot, C. J.: Evaluation of the global atmospheric moisture budget as seen from analyses, J. Climate, 8, 2255–2272, 1995.

Verdin, K.: ISLSCP II HYDRO1k elevation-derived products, in: ISLSCP Initiative II Collection, edited by: Hall, F. G., Collatz, G., Meeson, B., Los, S., Brown de Colstoun, E., and Landis, D., Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1007, 2011.

Wada, Y., van Beek, L. P., van Kempen, C. M., Reckman, J. W., Vasak, S., and Bierkens, M. F.: Global depletion of groundwater resources, Geophys. Res. Lett., 37, L20402, https://doi.org/10.1029/2010GL044571, 2010.

Wada, Y., Beek, L. P., Sperna Weiland, F. C., Chao, B. F., Wu, Y. H., and Bierkens, M. F.: Past and future contribution of global groundwater depletion to sea-level rise, Geophys. Res. Lett., 39, L09402, https://doi.org/10.1029/2012GL051230, 2012.

Wang, L., Dochartaigh, B., and Macdonald, D.: A literature review of recharge estimation and groundwater resource assessment in Africa, Natural Environment Research Council (NERC), Nottingham, England, 2010.

Wang, Z. and Thompson, B.: Is the Pearson r^2 biased, and if so, what is the best correction formula?, J. Exp. Educ., 75, 109–125, 2007.

Webb, R. W., Rosenzweig, C. E., and Levine, E. R.: Global soil texture and derived water-holding capacities (Webb et al.), Data set, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, available on-line at: http://www.daac.ornl.gov (last access: 15 July 2016), 2000.