Developing a pan-European high-resolution groundwater recharge map – Combining satellite data and national survey data using machine learning

Martinsen, Grith; Bessiere, Helene; Caballero, Yvan; Koch, Julian; Collados-Lara, Antonio Juan; Mansour, Majdi; Sallasmaa, Olli; Pulido-Velazquez, David; Williams, Natalya Hunter; Zaadnoordijk, Willem Jan

DOI
10.1016/j.scitotenv.2022.153464

Publication date
2022

Document Version
Final published version

Published in
Science of the Total Environment

Citation (APA)
Martinsen, G., Bessiere, H., Caballero, Y., Koch, J., Collados-Lara, A. J., Mansour, M., Sallasmaa, O., Pulido-Velazquez, D., Williams, N. H., Zaadnoordijk, W. J., & Stisen, S. (2022). Developing a pan-European high-resolution groundwater recharge map – Combining satellite data and national survey data using machine learning. *Science of the Total Environment*, *822*, 1-15. Article 153464. https://doi.org/10.1016/j.scitotenv.2022.153464

Science of the Total Environment 822 (2022) 153464

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Developing a pan-European high-resolution groundwater recharge map – Combining satellite data and national survey data using machine learning

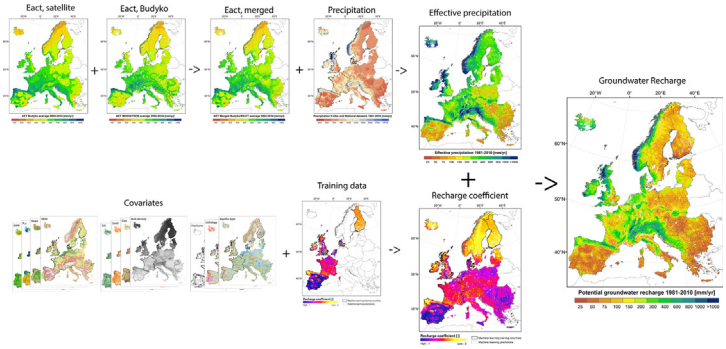
Grith Martinsen [a], Helene Bessiere [b], Yvan Caballero [c,d], Julian Koch [a], Antonio Juan Collados-Lara [e], Majdi Mansour [f], Olli Sallasmaa [g], David Pulido-Velazquez [e], Natalya Hunter Williams [h], Willem Jan Zaadnoordijk [i,j], Simon Stisen [a,*]

[a] Geological Survey of Denmark and Greenland, GEUS, Øster Voldgade 10, 1350 Copenhagen K, Denmark
[b] French Geological Survey, BRGM, 3 avenue Claude-Guillemin, BP 36009 45060 Orléans Cedex 02, France
[c] French Geological Survey, BRGM, Univ Montpellier, Montpellier, France
[d] G-eau, UMR 183, INRAE, CIRAD, IRD, AgroParisTech, Supagro, BRGM, Montpellier, France
[e] Spanish Geological Survey, IGME-CSIC, Urb. Alcázar del Genil, 4-Edif. Zulema, Bajo, 18006 Granada, Spain
[f] British Geological Survey, BGS–Keyworth, Nottingham NG12 5GG, United Kingdom
[g] Geological Survey of Finland - GTK, Vuorimiehentie 5, 96, FI-02151 Espoo, Finland
[h] Geological Survey Ireland, GSI, Booterstown Hall, Blackrock, Co. Dublin A94 N2R6, Ireland
[i] Geological Survey of the Netherlands, TNO, Princetonlaan 6, 3584 CB Utrecht, Netherlands
[j] Delft University of Technology, Faculty of Civil Engineering and Geosciences, Water Resources Section, Stevinweg 1, 2628 CN Delft, Netherlands

HIGHLIGHTS

- Combining national datasets, satellite data and machine learning
- A multilayer Pan-European dataset of E_{act}, P_{Eff} and groundwater recharge coefficients
- Long-term average groundwater recharge map for Europe in 1 km resolution
- A locally relevant, seam-less and harmonised Pan-European dataset

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 October 2021
Received in revised form 18 January 2022
Accepted 23 January 2022
Available online 29 January 2022

Editor: Christian Herrera

Keywords:
Groundwater recharge
Pan-European
Satellite data
Machine learning
Recharge coefficient
Effective precipitation

* Corresponding author.
E-mail address: sst@geus.dk (S. Stisen).

ABSTRACT

Groundwater recharge quantification is essential for sustainable groundwater resources management, but typically limited to local and regional scale estimates. A high-resolution (1 km × 1 km) dataset consisting of long-term average actual evapotranspiration, effective precipitation, a groundwater recharge coefficient, and the resulting groundwater recharge map has been created for all of Europe using a variety of pan-European and seven national gridded datasets. As an initial step, the approach developed for continental scale mapping consists of a merged estimate of actual evapotranspiration originating from satellite data and the vegetation controlled Budyko approach to subsequently estimate effective precipitation. Secondly, a machine learning model based on the Random Forest regressor was developed for mapping groundwater recharge coefficients, using a range of covariates related to geology, soil, topography and climate. A common feature of the approach is the validation and training against effective precipitation, recharge coefficients and groundwater recharge from seven national gridded datasets covering the UK, Ireland, Finland, Denmark, the Netherlands, France and Spain, representing a wide range of climatic and hydrogeological conditions across Europe. The groundwater recharge map provides harmonised high-resolution estimates across Europe and locally relevant estimates for areas where this information is otherwise not available, while being consistent with the

http://dx.doi.org/10.1016/j.scitotenv.2022.153464

existing national gridded datasets. The Pan-European groundwater recharge pattern compares well with results from the global hydrological model PCR-GLOBWB 2. At country scale, the results were compared to a German recharge map showing great similarity. The full dataset of long-term average actual evapotranspiration, effective precipitation, recharge coefficients and groundwater recharge is available through the EuroGeoSurveys' open access European Geological Data Infrastructure (EGDI).

1. Introduction

Groundwater plays a vital role for humans and for ecosystems worldwide. It is employed to supply irrigation, drinking water and industrial demands, and sustains wetlands, rivers and other groundwater-dependent ecosystems. However, groundwater resources are under pressure from both overexploitation and climate change (Cuthbert et al., 2019; Pulido-Velazquez et al., 2018b). The rate of groundwater recharge, expressing the annual renewal potential, is a key hydrological variable for quantifying available water resources, safe yields and vulnerability to climate variability and change (Riedel and Weber, 2020). It is largely controlled by: precipitation, evapotranspiration, surface infiltration, land cover and soil hydraulic characteristics (Mohan et al., 2018). Consequently, it is highly variable in time and space, mainly across climate gradients, but also as a result of patterns in topography, soil, land cover and geology (MacDonald et al., 2021; Moeck et al., 2020; Xu and Beekman, 2019).

Groundwater recharge can be measured locally using lysimeters (Voortman et al., 2015; Xu and Chen, 2005). However, these measurements are difficult to scale up, due to the impact of local soil and land use characteristics and the fact that natural conditions are disturbed when constructing the lysimeter (Vásquez et al., 2015). Alternative methods based on water table fluctuations use data that is naturally integrated over a larger scale (Jie et al., 2011), but require additional information such as storage coefficient and special conditions such as distinguishable precipitation events (Crosbie et al., 2005).

Several of the key factors controlling groundwater recharge are difficult to estimate and display significant spatial variability even within river basin scale. Consequently, accurate large-scale quantification of groundwater recharge is challenging, since local measurements are difficult to scale-up and coarse-scale modelling lacks spatial heterogeneity. Due to the timescales of groundwater movement, global to continental scale mapping of groundwater recharge typically focusses on long-term average rates, to quantify sustainable resources at decadal scales. Such large-scale assessments are typically based on either upscaling of local estimates (MacDonald et al., 2021; Moeck et al., 2020) or large-scale modelling (Döll and Fiedler, 2008).

A recent study for the Continental US approached the challenges of estimating large-scale groundwater recharge using a residual approach utilising stream gauge data from a large set of watersheds and estimating quick-flow and evapotranspiration to train regression equations based on continentally available data (Reitz et al., 2017).

For the African continent, (MacDonald et al., 2021) upscaled 134 long-term average groundwater recharge estimates using linear mixed models based on climate, vegetation, soil and geological covariates. However, the final prediction model was reduced to a simple linear function of long-term average rainfall, since no other covariates contributed to the performance, highlighting the potential for improvement particularly in more data rich regions. Hence, upscaling by interpolation remains sensitive to inaccuracy and misrepresentation by often few local estimates and lack of known continentally valid upscaling relations.

Mohan et al. (2018) compiled 715 recharge estimates from various parts of the world and identified potential explanatory factors that could influence those estimations to build multiple linear regression models between recharge estimates and explanatory factors and use the best one to build a global recharge map at a spatial resolution of 0.5°x0.5°.

A range of continental and global hydrological and land surface models enable simulation of groundwater recharge, however they commonly lack a systematic calibration and validation against independent groundwater recharge estimates at a relevant scale (Li et al., 2021; Reinecke et al., 2020; Samaniego et al., 2019). Continental and global models are typically evaluated against surface fluxes and river discharge, while the flow path separation between groundwater and surface water is highly conceptual, difficult to validate and consequently uncertain. Still, these models offer a unique spatial coverage and are useful for impact assessments, particularly in ensemble approaches (Reinecke et al., 2020; Wanders et al., 2019).

At the global scale (Döll and Fiedler, 2008), estimated long-term average groundwater recharge at 0.5° scale based on a heuristic recharge factor approach implemented in the global hydrological model WaterGAP (Müller Schmied et al., 2021) and dependent on globally available datasets on soil texture, relief, hydrogeology etc. The model was mainly validated against basin scale river discharge, while a validation against independent groundwater recharge estimates in semiarid regions revealed significant shortcoming.

For the Continental USA, large modelling efforts using coupled and integrated groundwater surface water models, have been conducted (Alattar et al., 2020; Maxwell et al., 2015). These enable calibration and validation against both stream flow records and groundwater level time series, in an attempt to reduce the uncertainty on groundwater recharge simulation (Alattar et al., 2020). Similarly, the PCR-GLOBWB model has been run both globally and at a European scale, producing groundwater recharge estimates (Sutanudjaja et al., 2018). Continental- to global-scale coupled groundwater surface water models remain coarse in scale and include largely simplified representations of hydrogeology and vertical discretisation, resulting in lack of local representation. This makes them inherently difficult to validate with independent estimates, and they have been shown inconsistent with groundwater storage trends when validated at large scale against observations by satellite gravimetry (Scanlon et al., 2018). However, the continuous progress on model development and data availability is closing the gap between continental and regional scale groundwater models also at the Pan-European scale (Jing et al., 2018; Trichakis et al., 2017).

Remote sensing methods have also been employed for large-scale groundwater recharge estimation, typically by combining groundwater storage fluctuations observed by the Gravity Recovery and Climate Experiment (GRACE) satellite (Richey et al., 2015; Wu et al., 2019) with land surface models. GRACE-based estimates are, however, limited to coarse spatial resolutions in the order of 300 km. Alternatively, satellite data at higher spatial resolution has been used for groundwater recharge estimation through a water budget approach (Healy and Scanlon, 2010). One such example is from a national mapping of groundwater recharge in New Zealand, including actual evapotranspiration and leaf area index estimates from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor in combination with climate and other land surface data (Westerhoff et al., 2018).

Satellite data can also provide valuable information for groundwater recharge through estimation of water balance components such as actual evapotranspiration. Satellite remote sensing methods for actual evapotranspiration estimation have enabled global mapping by exploiting the unique data availability and coverage provided by a suite of sensors and platforms. Satellite sensors do not measure actual evapotranspiration directly, but can provide detailed and accurate information on key variables linked to and controlling the evaporative processes on the land surface (Kalma et al., 2008). These include radiation, albedo, vegetation, land surface temperature etc. Based on these remotely-sensed variables, models have been developed for estimating actual evapotranspiration with a few other auxiliary data, without water balance accounting, and without the need for precipitation data. Instead, these models are typically based on an energy balance

approach, where the net incoming radiation is partitioned into sensible or latent heat. The advantage of the satellite-based methods is that they directly reflect observed patterns of key controlling states and fluxes at the Earth's surface without the need to formulate explicit relation to, e.g., soil and climate conditions.

Groundwater recharge maps at pan-European scale are needed for assessing groundwater resources and monitoring their quantitative and qualitative status (European Environmental Agency, 2018). Regional- to national-scale groundwater recharge estimates are available from several European countries, based on local data, validated methods and detailed coupled groundwater-surface water models. These gridded datasets have so far not been utilised in Pan-European groundwater recharge estimates and challenges remain on harmonizing between independent estimates and methods, and transferring knowledge to countries without national gridded estimates.

The current study exploits detailed national-scale groundwater recharge estimates from seven European countries, in combination with satellite remote sensing data and a machine learning approach to produce a harmonised Pan-European long-term average potential groundwater recharge map. The approach builds on existing databases and a recent study of long-term average evapotranspiration patterns across Europe (Stisen et al., 2021a) combined in a water budget approach, based on precipitation, actual evapotranspiration and a groundwater recharge coefficient. The goal of the study is to produce a harmonised Pan-European dataset on long-term average effective precipitation and groundwater recharge that is spatially consistent, locally relevant and honours the national datasets while providing new information outside their coverage.

2. Methods and data

2.1. Water budget approach

Several definitions of groundwater recharge exist, depending on the application, estimation method and reference scale (Healy and Scanlon, 2010). The current study aims at a large-scale (continental) and high-resolution (1 km) recharge quantification for which the water budget approach as described by Healy and Scanlon (2010) is appropriate. This approach is advantageous because it can utilise existing datasets and satellite remote sensing data. In addition, for long-term average estimates, storage changes (ΔS) in the upper soil layer can be ignored.

Consequently, groundwater recharge is defined in the context of large-scale application and a simple soil column water budget framework, where the four components of the water budget are: Precipitation (P), actual evapotranspiration (E_{act}), groundwater recharge R_{GW} and shallow runoff (Q). As such, the method does not directly account for the large scale river runoff, which is included indirectly through the national training data described in Section 2.2.

We refer to the term potential groundwater recharge ($R_{GW,pot}$) as the water percolating from the upper soil layer to the groundwater table, defined as the excess of P-E_{act}-Q. The difference between precipitation and actual evapotranspiration (P-E_{act}) is referred to as effective precipitation (P_{Eff}), which is partitioned into either groundwater recharge or shallow runoff. The fraction of the groundwater recharge is defined as the groundwater recharge coefficient ($C_{GWR} = R_{GW} / P_{Eff}$) and shallow runoff is therefore equal to $Q = P_{Eff} * (1 - C_{GWR})$. At 1 km resolution, the actual evapotranspiration refers to the total evaporative loss to the atmosphere from surface evaporation, transpiration, and groundwater and surface water evaporation. Likewise, the runoff at this scale refers to all shallow lateral flow above the groundwater table and therefore includes more than the Hortonian surface runoff. Note that our definition of groundwater recharge refers to the top of the saturated groundwater system. It does not include surface water interaction (apart from shallow runoff). Moreover, the groundwater recharge of underlying aquifers will be different, depending on the groundwater system and properties of the subsurface. In a final postprocessing step, we will evaluate the $R_{GW,pot}$ estimate in relation to aquifer properties. The equation for $R_{GW,pot}$ in our implementation of the water budget approach is:

$$R_{GW,pot} = (P - E_{act}) * C_{GWR} \quad (1)$$

with the C_{GWR} [−] expressing the fraction of effective precipitation that reaches the groundwater table.

Long-term average estimates of potential groundwater recharge for the period 1981–2010 are produced at 1 km scale across Europe, requiring equivalent inputs of precipitation [mm/year], actual evapotranspiration [mm/year], and groundwater recharge coefficients [−].

2.2. Existing gridded national estimates

Gridded national groundwater recharge estimates are available from seven European countries (UK, Ireland, Finland, France, Spain Netherlands and Denmark), originating from different approaches and relying on different input data. They are all regarded as long-term average estimates, and include separate estimates of effective precipitation and groundwater recharge including or enabling a grid-based estimate of groundwater recharge coefficient in line with Eq. 1. This is an important feature, since not only the national groundwater recharge estimates but also their associated effective precipitation and groundwater recharge coefficients will be used to evaluate and develop the Pan-European mapping.

The methods used for national scale groundwater recharge assessments can broadly be categorised into three groups. 1) Recharge/runoff coefficient-based methods combining climatologically-driven effective precipitation estimates with topography-, soil- and geology-driven mapping of recharge/runoff coefficients (UK Ireland, Finland and France), 2) Empirical models linking effective precipitation to independent point-scale estimates of groundwater recharge (Spain) and 3) Coupled 3D groundwater-surface water models calibrated against stream flow and groundwater levels, estimating groundwater recharge as an internal flux (Netherlands and Denmark). Each national gridded recharge estimate has been calibrated and or validated independently against river runoff, groundwater levels or groundwater formation datasets. An overview of the national datasets and their validation status is given in Table 1. Each of the national estimates are developed prior to the current study and constitute the official national groundwater recharge estimates by the respective national geological surveys. In selecting national recharge datasets, preference have been given to covering as much hydrogeological diversity across Europe as possible, over similarity in method or validation approach.

2.2.1. Recharge coefficient methods

For the UK, the annual average potential recharge values calculated for the period 1981–2010 are used to determine the long-term average recharge (Mansour et al., 2018). These values are calculated at a national scale using the modified EA-FAO method proposed by (Griffiths et al., 2006), which accounts for actual evaporation and soil moisture deficit, and calculates potential recharge as a fraction of the excess water. The model is driven by daily rainfall data (Tanguy et al., 2021) and monthly potential evaporation data obtained from the Met Office Rainfall and Evaporation Calculation System (MORECS) (Hough and Jones, 1997). Hydrogeological data used in the model include the soil map (Boorman et al., 1995), hydrogeological map, solid geology map and the superficial deposit map (BGS, 2021). The model also uses a land cover map LCM2000 (NERC, 2000), a digital topographical model and the UK river networks (NERC, 2003). The groundwater recharge estimates are used to drive national scale simulation of groundwater levels and river flows within the HYDRO-JULES modelling framework (Dadson et al., 2019; Pachocka et al., 2015).

For Ireland, the annual recharge to the deep groundwater system calculated over the period 1981–2010 is used. These values are estimated using guidelines outlined by the Irish Working Group on Groundwater I(IWGGW, 2005) and revised by Hunter Williams et al. (2013) and further in Hunter Williams et al. (2021). A number of hydrogeological maps including soil

Table 1
Overview of national recharge models and their evaluation. Method abbreviations are: recharge coefficient method (RCM), empirical method (EM) and coupled groundwater-surface water model (CGSM).

Country	Method	Period	Resolution	Evaluation
United Kingdom	RCM	1981–2010	2 km	Calibrated on long term average runoff values recorded at 56 gauging stations located at major rivers (Mansour et al., 2018) and monthly values from 41 stations. In addition, the map is compared to recharge values calculated at borehole scale using multiple recharge estimation tools (Seidenfaden et al., 2022).
Ireland	RCM	1981–2010	1.5 km	Recharge coefficients across the range of hydrogeological scenarios were established based on previous studies within Ireland (see Hunter Williams et al, (2013) for details). Recharge estimates validated at catchment scales by comparison with surface water flow hydrographs by Hunter Williams et al. (2011) and against hydrograph separation estimates and literature values within Geological Survey Ireland's GW3D project (unpublished).
Finland	RCM	1981–2010	1 km	Calibrated and validated against a database consisting of estimated values of groundwater formation and area provided by the Finnish Environment Institution (SYKE). Validation was made against roughly 4,850 groundwater areas from SYKE database. The average size was 2.5 km^2, ranging from 0.018 km^2 to 97.3 km^2.
France	RCM	1981–2010	8 km	The resulting potential groundwater recharge values were assessed at the river basin scales for the Rhone-Mediterranean and Corsica regions for with is was initially developed (Caballero et al., 2016) and compared to local recharge observations where available (Caballero et al., 2021). The computed effective precipitation was compared to that simulated by SURFEX (Le Moigne et al., 2020) at the Rhone Mediterranean river basin scale (Caballero et al., 2016; Le Cointe et al., 2019).
Spain	EM	1980–2010	10 km	Groundwater recharge estimates validated against 47 local recharge datasets (Alcalá and Custodio, 2014). Meteorological estimates are validated against 988 stations (Quintana-Seguí et al., 2017).
Netherlands	CGSM	2011–2018	250 m	Each version of the national model is validated against surface water fluxes at 25 sites in the main surface water system and 45 sites in regional surface waters as well as measurements of the phreatic water table and groundwater heads at thousands of locations (De Lange et al., 2014).
Denmark	CGSM	1990–2010	500 m	The national water resources model (DK-Model) was calibrated and validated against stream flow (305 stations), groundwater head levels (28,277 wells) and spatial patterns of ET_{act} from remote sensing (Soltani et al., 2021a).

drainage, generalised soil class, subsoil type, subsoil permeability, bedrock aquifer, and sand and gravel maps are combined in a GIS to give a particular hydrogeological scenario that is then related to a recharge coefficient. A map of recharge coefficients is combined (in a multiplication) with an effective rainfall map produced using daily rainfall, potential and actual evapotranspiration from MetÉireann's MÉRA model and Irish Centre for High End Computing calculations using the Schulte et al. (2005) method. This produces the groundwater recharge map. For the current study, recharge coefficients have been re-calculated as the ratio of deep groundwater recharge to effective rainfall.

For Finland, the annual recharge to the shallow groundwater system calculated over the period 1981–2010 is used. These values are estimated using a similar methodology as the Irish groundwater recharge map. Daily rainfall and temperature data obtained from the Finnish meteorological institute are used together with hydrogeological data mainly for the superficial deposits from the Finnish Geological Survey (GTK).

For France, the yearly potential groundwater recharge values calculated over the period 1981–2010 are used. The effective precipitation has been calculated using three simple water balance models (Dingman, 1994; Edijatno and Michel, 1989; Thornthwaite, 1948) driven by rainfall, temperature, and potential evapotranspiration from the SAFRAN database (Vidal et al., 2010) together with information about the maximum soil water content obtained from the French national soil map (DoneSol INRA, 2014). Effective precipitation recharge coefficients, are estimated by comparing base flow index values (Gustard et al., 1992) calculated for a set of French river basins to values built of the Network Development and Persistence Index (IDPR) map (Mardhel et al., 2021), averaged over the river basins area. The IDPR gridded map (50 m × 50 m) provides a qualification of the disparity between the theoretical drainage network produced by automatic analysis of a digital elevation model (DEM) and the actual presence of stream and river branches. Thereby, the index expresses the land surface infiltration potential spatial variation building on the assumption that, where there are no rivers (or, in contrast, where there are rivers), infiltration is dominant (or runoff is dominant). Using the resulting relationship between baseflow index values and mean IDPR values over the selected river basins, and considering the baseflow index equal to the effective precipitation infiltration ratio at the annual scale, values of the latter are calculated at the scale of the superficial and homogeneous groundwater bodies, as defined in the BDLISA_V2 database (Brugeron et al., 2018), from the mean IDPR value averaged over their surface. Finally, potential groundwater recharge is calculated by multiplying the effective precipitation infiltration ratio by the effective precipitation.

2.2.2. Empirical methods

For Spain, the annual recharge values calculated across continental Spain for the period 1980–2010 are used to determine the long-term average recharge. Recharge values are estimated by applying an empirical rainfall recharge model (Pulido-Velazquez et al., 2018a) at a resolution of 10 km. The model is defined by forcing such that a perturbation of the effective precipitation data produces a time series whose mean and standard deviation are equal to net recharge estimates derived from a chloride mass balance method (Alcalá and Custodio, 2014). The model was driven by precipitation and maximum and minimum temperature data obtained from Spain02 dataset (Herrera et al., 2016) with actual evapotranspiration calculated using Turc's model (Turc, 1954).

2.2.3. Coupled groundwater-surface water models

For the Netherlands, recharge values estimated for the period 2011–2018 obtained from the Dutch National Hydrological Instrument (NHI-LHM; http://www.nhi.nu) are used. The NHI contains a coupling of four sub-models at a resolution of 250 m, which together can simulate the groundwater (Vermeulen et al., 2021), surface water (De Lange et al., 2014), and the vadose zone (van Walsum and Groenendijk, 2008). The coupled models are driven by meteorological data including daily rainfall and potential evapotranspiration data (Royal Dutch Meteorological Institute, KNMI), and use hydrogeological data including subsurface schematisation based on geohydrological models of the Geological Survey of the Netherlands: REGIS II V2.2 (TNO-GSN, 2021a) and GeoTOP (TNO-GSN, 2021b), surface water information from Waterboards and Rijkswaterstaat land use maps, soil maps and a database of soil physical properties (Heinen et al., 2020; Schröder et al., 2021). Moreover, model input and schematization is continuously improved by the national and regional authorities and stakeholders united in the consortium of the Netherlands Hydrological Instrument.

For Denmark, recharge estimates are simulated using the National Water Resources Model (DK-Model) for the period 1990–2010 (Stisen et al., 2012). The model is based on the MIKE SHE code (Graham et al., 2005) and incorporates all major components of the hydrological cycle including 3D groundwater flow, unsaturated zone and stream routing. The model is run at 500 m resolution and calibrated against satellite-based

evapotranspiration and a large database of groundwater levels and stream gauges (Soltani et al., 2021a). For the current study, groundwater recharge is extracted as the downward flux between first and second geological/computational layers, and recharge coefficients are calculated as the ratio between groundwater recharge and effective precipitation. The DK-Model acts as the national Danish model for groundwater recharge and for reporting groundwater resources to the EU.

The gridded datasets from the seven countries, from here on referred to as the national pilots, will be utilised as evaluation and training data for the development of Pan-European datasets for effective precipitation, recharge coefficient and potential groundwater recharge.

2.3. Precipitation

Precipitation is the main driver of hydrological processes and reliable precipitation data are critical for any assessment of groundwater recharge (Mohan et al., 2018). The current study uses the latest version of the *E*-Obs precipitation dataset at 0.1° resolution with Pan-European coverage (Cornes et al., 2018). The open access E-Obs precipitation dataset is based on a European initiative to collect and harmonize precipitation records from national meteorological agencies in an ensemble of interpolations at a daily timescale. Here, the daily time series are aggregated to long-term averages for the period 1981–2010, and daily ensemble interpolation uncertainty is assumed to be negligible. Although *E*-Obs is a major step forward in collecting and harmonizing Pan-European precipitation data, the density of precipitation stations varies considerably across Europe. This is illustrated in Fig. 1a, where the stations included in *E*-Obs for the period 1981–2010 are mapped. Of the seven countries with available groundwater recharge estimates in this study, Denmark, UK, France and Spain have few rain gauges in *E*-Obs. Consequently, the long-term average precipitation data used for the National groundwater recharge estimates were merged into the E-Obs long-term average dataset by substitution. This will not only improve our precipitation dataset, but also secure that differences between national groundwater recharge estimates and Pan-European recharge estimates are not caused simply by differences in precipitation input. The resulting Pan-European long-term average precipitation map in 0.1°resolution is illustrated in Fig. 1b for the period 1981–2010 (an additional long-term average precipitation map from 2002 to 2014 is also produced for alignment with the AET estimates, see below).

2.4. Actual evapotranspiration

Actual evapotranspiration is a hydrological flux that is challenging to measure and upscale (Franssen et al., 2010). For long-term average actual evapotranspiration assessment, several methods exist including water balance approaches, the Budyko approach and satellite-based methods. The water balance approach and the Budyko approach provide estimates over longer periods and use long records of stream flow, precipitation and potential evapotranspiration to assess the actual evapotranspiration in a river basin or at coarse grid scales. While the water balance approach, which equates the actual evapotranspiration in a river basin to the difference between the precipitation in the basin and the river discharge is simple and entirely observation-based, it is challenged by lack of spatial coverage in river discharge observations, well-defined groundwater boundaries and spatial resolution (Soltani et al., 2021b). Stisen et al. (2021a) showed that applying gridded discharge data in the water balance approach at the European scale produced unrealistic actual evapotranspiration estimates that were not consistent with other methods. On the other hand, the same study showed strong agreement between actual evapotranspiration levels and spatial patterns across Europe derived by both the Budyko and satellite-based methods, suggesting a combination of these approaches as a robust method for Pan-European long-term average actual evapotranspiration mapping (Stisen et al., 2021a). In the current study, such a combination of Budyko and satellite-based estimates has been performed through a merging procedure described below. In addition, the Budyko based estimate has been rerun with updated climate input data consistent with the study.

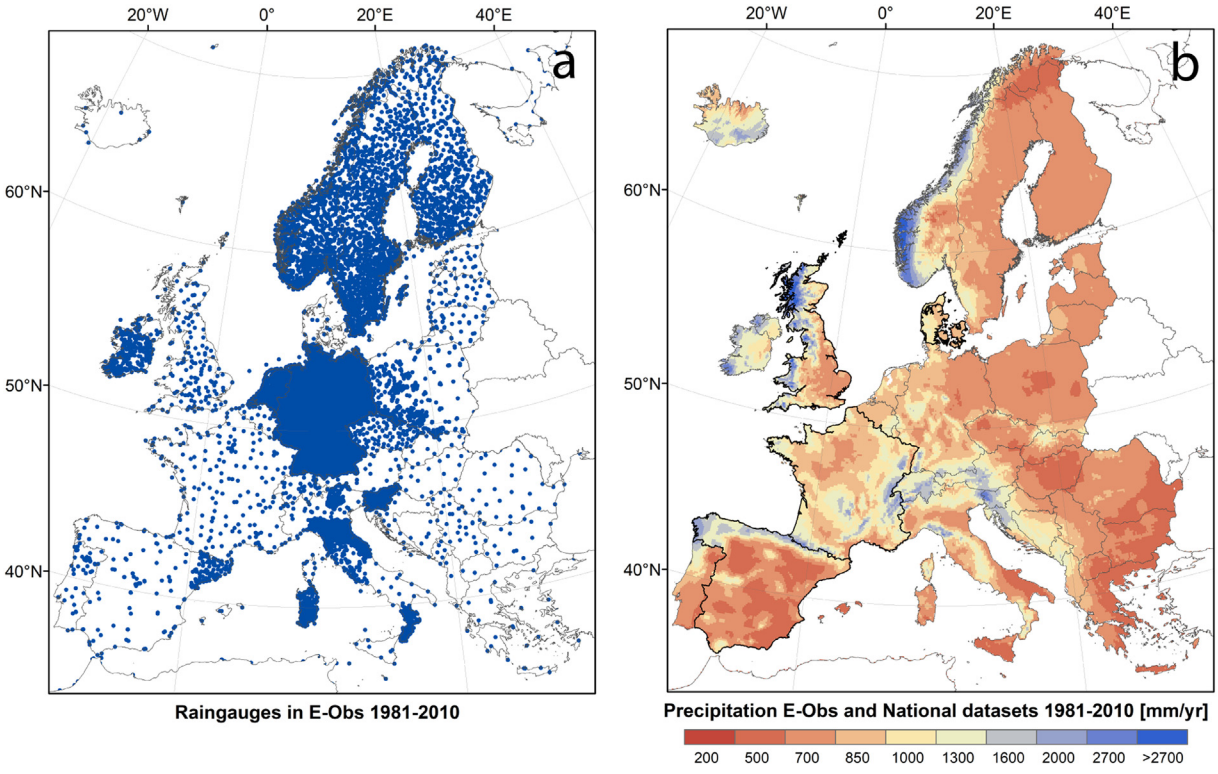


Fig. 1. Raingauges behind the *E*-Obs dataset 1981–2010 and E-Obs mean annual precipitation from E-Obs substituted with national gridded data for the British mainland, France, Spain and Denmark.

2.4.1. Budyko method

The Budyko approach (Budyko, 1974) describes how the ratio of long-term average precipitation and climatic water demand (potential evapotranspiration (E_{pot})) drives the partitioning of precipitation into actual evapotranspiration, and streamflow at catchment scale (Berghuijs et al., 2014), where actual evapotranspiration is limited by water in dry conditions, and by energy in wet conditions (Zhang et al., 2004). Although the framework was developed for river basin scales, it has also been applied to gridded data in numerous studies (e.g. Rouholahnejad Freund et al., 2020; Teuling et al., 2019). We use the following equation (Yang et al., 2008) to estimate long-term average actual evapotranspiration within the Budyko framework at a grid resolution of 0.1°.

$$E_{act,Budyko} = f(P, E_{pot}) = \frac{P}{\left(\left(\frac{P}{E_{pot}}\right)^{\omega} + 1\right)^{\frac{1}{\omega}}} \quad (2)$$

where the parameter ω [−] controls the partitioning of P into E_{act} and runoff, representing the integrated effects of catchment attributes such as vegetation, soil, climate and topography. (Donohue et al., 2007) highlighted the importance of vegetation cover, leading to the development of remote sensing-based algorithms to characterize ω, based on the normalised difference vegetation index (NDVI) (Li et al., 2013). (Li et al., 2013) suggested the following linear relation between ω and the normalised NDVI, which has been adopted in this study.

$$\omega = 2.36 \times NDVI_{norm} + 1.16 \quad (3)$$

Potential evapotranspiration data are freely accessible at https://wci.earth2observe.eu/ in 0.05° resolution based on different equations. For our implementation of the Budyko equation, for robustness, averages of the three products using the Penman–Monteith, Priestley–Taylor and Hargreaves approaches (Sperna Weiland et al., 2015) were calculated. For the NDVI, the MOD13A2 global 16-day values at a 1 km spatial resolution were used to produce an annual vegetation climatology for Europe. Potential evapotranspiration and normalised difference vegetation index data are processed for the period 2002–2014 and resampled to 0.1° resolution equivalent to the *E*-Obs dataset. For potential evapotranspiration, the period 1981–2010 is also processed in order to produce Budyko estimates for both periods 1981–2010 and 2002–2014.

2.4.2. Satellite-based methods

Several actual evapotranspiration products based on vegetation- driven algorithms are available globally due to their simplicity and timesteps of typically 8–16 days. On the contrary, land surface temperature-driven algorithms are more complicated due to higher sensitivity to cloud cover, impact of observation time and the resulting instantaneous nature of the estimates. Therefore, no global satellite based actual evapotranspiration estimates are available based on the thermal signal from satellites, even though the temperature of the surface is the most direct indication of evaporative cooling and energy partitioning. In a recent study, (Stisen et al., 2021a) developed two long-term averages of the actual evapotranspiration for Europe based on the two-source energy balance algorithm (Norman et al., 1995) and the Priestley-Taylor Jet Propulsion Laboratory algorithm ($PT\text{-}JPL_{thermal}$)(Moyano et al., 2018), and compared them to globally available vegetation driven remote sensing methods from MODIS16 (Mu et al., 2007) and PML_V2 (Zhang et al., 2019) plus the water balance and Budyko approaches. The conclusion of their evaluation was that while TSEB, $PT\text{-}JPL_{Thermal}$, MODIS16 and the Budyko approaches generated very similar spatial patterns across Europe, their implementation of the $PT\text{-}JPL_{Thermal}$ was biased. The PML_V2 and water balance approach resulted in distinctly different spatial patterns, that were not plausible or consistent with hydrological and climatological understanding.

2.4.3. Merged Budyko and satellite-based E_{act} estimate

Based on this knowledge, the TSEB, MODIS16 and Budyko are all considered appropriate methods for estimating long-term average actual evapotranspiration across Europe. The satellite-based methods produce estimates in a resolution of 1 km consistent with the resolution of the driving remotely sensed data, whereas the Budyko is developed for river basin scale and here applied at a 0.1°resolution (approx. 10 km). The upper limit of E_{act} in the Budyko approach is, in contrast to satellite based estimates, constrained by the precipitation data, which is advantageous for the water budget approach to groundwater recharge estimation, since actual evapotranspiration cannot exceed precipitation and result in a negative effective precipitation. Given the general similarity and advantages of the three methods, the current study deployed a simple data fusion approach in line with the ideas behind more advanced fusion methods (Shang et al., 2020). First the long-term average actual evapotranspiration from the two satellite-based methods TSEB and MODIS16 are averaged at 1 km resolution. In a second step, the coarse 0.1 deg. Budyko estimate is sharpened, by merging it with the satellite-based estimate. This is done very simply by aggregating the satellite-based estimate to 0.1°, subtracting it from the Budyko 0.1° estimate, resampling the difference to 1 km and adding the satellite-based estimate at 1 km resolution.

$$E_{act,Budyko/Satellite} = \left(\left(E_{act,Budyko} - E_{act,Satellite}\right)_{0.1deg}\right)_{1km} + \left(E_{act,Satellite}\right)_{1km} \quad (4)$$

Other, much more sophisticated, approaches have been developed for improved resolution of actual evapotranspiration estimates, by sharpening the input data to remote sensing-based actual evapotranspiration algorithms individually in a daily timestep based on high resolution satellite data (Guzinski and Nieto, 2019). However, the approach suggested here is applied to one long-term average actual evapotranspiration estimate and will ensure that the overall actual evapotranspiration magnitude at a 0.1°resolution from the Budyko is maintained, while exploiting the spatial detail of the 1 km resolution satellite-based estimates.

The merged Budyko/Satellite actual evapotranspiration estimate refers to long-term average for the period 2002–2014, for which the satellite data are available. Although these patterns are not assumed to vary considerably over time for long-term averages, an adjustment to the preferred reference period 1981–2010 is made. This adjustment is based on the ratio between $E_{act,Budyko1981–2010}/E_{act,Budyko2002–2014}$ which is applied as a gridded scaling factor to the $E_{act,Budyko/Satellite}$. The ratio remains close to one for all grids indicating no significant shifts in actual evapotranspiration levels.

For details on the MODIS16 product and the calculation of the thermal based TSEB dataset, the reader is referred to (Stisen et al., 2021a), and in addition all data are freely available at the GEUS data repository (Stisen et al., 2021b).

Effective precipitation (P_{Eff}), defined as the difference between long-term average P and AET, is calculated at a 1 km resolution, where the *E*-Obs long-term average P dataset is grid-refined to the scale of the AET. Prior to the calculation of P_{Eff}, the $E_{act,Budyko/Satellite}$ data are adjusted to the period 1981–2010 by the ratio of $E_{act,Budyko1981–2010}/E_{act,Budyko2002–2014}$.

2.5. Groundwater recharge coefficient

No generally-applicable equations exist for independent quantification of the recharge coefficient, which typically depends on meteorological factors, such as the precipitation intensity, and land surface characteristics such as land use, topography, soil, geology, all of which can vary greatly spatially, especially at the continental scale. The current study uses a data-driven machine learning approach, where the gridded C_{GWR} estimates from the national pilots are used as training data to develop a Random Forest (RF) regressor based on a range of covariates to predict C_{GWR} at the 1 km scale across Europe.

2.5.1. Covariates

The covariates for determining the recharge coefficient have been selected based on expert knowledge with the condition that they have European coverage in 1 km resolution (or similar). Covariates have been

selected based on their perceived explanatory power regarding spatial variations in recharge coefficients, meaning that they contain possible information on the runoff or recharge generating processes on the surface or in the shallow soil or geology. The covariates can be grouped into four classes related to: topography, soil, hydrogeology, and climate. The effective precipitation has been selected as the only covariate in the climate class, since it combines precipitation and actual evapotranspiration. It can be argued that it also includes vegetation information through the satellite-based actual evapotranspiration estimates. Three topography related covariates were selected: elevation, slope and IDPR, the Network Development and Persistence Index (Mardhel et al., 2021) as described in Section 2.2.1. The IDPR data is provided by the French Geological Survey at 50 × 50 m resolution and aggregated to 1 km resolution. Latitude and longitude has not been included since it was not considered likely to contain information on recharge processes or be closely correlated to such processes as compared to elevation derived co-variates.

The soil related covariates are four 1 km resolution soil texture maps: sand, silt, and clay fractions and the bulk density data representing soil depths from 5 to 200 cm, aggregated from the SoilsGrid250m dataset (Batjes et al., 2020; Hengl et al., 2017). Three hydrogeological covariates have been derived from the International Hydrological Map of Europe (IHME1500) in 1:1,500,000 (BGR, 2021) available through the EGDI database (http://www.europe-geology.eu/). The IHME1500 dataset consists of a hydrogeological classification of aquifer productivity (aquifer type), lithology, and fractures. These vector-based data are projected to the 1 km resolution, and some classifications are simplified (six aquifer types and 10 lithology classes). Fig. 2 shows the covariates and the outlines of the national pilots used for training and validation.

2.5.2. Random forest model

Proposed by (Breiman, 2001), a Random Forest (RF) model builds an ensemble of weak decision tree models, where each decision tree recursively splits the training data into more homogenous groups. RF has emerged as one of the most prevalent modelling tools covering a wide range of geophysical and environmental contexts, where it has been found especially useful for tasks relating to spatial modelling. These include, among others, mapping of soil properties (Hengl et al., 2017), water quality indicators (Erickson et al., 2021) or groundwater depth (Koch et al., 2021). A trained RF model can be interpreted with the help of the concept of feature importance with attest an importance to each covariate of the model. In this study, we have applied the mean decrease impurity (MDI) method, which analyses the contribution of each covariate in a trained model by quantifying the gain in loss-function through splitting the data using information from a specific covariate. We applied the Scikit-learn Python package (Pedregosa et al., 2011) to conduct the RF modelling for this study.

The training data based on the national gridded datasets constitutes a total of 1,560,000 datapoints, which is split into training and validation datasets. It should be noted that these points are not observations, but national scale model results. Prior to training, categorical variables need to be converted to numerical variables, which was done using the concept of one-hot encoding. Here, each category is assigned an individual binary feature. Applying this encoding to the six aquifer types and 10 lithology classes of the IHME1500 dataset, a total of 44 features were obtained. We expect a high degree of redundancy among the selected features, since many of them are correlated with each other, for example the soil properties for six depth layers. We wished to select only a subset of features to minimise resources needed for training and applying the RF model. For this, a hierarchical dendrogram based on the correlation matrix of the 44 features was processed. Features are clustered based on their cross-correlation, and a threshold was selected to group features, resulting in 23 features. Only one feature per cluster was used for the subsequent modelling and, for example, all six layers of sand content were found redundant and thus only a single layer was used for the modelling. The same was the case for clay and silt content, however, bulk density was split into a top (two layers) and bottom (four layers) feature. Further, in the IHME1500 dataset, two cases of identical aquifer type and lithology were identified.

3. Results

3.1. Actual evapotranspiration

The Pan-European actual evapotranspiration map derived from the averaged TSEB and MODIS16 estimates for the period 2002–2014 is displayed in Fig. 3a. It shows a pattern of annual actual evapotranspiration ranging from around 200 mm/year to 900 mm/year. The lowest values are found in the far north of Scandinavia, where the actual evapotranspiration is low due to low incoming radiation and temperatures and in central Iberia, where low precipitation limits the actual evapotranspiration. Highest actual evapotranspiration values are found along the mountain ranges on the Atlantic coast of Iberia and in the central parts of France, Slovenia, Croatia and central Italy, where the combination of high precipitation and available energy is present. Considering West-East transects, there is a trend from higher actual evapotranspiration values in the West (Ireland, the British Isles, and the Atlantic coast) to lower values in Eastern Europe (except the Carpathians). This tendency largely follows precipitation gradients.

The corresponding Budyko actual evapotranspiration estimate (Fig. 3b) is very similar in both magnitude and spatial pattern, as described in (Stisen et al., 2021a). As a consequence of the general similarity between the independent satellite-based and Budyko estimates, the merged actual evapotranspiration map (Fig. 3c) also displays the same patterns. From Fig. 3, it is can be seen that the overall actual evapotranspiration levels of the merged estimate follow the coarse Budyko estimate, while the fine-scale variability is inherited from the satellite-based estimate at 1 km resolution.

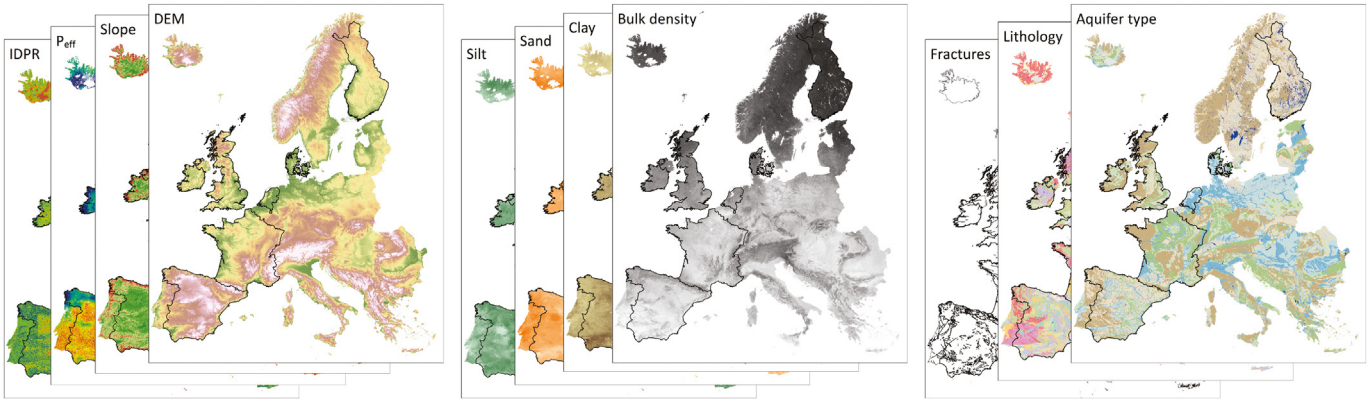


Fig. 2. Illustration of the range of covariates used for the machine learning predictions of recharge coefficients. The seven countries providing training data on recharge coefficients are outlined with black boundaries.

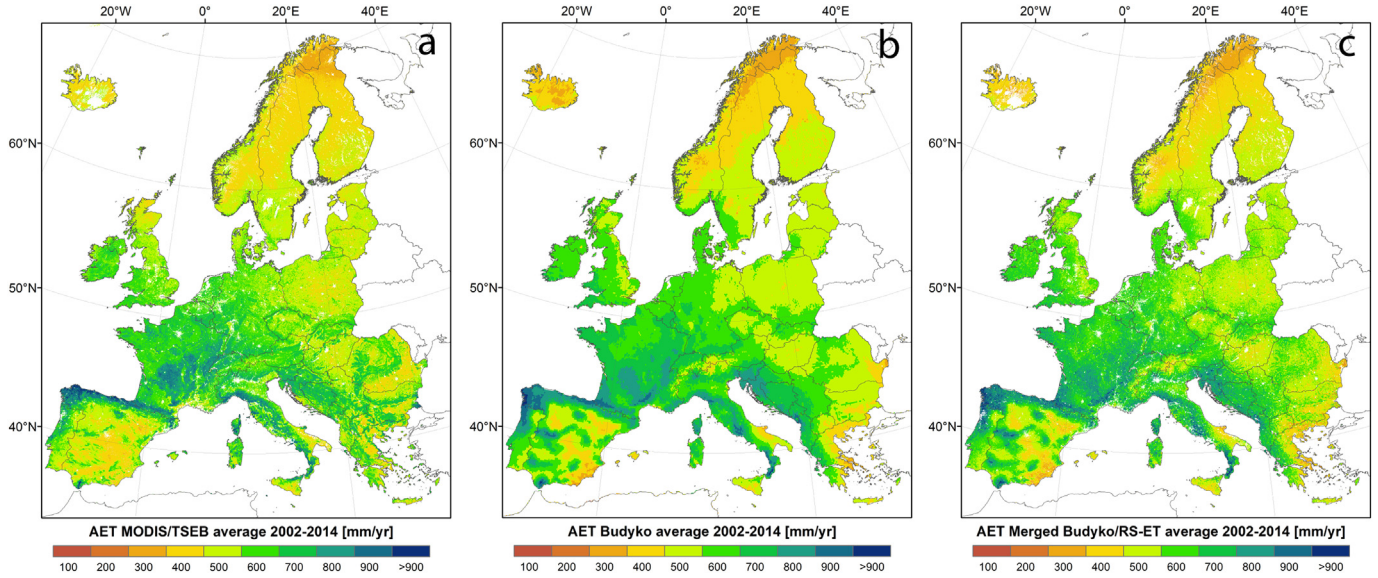


Fig. 3. Long-term average annual actual evapotranspiration (E_{act}) from a) satellite remote sensing, b) Budyko estimate and c) Merged remote sensing and Budyko estimate. White no-data areas in a) and c) represents grids classified as urban or open water.

3.2. Effective precipitation

The P_{Eff} map at 1 km resolution is compared to the corresponding P_{Eff} datasets from the seven national pilots (1,560,000 grid points). National pilots are also available at 1 km resolution, except for Spain (10 km). The evaluation of the Pan-European effective precipitation, P_{Eff}, estimate for 1981–2010 against national pilots reveals a consistent underestimation (Fig. 4a.) This underestimation is seen across all national pilots and suggests a systematic difference between the approaches (see supplementary material). Because of the systematic bias, a correction factor is applied to the entire Pan-European P_{Eff} dataset, corresponding to $1/0.79 = 1.26$ (Fig. 4a). The resulting P_{Eff} map for 1981–2010 (Fig. 4b) displays similar spatial patterns as the actual evapotranspiration map, however it is more strongly influenced by the high precipitation rates along the north Atlantic Coast and in the Alps.

3.3. Groundwater recharge coefficient

Groundwater recharge coefficients, C_{GWR}, from the seven national pilots were available either as explicit estimates at a national scale, or as the ratio of the potential groundwater recharge $R_{GW,Pot}$ and P_{Eff} from the national model. The national gridded C_{GWR} data are mapped in Fig. 5a, which illustrates the wide range in estimated values across Europe, from high coefficients close to 1 in most of Spain, low values in most of Finland and distinct patterns related to geology and topography in the UK and France. Ireland has areas of very low effective recharge coefficient values, in part

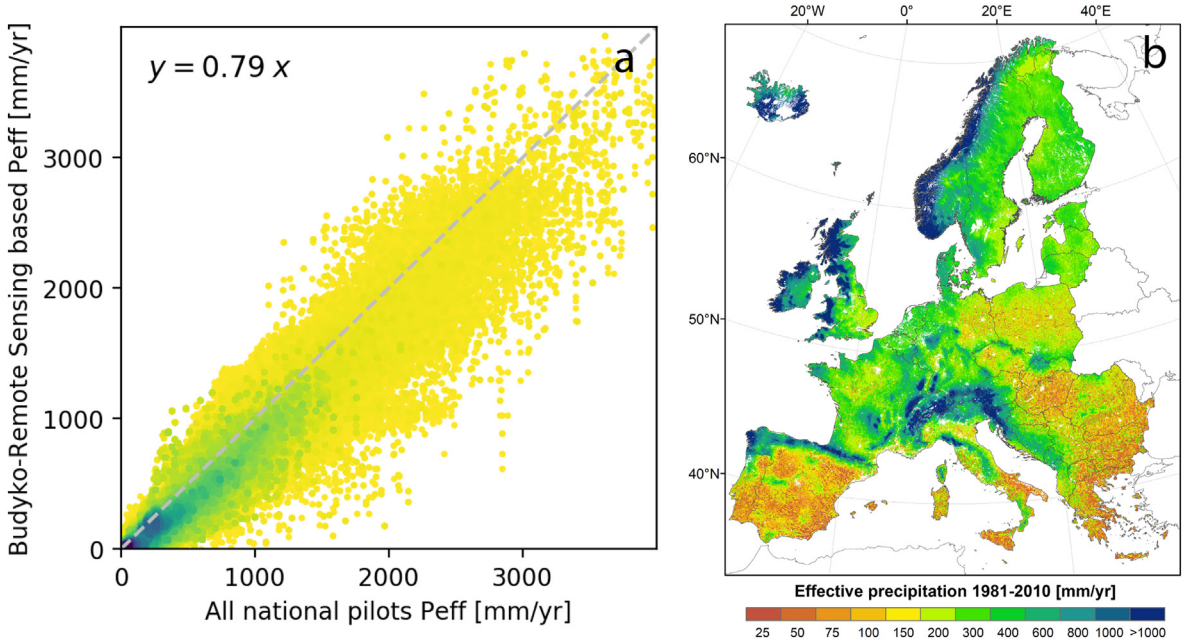


Fig. 4. a) Comparison of the effective precipitation from the gridded national pilots and from the initial pan-European estimate leading to the correction factor of 1.26 (=1/0.79), where the colours indicate the number of data points per pixel with blue colours for a higher density of the total of 1,560,000 data points, the grey dashed line is the 1:1 line, while the fit to data is given by the equation. b) Long-term average annual effective precipitation (1981–2010) after application of correction factor.

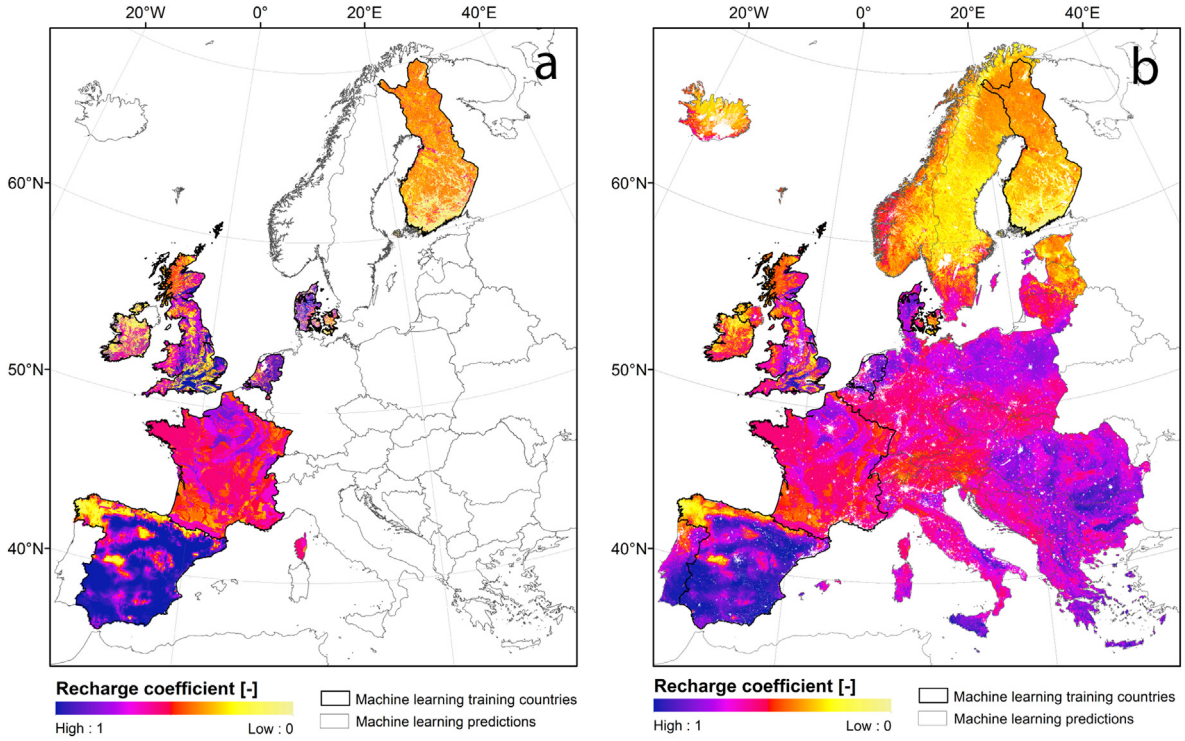


Fig. 5. Groundwater recharge coefficients. a) for seven national pilots, used for machine learning training and b) machine learning predicted pan-European groundwater recharge coefficients. (*Contains British Geological Survey material © UKRI 2022. All rights reserved)*

due to a threshold approach where aquifers have low storage and transmissivity, and therefore low recharge acceptance capacity. The Netherlands and Denmark, which are dominated by sedimentary deposits, tend to have relatively high recharge coefficients. The national pilots cover the entire range in possible C_{GWR} [0–1] and cover a large spread in aridity, soil types, geology and topography. This is an important feature since it strengthens the robustness of the algorithm when training the Random Forest regressor to provide predictions outside the national pilots.

The Random Forest regressor was trained using the set of covariates listed in Section 2.5. Initially, four different sampling fractions of the training dataset were applied. In principle all data could be utilised for training, but this will likely result in a direct replication of the national models. Instead, we aimed to use as little training data as possible in order to estimate a seamless and consistent C_{GWR} at pan-European scale, utilising the knowledge from the national models in an optimal way. The four sample sizes were 10, 30, 50 and 90%, with 10% and 90% regarded mostly as benchmarks for model performance. The validation results (based on data not included in the training) are given in Table 2. The overall performance when evaluated within the national pilots is around RMSE of 0.13 and the sampling sizes above 10% have limited impact on the performance statistics. A sampling size of 30% was selected, since it leaves a large validation dataset, and visual inspection of predictions showed that some artifacts of the 10 km grid size in training data over Spain created artifacts in the 1 km prediction when trained with large sample sizes above 30%. Further, 30% of the training data still reached a very comparable performance with respect to the 90% benchmark, where the national models are expected to be replicated. On the other hand, sampling 30% of the training data yielded a clearly improved performance with respect to the 10% baseline, which seems to be lacking relevant information on the C_{GWR} variability.

The Random Forest prediction of C_{GWR} at the Pan-European scale is shown in Fig. 5b. As indicated by the validation performance in Table 2, the prediction can reproduce the gridded national pilot estimates to a high degree. For the regions outside the national pilots, no similar evaluation data are available. Visual inspection shows C_{GWR} values in Central and Eastern Europe tend to be in the medium to high range, while Norway and Sweden look similar to Finland.

Table 2
Random Forest machine learning validation performance with varying training sample sizes.

Performance	Sampling size of training data			
	10%	30%	50%	90%
R^2	0.74	0.76	0.78	0.79
MAE	0.096	0.088	0.085	0.080
RMSE	0.14	0.13	0.13	0.12

3.3.1. Feature importance

The feature importance analysis allows insights into the trained Random Forest model. The most important covariate for the prediction of groundwater recharge coefficients is the effective precipitation. A clearly visible link is, for example, low effective precipitation grids that are connected to high recharge coefficients across the Iberian Peninsula and Eastern Europe. Bulk density is also identified as an important covariate in the model. Low recharge coefficients are generally co-located with low bulk density values found in Scandinavia, northern Spain and parts of the UK. The DEM is the third most important covariate where, for example, high elevation is often linked to low recharge coefficients. The aquifer types, lithologies and fractures from the IHME1500 dataset are found to be the least important covariates (Fig. 6).

3.4. Potential groundwater recharge

The potential groundwater recharge ($R_{GW,pot}$) is calculated as long-term average for the period 1981–2010 by multiplying the P_{Eff} (Fig. 3b) and the C_{GWR} maps (Fig. 4b). The resulting map (Fig. 7a) shows an overall pattern similar to the P_{Eff} map, where the most notable differences are the relatively lower recharge values in Scandinavia due to the low C_{GWR} values. The $R_{GW,pot}$ values plotted against the national pilots for all 1,560,000 data points in Fig. 7b show a generally good agreement. The mean absolute difference (MAD) is 70.5 mm/y, while the MAD in percent of the mean recharge from national pilots (MAD%) is 35%. Fig. 7b reveals two marked anomalies

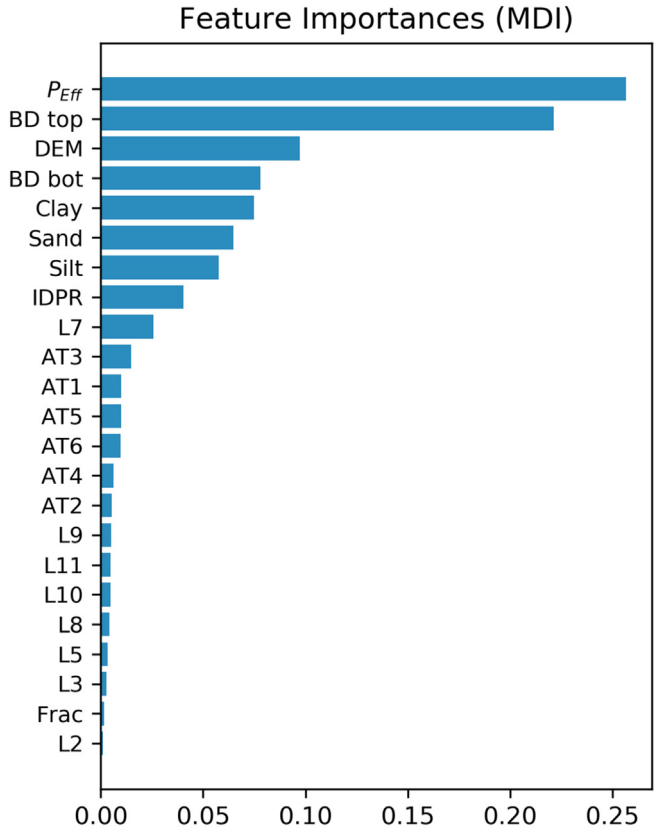


Fig. 6. Feature importance of covariates used for predicting groundwater recharge coefficients. AT stands for aquifer types and L for lithologies from the IHME1500 dataset. BD refers to bulk density.

at national pilot $R_{GW,pot}$ values around 100 and 200 mm/y. These are caused by a cap on the groundwater recharge amounts to certain aquifer categories in the Irish national pilot dataset, which are not consistent with the Pan-European potential groundwater recharge approach. Fig. 7c shows the equivalent scatterplot, but excluding the Irish domain, which removes the anomaly and reduces the mean absolute difference to 64.6 mm/y, with a MAD% of 31% and an R of 0.88. The Pan-European and national pilot $R_{GW,pot}$ estimates are generally well-correlated, also in the range from 0 to 500 mm/y with the most data points.

4. Discussion

Validation of the Pan-European effective precipitation, an intermediate step in the estimation of groundwater recharge, shows a systematic overestimation compared to similar estimates for the seven national pilots (see also supplementary material). The bias is not dominated by differences in precipitation input, since it also occurs for the national pilots, where precipitation data are shared between the Pan-European maps and the national pilots. Instead, the validation indicates that the actual evapotranspiration from the combination of Budyko and satellite data ($E_{act,Budyko/Satellite}$) is higher than the E_{act} estimates used for calculating the effective precipitation for the national pilots. The main difference between the methods is that $E_{act,Budyko/Satellite}$ is solely based on long-term average data, while the national pilots utilise daily to monthly data that are subsequently aggregated to long-term average estimates. This could introduce a bias, since the $E_{act,Budyko/Satellite}$ estimate does not account explicitly for effects of seasonality. Such seasonality effects are expected to be small in climates where the actual evapotranspiration is always limited by either water (P) or energy (E_{Pot}), i.e. when P and E_{Pot} are seasonally out of phase (Donohue et al., 2007). They could, however, be important when conditions change between being water-limited and energy-limited. Seasonality could potentially be built into the ω parameter of the Budyko framework, as suggested by Ning et al. (2020).

The systematic bias of the effective precipitation in the current study could also be attributed to other effects such as non-linear behaviour during intense rainfall or extreme drought. Because the cause of the bias was unclear, a simple bias correction factor was applied across all of Europe.

The presented approach to large-scale high-resolution mapping of long-term average groundwater recharge utilises a set of national scale gridded estimates developed by the respective national geological surveys. The approach builds on the assumption that these national estimates, based on regional knowledge and best available data, are superior to existing global estimates. Thus, the challenges associated with upscaling point estimates can be circumvented by learning from national mapping experiences covering a range of climatologically and geologically different regions.

One challenge associated with the proposed approach is the multitude of methods used in the various national calculations, which will impact not only the recharge estimates themselves but also what they represent. Most of the national recharge estimates represent potential groundwater recharge, which focus on the shallow hydrology without explicitly considering the storage capacity and properties of the underlying aquifer. However, some estimates like the Spanish model, are empirically adjusted to give estimates of actual groundwater recharge to the deeper aquifers. The Irish estimate includes a cap on recharge coefficients over poorly productive bedrock aquifers, which are limited to a specified upper limit on recharge rates. Recharge estimates based on integrated groundwater surface water models (Netherlands and Denmark) have accuracy due to the calibration against groundwater level fluctuations and surface water flows implicitly containing information on recharge rates. French estimates of recharge coefficients rely on the baseflow component of streamflow that results from, but is not limited to, drainage of groundwater (Stoelzle et al., 2020). These differences would mainly impact the recharge coefficients that determine the partitioning of the effective precipitation between recharge and runoff. However, validation of the machine learning-based recharge coefficient estimates for the national pilots showed very good agreement, except for the Irish locations with low transmissivity and storage bedrock water table aquifers with limited recharge acceptance capacity. This local mismatch in Ireland originates from differences in definitions and upper limits to the recharge coefficient. Similarly, the recharge coefficient estimates in southern Finland shows some deviation with both over and underestimation compared to the national Finish estimate. Elsewhere, the Random Forest algorithm was able to predict local coefficients, even though the training data might reflect different recharge definitions. How the recharge estimation method influences predictions outside the national pilots remains unknown. Some patterns in the predicted recharge coefficient map (Fig. 5) show similarity to valuables such as elevation and bulk density that are not necessarily physically intuitive. However, such correlations might not be true dependencies but correlation to other controlling variables, e.g. elevation and bulk density could be related to soil depth and hard rock geology, or compensate for possible biases in climate or training data.

To evaluate the Pan-European potential groundwater recharge estimate against other estimates and outside the seven national pilots, groundwater recharge from the global hydrological model PCR-GLOBWB 2 in 5 arc minute resolution has been used (Sutanudjaja et al., 2018). In contrast to most global models, PCR-GLOBWB 2 includes a two-layer groundwater flow module in combination with a river routing scheme and land surface and water use modules. PCR-GLOBWB is a water resources model and has been evaluated extensively against global river runoff databases. The recharge from PCR-GLOBWB (Fig. 8) is considered a good benchmark for comparison to the Pan-European potential groundwater recharge estimate, with its complete coverage and similar resolution (approx. 5 km). There is good agreement on the overall spatial pattern across Europe between the Pan-European recharge map (Fig. 7a) and PCR-GLOBWB (Fig. 8). Most notable differences are on regional patterns such as in Italy, Portugal, the Carpathians, Sweden and Finland. Spatial pattern similarity, expressed through the bias insensitive spatial pattern metric SPAEF (Koch et al.,

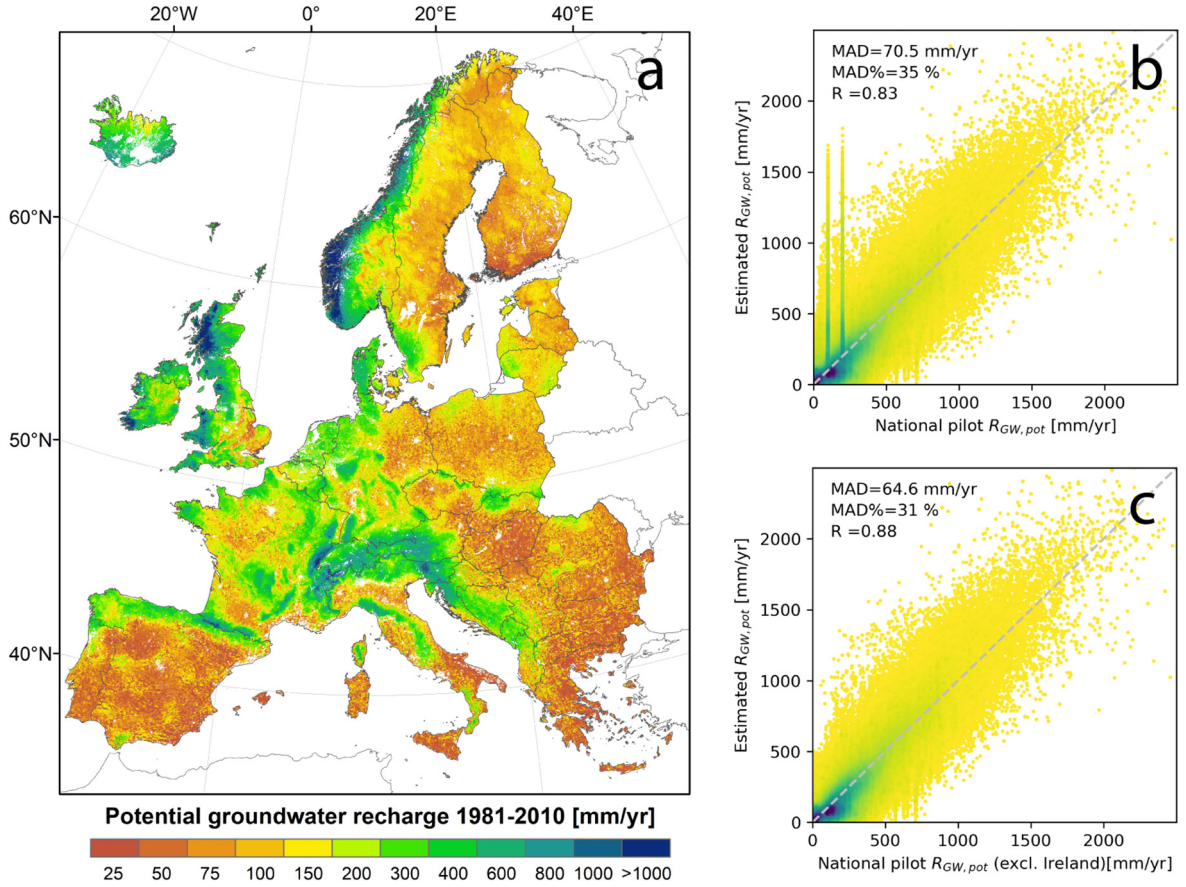


Fig. 7. a) Long-term average Pan-European potential groundwater recharge map (1981–2010). Right, comparison to gridded national pilots, b) including all national pilot data (with a mean absolute difference (MAD) of 70.5 mm/yr and c) the national pilots excluding Ireland with a mean absolute difference of 64.6 mm/yr. Blue colours (in b and c) indicate more data points per pixel.

2018) with an optimal value of 1, give a value of 0.62 for the European scale comparison of PCR-GLOBWB and the current study. An explanation of SPAEF and a table of its subcomponents is given the supplementary material (Table S1).

For Germany, a national average groundwater recharge map is available for the period 1961–1990 in 1 km resolution as produced by the German Geological Survey (BGR, 2019). The German groundwater recharge estimate by BGR is based on a multi-stage regression, where the baseflow index (BFI = baseflow / total runoff) was determined as a regression target depending on slope, water network density, soil cover, field capacity, depth of the groundwater below the surface and the share of direct runoff in total runoff (Neumann, 2005).

Although for a different reference period, the Pan-European potential groundwater recharge estimate for Germany (Fig. 9a) is very similar to the German national groundwater recharge map (Fig. 9b), with patterns of lower recharge rates in Eastern Germany, and higher in the southern alpine region and intermediate values in the North-West resulting in a SPAEF value of 0.72 (Table S1). In comparison, the PCR-GLOBWB recharge estimate for Germany (Fig. 9c) displays a similar overall pattern, but with more notable differences, with a zone of high recharge rates in Eastern Germany and generally higher values in the North-West resulting in a SPAEF of 0.32 compared to the BGR map (Fig. 9b) and 0.43 compared to the German subset of the pan-European map (Fig. 9a).

A comparison of the pan-European recharge map to point scale estimates has not been included in the current study, since it is part of parallel work within the same project. That work compared both point scale estimates derived across a range of methods and a comparison to the Pan-European recharge map. Their point scale estimates varied greatly depending on estimation method challenging a robust validation of the Pan-European map (Seidenfaden et al., 2022).

Another challenge for the applied methodology is the lack of training data from large regions in Europe, where no data are available for evaluation of the effective precipitation and the potential groundwater recharge. The lack of data representing specific regions is probably most challenging for the machine learning prediction of recharge coefficients, which in contrast to precipitation and evapotranspiration data sets, has no information in some regions. This could make the training data for the Random Forest algorithm less representative, although the recharge coefficients from the seven national pilots do cover the full range of possible values. Lack of representativeness is a well-known problem in machine learning methods and a subject of increased attention, giving rise to new methods assessing the likely areas of applicability and related uncertainties (Meyer and Pebesma, 2021). Such an analysis is however outside the scope of the current study.

The goal of the Pan-European groundwater recharge mapping has been to utilise and honour the national pilots, in a method that can provide harmonised high-resolution estimates across Europe and provide locally relevant estimates for areas where this information is not available. The advantage of this approach is that local knowledge and experience from the seven national pilots can be utilised both for selecting appropriate covariates and for guiding the machine learning. The geographical coverage and range of methods behind the seven pilots will strengthen the robustness of the estimate compared to selecting a single national approach and applying it to the Pan-European scale. In addition, the data generated provide gridded estimates of each of the components (E_{Act}, P_{Eff} and C_{GWR}) to calculate the groundwater recharge as separate datasets so users can substitute one component if better local or regions data are available.

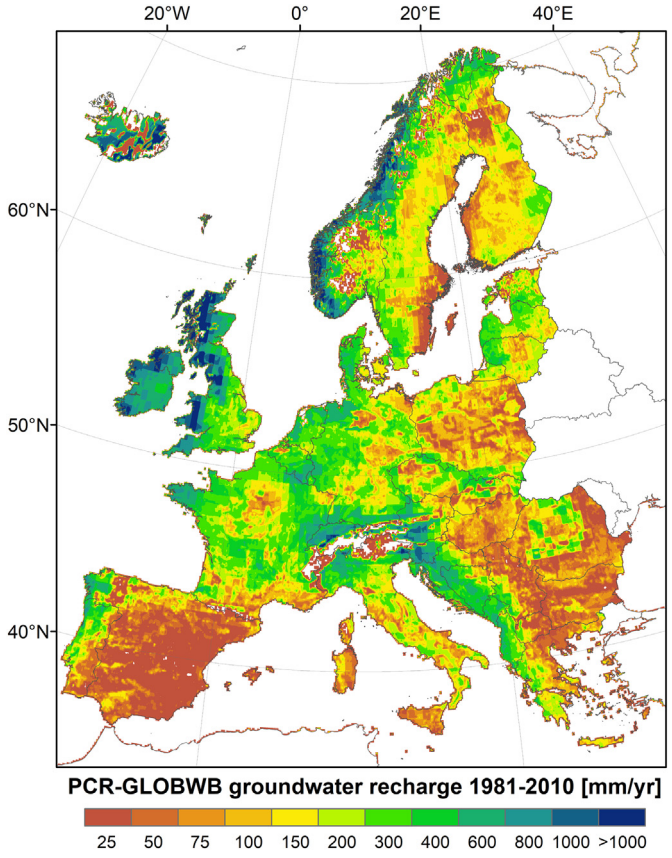


Fig. 8. PCR-GLOBWB (Utrecht University) groundwater recharge estimate for 1981–2010 for the Pan-European domain.

Like any large-scale dataset, the dataset should be used with caution at local scale, and it is encouraged to validate, especially the P_{Eff} values, to ensure that it is consistent with local information such as precipitation and river runoff at catchment scale. The dataset aims mainly at regional to continental analysis and as a robust groundwater recharge estimate for regional to national scale in cases where this data does not already exist.

The comparison to the PCR-GLOBWB 2 simulations provided a pan-European assessment of the $R_{GW,pot}$ results, which validates the general pattern. The difference between the PCR-GLOBWB 2 estimate and the national pilots is too large to also evaluate regional patterns. For this, an independent National estimate for Germany provides a more useful means. This comparison is very encouraging, since no German data were used for bias-correction or training and, as such, Germany could be representative of any other region outside the seven national pilots. It could be argued that Germany shares hydrogeological features with France, Netherlands and Denmark, which might improve the performance. In this regard, the Eastern European region is not as well represented from a hydrogeological perspective and the Pan-European recharge estimate will likely be associated with the highest uncertainty in this region, which is also poorly covered by rain gauges in the *E*-Obs dataset.

5. Conclusion

A Pan-European groundwater recharge map has been created using a variety of information sources. The map represents a long-term average recharge defined in the context of large-scale application and a soil column water budget framework. It is equal to the precipitation minus the actual evapotranspiration and minus the shallow runoff. The precipitation has been obtained from the E-Obs Pan-European dataset. The actual evapotranspiration has been determined using a mixture of satellite data and Budyko approach. The portion of the precipitation minus actual evapotranspiration that recharges the groundwater has been modeled using a recharge coefficient. This coefficient has been determined using a Random Forest regressor trained using data from seven national pilots and eleven Pan-European covariate maps. The general pattern of the resulting recharge map compares well with results from the PCR-GLOBWB 2 global hydrological model. The comparison with an independent German recharge map shows that the regional pattern in Germany is reproduced well. This provides confidence that the applied method has provided a harmonised high-resolution Pan-European recharge map.

CRediT authorship contribution statement

Grith Martinsen: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – review & editing. **Helene Bessiere:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Yvan Caballero:** Conceptualization, Data curation, Methodology,

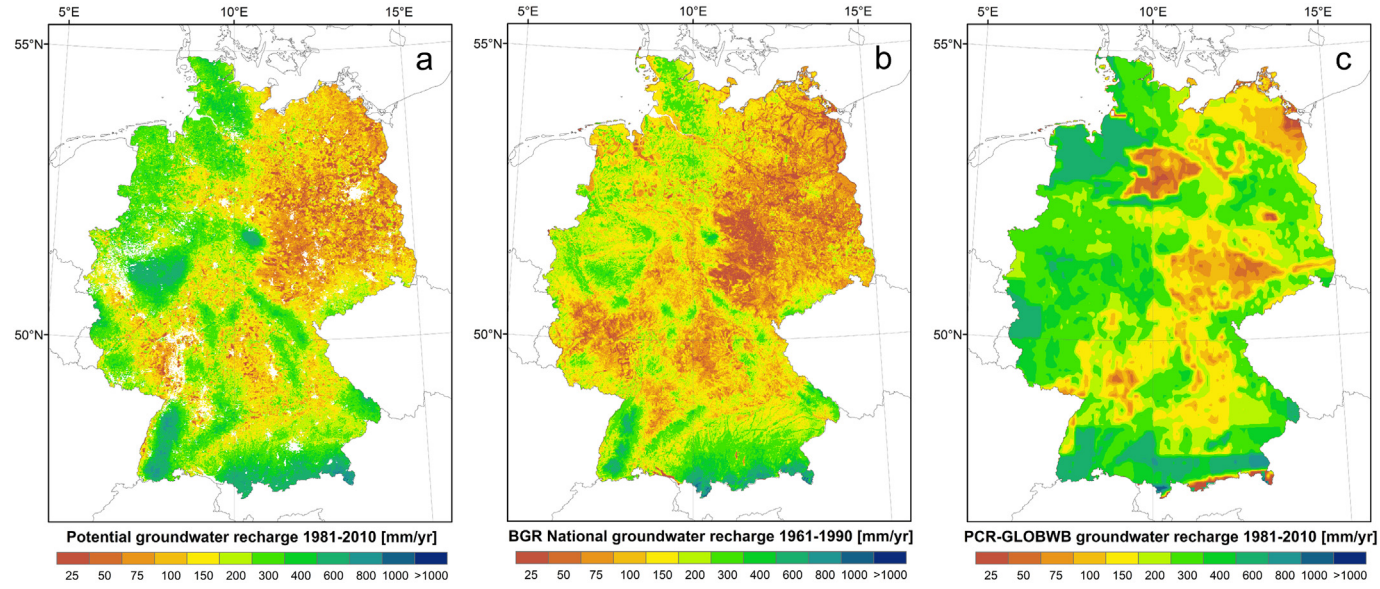


Fig. 9. a) Subset of Germany from the Pan-European potential groundwater recharge produced in this study (1981–2010), b) German National groundwater recharge map (BGR - 1961–1990) and c) subset of Germany from PCR-GLOBWB groundwater recharge estimate (Utrecht University, 1981–2010).

Writing – review & editing. **Julian Koch:** Formal analysis, Methodology, Visualization. **Antonio Juan Collados-Lara:** Data curation, Writing – review & editing. **Majdi Mansour:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Olli Sallasmaa:** Conceptualization, Data curation, Methodology, Writing – review & editing. **David Pulido Velázquez:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Natalya Hunter Williams:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Willem Jan Zaadnoordijk:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Simon Stisen:** Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research has been carried out within the TACTIC project, part of the GeoERA programme, which received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no 731166, the participating geological survey organisations and Innovation Fund Denmark under agreement no 8055-00073B. In addition, some French partners were also supported by the INDECIS project financed by the European ERA4CS Joint Call for Transnational Collaborative Research Projects. Mansour publishes with the permission of the Executive Director of the British Geological Survey (UKRI).

Timo Kroon of Deltares is thanked for providing the Dutch model results and Edwin Sutanudjaja and Niko Wanders from Utrecht University for providing Pan-European PCR-GLOBWB 2 groundwater recharge results. Sandra Lanini and Stéphanie Pinson from BRGM are acknowledged for their respective participation in the French case for potential groundwater recharge and European IDPR computation.

Access to the dataset will be provided through the European Geological Data Infrastructure (EGDI) via URL: http://www.europe-geology.eu/groundwater/.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.153464.

References

Alattar, M.H., Troy, T.J., Russo, T.A., Boyce, S.E., 2020. Modeling the surface water and groundwater budgets of the US using MODFLOW-OWHM. Adv. Water Resour. 143, 103682. https://doi.org/10.1016/j.advwatres.2020.103682.

Alcalá, F.J., Custodio, E., 2014. Spatial average aquifer recharge through atmospheric chloride mass balance and its uncertainty in continental Spain. Hydrol. Process. 28, 218–236. https://doi.org/10.1002/hyp.9556.

Batjes, N.H., Ribeiro, E., van Oostrum, A., 2020. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). Earth Syst. Sci. Data 12, 299–320. https://doi.org/10.5194/essd-12-299-2020.

Berghuijs, W.R., Woods, R.A., Hrachowitz, M., 2014. A precipitation shift from snow towards rain leads to a decrease in streamflow. Nat. Clim. Chang. 4, 583–586. https://doi.org/10.1038/nclimate2246.

BGR, 2019. Mean Annual Groundwater Recharge of Germany 1:1,000,000 (GWN1000) (WMS) [WWW Document]. https://gdk.gdi-de.org/geonetwork/srv/api/records/40E14FF1-99D4-43DA-AF7B-C039F0463BF8.

BGR, 2021. IHME1500 [WWW Document]. https://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/ihme1500_projektbeschr_en.html.

BGS, 2021. BGS maps portal — maps and sections 1832 to 2014 [WWW Document]. https://webapps.bgs.ac.uk/data/maps/?_ga=2.199834222.1159015494.1631094168-1723672873.1631094168.

Boorman, D.B., Hollist, J.M., Lilly, A., 1995. Hydrology of Soil Types: A Hydrologically-based Classification of the Soils of the United Kingdom. Institute of Hydrology Report No. 126.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Brugeron, A., Paroissien, J.B., Tillier, L., 2018. Référentiel hydrogéologique BDLISA version 2: Principes de construction et évolutions. Rapport Final BRGM/RP-67489-FR.

Budyko, M.I., 1974. Climate and Life. International Geophysics Series. 18. Academic, New York, p. 508.

Caballero, Y., Lanini, S., Zerouali, L., Bailly-Comte, V., 2016. Caractérisation de la recharge des aquifères et évolution future en contexte de changement climatique. Application au bassin Rhône Méditerranée Corse. Rapport final. BRGM/RP-65807-FR. 188 p. http://ficheinfoterre.brgm.fr/document/RP-65807-FR.

Caballero, Y., Lanini, S., Lechevalier, J., Maréchal, J.-C., Le Cointe, P., Pinson, S., Arnaud, L., Ferront, L., 2021. Recharge des aquifères et évolution future en contexte de changement climatique - Application au bassin Rhône - Méditerranée-Corse - Phase 2. Rapport final. BRGM/RP-69217-FR. 434 p., 10 ann. http://ficheinfoterre.brgm.fr/document/RP-69217-FR.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M., Jones, P.D., 2018. An ensemble version of the E-OBS temperature and precipitation data sets. J. Geophys. Res. Atmos. 123, 9391–9409. https://doi.org/10.1029/2017JD028200.

Crosbie, R.S., Binning, P., Kalma, J.D., 2005. A time series approach to inferring groundwater recharge using the water table fluctuation method. Water Resour. Res. 41. https://doi.org/10.1029/2004WR003077.

Cuthbert, M.O., Taylor, R.G., Favreau, G., Todd, M.C., Shamsudduha, M., Villholth, K.G., MacDonald, A.M., Scanlon, B.R., Kotchoni, D.O.V., Vouillamoz, J.M., Lawson, F.M.A., Adjomayi, P.A., Kashaigili, J., Seddon, D., Sorensen, J.P.R., Ebrahim, G.Y., Owor, M., Nyenje, P.M., Nazoumou, Y., Goni, I., Ousmane, B.I., Sibanda, T., Ascott, M.J., Macdonald, D.M.J., Agyekum, W., Koussoubé, Y., Wanke, H., Kim, H., Wada, Y., Lo, M.H., Oki, T., Kukuric, N., 2019. Observed controls on resilience of groundwater to climate variability in sub-Saharan Africa. Nature 572, 230–234. https://doi.org/10.1038/s41586-019-1441-7.

Dadson, S., Blyth, E., Clark, D., Hughes, A., Hannaford, J., Lawrence, B., Polcher, J., Reynard, N., 2019. Hydro-JULES: next generation land-surface and hydrological predictions. EGU General Assembly 2019. 21 EGU2019-14578.

De Lange, W.J., Prinsen, G.F., Hoogewoud, J.C., Veldhuizen, A.A., Verkaik, J., Oude Essink, G.H.P., van Walsum, P.E.V., Delsman, J.R., Hunink, J.C., Massop, H.T.L., Kroon, T., 2014. An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: the Netherlands hydrological instrument. Environ. Model Softw. 59, 98–108. https://doi.org/10.1016/j.envsoft.2014.05.009.

Dingman, S.L., 1994. Physical Hydrology. Prentice Hall.

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

DoneSol INRA, 2014. Dictionnaire de données, INRA, Version du 1er mai 2014 [WWW Document]. https://dw3.gissol.fr/login.

Donohue, R.J., Roderick, M.L., McVicar, T.R., 2007. On the importance of including vegetation dynamics in Budyko's hydrological model. Hydrol. Earth Syst. Sci. 11, 983–995. https://doi.org/10.5194/hess-11-983-2007.

Edijatno, Michel, C., 1989. Un modèle pluie-débit à trois paramètres (A three-parameter daily rainfall-runoff model, in French). La Houille Blanche 2, 113–121.

Erickson, M.L., Elliott, S.M., Brown, C.J., Stackelberg, P.E., Ransom, K.M., Reddy, J.E., 2021. Machine learning predicted redox conditions in the glacial aquifer system, northern continental United States. Water Resour. Res. 57. https://doi.org/10.1029/2020WR028207.

European Environmental Agency, 2018. European waters assessment of status and pressures 2018. Parents and Children Communicating with Society: Managing Relationships Outside of Home. EEA Report No 7/2018 https://doi.org/10.2800/303664.

Franssen, H.J.H., Stöckli, R., Lehner, I., Rotenberg, E., Seneviratne, S.I., 2010. Energy balance closure of eddy-covariance data: a multisite analysis for European FLUXNET stations. Agric. For. Meteorol. 150, 1553–1567. https://doi.org/10.1016/j.agrformet.2010.08.005.

Graham, D.N., Butts, M.B., Frevert, D.K., 2005. Flexible, integrated watershed modelling with MIKE SHE. In: Singh, V.P. (Ed.), Watershed Models. Taylor and Francis, Boca Raton, Fla., pp. 245–272.

Griffiths, J., Young, A.R., Keller, V., 2006. Model Scheme for Representing Rainfall Interception and Soil Moisture. Environment Agency R&D Project W6-101 Continuous Estimation of River Flows (CERF).

Gustard, A., Bullock, A., Dixon, J.M., 1992. Low Flow Estimation in the United Kingdom. Institute of Hydrology Report No. 108.

Guzinski, R., Nieto, H., 2019. Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations. Remote Sens. Environ. 221, 157–172. https://doi.org/10.1016/j.rse.2018.11.019.

Healy, R.W., Scanlon, B.R., 2010. Estimating Groundwater Recharge. Cambridge University Press, Cambridge https://doi.org/10.1017/CBO9780511780745.

Heinen, M., Bakker, G., Wösten, J.H.M., 2020. Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks: update 2018 (Waterretention and Permeability Characteristics of Soils in the Netherlands, in Dutch).

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. PLoS ONE 12, e0169748. https://doi.org/10.1371/journal.pone.0169748.

Herrera, S., Fernández, J., Gutiérrez, J.M., 2016. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. Int. J. Climatol. 36, 900–908. https://doi.org/10.1002/joc.4391.

Hough, M.N., Jones, R.J.A., 1997. The United Kingdom meteorological office rainfall and evaporation calculation system: MORECS version 2.0-an overview. Hydrol. Earth Syst. Sci. 1, 227–239. https://doi.org/10.5194/hess-1-227-1997.

Hunter Williams, N., Misstear, B., Daly, D., Johnston, P., Lee, M., Cooney, P., Hickey, C., 2011. A national groundwater recharge map for Ireland. Proceedings National Hydrology Conference. Irish National Committees for the IHP and ICID, Athlone, Ireland.

Hunter Williams, N.H., Misstear, B.D.R., Daly, D., Lee, M., 2013. Development of a national groundwater recharge map for the republic of Ireland. Q. J. Eng. Geol. Hydrogeol. 46, 493–506. https://doi.org/10.1144/qjegh2012-016.

Hunter Williams, N.H., Carey, S., Werner, C., Nolan, P., 2021. Updated National Groundwater Recharge Map. Irish Groundwater Newsletter, pp. 31–34.

IWGGW, 2005. Guidance on the Assessment of the Impacts of Groundwater Abstractions. Irish Working Group on Groundwater, Guidance Document, GW5.

Jie, Z., van Heyden, J., Bendel, D., Barthel, R., 2011. Combination of soil-water balance models and water-table fluctuation methods for evaluation and improvement of groundwater recharge calculations. Hydrogeol. J. 19, 1487–1502. https://doi.org/10.1007/s10040-011-0772-8.

Jing, M., Heße, F., Kumar, R., Wang, W., Fischer, T., Walther, M., Zink, M., Zech, A., Samaniego, L., Kolditz, O., Attinger, S., 2018. Improved regional-scale groundwater representation by the coupling of the mesoscale hydrologic model (mHM v5.7) to the groundwater model OpenGeoSys (OGS). Geosci. Model Dev. 11, 1989–2007. https://doi.org/10.5194/gmd-11-1989-2018.

Kalma, J.D., McVicar, T.R., McCabe, M.F., 2008. Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data. Surv. Geophys. 29, 421–469.

Koch, J., Demirel, M.C., Stisen, S., Cüneyd Demirel, M., Stisen, S., 2018. The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models. Geosci. Model Dev. 11, 1873–1886. https://doi.org/10.5194/gmd-11-1873-2018.

Koch, J., Gotfredsen, J., Schneider, R., Troldborg, L., Stisen, S., Henriksen, H.J., 2021. High resolution water table modelling of the shallow groundwater using a knowledge-guided gradient boosting decision tree model. Front. Water 3, 701726.

Le Cointe, P., Arnaud, L., Béranger, S., Caballero, Y., Lanini, S., Bertin, C., Pinson, S., Thinon-Larminach, M., Tilloloy, F., 2019. Réponse des Eaux souterraines au CHangement climatique dans le bassin AdouR-GaronnE (RECHARGE). /RP-67149-FR. 155 p., 7 ann. dont 6 pl. HT. http://ficheinfoterre.brgm.fr/document/RP-67149-FR.

Le Moigne, P., Besson, F., Martin, E., Boé, J., Boone, A., Decharme, B., Etchevers, P., Faroux, S., Habets, F., Lafaysse, M., Leroux, D., Rousset-Regimbeau, F., 2020. The latest improvements with SURFEX v8.0 of the Safran–Isba–Modcou hydrometeorological model for France. Geosci. Model Dev. 13, 3925–3946. https://doi.org/10.5194/gmd-13-3925-2020.

Li, B., Rodell, M., Peters-Lidard, C., Erlingis, J., Kumar, S., Mocko, D., 2021. Groundwater recharge estimated by land surface models: an evaluation in the conterminous United States. J. Hydrometeorol. 22, 499–522. https://doi.org/10.1175/JHM-D-20-0130.1.

Li, D., Pan, M., Cong, Z., Zhang, L., Wood, E., 2013. Vegetation control on water and energy balance within the budyko framework. Water Resour. Res. 49, 969–976. https://doi.org/10.1002/wrcr.20107.

MacDonald, A.M., Lark, R.M., Taylor, R.G., Abiye, T., Fallas, H.C., Favreau, G., Goni, I.B., Kebede, S., Scanlon, B., Sorensen, J.P.R., Tijani, M., Upton, K.A., West, C., 2021. Mapping groundwater recharge in Africa from ground observations and implications for water security. Environ. Res. Lett. 16. https://doi.org/10.1088/1748-9326/abd661.

Mansour, M.M., Wang, L., Whiteman, M., Hughes, A.G., 2018. Estimation of spatially distributed groundwater potential recharge for the United Kingdom. Q. J. Eng. Geol. Hydrogeol. 51, 247–263. https://doi.org/10.1144/qjegh2017-051.

Mardhel, V., Pinson, S., Allier, D., 2021. Description of an indirect method (IDPR) to determine spatial distribution of infiltration and runoff and its hydrogeological applications to the French territory. J. Hydrol. 592. https://doi.org/10.1016/j.jhydrol.2020.125609.

Maxwell, R.M., Condon, L.E., Kollet, S.J., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geosci. Model Dev. 8, 923–937. https://doi.org/10.5194/gmd-8-923-2015.

Meyer, H., Pebesma, E., 2021. Predicting into unknown space? Estimating the area of applicability of spatial prediction models. https://doi.org/10.1111/2041-210X.13650 2041–210X.13650.

Moeck, C., Grech-Cumbo, N., Podgorski, J., Bretzler, A., Gurdak, J.J., Berg, M., Schirmer, M., 2020. A global-scale dataset of direct natural groundwater recharge rates: a review of variables, processes and relationships. Sci. Total Environ. 717, 137042. https://doi.org/10.1016/j.scitotenv.2020.137042.

Mohan, C., Western, A.W., Wei, Y., Saft, M., 2018. Predicting groundwater recharge for varying land cover and climate conditions – a global meta-study. Hydrol. Earth Syst. Sci. 22, 2689–2703. https://doi.org/10.5194/hess-22-2689-2018.

Moyano, M.C., Garcia, M., Palacios-Orueta, A., Tornos, L., Fisher, J.B., Fernández, N., Recuero, L., Juana, L., 2018. Vegetation water use based on a thermal and optical remote sensing model in the Mediterranean region of Doñana. Remote Sens. 10. https://doi.org/10.3390/rs10071105.

Mu, Q., Heinsch, F.A., Zhao, M., Running, S.W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens. Environ. 111, 519–536. https://doi.org/10.1016/j.rse.2007.04.015.

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T.A., Popat, E., Portmann, F.T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C.-E., Trautmann, T., Döll, P., 2021. The global water resources and use model WaterGAP v2.2d: model description and evaluation. Geosci. Model Dev. 14, 1037–1079. https://doi.org/10.5194/gmd-14-1037-2021.

NERC, 2000. Natural Environment Research Council (NERC) Countryside Survey 2000 Module 7. Land Cover Map 2000 Final Report.

NERC, 2003. Natural Environment Research Council (NERC), Hydrological Data United Kingdom, Hydrometric Register and Statistics 1996–2000.

Neumann, J., 2005. Area-differentiated groundwater recharge in Germany - development and application of the macroscale method HAD-GWNeu. Martin Luther University Halle-Wittenberg, Halle/Saale Dissertation.

Ning, T., Liu, W., Li, Z., Feng, Q., 2020. Modelling and attributing evapotranspiration changes on China's Loess Plateau with Budyko framework considering vegetation dynamics and climate seasonality. Stoch. Env. Res. Risk A. 34, 1217–1230. https://doi.org/10.1007/s00477-020-01813-0.

Norman, J.M.M., Kustas, W.P.P., Humes, K.S.S., 1995. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. Agric. For. Meteorol. 77, 263–293. https://doi.org/10.1016/0168-1923(95)02265-Y.

Pachocka, M., Mansour, M., Hughes, A., Ward, R., 2015. Challenges of modelling a complex multi-aquifer groundwater system at a national scale: case study from the UK. MODFLOW and More 2015: Modeling a Complex World, Colorado, USA, 31 May - 6 June 2015. Colorado School of Mines, Colorado, USA, pp. 216–220.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Pulido-Velazquez, D., Collados-Lara, A.-J., Alcalá, F.J., 2018a. Assessing impacts of future potential climate change scenarios on aquifer recharge in continental Spain. J. Hydrol. 567, 803–819. https://doi.org/10.1016/j.jhydrol.2017.10.077.

Pulido-Velazquez, D., Renau-Pruñonosa, A., Llopis-Albert, C., Morell, I., Collados-Lara, A.-J., Senent-Aparicio, J., Baena-Ruiz, L., 2018b. Integrated assessment of future potential global change scenarios and their hydrological impacts in coastal aquifers – a new tool to analyse management alternatives in the Plana Oropesa-Torreblanca aquifer. Hydrol. Earth Syst. Sci. 22, 3053–3074. https://doi.org/10.5194/hess-22-3053-2018.

Quintana-Seguí, P., Turco, M., Herrera, S., Miguez-Macho, G., 2017. Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-interim. Hydrol. Earth Syst. Sci. 21, 2187–2201. https://doi.org/10.5194/hess-21-2187-2017.

Reinecke, R., Müller Schmied, H., Trautmann, T., Burek, P., Flörke, M., Gosling, S., Grillakis, M., Hanasaki, N., Koutroulis, A., Pokhrel, Y., Seaby, L., Thiery, W., Wada, Y., Yusuke, S., Döll, P., 2020. Uncertainty of simulated groundwater recharge at different global warming levels: a global-scale multi-model ensemble study. Hydrol. Earth Syst. Sci. Discuss. 1–33. https://doi.org/10.5194/hess-2020-235.

Reitz, M., Sanford, W.E., Senay, G.B., Cazenas, J., 2017. Annual estimates of recharge, quickflow runoff, and evapotranspiration for the contiguous U.S. using empirical regression equations. J. Am. Water Resour. Assoc. 53, 961–983. https://doi.org/10.1111/1752-1688.12546.

Richey, A.S., Thomas, B.F., Lo, M., Reager, J.T., Famiglietti, J.S., Voss, K., Swenson, S., Rodell, M., 2015. Quantifying renewable groundwater stress with GRACE. Water Resour. Res. 51, 5217–5238. https://doi.org/10.1002/2015WR017349.

Riedel, T., Weber, T.K.D., 2020. Review: the influence of global change on Europe's water cycle and groundwater recharge. Hydrogeol. J. 28, 1939–1959. https://doi.org/10.1007/s10040-020-02165-3.

Rouholahnejad Freund, E., Fan, Y.W., Kirchner, J., 2020. Global assessment of how averaging over spatial heterogeneity in precipitation and potential evapotranspiration affects modeled evapotranspiration rates. Hydrol. Earth Syst. Sci. 24, 1927–1938. https://doi.org/10.5194/hess-24-1927-2020.

Samaniego, L., Thober, S., Wanders, N., Pan, M., Rakovec, O., Sheffield, J., Wood, E.F., Prudhomme, C., Rees, G., Houghton-Carr, H., Fry, M., Smith, K., Watts, G., Hisdal, H., Estrela, T., Buontempo, C., Marx, A., Kumar, R., 2019. Hydrological forecasts and projections for improved decision-making in the water sector in Europe. Bull. Am. Meteorol. Soc. 100, 2451–2471. https://doi.org/10.1175/BAMS-D-17-0274.1.

Scanlon, B.R., Zhang, Z., Save, H., Sun, A.Y., Schmied, H.M., Van Beek, L.P.H., Wiese, D.N., Wada, Y., Long, D., Reedy, R.C., Longuevergne, L., Döll, P., Bierkens, M.F.P., 2018. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. Proc. Natl. Acad. Sci. U. S. A. 115, E1080–E1089. https://doi.org/10.1073/pnas.1704665115.

Schröder, J.J., Hilhorst, G.J., Oenema, J., Verloop, J., van den Berg, W., 2021. BOFEK2020 - Bodemfysische schematisatie van Nederland: update bodemfysische eenhedenkaart. https://doi.org/10.18174/541544.

Schulte, R.P.O., Diamond, J., Finkele, K., Holden, N.M., Brereton, A.J., 2005. Assessing the soil moisture conditions of irish grasslands. Irish J. Agric. Food Res. 44, 95–110.

Seidenfaden, I.K., Mansour, M., Bessiere, H., Pulido-Velázquez, D.P., Atanskovic Samolov, K., Baena-Ruiz, L., Bishop, H., Dessì, B., Hinsby, K., Hunter Williams, N.H., Højbjerg, A., Larva, O., Martarelli, L., Mowbray, R., Nielsen, A., Ohman, J., Petrovic Pantic, T., Stroj, A., van der Keur, P., Zaadnoordijk, W.J., 2022. Evaluating recharge estimates based on groundwater level time series from different lumped models across Europe. Hydrogeol. J. Under review.

Shang, K., Yao, Y., Li, Y., Yang, J., Jia, K., Zhang, X., Chen, X., Bei, X., Guo, X., 2020. Fusion of five satellite-derived products using extremely randomized trees to estimate terrestrial latent heat flux over Europe. Remote Sens. 12. https://doi.org/10.3390/rs12040687.

Soltani, M., Bjerre, E., Koch, J., Stisen, S., 2021a. Integrating remote sensing data in optimization of a national water resources model to improve the spatial pattern performance of evapotranspiration. J. Hydrol. 603, 127026. https://doi.org/10.1016/j.jhydrol.2021.127026.

Soltani, M., Koch, J., Stisen, S., 2021b. Using a groundwater adjusted water balance approach and copulas to evaluate spatial patterns and dependence structures in remote sensing derived evapotranspiration products. Remote Sens. 13, 1–25. https://doi.org/10.3390/rs13050853.

Sperna Weiland, F., Lopez, P., van Dijk, A., Schellekens, J., 2015. Global high-resolution reference potential evaporation. 21st International Congress on Modelling and Simulation, 29 November–4 December 2015. Broadbeach, Queensland, Australia, pp. 2548–2554.

Stisen, S., Hojberg, A.L.L., Troldborg, L., Refsgaard, J.C.C., Christensen, B.S.B.S.B., Olsen, M., Henriksen, H.J.J., 2012. On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes. Hydrol. Earth Syst. Sci. 16, 4157–4176. https://doi.org/10.5194/hess-16-4157-2012.

Stisen, S., Soltani, M., Mendiguren, G., Langkilde, H., Garcia, M., Koch, J., 2021a. Spatial patterns in actual evapotranspiration climatologies for Europe. Remote Sens. 13, 2410. https://doi.org/10.3390/rs13122410.

Stisen, S., Soltani, M., Mendiguren, G., Langkilde, H., Garcia, M., Koch, J.A.-V.F., 2021. Gridded European Evapotranspiration Climatologies. https://doi.org/10.22008/FK2/T6NBHH.

Stoelzle, M., Schuetz, T., Weiler, M., Stahl, K., Tallaksen, L.M., 2020. Beyond binary baseflow separation: a delayed-flow index for multiple streamflow contributions. Hydrol. Earth Syst. Sci. 24, 849–867. https://doi.org/10.5194/hess-24-849-2020.

Sutanudjaja, E.H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J.H.C., Drost, N., Van Der Ent, R.J., De Graaf, I.E.M., Hoch, J.M., De Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M.W., Vannametee, E., Wisser, D., Bierkens, M.F.P., 2018. PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geosci. Model Dev. 11, 2429–2453. https://doi.org/10.5194/gmd-11-2429-2018.

Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D.G., Keller, V.D.J., 2021. Gridded Estimates of Daily and Monthly Areal Rainfall for the United Kingdom (1890-2012) [CEH-GEAR] [WWW Document]. https://doi.org/10.5285/5dc179dc-f692-49ba-9326-a6893a503f6e.

Teuling, A.J., De Badts, E.A.G., Jansen, F.A., Fuchs, R., Buitink, J., Van Dijke, A.J.H., Sterling, S.M., 2019. Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe. Hydrol. Earth Syst. Sci. 23, 3631–3652. https://doi.org/10.5194/hess-23-3631-2019.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev. 38, 55–94.

TNO-GSN, 2021. REGIS II: The Hydrogeological Model.

TNO-GSN, 2021. Detailing the Upper Layers With GeoTOP.

Trichakis, I., Burek, P., de Roo, A., Pistocchi, A., 2017. Towards a pan-european integrated groundwater and surface water model: development and applications. Environ. Process. 4, S81–S93. https://doi.org/10.1007/s40710-017-0216-0.

Turc, L., 1954. Le bilan d'eau des sols: relation entre la precipitations, l'evaporation et l'ecoulement. Ann. Agron. 5, 491–569.

Vásquez, V., Thomsen, A., Iversen, B.V., Jensen, R., Ringgaard, R., Schelde, K., 2015. Integrating lysimeter drainage and eddy covariance flux measurements in a groundwater recharge model. Hydrol. Sci. J. 60, 1520–1537. https://doi.org/10.1080/02626667.2014.904964.

Vermeulen, P.T.M., Minnema, B., Roelofsen, F.J., 2021. iMOD User Manual. Version 5.3 July 30, 2021.

Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., Soubeyroux, J.-M., 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. Int. J. Climatol. 30, 1627–1644. https://doi.org/10.1002/joc.2003.

Voortman, B.R., Bartholomeus, R.P., van der Zee, S.E.A.T.M., Bierkens, M.F.P., Witte, J.P.M., 2015. Quantifying energy and water fluxes in dry dune ecosystems of the Netherlands. Hydrol. Earth Syst. Sci. 19, 3787–3805. https://doi.org/10.5194/hess-19-3787-2015.

van Walsum, P.E.V., Groenendijk, P., 2008. Quasi steady-state simulation of the unsaturated zone in groundwater modeling of lowland regions. Vadose Zone J. 7, 769–781. https://doi.org/10.2136/vzj2007.0146.

Wanders, N., Thober, S., Kumar, R., Pan, M., Sheffield, J., Samaniego, L., Wood, E.F., 2019. Development and evaluation of a pan-european multimodel seasonal hydrological forecasting system. J. Hydrometeorol. 20, 99–115. https://doi.org/10.1175/JHM-D-18-0040.1.

Westerhoff, R., White, P., Rawlinson, Z., 2018. Incorporation of satellite data and uncertainty in a nationwide groundwater recharge model in New Zealand. Remote Sens. 10, 1–25. https://doi.org/10.3390/rs10010058.

Wu, Q., Si, B., He, H., Wu, P., 2019. Determining regional-scale groundwater recharge with GRACE and GLDAS. Remote Sens. 11. https://doi.org/10.3390/rs11020154.

Xu, C.Y., Chen, D., 2005. Comparison of seven models for estimation of evapotranspiration and groundwater recharge using lysimeter measurement data in Germany. Hydrol. Process. 19, 3717–3734. https://doi.org/10.1002/hyp.5853.

Xu, Y., Beekman, H.E., 2019. Review: groundwater recharge estimation in arid and semi-arid southern Africa. Hydrogeol. J. 27, 929–943. https://doi.org/10.1007/s10040-018-1898-8.

Yang, H., Yang, D., Lei, Z., Sun, F., 2008. New analytical derivation of the mean annual water-energy balance equation. Water Resour. Res. 44. https://doi.org/10.1029/2007WR006135.

Zhang, L., Hickel, K., Dawes, W.R., Chiew, F.H.S., Western, A.W., Briggs, P.R., 2004. A rational function approach for estimating mean annual evapotranspiration. Water Resour. Res. 40. https://doi.org/10.1029/2003WR002710.

Zhang, Y., Kong, D., Gan, R., Chiew, F.H.S., McVicar, T.R., Zhang, Q., Yang, Y., 2019. Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. Remote Sens. Environ. 222, 165–182. https://doi.org/10.1016/j.rse.2018.12.031.