Understanding process controls on groundwater recharge variability across Africa through Recharge Landscapes

Charles West¹, Rafael Rosolem¹,², Alan M. MacDonald³, Mark O. Cuthbert⁴,⁵ and Thorsten Wagener¹,⁶

¹ Civil Engineering, University of Bristol, Bristol, United Kingdom
² Cabot Institute for the Environment, University of Bristol, Bristol, United Kingdom
³ British Geological Survey, Lyell Centre, Edinburgh EH14 4AP, United Kingdom
⁴ School of Earth and Environmental Sciences, Cardiff University, Park Place, Cardiff, CF10 3AT, United Kingdom
⁵ School of Civil and Environmental Engineering, The University of New South Wales, Sydney, Australia
⁶ Institute for Environmental Science and Geography, University of Potsdam, 14476 Potsdam, Germany

Abstract

Groundwater is critical in supporting current and future reliable water supply throughout Africa. Although continental maps of groundwater storage and recharge have been developed, we currently lack a clear understanding on how the controls on groundwater recharge vary across the entire continent. Reviewing the existing literature, we synthesize information on reported groundwater recharge controls in Africa. We find that 15 out of 22 of these controls can be characterised using global datasets. We develop 11 descriptors of climatic, topographic, vegetation, soil and geologic properties using global datasets, to characterise groundwater recharge controls in Africa. These descriptors cluster Africa into 15 Recharge Landscape Units for which we expect recharge controls to be similar. Over 80% of the continents land area is organized by just nine of these units. We also find that aggregating the Units by similarity into four broader Recharge Landscapes (Desert, Dryland, Wet tropical and Wet tropical forest) provides a suitable level of landscape organisation to explain differences in ground-based long-term mean annual recharge and recharge ratio (annual recharge / annual precipitation) estimates. Furthermore, wetter Recharge Landscapes are more efficient in converting rainfall to recharge than drier Recharge Landscapes as well as
having higher annual recharge rates. In Dryland Recharge Landscapes, we found that annual
recharge rates largely varied according to mean annual precipitation, whereas recharge ratio
estimates increase with increasing monthly variability in P-PET. However, we were unable to
explain why ground-based estimates of recharge signatures vary across other Recharge
Landscapes, in which there are fewer ground-based recharge estimates, using global datasets
alone. Even in dryland regions, there is still considerable unexplained variability in the
estimates of annual recharge and recharge ratio, stressing the limitations of global datasets for
investigating ground-based information.

Keywords: Groundwater recharge, Africa, recharge controls, ground-based estimates,
landscapes, comparative hydrology

1 Introduction

With an estimated storage of 0.66 million km$^3$, groundwater is the largest store of freshwater
in Africa, exceeding annual volumes of streamflow by a factor of 100, and its development is
fundamental for securing current and future water supply (MacDonald et al., 2012).
Understandably it is often regarded as important for both urban (Foster et al. 2020; Oiro et al.
2020) and rural (Calow et al. 1997; Lapworth et al. 2013; MacDonald and Calow
2009) communities, though quantifying its role in water supply remains challenging for many
parts of the continent (Chávez García Silva et al. 2020). In Northern Africa, groundwater
irrigation practices (Siebert et al. 2010) have potentially led to the depletion of groundwater
resources (Aeschbach-Hertig and Gleeson 2012; Rodell et al. 2018). In contrast, groundwater
use in sub-Saharan Africa is primarily for domestic supply and sanitation services (Braune
and Xu 2010) and is potentially being under-utilized for crop production (Giordano 2006;
Siebert et al. 2010). In dryland river basins, securing water supply from surface water is
challenging due to the high inter-annual variability of streamflow and persistent dry periods
Therefore increasing groundwater abstraction for more conjunctive use of surface water and groundwater could reduce vulnerability to climate driven surface water shortages, particularly in rural communities (Calow et al., 1997; Lapworth et al., 2013; MacDonald & Calow, 2009) and generally improve water accessibility (Robins et al., 2006). In the future, rapid population growth (Gerland et al. 2014; Parnell and Walawege 2011) and climate change could further enhance the value of African groundwater resources (Taylor et al. 2013).

Yet, our understanding of the spatial variability of groundwater recharge processes across Africa remains limited, constraining our ability to plan for the sustainable use of this resource (MacDonald et al., 2021), though recharge rates should not be regarded as a safe-yield for groundwater use (Aeschbach-Hertig and Gleeson 2012). Recent studies have tried to overcome this problem in multiple ways: [1] Scaling up knowledge from a limited number of detailed local studies. Cuthbert et al. (2019b) used multi-decadal groundwater level timeseries in conjunction with local knowledge to develop site specific conceptual models which allowed the authors to highlight a relationship between climate and recharge frequency, sensitivity to precipitation and dominant recharge mechanisms. However, this approach relies heavily upon rare long-term data as well as local knowledge and therefore it is challenging to transfer findings to larger scales or different regions. [2] Recently, MacDonald et al. (2021) used 134 ground-based annual recharge estimates compiled from the literature along with global datasets to develop a continental statistical model. This model enabled them to estimate long-term groundwater recharge rates across Africa using mean annual precipitation without qualitative inclusion of different recharge processes. [3] Most studies have based their continental scale estimates on process-based models. Global scale hydrological models and land surface models can estimate groundwater recharge rates across
large spatial domains (Reinecke et al. 2021). However, recharge outputs from these models have not yet been thoroughly evaluated (Bierkens, 2015; Telteu et al., 2021; Wagener et al., 2021). Furthermore, global models thus far only include a limited number of process representations and neglect regionally dominant controls which could be important for Africa, such as karst (Hartmann et al., 2015; Hartmann et al., 2014) or dryland-specific hydrological processes (Quichimbo et al. 2021). This is likely because most (if not all) available continental to global scale models available to estimate recharge have their origin from regions outside of Africa.

In this study we want to take a step back to review what dominant controls should be present in a model across Africa and investigate how well we can quantify these process controls given the available data. In doing so, we specifically aim to answer three questions: (i) What are the dominant controls on groundwater recharge already identified across Africa in previous studies? (ii) Using global datasets only, what descriptors of controlling processes can we define, and which regions of Africa should have similar recharge controls when clustered using these descriptors? (iii) How do these regions for which we expect similar controls compare to ground-based recharge observations? Due to the limited amount of ground-based data on groundwater recharge in Africa, we adopt an approach which builds strongly on our a priori understanding of recharge controls in Africa identified from the literature. In doing so we build on previous efforts by Scanlon et al. (2006) who synthesized qualitative local knowledge of recharge processes for the world’s dry regions. In keeping with the database compiled by (MacDonald et al., 2021), we only review the controls on recharge which is distributed throughout the landscape. MacDonald et al. (2021) define distributed recharge as both diffuse and focussed recharge but exclude focussed recharge from large discrete features such as rivers or lakes. Where focussed recharge is widely distributed through ephemeral rivers, depressions or rock fractures which are common over a
large area and contribute to regional recharge, they include this in their definition of distributed recharge which we use. We follow the ideas of Winter's concept of hydrological landscapes (Winter 2001) and define Recharge Landscape Units to represent areas for which we expect similar recharge controls. We then compare these areas against an openly available, comprehensive and thoroughly quality assured dataset of ground-based recharge estimates in Africa, recently published by MacDonald et al. (2021). Although we use their database in our analysis, this work has some key differences to the previous work by MacDonald et al. (2021). Firstly, we attempt to explicitly link our analysis to both the qualitative (understanding of recharge controls) and quantitative (i.e., ground-based recharge estimates) findings in the literature, whereas MacDonald et al. (2021) only investigate the quantitative data. Furthermore, our classification approach allows us to explore whether relationships between environmental controls and recharge signatures vary between different environmental settings. In contrast, the statistical approach taken by MacDonald et al. (2021) only allowed them to investigate relationships which applied to the whole continent. Finally, we investigate both long-term mean annual recharge rates of groundwater recharge and recharge ratios. This allows us to understand how different recharge signatures vary and interact in space, furthering our understanding of groundwater recharge beyond just looking at annual rates.

2. Review of process controls on groundwater recharge across Africa

Most of the existing knowledge base on groundwater recharge processes, controls and rates in Africa comes from a relatively small number of case studies investigating recharge at the field, catchment, or sometimes regional scale. These studies use a wide range of methods to understand recharge processes throughout the continent, with approaches often varying according to environmental setting, data availability and the objective of the individual studies (MacDonald et al. 2021). Details of the strengths and weaknesses of the different
methods can be found in Scanlon et al. (2002) and Healy (2010). We organize the review of controls into four domains: climate and weather, topography, landcover/use, and soils and geology. The aim of this review is firstly to identify dominant controls on groundwater recharge, and secondly to understand whether these controls have clear positive or negative relationships with groundwater recharge, or if their relationship with recharge is ambiguous. We are considering processes that govern the potential recharge of an aquifer, which can be more than the actual recharge due to interflow processes or if the potential recharge rate is so large that it exceeds the rate at which water can flow laterally through the aquifer (Theis 1940). In the latter case, the aquifer can become over-full such that available recharge is rejected. We show a summary of this review in Figure 1. An extended version of the review can be found in the supplemental information.

**Climate and weather**

Annual scale components of the water-energy balance are a first order control on the spatial variability of groundwater recharge (Kim and Jackson, 2012; Mohan *et al.*, 2018; Cuthbert *et al.*, 2019b; MacDonald *et al.*, 2021), as they control the quantity of water available to be partitioned into groundwater recharge, as well as the energy available to partially control atmospheric losses (Budyko, 1974). Hence studies in Africa show variability of annual recharge rates along a climate gradient, largely defined by precipitation due to the generally high levels of energy available (MacDonald *et al.* 2021). In an upland catchment of Cameroon where rainfall exceeds 3000 mm/year, estimated recharge rates exceed 900 mm/year (Kamtchueng *et al.* 2015), in comparison to recharge rates between 160 mm/year and 330 mm/year in the Ethiopian Highlands where annual rainfall is approximately 1300 mm/year (Azagegn *et al.* 2015; Banks *et al.* 2021; Demlie 2015). Groundwater resources throughout the deserts, which receive very little annual rainfall (Nicholson 2000), are recharged at rates below 5 mm/year (Foster *et al.*, 1982; Dabous and Osmond, 2001; Zouari...
et al., 2011), or may not even be actively recharged (Befus et al. 2017). In these regions deep
‘fossil’ groundwaters recharged prior to the Holocene dominate aquifer stores (Sturchio et al.,
2004; Guendouz et al., 2006; Abotalib et al., 2016; Jasechko et al., 2017).

Groundwater recharge volumes are often biased towards the rainy season as elevated rainfall
is required to overcome high rates of evapotranspiration (Bromley et al., 1997; Demlie et al.,
2007; Walraevens et al., 2009; Mechal et al., 2015), and greater monthly and daily
precipitation intensity leads to a more efficient conversion of rainfall to recharge (Jasechko
observations in the Makutapora wellfield, Tanzania, suggest that recharge is dependent upon
months with the most extreme (> 95th percentile) rainfall (Taylor, Todd, et al. 2013) often
enhanced by the El Nino Southern Oscillation and the Indian Ocean Dipole. However, the
multiple climate oscillations known to affect climate patterns in Africa (Brown et al., 2010)
can have opposing effects in different parts of the continent (Nicholson and Kim 1997).
Nonetheless, wetting and drying cycles are being reflected in observed groundwater
hydrographs throughout Africa (Taylor et al., 2013; Cuthbert et al., 2019b; Kolusu et al.,
2019), showing both seasonally extreme recharge events as well as recharge events which are
more episodic in nature.

Episodic rainfall events are particularly important in arid landscapes where recharge often
depends upon a small number of days of intense rainfall (Vogel and Van Urk, 1975; Mazor et
al., 1977; Van Tonder and Kirchner, 1990; Nkotagu, 1996; De Vries et al., 2000; Xu and
Beekman, 2003; Wanke et al., 2008). Döll and Fiedler (2008) stressed the importance of
heavy rainfall events in semi-arid and arid regions as they modelled groundwater recharge
globally, applying a rainfall threshold of 10 mm/day to drylands, below which they assumed
recharge would not occur. They identified this threshold via an independent analysis of 25
chloride profile estimates of annual recharge distributed throughout the world as well as
regional model estimates of recharge in Death Valley, California.

In summary, annual and seasonal precipitation as well as heavy rainfall events have a positive
relationship with groundwater recharge in Africa – largely driving inter- and intra-annual
recharge variability, while the amount of energy available from radiation has a negative
relationship with groundwater recharge. However, the influence of large-scale climate
oscillations on groundwater recharge in Africa is less clear as their effect on climate patterns
vary regionally.

**Topography**

Topographic slope controls the movement of water across the land surface and therefore
controls water infiltration into the subsurface and groundwater recharge, with gentler slopes
promoting more recharge than steeper slopes (Simmers 1990). The role of slope in
controlling groundwater recharge has been discussed throughout many different regions of
Africa, including Ethiopia (Gebreyohannes et al. 2013), Nigeria (Abdulateef et al. 2021;
Fashae et al. 2014), Botswana (Lentswe and Molwalefhe 2020) and Algeria (Boufekane et
al., 2020). Yet interestingly, McKenna and Sala (2018) found that recharge beneath flat
playas in the south-western United States is greater when they are surrounded by steeper
slopes which promote greater run-on onto the playa.

In dry regions, intense rainfall events are important drivers of focused recharge through flash
flooding (Sultan et al. 2000) and the formation of ephemeral water bodies and depression
storage (Lehner and Döll, 2004), i.e. in areas where water accumulates on the land surface.
In Africa’s dry regions, alluvial aquifers underlying dry riverbeds are recharged episodically
or perhaps seasonally by river transmission losses following heavy rainfall (Tantawi, El-
Sayed and Awad, 1998; Sultan et al., 2000; Gheith and Sultan, 2002; Benito et al., 2010;
Walker et al., 2019; Seddon et al., 2021). These storms can activate focused recharge mechanisms despite negligible diffuse recharge in interfluve regions due to high evaporation (Favreau et al. 2009). In endoreic arid basins, surface water can also accumulate in salt pans which typically occupy topographic depressions (Lehner and Döll 2004). (De Vries et al., 2000) use chloride profiles to show that in the eastern fringes of the Kalahari Desert, recharge is enhanced under these pans, with estimated annual rates of 50mm in comparison to 7mm for the surrounding landscape.

Therefore, slope generally has a negative relationship with groundwater recharge since it will provide an easier flow path for water to move downhill, whereas topographic depressions have a positive relationship with (focused) groundwater recharge because they allow water to accumulate.

**Landcover/use**

Landcover and use varies considerably across the African continent. Bare soils (33% of Africa’s land area) occupy most of northern Africa as well as parts of southern and eastern Africa, whilst grasslands (15.4%), shrublands (13.4%) and agriculture (11.6%) are largely distributed throughout the Sahel and Southern and Eastern Africa, and forests and woodland (26%) spread across western, central and south-eastern regions (Mayaux et al., 2004; Tsendbazar et al., 2017; Xiong et al., 2017). These vegetation patterns influence the spatial variability of groundwater recharge (Kim and Jackson 2012) through their control over transpiration, interception and soil evaporation fluxes (Gordon et al., 2005; Schlesinger and Jasechko, 2014; Good et al., 2015).

An estimated 7% of the continent’s precipitation returns to the atmosphere via interception evaporation, mostly occurring in the densely forested regions of Central Africa where this flux can exceed 10% of the precipitation input (Miralles et al. 2010; Zhang et al. 2016; Zheng
et al. 2017). Globally, we could not find any studies directly discussing the relationship between rainfall interception and groundwater recharge. However, it seems reasonable to assume that by limiting the amount of precipitation reaching the land surface, interception consequently reduces groundwater recharge.

An estimated 49% and 21% of precipitation over Africa returns to the atmosphere via transpiration and bare soil evaporation, respectively (Zhang et al. 2016). The bulk of continental transpiration is associated with the tropical forests (Gordon et al., 2005; Good et al., 2015), where tall vegetation with deep rooting systems increases the capacity of root-zone moisture storage (Nijzink et al. 2016) and the access to deeper groundwater (Barbeta and Peñuelas 2017). When investigating groundwater recharge at regional and catchment scales, studies often find that recharge rates are lower in areas which are forested than in areas which are unforest or have bare soils (Gebreyohannes et al. 2013; Houston 1982; Howard and Karundu 1992; Stone and Edmunds 2012). Furthermore, the presence of woodland or forest can restrict groundwater recharge to years of particularly high rainfall, even when recharge in grass, crop or unvegetated parts of the catchment occurs annually (Houston 1982; Howard and Karundu 1992). In the Kalahari Desert, dense bush and tree savannah is believed to transpire much of the annual rainfall during the long dry season, leading to very little recharge (De Vries et al., 2000; Sibanda et al., 2009). Similarly, chloride profiles in Senegal, suggest that groundwater recharge rates decline as vegetation density increases (Edmunds and Gaye 1994). Land clearing, often for agricultural expansion, can also enhance groundwater recharge rates by reducing evapotranspiration (Taylor and Howard 1996; Været et al. 2009).

Land clearing for agriculture does not only affect recharge through changes to evapotranspiration, it can also alter the mechanisms through which recharge occurs, by altering soil surface properties (Wirmvem et al. 2015) as well as runoff run-on processes.
(Leduc et al., 2001; Leblanc et al., 2008; Favreau et al., 2009; Ibrahim et al., 2014; Wirmvem et al., 2015). Agricultural land adjacent to many of Africa’s largest lakes and rivers is regularly equipped for irrigation (Siebert et al. 2015). Excess irrigation water can infiltrate into the soil and percolate to the aquifer, increasing groundwater recharge rates (Bouimouass et al. 2020; Scanlon et al. 2007). Nonetheless, as irrigation technologies become more efficient, recharge via irrigation excesses is expected to decline (Scanlon et al. 2007). Urban settings only account for less than 0.01% of the African landscape (Zhou et al. 2015). Although, urbanisation is typically perceived as reducing groundwater recharge by reducing the permeable surface area, recharge rates in urban areas can be as high as or even higher than nearby rural areas (Lerner 2002; Sharp 2010). Urbanization can dampen existing recharge mechanisms, but it can also introduce new mechanisms such as localised recharge where there is little drainage infrastructure (Lerner 2002; Sharp 2010), as well as leakages from on-site sanitation (Foster et al., 1999; Diouf, 2012; Lapworth et al., 2017) and piped distribution networks if such water supply is available.

In short, we find that the transpiration and canopy storage controls of different landcovers show a negative relationship with groundwater recharge, whereas the additional supply of water to agricultural land through irrigation has a positive relationship with recharge. Effects of urbanisation on groundwater recharge on the other hand are more ambiguous.

**Soils and Geology**

Soils with larger sand fractions are more permeable and support higher recharge rates than finer clay soils. In a global scale meta-analysis of recharge estimates, Kim and Jackson (2012) show that on average sandy soils are 50% more efficient in converting water input into groundwater recharge. Similar results are found at regional and catchment scales in Senegal, Sudan and Zimbabwe, whereby higher recharge rates are estimated in areas where
the sand fraction is a more dominant component of the soil (Abdalla 2009; Butterworth et al. 1999; Edmunds and Gaye 1994). Lower recharge rates are found in clayey soils as the vertical percolation of water through the soil profile is restricted (Attandoh et al. 2013; Edmunds et al. 1992) and soil moisture is more exposed to evapotranspiration (Mensah et al. 2014; Yidana and Koffie, 2014; Kotchoni et al., 2018).

However, soil texture alone fails to recognise structural soil properties which enable infiltration via preferential flow paths which bypass the soil matrix (Beven and Germann 1982). Macropores in the soil structure allow infiltration to bypass vegetation rooting zones and impermeable soil layers (De Vries et al., 2000; Mazor, 1982; Van Tonder & Kirchner, 1990; Xu & Beekman, 2003) and facilitate recharge in conditions which would otherwise be prohibitive. These preferential flow paths are an important mechanism for groundwater recharge across a range of contrasting environmental settings. In the Botswana Kalahari Desert, semi-arid Tanzania and the tropical highlands of Ethiopia, the contribution of preferential flows to groundwater recharge is approximately 24%, 60% and 36%, respectively (Demlie et al. 2007; Nkotagu 1996; de Vries and Gieske 1990).

Rock fracturing (Nkotagu, 1996; Xu and Beekman, 2003; Adams et al., 2004; Kebede et al., 2005; Kamtchueng et al., 2015) and vertical conduits in karstic rock (Farid et al., 2014; Hartmann et al., 2014, 2017; Chemseddine et al., 2015; Ayadi et al., 2018; Leketa et al., 2019) also provide preferential flow paths for groundwater recharge. In dry landscapes such as the Kalahari Desert, rock fracturing at bedrock outcrops and isolated rock formations called inselbergs (Burke 2003) can locally enhance groundwater rates (Mazor, 1982; Butterworth et al., 1999; Brunner et al., 2004; Wanke et al., 2008). The distribution and geometry of the superficial geology can also have a marked impact on recharge pathways and rates in conjunction with the underlying bedrock and distribution of stream networks (Zarate
et al. 2021). Similar observations have been made regarding focused recharge opportunities for water in karstic regions (Hartmann et al. 2017).

Soil perturbations such as crusting, cementation, compaction, weathering, and tillage can also have a significant impact on recharge rates. Whilst studies mostly find that soil crusting (Favreau et al. 2009; Jacks and Traoré 2014; Wakindiki and Ben-Hur 2002), cementation (Nash et al., 1994; De Vries et al., 2000; Xu and Beekman, 2003; Francis et al., 2007) and compaction (Hamza and Anderson, 2005; du Toit et al., 2009) reduce the permeability of soil layers and hence reduce groundwater recharge, the effects of deeply weathered soils known as laterites (Bromley et al., 1997; Rueedi et al., 2005; Cuthbert and Tindimugaya, 2010; Bonsor et al., 2014) and agricultural tilling practices (Abu-Hamdeh, 2004; Osunbitan et al., 2005; Spaan et al., 2005; Strudley et al., 2008; Thierfelder and Wall, 2009; Abidela Hussein et al., 2019) on recharge are much less clear.

Therefore, in summation we find that, soil grain sizes, bedrock outcrops and properties which promote preferential flow paths, such as soil macropores, rock fractures and karst geology, have a positive relationship with groundwater recharge. Some soil perturbations such as compaction, cementation and crusting have a negative relationship with groundwater recharge, whereas others, including tilling and soil laterization, have a less clear relationship with recharge.

Interactions between controls

Up to now we have largely looked at landscape properties and their control over recharge processes independently, in reality, groundwater recharge is a function of the interactions between these controls. Hence at the continental scale, we would typically expect to find some of the lowest recharge rates in areas with the most freely draining soils, as these regions also have the lowest precipitation volumes. By identifying patterns in the landscape, i.e.
climate, topography, vegetation, soils and geology, we can begin to conceptualise recharge processes of different environmental settings found in Africa. We can find these patterns as landscapes are continuously co-evolving (Troch et al. 2013) via an array of physical and biological processes which effect the uplift and deformation of bedrock and the erosion, transportation and deposition of sediments (Dietrich and Perron 2006; Reinhardt et al. 2010).

This co-evolution, explains why we typically expect to find certain landscapes throughout the continent, including rainforests, tropical woodlands and savannas and deserts.

We often regard climate as an external force driving the hydrological system, but it also controls the spatial and temporal patterns of landcover (Zhou et al., 2014; Hawinkel et al., 2016; Bouvet et al., 2018; Measho et al., 2019; Ndehedehe et al., 2019) and soils (Jenny 1941; Towett et al. 2015). Climate and vegetation patterns as well as soil properties are also strongly affected by local topography. In mountainous areas we see vegetation becoming shorter and less dense above the treeline, as temperatures decline and thinning soils make ground conditions less stable (Harsch et al., 2009; Egli and Poulenard, 2016). Increased precipitation and runoff due to orographic forcing as well as steeper slopes, promote more active erosion and sediment transport fluxes at elevation and therefore prevents the accumulation of soils (Acosta et al. 2015). In contrast, at lower elevations, vegetation can assist the accumulation of soils by reducing surface water erosion and promoting infiltration (Acosta et al. 2015; Descheemaeker et al. 2006; Descroix et al. 2009; Thompson et al. 2010).

In water limited regions, vegetation density often increases in topographic depressions such as ephemeral streams, as accessibility to groundwater may be locally improved (Morin et al., 2009; Steward et al., 2012; Ndehedehe et al., 2019; Grodek et al., 2020).
Summary

Figure 1. Summary of groundwater recharge controls for Africa identified in the literature. Controls are colour coded according to their relationship with recharge with red and blue representing negative and positive relationships, respectively. Bold font highlights controls which we can characterise using global datasets.

3. Materials and methods

To consolidate our understanding of groundwater recharge controls taken from reviewing the literature, we take a classification approach which we can then use as a tool to investigate why ground-based estimates of groundwater recharge vary spatially across the continent. We acknowledge their will be uncertainty in the classification due to the limitations of our own understanding and of global datasets. However, we aim to connect the qualitative and quantitative information obtained from local/regional findings to large scale regionalization approaches.
3.1 Global Datasets

We used nine global datasets to characterize the previously identified groundwater recharge controls. Furthermore, controls were only integrated into our classification if the literature indicated it had a clear positive or negative relationship with groundwater recharge and it could be characterized using global datasets. The datasets used and the indices calculated are summarized in Table 1.

Indices describing annual and seasonal climate attributes mostly characterize first-order estimates of the water potentially available for groundwater recharge (P-PET) annually and seasonally as well as its variability. This also builds on previous work by Wolock et al. (2004) who used P-PET as the climatic index to delineate hydrological landscapes in the United States. We characterized heavy rainfall across Africa using a threshold of 10 mm/day. Several studies in Africa (Döll and Fiedler 2008; Owor et al. 2009; Taylor and Howard 1996) have found annual recharge has a stronger correlation with the average volume of rainfall per year on days with at least 10 mm of rain, than with mean annual precipitation and hence we selected this as threshold for heavy rainfall in Africa. Though we acknowledge the rainfall threshold for recharge occurrence likely varies across the continent. We characterized the influence of landcover on groundwater recharge via transpiration and canopy storage processes, by attributing vegetation specific transpiration coefficients to a landcover dataset and by looking at the Leaf Area Index, respectively. This approach is also often taken when parameterizing these processes in continental scale hydrological modelling (Telteu et al., 2021). To avoid having multiple indices to describe soil textures we instead calculated the ratio of soils which promote infiltration (i.e., sand) to those which restrict infiltration (i.e., silt and clay) (Saxton et al., 1986; Wösten et al., 2001). We used the depth to bedrock dataset of (Pelletier et al. 2016) to highlight bedrock outcrop regions and the world map of carbonate rock outcrops (Williams and Ford 2006) to highlight the extent of carbonate rock outcrops.
Table 1. Details of the recharge control indices we defined to characterise recharge controls across Africa and the global datasets we used to calculate them.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
<th>Units</th>
<th>Period</th>
<th>Data source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Climate attributes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-PET</td>
<td>Mean annual precipitation minus mean annual PET.</td>
<td>mm/year</td>
<td>1979-2015</td>
<td>1. MSWEP v1.2 (Precipitation)</td>
<td>1. (Beck et al., 2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 0.25°</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Temporal res.: Daily</td>
<td></td>
</tr>
<tr>
<td>P-PET in season</td>
<td>Mean annual volume of precipitation in excess to PET in months considered in-season. A month is considered in-season when P exceeds PET.</td>
<td>mm/year</td>
<td>1979-2015</td>
<td>2. CRU v4 (PET)</td>
<td>2. (Harris et al., 2020)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 0.5°</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Temporal res.: Monthly</td>
<td></td>
</tr>
<tr>
<td>σ(P-PET)</td>
<td>The standard deviation of monthly P-PET</td>
<td>mm/month</td>
<td>1979-2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P10</td>
<td>The average volume of rainfall per year on days with at least 10 mm of rain.</td>
<td>mm/year</td>
<td>1979-2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Topography attributes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td>Geodesic slope of the DEM using a 3 by 3 moving window.</td>
<td>Degrees</td>
<td></td>
<td>HydroSHEDS</td>
<td>(Lehner et al., 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 15 arc seconds</td>
<td></td>
</tr>
<tr>
<td><strong>Landcover/use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 300m</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Temporal res.: Yearly</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 0.25°</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Temporal res.: Monthly</td>
<td></td>
</tr>
<tr>
<td>Irrigation</td>
<td>Area equipped for irrigation multiplied by the fractional area actually irrigated.</td>
<td>km²</td>
<td>2005</td>
<td>Global Map of Irrigation Areas</td>
<td>(Siebert et al., 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 5 arc minutes</td>
<td></td>
</tr>
<tr>
<td><strong>Soil attributes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sand / (Clay + Silt)</td>
<td>The ratio of sand (&gt;0.05mm) to silt (0.002-0.05mm) and clay (&lt;0.002mm) in the fine earth fraction of the top 2m of the</td>
<td>-</td>
<td>-</td>
<td>SoilGrids250m</td>
<td>(Hengl et al. 2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spatial res.: 250m</td>
<td></td>
</tr>
</tbody>
</table>
soil profile. Proportions of each soil texture are by weight. Take the depth weighted harmonic mean across intervals of 0-5cm, 5-15cm, 15-30cm, 30cm-60cm, 60-100cm, 100-200cm.

<table>
<thead>
<tr>
<th>Geology attributes</th>
<th>Depth to bedrock</th>
<th>Average soil and sedimentary deposit thickness. Maximum of 50m.</th>
<th>m</th>
<th>-</th>
<th>Gridded Thickness of Soil, Regolith and Sedimentary Deposit Layers Spatial res.: 30 arc seconds (Pelletier et al. 2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karst</td>
<td>Extent of carbonate rock outcrop areas.</td>
<td>-</td>
<td>-</td>
<td>World Map of Carbonate Rock Outcrops V3.0 (Williams and Ford 2006)</td>
<td></td>
</tr>
</tbody>
</table>

### 3.2 Ground-based annual recharge and recharge ratio estimates

We used the database compiled by MacDonald et al. (2021) of long-term mean annual recharge estimates compiled from case studies in the literature. We selected this database above other meta-datasets (Moeck et al. 2020; Mohan et al. 2018) because of its focus on Africa, the thorough quality assurance conducted throughout its compilation, and the additional meta-data provided, such as recharge estimate uncertainty ranges. Through quality assurance steps, MacDonald et al. (2021) removed 182 datapoints (from an initial 316), due to duplicative studies in the same location and findings which were solely dependent upon hydrological modelling. Additional screening removed data points where the site co-ordinates and date of the study period were not provided. Finally, we removed estimates dated prior to 1979 or after 2015, as they would not correspond to the timing of the climate datasets we used. Ultimately, we were left with 129 ground-based estimates of annual groundwater recharge distributed across Africa. 111 of these sites/studies also reported corresponding mean annual precipitation rates, so we could estimate long-term mean recharge ratios at these
locations (Figure 2). Spatially, 31 of these estimates reflect recharge rates over spatial scales less than 100 km², a further 41, 29, and 28 are for spatial scales of 100-2500 km², 2500-62500 km² and greater than 62500 km², respectively.

Figure 2. The remaining annual recharge and recharge ratio estimates collected from case studies in the literature by MacDonald et al. (2021), after initial screening of the data. The recharge ratio is defined as the fraction of precipitation being converted to recharge (recharge / precipitation).

3.3 Clustering

To delineate regions with expected similar recharge control indices (i.e., Recharge Landscape Units) we use a fuzzy c-means clustering algorithm (Bezdec 1981). This fuzzy clustering algorithm allows for pixels to belong to multiple units simultaneously, albeit with varying degrees of membership, thus enabling us to study the gradual transition between units (e.g., reflecting different landscapes). The degree of overlap in membership allowed us to determine the uniqueness of each delineated Recharge Landscape Unit. The degree of membership is dependent upon how close in value each pixel’s recharge control indices are to the centroid of each unit, which is regarded as being representative for a unit. Membership scores vary from 0 to 1, with 0 representing no similarity and 1 suggesting the pixel’s
recharge control indices are equal to the values of the unit's centroid. Further details on the
algorithm and on application details are provided in the supplemental material. Ultimately,
we attributed each pixel to the unit with which it has the highest degree of membership,
which we refer to as its primary unit.

3.4 Random Forests

We used classification-based Random Forests to expand our classification for recharge
controls in Africa to the rest of the world. Random Forests is a machine learning algorithm
which combines multiple trees to produce an ensemble of predictions (Breiman 2001;
Breiman et al. 1984), which link predictor variables (recharge control indices) to a response
(Recharge Landscape Units). Each individual tree develops rules for predicting responses
which are structured as a binary decision tree composing of nodes and branches. At each
node a conditional binary split is applied to one of the predictor variables. The split forms
two branches which link to nodes in the overlying stratum. This splitting continues until the
terminal node (the leaf) is met and the outcome is predicted. Each classification tree in the
ensemble model is trained on observations (Pixels of classification for recharge controls in
Africa) which were randomly selected with replacement from a sub-sample of 70% of the
total observations (‘in-bag’ observations). The random forest model consists of 25 trees each
with a maximum of 400 decision splits. Increasing the number of trees or decision splits did
not significantly improve model performance. Addor et al., (2018) previously used Random
Forests to predict observed streamflow signatures across the USA and Stein et al., (2021)
used random forests to explore how climate and catchment attributes influence flood
generating processes.
4 Results

4.1 Recharge Landscape Units outline regions with similar recharge controls in Africa

Based on our review in section 2, we defined and calculated 11 indices to characterise the different controls on distributed groundwater recharge we identified in our review (Figure 1). To avoid using redundant information for each control, we checked the correlations between each of the indices initially considered and removed indices such that none of the indices for a given control had Pearson correlation coefficients greater than or equal to 0.7 with one another (see supplemental information) (Dormann et al. 2013).
Figure 3. 11 recharge control indices characterising controls identified in the literature using global datasets. a) P-PET; b) P-PET in-season; c) $\sigma$(P-PET); d) P10; e) Slope; f) Kveg; g) Leaf Area Index; h) Irrigated area; i) Sand / (Clay + Silt); j) Depth to bedrock; k) Karst. The definitions of each index the datasets used for their characterisation are stated in Table 1.
The cluster analysis combines the 11 indices into 15 Recharge Landscape Units with similar recharge control indices of which 9 cover over 80% of the African land area (Figure 4). We initially identified 14 units using fuzzy clustering, as additional units did not greatly reduce the dissimilarity within individual units. The 15th unit which delineates potential karst regions was manually superimposed. Even though we expect recharge to vary significantly between the different settings in which karst is found, we delineate the group as a whole, because we expect the recharge mechanism associated to karst environments to be a dominant control on recharge processes. We can see the continent has been roughly organised into very dry regions in the north and south of the continent and wetter regions spanning from West Africa down through Central Africa towards Mozambique and Madagascar. Even though the spatial organisation of the units suggest proximity is a reasonable indicator for similarity, we do find regions with similar recharge control indices which are also far away from each other. For example, hyper arid regions with shallow soils can be found along Namibia's coastline as well as the coastlines of Egypt and Sudan and throughout the Sahara Desert (unit 5) and extremely wet regions can be found on the coast of West Africa and eastern Madagascar (unit 7). Likewise dry highland regions with high slope can be found in South Africa, the East African Rift, Ethiopian Highlands and in the Atlas Mountains (unit 4) and flat regions with thick soil profiles can be found throughout the Sahel, South Sudan and the Kalahari basin (unit 13). In contrast, we also find Recharge Landscape Units which appear to represent unique and spatially concentrated areas, such as the Congo Basin Rainforest (units 9 and 14), as well as regions where properties appear more diverse with multiple units appearing within smaller areas, such as Madagascar and Ethiopia.
Figure 4. Map of the 15 Recharge Landscape Units of our classification for a priori understanding of recharge controls in Africa. We group the Recharge Landscape Units into broader groups of similar units which we call Recharge Landscapes.
We found that grouping Recharge Landscape Units into broader Recharge Landscapes suitably organises the African landscape into regions with noticeably different distributions of long-term average annual recharge and recharge ratio (Figure 6). Further disaggregation of the landscape did not allow us to explain any further variability in the ground-based estimates of recharge signatures when using global datasets. These broader Recharge Landscapes also aggregate Recharge Landscape Units with similar recharge control indices, as shown by the boxplots in Figure 5. For each index, boxplots are organized by the median values of each unit, ordered from left to right in descending order. In Dryland and Wet tropical Recharge Landscapes, we see that climate and weather, landcover and soil texture indices transition smoothly across all units. Units within Wet tropical forest Recharge Landscape are typically associated to high Kveg and Leaf Area index values and fine soil textures, whilst units of the Desert Recharge Landscape have low Kveg and Leaf Area values as well as predominantly sandy soils. Similarly, most units have similar topographic slopes except for unit 3, 4 and 13 which represent highland and flat plain regions. There is a clear divide in the depth of soils in each of the units, with six of the units showing deeper soil profiles and 8 showing a tendency towards shallow soils. We can see that unit 15 which represents karst regions occurs in a wide range of different climate, topographical, landcover and soil settings. Irrigated areas do not contribute to large areas of any of our Recharge Landscape Units.

Desert Recharge Landscapes could only be further differentiated by their depth to bedrock, while other landscape types were dis-aggregated by climate seasonality, slope, landcover and slope, as well as the depth to bedrock. Desert Recharge Landscape Units are differentiated according to where depth to bedrock is less than 13.5 m (unit 5), where the bedrock depth is between 13.5 m and 33.9 m (unit 6) and where the depth to bedrock is greater than 33.9 m (unit 3). This reflects differences in topography throughout Desert Recharge Landscapes, as mountainous Desert Recharge Landscapes with greater slopes also have smaller bedrock.
depths. Dryland Recharge Landscapes are also largely dis-aggregated according to the depth
to bedrock, with unit 13 representing where bedrock depth is greater than 37m, unit 10 where
bedrock depth is between 16.3m and 37m and units 11 and 12 where the bedrock depth is less
than 16m.

Figure 5. Boxplots showing the index values in each of the Recharge Landscape Units we identified. Boxplots are organised
from left to right in descending order of the median values in each unit. We show irrigated area as both the total area
irrigated within the Recharge Landscape Units (h) and as a percentage of the areas for each Recharge Landscape Unit (i).

Ground-based estimates of annual recharge (recharge ratio) are bias towards drier settings
with 20 (15), 66 (58), 28 (25) and 3 (3) data points in Desert, Dryland, Wet tropical and Wet
tropical forest Recharge Landscapes, respectively. Recharge Landscapes which have high
annual recharge rates also have higher recharge ratios suggesting that as well as being
generally wetter, they are more efficient in converting that rainfall into recharge (Figure 6).
The variability of ground-based annual recharge estimates within Landscapes is greatest in
wetter settings, as shown by standard deviations of 5.6 mm/year, 66.0 mm/year, 84.0
mm/year and 400.1 mm/year for Desert, Dryland, Wet tropical and Wet tropical forest
Recharge Landscapes, respectively. The standard deviation between the mean annual
recharge estimates of each Recharge Landscape is 217.2 mm/year in contrast to a standard deviation of 113.7 mm/year when looking across the whole dataset population. Similarly for ground-based estimates of recharge ratio, the standard deviations within Desert, Dryland, Wet tropical and Wet tropical forest Recharge Landscapes are, 0.033, 0.059, 0.070 and 0.092, respectively. Again, standard deviation between the mean recharge ratio estimates of each Recharge Landscape is greater than across the whole population, each being 0.079 and 0.072 respectively.

We also investigated the possible influence of the different groundwater recharge estimation methods to see whether this explained any of the variability in annual recharge and recharge ratio estimates within the individual spatial units (see supplemental information). However, in agreement with (MacDonald et al. 2021) we did not find a relationship between the estimation methods used and the recharge signatures. Additionally, one of the benefits of the database compiled by MacDonald et al. (2021) is that they provide uncertainty ranges for each of their ground-based estimates. Although figure 6 does not show these uncertainty ranges, we found that these uncertainty ranges were lowest for both annual recharge and recharge ratio in Desert Recharge Landscapes and were largest in Wet tropical and Wet tropical forest Recharge Landscapes (see supplemental). However, uncertainty ranges relative to the ground-based estimates were largest in Desert Recharge Landscapes and lowest in Wet tropical forest Recharge Landscapes. Below we discuss the larger Recharge Landscapes.
Figure 6. a) Map of ground-based estimate data points distributed across the Classification of recharge controls in Africa. Boxplots of the ground-based estimates of long-term mean annual recharge (b) and recharge ratio (c) found in each of the Recharge Landscape Units. No data points are located within Unit 9 and hence it is not shown. Only one data point is located within Unit 14. Unit 15 representing karst does not have its own boxplot. Instead, we have superimposed (red dots) these data points above the units which they would have otherwise been attributed to. For this plot we simply use the average ground-based estimates of annual recharge and recharge ratio at each data point and ignore the uncertainty ranges.

Desert (RLU 3, 5, 6)

Desert Recharge Landscapes are characterised by low moisture availability (P-PET), low vegetation cover (kveg) and very high sand content in its soils (Figure 5). These properties lead to the lowest annual recharge and recharge ratio estimates occurring in Africa, as 80% of annual recharge (recharge ratio) estimates in Desert Recharge Landscapes are below 5mm/year (4%). Low recharge ratios suggest that even when rain does fall, only a small fraction is converted to recharge, despite the sandy soils, owing to the very high potential evapotranspiration demand. In these regions evapotranspiration can draw on moisture from substantial depths which prohibits the downward flux of moisture towards to water table (Lehmann et al. 2019; Scanlon et al. 2006). We also find ground-based recharge estimates in Desert Recharge Landscapes show very little variability. Although we find marginally greater annual recharge rates and recharge ratios in unit 5, we cannot explain why, and differences may not be significant as there are only 20 data points across this region.
Dryland (10, 11, 12, 13)

About 51% of the 129 ground-based estimates are sited in Dryland Recharge Landscapes where water is generally only available for recharge seasonally (units 10, 11, 12 and 13). 70% of these sites have annual recharge rates between 3-30 mm/year and a further 18% of these sites have rates between 30-100 mm/year. Typically, in these regions less than 10% of rainfall is converted to recharge, with only 9 of the 58 sites recording higher recharge ratios.

In this Recharge Landscape, we find that long-term estimates of annual recharge vary according to mean annual precipitation, whereas recharge ratios are greater at sites with greater monthly variability in P-PET (Figure 7).

Figure 7. Boxplots showing how ground-based estimates of mean annual recharge (a, c) and recharge ratio (b, d) vary according to monthly variability of P-PET and Mean Annual Precipitation in Dryland Recharge Landscapes. Recharge signatures are binned according to percentiles (0-20; 20-40; 40-60; 60-80; 80-100) of the controlling variable. In the top left corner of each sub-plot, we show the spearman rank correlation and the p-value for testing the hypothesis of no correlation.
18 (26) out of 28 annual recharge estimates in the Wet tropical landscapes (units 1, 2, 7, 8) exceed 100mm/year (50mm/year). These sites are also the more efficient in converting rainfall to recharge with 56% (92%) of them having recharge ratios greater than 10% (5%). The wetter conditions as well as seasonal periods of heavy monsoon rain allows deeper drainage, despite increased partitioning of rainfall at the land surface by vegetation, steeper terrain, and less permeable soils. Most of the variability between and within Wet tropical landscape units is attributed to differences in annual and seasonal scale water excess (P-PET) and heavy rainfall events (P10).

Differences in annual recharge and recharge ratio estimates of units 1 (median annual recharge 115mm/year; median recharge ratio 9%) and 2 (median annual recharge 148mm/year; median recharge ratio 14%) could be attributed to greater LAI and Kveg properties in unit 2. However, when comparing the properties of the individual sites we do not find this relationship. Highland areas (unit 4) show a particularly large variability in the fraction of precipitation being converted to recharge. This perhaps reflects the high degree of variability we can expect in highland regions depending upon landscape positioning.

These areas are characterised by the highest vegetation cover (LAI) and moisture availability (P-PET). We only have three ground-based estimates of annual recharge and recharge ratio within this Recharge Landscape: two in unit 7 and one in 14. The highest annual recharge estimate in our database is located in unit 7, with 31% of rainfall being converted to recharge to allow a rate of 941 mm/yr. Referring to existing literature, we find that in addition to high annual precipitation rates (3050 mm/yr) extensive bedrock fracturing near the land surface enables rapid infiltration and recharge (Kamtchueng et al. 2015).
We do not find a clear pattern whereby the presence of karst at a site indicates higher annual recharge rates or recharge ratios than other sites within a similar setting (Figure 6). When investigating the individual studies, some studies reported karst despite not being identified as such by the global dataset, and vice versa. Within settings defined as karst by global datasets, annual recharge rates and recharge ratios increase with increasing annual scale P-PET (see supplemental information).

5 Discussion

5.1 Which regions of Africa show similar recharge controls when clustered using descriptors derived from global datasets?

We find 15 Recharge Landscape Units within which we expect recharge processes to be similar, according to our clustering result. Only 9 Recharge Landscape Units are needed to characterize over 80% of the continent’s land area. Although we initially wanted to investigate the potential membership of pixels to multiple Recharge Landscape Units, more than 80% of pixels have a membership greater than 0.8 to their primary unit (please see supplemental information). This is likely due to the high dimensionality of the dataset we are using in the cluster analysis. In light of this, and because we further aggregated our 14 (out of 15) Recharge Landscape Units into broader Recharge Landscapes (largely according to climate), we simply used the primary unit membership. The Recharge Landscapes we identify are Desert, Dryland, Wet tropical and Wet tropical forest, which account for 32.5%, 26.9%, 24.6% and 8.4% of Africa's land area respectively (total of 92.4%). An additional 7.25% of the continent’s land area is defined by its geology (i.e. karst) and can be found distributed across each of the four previously mentioned Recharge Landscapes (as we would expect according to previous studies, e.g. Hartmann et al., 2017). At the resolution of our
classification, climate indices have strong positive correlations with landcover indices (pearson correlation coefficient > 0.7). It is not surprising that our Recharge Landscapes strongly resemble previous climate classifications (Peel et al., 2007; Knoben et al., 2018), because climate is a dominant control on the long-term evolution of land surface and near surface landscape characteristics including topography (Chen et al. 2019), soils and vegetation (Pelletier et al. 2013). It is important to recognise that the classification of places may vary temporally (Aleman et al. 2020; Tierney et al. 2017), however as existing continental scale datasets for ground-based recharge estimates only provide long-term mean annual rates, we were not currently able to investigate the temporal variability of groundwater recharge and how this relates to changing landscape classification (Sawicz et al. 2014). Furthermore, we regard the classification as a tool for analysis rather than something unchanging in time.

Our Recharge Landscapes broadly resemble the ecozones in classifications by Olson et al. (2001) and Jasechko et al. (2014), which identify five and three different regions across Africa respectively. They are also similar to the five regions delineated by MacDonald et al. (2021) when using aridity classes to investigate the spatial variability of recharge across Africa. Unlike Olson et al. (2001) and Jasechko et al. (2014) we do not aggregate deserts and xeric shrublands, which we instead include in our Dryland Recharge Landscapes. Hence our Desert Recharge Landscapes more closely align with the hyper-arid regions delineated by MacDonald et al. (2021), whilst our Dryland Recharge Landscapes also align with their arid and semi-arid regions. By separating dry systems according to the occurrence of vegetation, we differentiate between regions where transpiration has a greater effect on recharge processes (Scott et al., 2006; Cavanaugh et al., 2011; Gebreyohannes et al., 2013). Consequently, we organise the Kalahari Desert as a Dryland, as it is affected by transpiration (Foster et al. 1982). Our Dryland Recharge Landscapes can be found throughout the desert,
shrubland and tropical biomes of classifications by Olson et al. (2001) and Jasechko et al. (2014). Thus, previous ecozone classifications may have delineated these regions too broadly. We also see that by identifying Dryland Recharge Landscapes with low slope and high bedrock depths (RLU 13), we identified a landscape unit where large seasonal wetlands are likely to occur (Olson et al. 2001). These wetlands include the Okavango delta, the Kafue and Barotse floodplains in Southern Africa; the Sudd Swamps in Eastern Africa; and the inland Niger delta, Hadejia-Nguru wetlands and wetlands of Southern Chad in the Sahel. Such wetlands can be significant sources of annually occurring focused groundwater recharge, given soil conditions do not restrict infiltration (Edmunds et al., 1999; Wolski et al., 2006). Unlike the classifications of Olson et al. (2001), Jasechko et al. (2014) and MacDonald et al. (2021), we further disaggregate Desert Recharge Landscapes according to depth to bedrock. In Desert Recharge Landscapes, shallow bedrock depths largely align with mountainous regions, which are often regarded as important recharge zones for current episodic recharge events (Gheith and Sultan 2002; Sultan et al. 2007) and more regular recharge events in previous paleoclimate periods (Sturchio et al. 2004). Our Wet tropical forest Recharge Landscapes largely align with the tropical and subtropical moist forests shown in Olson et al. (2001). Though further disaggregation into units identifies unique regions such as the Swamp forests of the Congo Basin and regions with extreme monsoonal rainfall in the Gulf of Guinea. In contrast, neither Jasechko et al. (2014) nor MacDonald et al. (2021) identify the forested regions of their tropical and humid classes, respectively.

5.2 How do regions with similar controls compare to ground-based recharge estimates?

In Africa, Recharge Landscapes with greater long-term mean annual recharge rates are also more efficient in converting precipitation to recharge, as shown by the higher long-term mean
recharge ratio estimates. We do not know whether this relationship is found across other
continents or regions as previous studies investigating the controls on ground-based recharge
estimates across large spatial scales assess the spatial variability of annual recharge rates only
(Moon et al., 2004; Mohan et al., 2018; Moeck et al., 2020; MacDonald et al., 2021).
Investigating how recharge signatures interact in space allowed us to advance our
conceptualisations of recharge processes across Africa. Though comparative hydrology is
only just starting to be recognised by observational investigations within the groundwater
community (Haaf et al. 2020; Heudorfer et al. 2019), it is well established within the surface
water community (Addor et al. 2018; Sawicz et al. 2011, 2014) and has already been used in
global scale groundwater investigations using global scale modelling products (Cuthbert et
al., 2019a).
Even though we can explain the variability of ground-based estimates of annual recharge and
recharge ratio between different Recharge Landscapes, we have very limited ability to
explain why they vary within Recharge Landscapes using global datasets. Wet tropical and
Wet tropical forest Recharge Landscapes receive higher rates of annual recharge and are also
more efficient in converting precipitation to recharge than Dryland and Desert Recharge
Landscapes, as shown by the higher recharge ratio estimates in these places. This is not
surprising, as heavy seasonal, monthly and daily rainfall is already known to be important for
recharge processes in both tropical and dry regions of Africa (Döll and Fiedler 2008;
Jasechko and Taylor 2015; Owor et al. 2009; Taylor, Todd, et al. 2013). Furthermore, in
agreement with Taylor et al. (2013), we find that mean annual recharge ratios in Dryland
Recharge Landscapes, increase with monthly variability in P-PET. However, interactions
with other large-scale physical or biological indices offer little further explanation for why
ground-based estimates of annual recharge and recharge ratio vary within individual
Recharge Landscapes. For the most part, our inability to explain the spatial variability of
ground-based recharge estimates within Recharge Landscapes stresses the limitations of
global datasets for describing the complex interactions between landscape properties and how
they control more local recharge processes. It could also be attributed to the scale differences
in the resolution of our classification and the representative areas of each of our ground-based
estimates. Previous studies trying explain the spatial variability of recharge processes at
continental and global scales also mostly establish relationships with broad climate and eco-
hydrological patterns (Jasechko et al., 2014; Cuthbert et al., 2019b; MacDonald et al., 2021).
Furthermore, MacDonald et al. (2021) were also unable to explain the more regional/local
spatial variability in ground-based estimates of recharge using global datasets. More
specifically, they found that there are spatial correlations in long-term average recharge rates
across Africa up to distances of 900 km, which cannot yet be explained by environmental
properties. Ultimately, this suggests a gap between what we can learn from local insight and
from large scale regionalization, regarding the interaction of environmental properties and
their control over recharge processes. This potentially has wider implications as global-scale
models which are frequently used to estimate groundwater recharge at these scales (Döll and
Fiedler 2008; Wada et al. 2010), typically rely upon assimilating global datasets for climate
forcing, characterising the land surface and model parameterisations (Telteu et al. 2021).
Nonetheless, understanding which recharge controls are currently identifiable using these
methods could help guide the evaluation and future development of continental scale models
(Gleeson et al. 2021). For example, recharge ratio estimates across the Sahara from PCR-
GLOBWB are typically greater than 0.2 (Jasechko et al. 2014), whilst our analysis shows
they are mostly below 0.04.

5.3 Looking ahead

Given the limited explanatory power of global datasets as shown in our and other previous
studies, it is likely that continental and global scale modelling of groundwater recharge can
benefit from the implementation of landscape-based conceptualisations of recharge processes and controls (Gao et al. 2018). Hartmann et al. (2015) showed (for carbonate rock regions across Europe and Northern Africa) that even relatively simple process conceptualizations capture main differences in recharge dynamics between different large landscape groups. Such conceptual models characterize largely our prior understanding of groundwater recharge in different landscapes. This is likely to be particularly important in data sparse regions where we cannot reasonably rely upon model parameterisation schemes that rely heavily on the reliability of soils and other data (Wagener et al. 2021). Adding information through the definition of simple system conceptualizations, would enable us to further combine expected hydrologic behaviour of the landscape with widely available datasets (e.g. Cuthbert et al., 2019b). By focussing on regionally dominant recharge controls, we can develop more parsimonious mathematical models that are also more appropriate for the data scarcity found in many places (Sarrazin et al., 2018), or specific hydrologic processes of most relevance (Quichimbo et al. 2021).

Figure 8. Application of the recharge landscape classification framework to domains outside of the study region. We used a random forest to transfer our Recharge Landscape Units across the rest of the world, with the previously discussed recharge control indices acting as predictor variables. The random forest model is an ensemble of 25 classification trees each with a maximum of 400 decision splits. The model was trained on data points in Africa which were randomly selected with replacement from a sub-sample of 70% of the Africa data points (‘in-bag’). Model testing on ‘out of bag’ data points found a misclassification rate of just 4%. Areas shown in white are significantly dissimilar to the study region. The criterion for this separation was having mean temperatures below 13.5°C or above 35.5°C and snow fractions above 0.1. We estimated snow fractions by using a simple temperature threshold. Precipitation on days with an average temperature below 1°C is regarded as entirely snowfall whereas it is entirely rainfall on days with an average temperature above 1°C (Berghuijs et al., 2014).
We use a global gridded dataset of daily temperature provided by the Climate Prediction Center, NOAA (NOAA/OAR/ESRL PSD). Further details are provided in the supplemental information.

The value of comparative hydrology in this context could lie in identifying regions of similarity beyond the direct study domain. As discussed here, specific studies with ground-based estimates of groundwater recharge are rare – certainly across Africa. Figure 8 shows how the classification approach introduced here would classify other regions of the world if applied globally. All areas shown in white are significantly dissimilar to our study domain and hence unsuitable for comparison. However, areas in colour map onto some areas in our direct study domain. For example, studies in karst regions (shown in red) might complement the rather sparse ground-based measurements available inside Africa, thus offering an opportunity to expand on existing datasets like that compiled by MacDonald et al. (2021).

6 Conclusions

We set out to study the variability of groundwater recharge across Africa through the use of a classification of groundwater recharge controls as landscape elements, utilising global datasets to characterize our a priori understanding following an extensive literature review. Our final classification consists of 15 recharge landscape units which are similar across the 11 indices we used to describe recharge controls across the continent. We aggregated these Recharge Landscape Units into four larger Recharge Landscapes, including Desert, Dryland, Wet tropical, and Wet tropical forest, which broadly agrees with classifications by Olson et al. (2001) and Jasechko et al. (2014). Karstic environments are treated separately, scattered across each of the Recharge Landscapes we have found.

A classification approach has allowed us to consolidate most of the findings from previous studies into a spatial representation of expected recharge controls across the African...
continent. Much of our previous understanding of recharge processes in Africa was point or
plot based, originating from the case studies which have assessed recharge processes and
controls throughout the region. We hypothesize that the small number of Recharge
Landscapes needed to characterize the broader recharge controls of the African landscape, is
explained by the dominance of climatic controls, likely connected with the co-evolution of
vegetation, soils, and topography. These Recharge Landscapes were useful in organising
ground-based estimates of annual recharge and recharge ratio. Yet, in exception of Dryland
Recharge Landscapes, we were not able to explain the variability of estimated recharge
signatures within each of the Recharge Landscapes using global datasets alone.

There is still considerable variability in ground-based estimates of groundwater recharge
which cannot yet be explained using global datasets. This result highlights the limits of using
global datasets to decipher the complex interactions of landscape properties in controlling
recharge processes. Nonetheless, future data-based modelling of groundwater recharge at
continental scales could be advanced by using methods which explore the relationships
between controls and recharge within regions of similarity, instead of across the entire
continent (MacDonald et al. 2021). Further advancement is also likely to come from the
development of system conceptualizations which allow us to add more information than that
embedded in global datasets (Wagener et al. 2021). This would lead to a convergence of top-
down strategies (such as classification) with other more bottom-up approaches like the one
taken by Cuthbert et al. (2019b). Further expanding the study domain using similarity
principles might offer a strategy for expanding existing strategies. Furthermore, considering
the co-evolution of multiple landscape properties could help further separate the
hydrologically relevant behaviour of different places (Troch et al. 2013), which in turn could
help the predictive ability of global datasets used in model parameterisations. Currently such
expected hydrologic behaviour (derived from literature reviews), is only considered through
the definition of appropriate predictor variables.

Finally, as meta-analysis databases become more common in continental and global scale
hydrological studies (Moek et al. 2020; Wang et al. 2020), we would like to stress the
importance of thorough quality assurance in their initial development. Our findings from
these studies depend upon strong underlying datasets and it is unlikely future studies will
assess the quality of these datasets when investigating or expanding upon them. For the same
reasons, the initial development of these databases should also ensure that additional meta-
information is comprehensive.

Acknowledgements

CW is funded as part of the WISE CDT under a grant from the Engineering and Physical
Sciences Research Council (EPSRC), grant EP/L016214/1. MOC gratefully acknowledges
funding for an Independent Research Fellowship from the UK Natural Environment Research
Council (NE/P017819/1).

References

from Isotope and Chloride Concentrations in North White Nile Rift, Sudan.” Hydrogeology
“Assessment of Groundwater Recharge Potential in a Typical Geological Transition Zone in
Abidela Hussein, Misbah, Habtamu Muche, Petra Schmitter, Prossie Nakawuka, Seifu A. Tilahun,
Degraded Soils in the (Sub) Humid Ethiopian Highlands.” Land 8(11):159. doi:
10.3390/land8110159.
Abotalib, Abotalib Z., Mohamed Sultan, and Racha Elkadiri. 2016. “Groundwater Processes in
Saharan Africa: Implications for Landscape Evolution in Arid Environments.” Earth-Science
Soil Physical Properties.” 13th International Soil Conservation Organisation Conference
(669):1–6.
Acosta, Verónica Torres, Taylor F. Schildgen, Brian A. Clarke, Dirk Scherler, Bodo Bookhagen,


Benito, Gerardo, Rick Rohde, Mary Seely, Christoph Külls, Ofer Dahan, Yehouda Enzel, Simon


Demlie, Molla, Stefan Wohnjich, Birhanu Gizaw, and Willibald Stichler. 2007. “Groundwater Recharge in the Akaki Catchment, Central Ethiopia: Evidence from Environmental Isotopes


Nijzink, Remko, Christopher Hutton, Iliax Pechliyanidis, René Capell, Berit Arheimer, Jim Freer, Dawei Han, Thorsten Wagener, Kevin McGuire, Hubert Savenije, and Markus Hrachowitz.


1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476

Introduction and a Standard Representation of 16 Global Water Models to Support
Intercomparison, Improvement, and Communication.” Geoscientific Model Development


Infiltration and Soil Water Content in Zambia and Zimbabwe.” Soil and Tillage Research

Relationships across Climatic and Soil Type Gradients.” Journal of Geophysical Research:

the Green Sahara.” Science Advances 3(1). doi: 10.1126/sciadv.1601503.

du Toit, G. van N., H. A. Snyman, and P. J. Malan. 2009. “Physical Impact of Grazing by Sheep on
Soil Parameters in the Nama Karoo Subshrub/Grass Rangeland of South Africa.” Journal of


Towett, Erick K., Keith D. Shepherd, Jerome E. Tondoh, Leigh A. Winowiecki, Tamene Lulseged,
Mercy Nyambura, Andrew Sila, Tor G. Vågen, and Georg Cadisch. 2015. “Total Elemental
Composition of Soils in Sub-Saharan Africa and Relationship with Soil Forming Factors.”

the Green Sahara.” Science Advances 3(1). doi: 10.1126/sciadv.1601503.

du Toit, G. van N., H. A. Snyman, and P. J. Malan. 2009. “Physical Impact of Grazing by Sheep on
Soil Parameters in the Nama Karoo Subshrub/Grass Rangeland of South Africa.” Journal of


Towett, Erick K., Keith D. Shepherd, Jerome E. Tondoh, Leigh A. Winowiecki, Tamene Lulseged,
Mercy Nyambura, Andrew Sila, Tor G. Vågen, and Georg Cadisch. 2015. “Total Elemental
Composition of Soils in Sub-Saharan Africa and Relationship with Soil Forming Factors.”

the Green Sahara.” Science Advances 3(1). doi: 10.1126/sciadv.1601503.

du Toit, G. van N., H. A. Snyman, and P. J. Malan. 2009. “Physical Impact of Grazing by Sheep on
Soil Parameters in the Nama Karoo Subshrub/Grass Rangeland of South Africa.” Journal of


Towett, Erick K., Keith D. Shepherd, Jerome E. Tondoh, Leigh A. Winowiecki, Tamene Lulseged,
Mercy Nyambura, Andrew Sila, Tor G. Vågen, and Georg Cadisch. 2015. “Total Elemental
Composition of Soils in Sub-Saharan Africa and Relationship with Soil Forming Factors.”

the Green Sahara.” Science Advances 3(1). doi: 10.1126/sciadv.1601503.

du Toit, G. van N., H. A. Snyman, and P. J. Malan. 2009. “Physical Impact of Grazing by Sheep on
Soil Parameters in the Nama Karoo Subshrub/Grass Rangeland of South Africa.” Journal of


Towett, Erick K., Keith D. Shepherd, Jerome E. Tondoh, Leigh A. Winowiecki, Tamene Lulseged,
Mercy Nyambura, Andrew Sila, Tor G. Vågen, and Georg Cadisch. 2015. “Total Elemental
Composition of Soils in Sub-Saharan Africa and Relationship with Soil Forming Factors.”

the Green Sahara.” Science Advances 3(1). doi: 10.1126/sciadv.1601503.


