DELFT UNIVERSITY OF TECHNOLOGY

MASTER THESIS

Is riparian vegetation an adequate proxy for water storage in sand rivers?

Author:
Luuk STREEFKERK

Supervisors:
Prof. Pieter VAN DER ZAAG,
Prof. Thom BOGAARD,
Dr. Claire MICHAILOVSKY

January 24, 2025

DELFT UNIVERSITY OF TECHNOLOGY

Abstract

Civil Engineering and Geosciences
Department Water Resources

Is riparian vegetation an adequate proxy for water storage in sand rivers?

by Luuk STREEFKERK

Southern Zimbabwe is an illustrative example where (increasing) water scarcity can lead to food- and financial insecurity and hinder socio-economic development. Sand rivers are a nature-based alternative to reservoirs, hosting shallow aquifers in their sandy beds with potential for decentralized and clean water abstraction. However, their extent and storage potential remain poorly understood, limiting awareness and investment. Riparian vegetation, which depends on sand rivers and remains green during the dry season, could serve as a valuable proxy for water storage.

This study uses a remote sensing approach to estimate and map the capacity of sand rivers to store water across three sub-catchments of the Limpopo River Basin (Shashani, Mzingwane, and Shashe), with two methods. Firstly, the sand river and riparian zones were delineated. The minimum storage capacity was estimated and mapped by summing WaPOR v3 evaporation data over the dry season. Secondly, 35 depth measurements were combined with spatial analyses to empirically predict depth and geometric storage capacity for the whole study area.

The results showed that there is a significant water storage potential in the channels. The two estimates can be combined for an unconsumed water availability, showing a significant sustainable potential (totaling 83 x10^6 m^3 potentially irrigating 8300ha). Dry season evaporation was consistently exceeded by geometric storage at medium to large evaporation rates, suggesting that these rivers unconsumed storage potential. The maps suggest that most potential water storage is located in the main river stems (the largest 10% hold 45-55% of water). Most tributares seemed to evaporate all water. Using water for irrigation would directly compete with vegetation. However, some local hotspots were still observed at smaller rivers.

Although significant uncertainties remain and field validation is still needed, the models show promise for wide-scale planning and development. The findings highlight that sand rivers, with rough estimates of their decentralized, cost-effective, and sustainable water storage, can be mapped remotely with minimal effort or field data. This could open up possibilities to offer support to farmers in semi-arid regions and pave the way for farmer-led irrigation initiatives.

Acknowledgements

First of all, I want to thank Julie, my girlfriend, for drawing the illustrations for the legends and the presentation, and for supporting me through some stressful (and boring) work sessions, as well as putting up with my endless dreaming about sand rivers.

I'd also like to thank my family — Rob, Sandra, and Ileen — for their emotional support, as well as their help with planning and content. Especially Ileen for her input on the content, Sandra for always being there emotionally, and Rob for his trust and constant positivity.

A big thank you to Pieter, my first supervisor, for his enthusiasm and for giving me the chance to work on such an exciting topic with a goal that felt meaningful to work towards.

I'm also really grateful to Thom for making this all possible by taking me on as his 100th graduation student with such flexibility.

Lastly, thanks to Claire for her quick, detailed, and consistent feedback — it's been incredibly helpful.

Contents

Introduction

1.1 Background

Water scarcity in dryland regions is becoming increasingly critical due to rising water demand, population growth, urbanization and climate change (e.g., Scanlon et al., 2006; Cuthbert et al., 2016). Water scarcity hinders the socio-economic development of semi-arid to arid regions, where smallholders cannot afford the losses caused by an unreliable water supply (A4store, 2024). An illustrative example is the Northern Limpopo River Basin in Zimbabwe, where water resource availability declined between 1980 and 2010, a trend that likely continues in terms of both total annual availability and frequency (Love et al., 2010). Prolonged droughts can lead to crop failure and livestock deaths (Blok et al., 2017). To increase the food- and economic security of farmers, and to rehabilitate adjacent land, water supply must become more reliable.

Currently, groundwater is the primary water source for socio-economic development in semi-arid regions (Saveca et al., 2022). Deep wells typically provide clean water during dry seasons. However, they can be costly, unsustainable, unreliable, yield low outputs, may be difficult to access or get contaminated (Hussey, 2007, Mpala et al., 2020). Dams and reservoirs can store large amounts of water, but they face high evaporation rates and can negatively affect the ecological environment and communities near them. In areas with hills or elevation changes, nature-based solutions such as rainwater harvesting can be particularly effective due to efficient collection of downhill runoff. This report focuses on a nature based solution in flatter areas, namely sand rivers, which have untapped, distributed potential for water storage and cheap abstraction (e.g. Moyce et al., 2006, Mpala et al., 2016).

Sand rivers are ephemeral rivers formed by sediment erosion, a process increased by upstream land use changes during the 21^{st} century. They host aquifers in their sandy beds, underlain by bedrock, protected from evaporation after a depth of around 0.7m. Sand rivers flow only in response to rainfall, typically recharging annually during floods (Mpala et al., 2020). Despite the surface being dry for most of the year, a sub-surface water storage is present, which can continuously flow downstream and supply clean water to well points (Love, 2013). The water is typically of good quality as it is naturally filtered by sand, although exceptions exist, such as salinity and contamination from animals (Hussey, 2007). Abstraction methods of the shallow aquifers are increasingly cheap and developed. A well-point with hand pump can cost less than $100 to irrigate at least 1 ha (A4store, 2024). The use of sand river aquifers is a form of private smallholder irrigation, which is acknowledged for its capacity to adapt to local circumstances (Duker et al., 2020b). The sense of ownership can intrinsically motivate farmers(Chauruka, 2022). Additionally, the fact that farmers can reset and learn from mistakes and every year, is a nice finishing touch of sand rivers as an option for water availability to address water scarcity, in particular for irrigation.

Sand aquifers are increasingly being used, as exemplified by the evaluation of sand dams in Kitui, Kenya, which are essentially artificially created sand rivers. According to Lasage et al., 2008, an estimated 100,000 people gained better access to water through these relatively low-cost sand dams between 1998 and 2008, and the average income of farmers near the dams increased by 60%. Also in Zimbabwe, NGO's like Dabane (A4store, 2024), are increasingly using sand rivers to provide water to stressed communities (Mpala et al., 2016). As of yet, still a large portion of sand river storage remains untapped (Moyce et al., 2006). Expensive, deep groundwater wells are being installed, when often sand rivers are in walking range.

The extent and capacity of sand rivers for irrigation, however, are not fully understood. The irrigation potential of sand rivers in Sub-Saharan Africa is only roughly estimated. It is 500 000 ha for 5 million people, when assuming 15% of Sub-Saharan Africa is criss-crossed with sand rivers and 5% of households live close enough to sand rivers (A4store, 2024). Other limitations to using sand rivers are the limited investment options for farmers to afford e.g. fences Chauruka, 2022, the cost of site selection due to long driving hours and the need to maintain abstraction systems.

Remote sensing offers an opportunity to map nature based water storage, promoting the potential of sand rivers' potential to funders and supporting irrigation planning. However, sand river depths are not directly visible from space, making storage volumes difficult to estimate. One previous study, Moyce et al. (2006), estimated the potential of a catchment based on sparse depth measurements and channel width. Kenyon (2022) analyzed spatial environmental patterns of sand rivers to gain insights on depth, but excluded vegetation. Pastora Estrada (2023) incorporated NDVI and canopy height to quantify transpiration in riparian vegetation using WaPOR, but coarse resolution limited accuracy. Recent advances such as autmatic sand river delineation by Bremer (2022), provide new opportunities to improve these estimates.

Sand rivers are not random, as they are shaped by climatic, topological, and geomorphological factors (Hussey, 2007). They exhibit distinct spatial patterns, with riparian vegetation being one of the most observable. The ability of vegetation to survive harsh dry seasons of up to 8 months highlights its reliance on water from the sand river alluvium, making it a potential proxy for water storage. However, the strength of this proxy remains unassessed. The extent and greenness of riparian vegetation vary with environmental factors such as climate, lithology, and topography, all of which - including vegetation - may indicate storage depth and size. Additional spatial patterns formed under these environmental factors could further clarify the relationship between riparian vegetation and water storage capacity.

This MSc thesis focuses on identifying and understanding the water storage capacity of sand rivers through its relationship with remotely sensed riparian vegetation and other spatial patterns in two sub-catchments of the Limpopo River basin in Zimbabwe; the Shashane and Mzingwane. The Shashe river was also included to have a variety of river widths. Actual WaPOR evaporation over the dry season result in map with conservative estimates of the water storage capacity. A second method resulted in estimates of the geometric storage capacity. Water quality is beyond the scope of this research; it is limited to water quantity. Understanding and mapping water storage capacity could help targeting potential irrigation sites, decrease the cost of site selection and encourage utilizing the potential of sand river aquifers for irrigation.

1.2 Objectives

The main objectives of this research are to:
1. Identify influences on the extent of riparian vegetation.
2. Map sand rivers.
3. Estimate water storage capacity (with two methods) to increase understanding of the total potential and distribution of water storage.

The mapping methods can support irrigation project planning by identifying suitable sites, reducing costs, and serving as promotional tools for sand rivers as nature-based solutions.

1.3 Research questions

The main research question of this study is as follows:

Can remotely observed spatial patterns like riparian vegetation serve as a proxy for the capacity of sand rivers to naturally store water, for irrigation and other human uses?

The research is subdivided into the following sub-questions:
1. What are key climatological, hydrological and geo-morphological factors that influence riparian vegetation?
2. How does riparian vegetation respond to rainfall patterns?
3. How can sand rivers and riparian vegetation be delineated?
4. How and can a minimum storage capacity be estimated with evaporation?
5. How can geometric water storage capacities be predicted with empirical relationships?
6. What can the estimates teach about total water availability and suitable locations for water abstraction?
7. What are the biggest uncertainties and can understanding them help us to reduce them?

1.4 Report structure

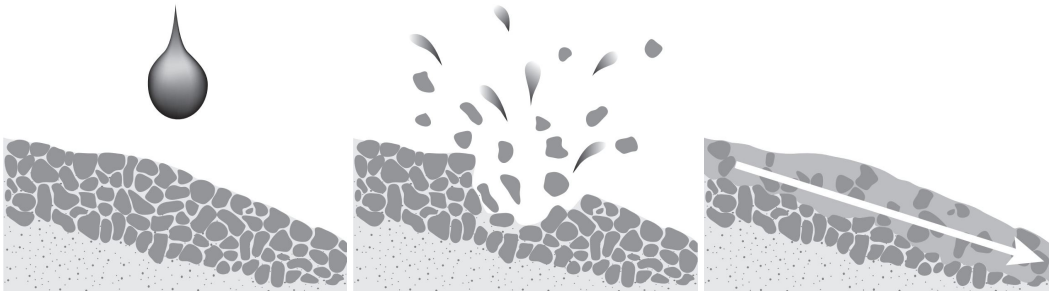
The report begins hereafter with a literature review explaining the fundamental concepts of sand rivers. The chapter Study Area presents relevant environmental characteristics. The methodology explains the set of data necessary to delineate sand rivers, perform statistical analysis and finally map water storage capacity. The results illustrate the behavior of riparian vegetation, statistical relationships between environmental features, and includes the maps with estimates of water storage capacity. In the discussion chapter, mainly the usefulness and limitations of the maps are discussed as well as broader implications. The report ends with a conclusion that answers the research questions.

Literature review

This chapter reviews the contextual and theoretical frameworks necessary to analyze the environmental factors influencing riparian vegetation and the water storage capacity of sand rivers.

2.1 Formation of Sand Rivers

Sand rivers are dynamic systems shaped by erosional and depositional forces, influenced by topography, lithology, sediment transport, and hydrological processes. Over time, erosion carves channels that fill with sediment, particularly in regions with high susceptibility to erosion. In arid environments, little rain and harsh sunshine can increase erodability in two ways. First is the killing of organic matter, leaving soil particles unbinded and without impact-absorbing protection layer against raindrops. Secondly is compaction of soil, which reduces infiltration capacity resulting in faster runoff (Hussey, 2007). Especially in steeper areas, soil particles can be loosened by the impact of raindrops (figure 2.1 b) and transported to the river channel by overland flow (figure 2.1 c) (Hussey, 2007). At higher rainfall intensities (40–70 mm/h), runoff discharge (/sheet flow) becomes the dominant driver of soil erosion, protecting the soil from direct raindrop impact (Vaezi et al., 2017).



(a) (b) (c)

FIGURE 2.1: Erosion process by raindrop on soil with low organic content, adapted from Hussey, 2007. (a) A raindrop is about to hit a dry, compacted, unbinded and unprotected surface (b) The raindrop impacts on the soil, it does not penetrate the soil but rebounds and dislodges soil particles (c) Overland flow transports loosened and saturated soil particles to a river

Sand deposition occurs in river channels (and floodplains) when the stream's transport capacity diminishes relatively to sediment supply, typically due to decreasing channel slope or flow velocity (Lane, 1955). Flow velocity diminishes further due to high permeability of sand and due to meandering, in turn increasing deposition (As cited by Kenyon, 2022; Nord, 1985). As slope decreases, first rocks and boulders remain, followed by deposition of gravel, then coarse sand, fine sand, silt and lastly clay. At a slope of around 1:250, sandy sediments begin to build up in the depressions of the riverbed and behind boulders or bedrock outcrops(Hussey,

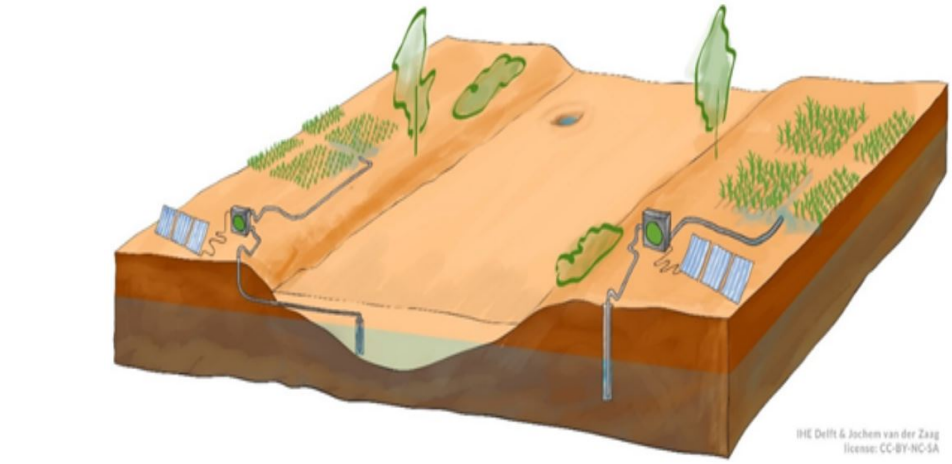
2007). Thick sand sediment accumulation in the Mzingwane sub-catchment typically occurs at gradients of 1:500 or gentler Owen, 1989. The study area's slopes average 1:400, with some locations reaching gradients as steep as 1:1000 (Moyce et al., 2006). Natural levees form near river edges where heavier sediments, such as sand and gravel, settle rapidly. In contrast, finer sediments like clay tend to accumulate in the lower floodplains.

As cited by Moyce et al. (2006), the distribution of sand river aquifers is influenced by several factors, including the river gradient, channel geometry, stream power fluctuations (linked to evaporation and infiltration losses), and sediment input rates due to erosion (**richards1982geomorphology**). Ephemeral rivers can meander depending on local slope and geology, forming point bars that are sometimes densely vegetated. Grain size distribution and hydraulic conductivity vary significantly with slope and sediment depth (Hussey, 2007). Seasonal sediment layers further stratify these deposits, creating non-uniform alluvial structures. According to Lane's balance (Lane, 1955), sediment thickness in sand rivers remains relatively stable over time if sediment supply and transport capacity are in equilibrium.

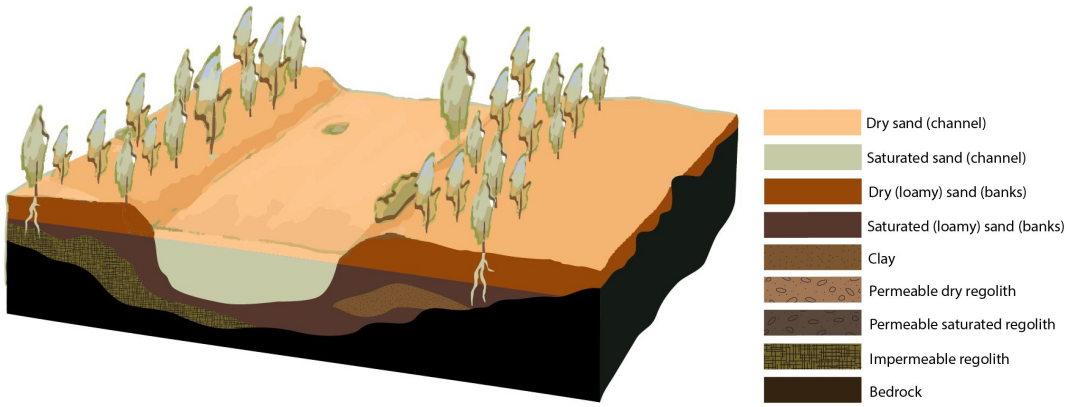
Human activities at the sand rivers can alter the formation and morphology of river channels, for example by clearing riparian vegetation and constructing dams. Removing vegetation exposes the soil, making it more susceptible to erosion, while dams affect discharge and thus sediment dynamics, often limiting sediment connectivity (Knight, 2022; Love et al., 2011).

2.2 Conceptual Cross Section and Abstraction Methods

Two conceptual models of sand rivers are illustrated in Figure 2.2. Water can be abstracted from the channel, which may have porosities of up to 50% (Nissen-Petersen, 1998), and occasionally from the banks, provided they are wide and sufficiently permeable (Figure 2.2(a)).



(a)



(b) (c)

FIGURE 2.2: Cross section examples of a sand river. (a) Adapted from Duker et al., 2020a, simplified cross section of an alluvial aquifer including well point systems on both the channel and banks. (b) Modified from Duker et al., 2020a simplified possible cross section showing a bowl-like configuration of the bedrock layer with deep channel and large RV extent. (c) The legend of the cross section (b)

The amount of abstractable water is determined by the aquifer's depth, width whether it is supplied by upstream subsurface flow and specific yield - the fraction of water drained by gravity - typically ranging between 0.10 and 0.20. The hydraulic conductivity in sand rivers is generally high, ranging between 25 and 250 m/d (Owen, 1994), allowing for high pumping rates. Aquifer depths vary from 0 to 40 m (Kenyon, 2022), with widths reaching up to 900 m. Even shallow aquifers (2 m) can suffice for irrigation purposes.

The banks and floodplains may be impermeable when composed of clay, confined by bedrock, or permeable when consisting of loamy sand. The permeability depends on variable factors such as bedrock configurations and clay layers, which are often local, unpredictable, and unobservable. This heterogeneity requires local investigations to determine suitable abstraction methods (e.g., Love, 2013). The thickness of underlying regolith is also heterogeneous. Vegetation can access water in regolith while well-points may not.

Figure 2.2(b) depicts one possible bedrock configuration. Over time, rivers erode and carve a certain shape out of bedrock, before its channel and banks are filled

with alluvial sediments. Although the shape of this bedrock varies unpredictably, one possibly reoccurring shape is the depicted bowl-shaped valley. Such steadily sloped configurations may allow substantial lateral water connectivity, benefiting vegetation, provided the upper layers are permeable. This supports an hypothesis of a correlation between lateral vegetation extent and channel thicknesses.

Water abstraction involves two main phases: filtering and lifting. One basic system is a well-point with suction pump (figure 2.2 a). Every system has a screen that controls and blocks sand particles. Screens for sand-abstraction come in the form of infiltration galleries, well-points or part of a caisson or well shaft (Hussey, 2007). Water is then brought to the surface using methods ranging from a simple bucket to a solar-powered mechanical pump. A well-point can be easily installed by digging, driving or jetting it into the sediment. A small well-point system of under $100 can be constructed in a day and the supplied water can be used to irrigate small agricultural plots up to around 1 ha (A4store, 2024).

2.3 Hydrology and Losses

The water availability, the extent riparian vegetation and its connectivity to the aquifer are influenced by various fluxes. While most researches has focused on abstraction systems and social integration, some studies have concentrated on understanding these fluxes through hydrological modeling (see fig.1), including transmission losses and recharge rates, seepage losses and evaporation losses (e.g. Blok et al., 2017; De Hamer et al., 2008; Love, 2013; Mansell et al., 2005,Mpala et al., 2020; Benito et al., 2010; Owen, 1994.
An alluvial aquifer can be described as a groundwater unit, generally unconfined, that is hosted in laterally discontinuous layers of sand, silt and clay, deposited by a river in a river channel, banks or flood plain (Love, 2013). Aquifer dimensions are determined by the extent and thickness of the alluvial fill in the river channel and, if connected, the lateral alluvial plains. Figure 2.3 schematizes the in- and outflows of aquifers, including recharge, downstream subsurface flow, and losses such as evaporation, vertical seepage, lateral seepage, and transpiration. These flows can be characterized as pressure redistributions in three time scales; longitudinal/ downstream flow (years to decades), transverse/lateral flow (weeks to months) and vertical flow (days to weeks) Cuthbert et al., 2016.

Recharge to the channel alluvium occurs during the early rainy season. High permeability of the sand allows little water to flow until the sediments are saturated (Moyce et al., 2006). Although the river channel may be fully replenished, water tables can lower quickly due to lateral leakage into the floodplains, especially if they are not also recharged (Hussey, 2007). Therefore, surface flows from small catchment areas are less likely to sufficiently recharge a river aquifer to maintain a year-round water supply. Recharge of aquifers of the lateral plains depends on the permeability of the aquifer, the distance from the channel and the duration of river flow (Owen, 1994). Alluvial aquifers can also be recharged by intermittent rainfall and during the dry season by dam releases.

Evaporation is a major flux but only occurs at the near-surface; the extinction depth lies between 0.6 and 0.9 m below surface (Mansell et al., 2005; Aerts et al., 2007).

Other fluxes differ unpredictably; vertical and lateral seepage depend on heterogeneous subsurface materials and their configurations. Vertical seepage, or bed

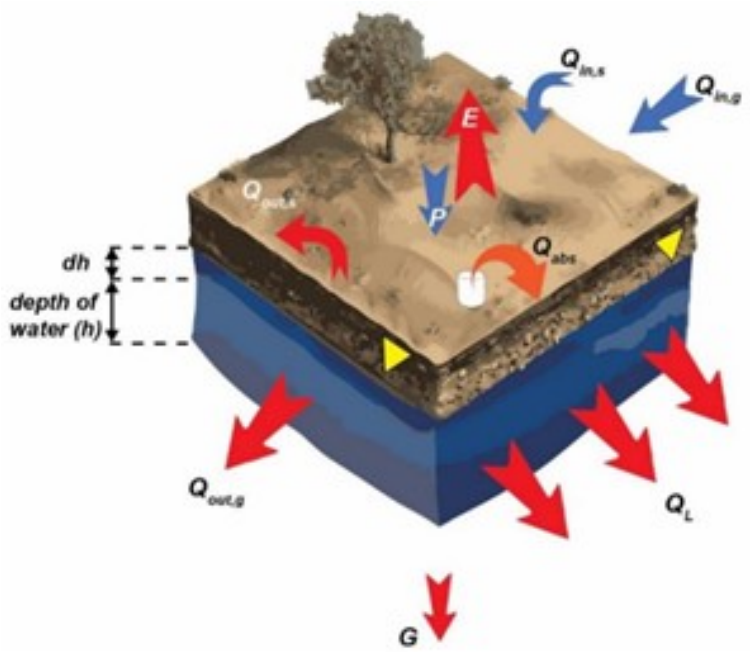


FIGURE 2.3: Idealized schematization of fluxes, adopted from Moulahoum ().

infiltration, is highly dependent on local geology, such as the permeability of sediments and the presence of confining layers, making it challenging to predict (see Section Geology and Soils). For instance, De Hamer et al. (2008) estimated that vertical seepage accounted for 86% of total losses, with the remaining 14% attributed to evapotranspiration. Lateral seepage, on the other hand, is influenced by factors such as clay content and the presence of heterogeneously configured rising bedrock and bedrock outcrops. Saveca et al. (2022) highlighted the significance of lateral groundwater flow, while other studies, such as Moulahoum, 2018 and Mansell et al., 2005, assume aquifers to be fully bounded, excluding lateral seepage from their models.

The horizontal, downstream flow can influence riparian vegetation to receive water from upstream and therefore influence spatial statistics if done on too small of a scale. However, as it depends on hydraulic conductivity and water gradient, it varies among studies, ranging from approximately 15-600m per 200 days dry season. It is acknowledged small by an extensive watershed-water balance coupled model Mpala et al. (2020) (85m3/d), less than 3m/d by (Owen, 2000), and 0.07 to 0.33m/d (Mansell et al., 2005). However, Love et al. (2010), finds that alluvium outflow in the studied river was of the same order of magnitude as average abstraction and is thus an important component of the water balance. Using average values of slope (0.29%), hydraulic conductivity (30 m/d) and a porosity (0.35), with Darcy's equation result in 0.23 m/d or 46 m of traveling distance over 200 days.

Outcrops and subcrops can create barriers to downstream subsurface flow, splitting the alluvium into compartments (e.g. Morin et al., 2009; Love et al., 2011; Hussey, 2007). This could quickly deplete a small alluvial aquifer unit when abstraction equipment is placed herein (Walker et al., 2018). Sub-surface flows are argued to pressure downstream water up against these barriers, so that it flows over the barrier further downstream, or raises the water level closer to the surface and increasing evaporation rates. The result is flat water levels in compartments, as shown

in figure (2.4).

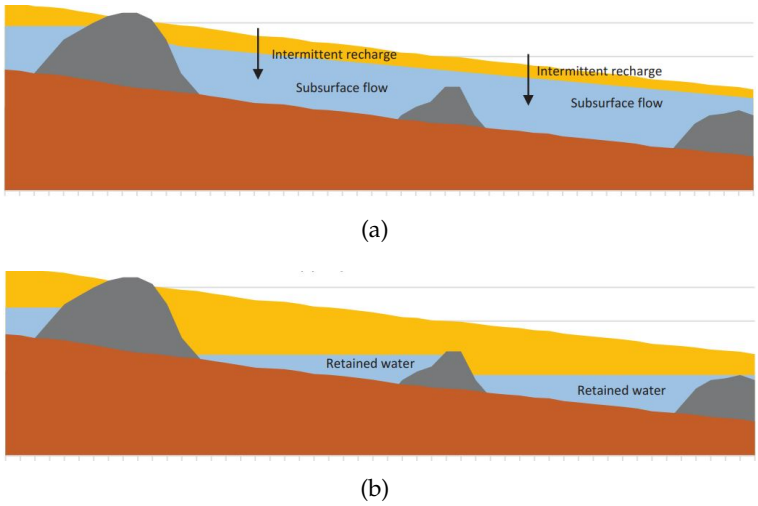


FIGURE 2.4: Adapted from Mpala et al., 2020. Longitudinal side view of the active channel of a sand river showing the influence of dikes and rock sills on the flow and evaporation, (a) a few weeks after the rainy season, and (b) during the dry season

Gaining or losing conditions can influence the water quantity in the channel and floodplains. While gaining conditions can occur when water levels in the channel are lowered by abstraction, rivers are (arguably) more commonly losing systems. This is because the sand channel fills quickly after surface flow, while the banks take longer to recharge (Hussey, 2007), and the surrounding terrain is often too flat for significant lateral inflow. Additionally, water losses from extensive floodplains — via transpiration and evaporation — can exceed those from the channel. Riparian vegetation contributes to these losses through deep tap roots and extensive lateral roots, while floodplains experience higher evaporation due to capillary rise in fine-grained sediments like silt and clay, which have greater adhesion and cohesion (Mpala et al., 2020). Despite these losses, studies often find that abstractable water remains at the end of the dry season (e.g. Blok et al., 2017, Moulahoum, 2018).

2.4 Geology and Soils

Geology and soils influence e.g. seepage, formation, and the thickness and extent of the alluvial channel and floodplains.
The formation of sand rivers is affected by weathering of lithology and erodability of soils, which govern sediment transport. For example, weathering of parent materials like sandstone can provide more sediment in the same semi-arid environment than granite (intrusive igneous, felsic rock) or gneiss (high-grade metamorphed granite), with basalt (extrusive volcanic, mafic rock) providing slightly more (Kenyon, 2022). Soil stability and erosion resistance depend on organic matter and vegetation cover (European Commission, Joint Research Centre, 2013). Soils which are very shallow, stony and without much water-holding clay (leptosols), or soils which are more acidic (acrisols) might host less organic matter than soils like luvisols (European Commission, Joint Research Centre, 2013). Even soils with clay content, such as lixisols, can lack organic matter and structure, making them prone to erosion under intense rainfall. Leaving the soils unprotected may result in crusts which prevents rain from entering the soil, resulting in overland flow eroding the topsoil, which is

the most fertile.
Soils are derived from their parent lithology through physical erosion (e.g. thermal expansion and contraction), chemical erosion (e.g. hydrolisis, oxidation, carbonisation) and biological erosion (e.g. root weathering, microbial erosion). Heating/cooling cycles are the main erosion type when bedrock is close to the surface (European Commission, Joint Research Centre, 2013). Resistant rock types can contribute to straighter, less meandering rivers (Twidale, 2004). Geologic structures such as faults, dikes or folds can, in addition to serving as preferential pathways for water flow, shape and align rivers. This influences sinuosity, as for example a fault can force a river in a bend due to sudden uplift or subsidence induced by groundwater erosion (...).
At geological boundaries or lithology transitions, alluvial fill can be augmented in width and thickness (Owen, 2000; Owen et al., 2005. When relatively soft rock is upstream of more resistant rock, a wide and shallow meandering river channel can occur (figure 2.5). Downstream of resistant rock, a waterfall and associated downstream plunge pool are buried and thus become focal points where alluvial sediment is thicker (figure 2.6) (Owen et al., 2005).

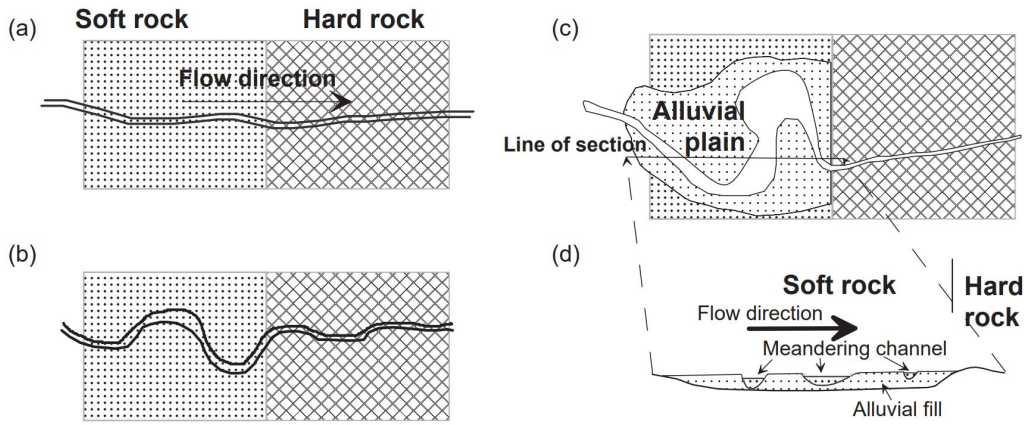


FIGURE 2.5: (Lithology transition with downstream resistant lithology, adapted from Owen et al., 2005. (a) The beginning of a narrow river channel (b) Meanders develop upstream of the transition (c) The meandering produces a wide and shallow valley which is eventually filled with alluvial sediments

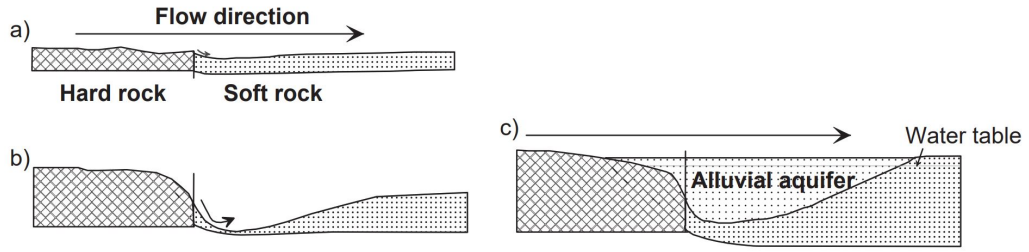


FIGURE 2.6: Lithology transition with upstream resistant lithology, adapted from Owen and Dahlin (2005). (a) A small depression forms in the softer rock with subsequent eddy currents further scouring out the softer rock (b) Further scouring may result in a waterfall (c) After erosional processes, sedimentation can result in a thick alluvial fill

In older terrains that are more deeply weathered, seepage can be a substantial

flux along boulders, fractures, and faults, creating significant groundwater loss pathways (De Hamer et al., 2008, Love et al., 2011). For example, granite and greenstone, typically impermeable, can develop secondary porosity through weathering and fracturing, creating preferential flow paths (Love, 2013, Blok et al., 2017). While younger crystalline bedrock has minimal seepage, rocks like limestone and coarse-grained sedimentary facies support high seepage losses (Moyce et al., 2006).

2.5 Quality of water

Water quality of sand rivers is generally high due to natural filtering processes in the sand and regular flushing of recharge and dam releases (Hussey, 2007, Moyce et al., 2006). It is typically considered safe for households uses and other uses like drinking water for livestock. However, contamination can be an exception, for example due to excessive fouling of animals near abstraction points.

2.6 Riparian vegetation

Definitions of riparian zones vary (Dufour et al., 2019; Zaimes et al., 2007). It is agreed that they are transitional zones between aquatic and terrestrial ecosystems, adjacent to a body of water and dependent on perennial or intermittent water. In this study, the riparian zone is defined as vegetation that remains green in the dry season. Temporal NDVI analysis, as demonstrated by Pace et al., 2021, provides insights into vegetation health, including indicators of water stress such as reduced leaf size, crown dieback, and limited root extension (Stella et al., 2013).

Riparian vegetation affects sand rivers by stabilizing soil and reducing erosion. Furthermore, deep tap roots plants draw alluvial groundwater, consuming approximately 1 meter over 200 days (5 mm/day) (Benito et al., 2010).

Vegetation growth is limited by the scarcest resource. This can be for example water availability, nutrient availability, soil structure, salinity, pests and light. In (semi-) arid regions like Zimbabwe, water is likely the scarcest resource, as observed in patterns of self-organization across the globe, caused by water-vegetation feedbacks (Rietkerk et al. (2002)). This assumption is supported by the little rainfall, long dry seasons and the evolution of vegetation to adapt to low nutrient levels European Commission, Joint Research Centre, 2013. However, nutrient bottlenecks, such as in Miombo woodlands, require further study (Frost, 1996). Especially if forests are cleared, the fragile nutrient cycle between e.g. the decay of litter, the roots of plants and a type of fungus called mycorrhiza can quickly be destroyed European Commission, Joint Research Centre, 2013.

Still, vegetation growth and its lateral extent depend greatly on water quantity and permeability of the banks and floodplains. As discussed, three observable factors influencing permeability are soil type, sinuosity and lithology type.

Study area

3.1 Topography and Climate

The spatial analysis will be performed on three rivers and some randomly selected tributaries in the limpopo river basin; the Shashani and Mzingwane sub-catchment in Zimbabwe, and part of the Shashe river in Botswana and Zimbabwe for wide river data points (fig. 3.1). These three rivers capture 35 usable depth measurements out of a total of 69 collected across 8 rivers researched in the Limpopo river basin (LRB), compiled by Kenyon (2022)), see appendix A. The tributaries of tributaries were cut off due to difficulties in calculating sinuosity and slope for short rivers, as well as time limitations.

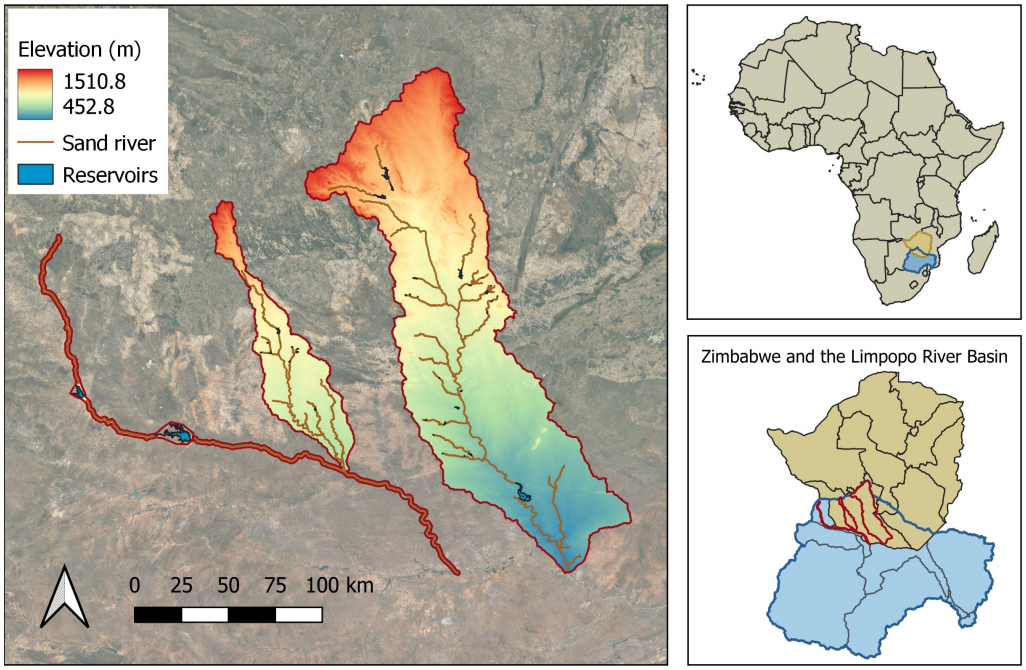
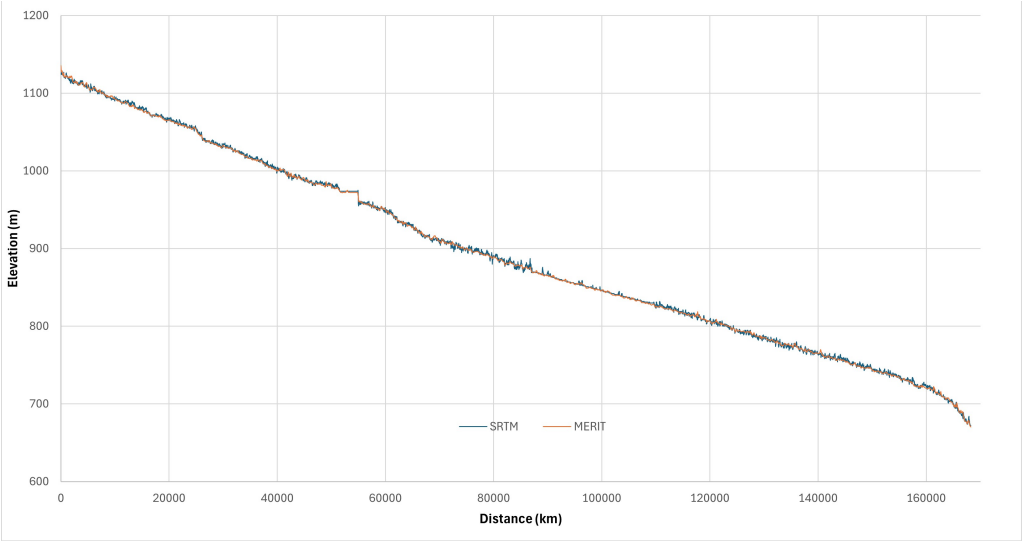
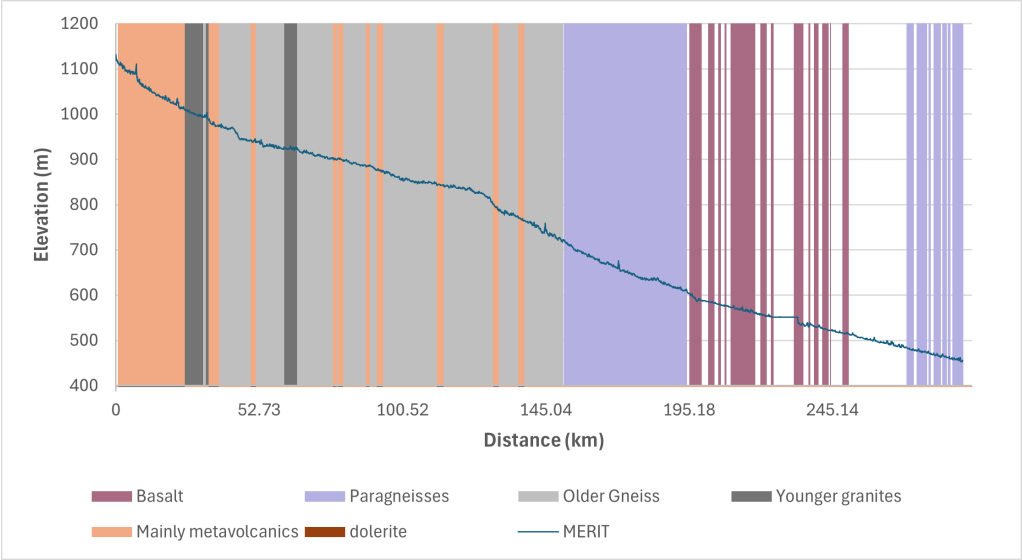


FIGURE 3.1: Study area. The Shashe river (west), Shashani river and some randomly selected tributaries (middle) and the Mzingwane river and some randomly selected tributaries (east)

As illustrated in Figure 3.2, the average slopes of the main stem of the Shashani and the Mzingwane are 0.00276 (1:362) and 0.00232 (1:431), respectively. While the Shashani has a relatively consistent slope, the Mzingwane shows more variety. For example, the last stretch of the Mzingwane from 230 km onwards has a gentler slope of 0.00159 (629). Large reservoirs can be spotted where the graph is flat. The variety of the Mzingwane is more clearly illustrated in Figure 3.2(b), where lithology types are added in the background. From high to low elevation, steeper slopes occur at the headwaters (underlain by mainly metavolcanics), followed by very gentle slopes (mainly older gneiss), a sudden steepening (metavolcanics and older gneiss), a medium steep slope (paragneisses) and finally a more gentle slope (basalt, sedimentary and paragneiss). Elevation ranges from around 1500m to 450m.



(a)



(b)

FIGURE 3.2: (a) Elevation profile of the Shashani river by SRTM and MERIT. (b) Elevation profile of the Mzingwane river by MERIT with georeferenced lithology types in the background. The lithology types that are blank are mostly alluvial deposits (gravel, sand and silts). Lithology data adapted from ()

The three rivers are ephemeral with flow generally restricted to periods of rain from November to March, though not continuous. Most flow was recorded between December and February, except where it has been controlled by dam operations. The Mzingwane river contributes 9.3% of the mean annual runoff of the Limpopo Basin, making it the third largest tributary to the Limpopo basin. The total sub-catchment area of the Shashani and Mzingwane are 2.837 km2 and 15.987 km2, respectively. The delineated main stems of the Shashe, Shashani and Mzingwane are 360 km, 166 km and 310 km long, respectively. The total length of the delineated rivers and streams is 1550 km. Widths of the main stems range widely from around 15m to 850m. Upstream areas contain hills with slopes of valleys up to around 12%. Further downstream, the terrain transitions into broad, flat-bottomed valleys Ashton et al., 2001.

Rainfall of Zimbabwe is highly erratic and varies between about 300 mm in a dry year and 800 in a wetter year, and generally falls only between November/December and February/March (Hussey, 2007; Rwasoka et al., 2011; love et al., 2006a). Intensities reach over 100 mm/h. This is particularly evident in the arid and semi-arid south and west. Like all semi-arid regions, potential evaporation rates greatly surpass precipitation rates. The driest duo years of the Mzingwane are 1982 and 1983 (sum of 509mm). The sum of Mzingwane's hydrological years 2022 and 2023 is (716mm).

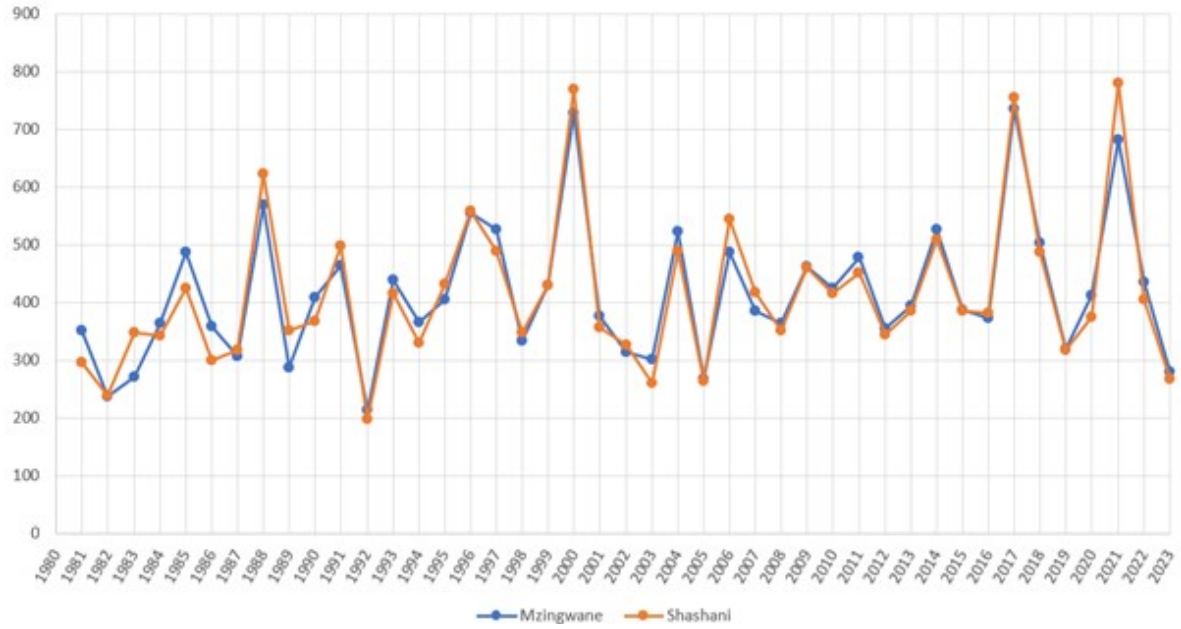


FIGURE 3.3: The sums of the hydrological years, summed over daily means of the Mzingwane and Shashani sub-catchments by CHIRPS. The hydrological year of for example 2023 starts 1 Oktober 2022 (start of the rainy season) and ends at 30 September 2023.

3.2 Geology and soils

The study area lies mostly in the Zimbabwe Craton and partly the Northern Marignal Zone (NMZ) of the Limpopo Mobile Belt (LMB) in the south, see Appendix (A.1). The study area's lithology consists mainly of a basement complex (Appendix A), that consists of granites and gneisses, with smaller regions of basalt and paragneiss in the lower Shashe and Mzingwane sub-catchment (see figure 3.4). The oldest gneisses in the Zimbabwon craton are dated at $\tilde{3}.5$ Ga (Wilson, 1990). The primary porosity of these lithology types - important as they can cause significant vertical seepage (**\cite {love2011water}**; De Hamer et al., 2008 - are very low. However, they are occasionally accompanied with secondary porosity like weathered rock and fractures, of which mostly faults and shear zones. Fracture lines occur mostly in the lower regions of the Mzingwane and Shashe.

Soil formations across the Limpopo basin reflect the underlying parent lithology, climate and biological activity (Ashton et al., 2001). The study area contains mostly shallow, poorly developed, well-drained soils, like leptosols (Anderson et al., 1993). Downstream of the Mzingwane, the leptosols become eutric, meaning a base saturation of 50% or more in the major part between 20 and 100cm depth (European Commission, Joint Research Centre, 2013). Mainly the depth and permeability of

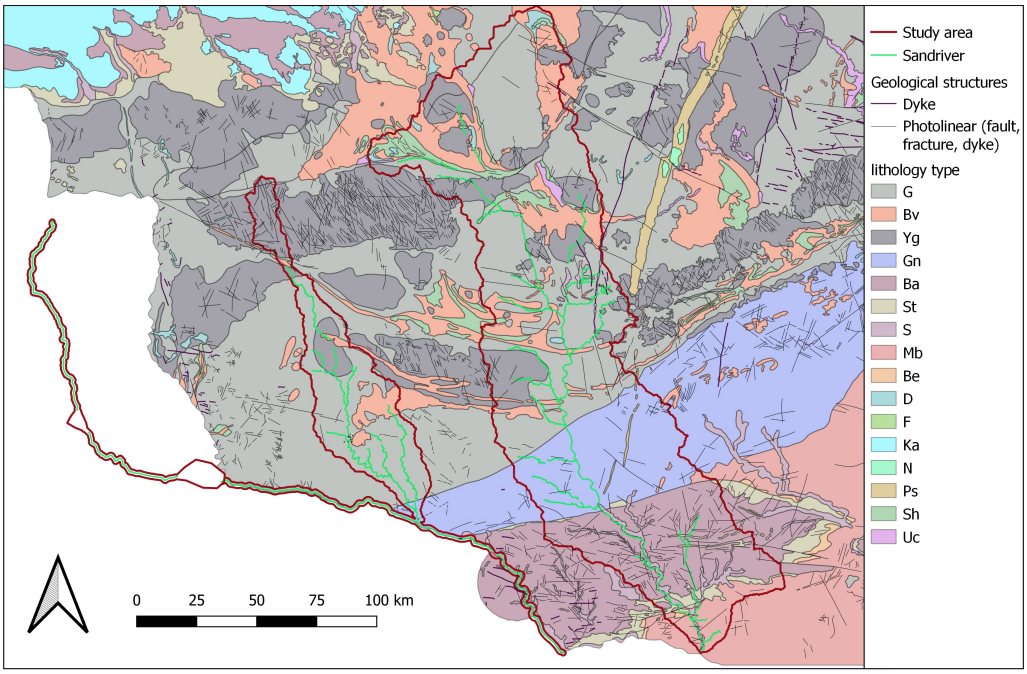


FIGURE 3.4: Georeferenced geology map, data adapted from IGRAC (International Groundwater Resource Assessment Centre). The study area mainly consists of gneiss (G), granite (Yg), mafic metavolcanics (Bv), paragneiss (Gn) and basalt (Ba)

the material that host riparian vegetation have an important influence on the aquifer volume and the extent of riparian vegetation.

Permeability of soils found in the study area are all considered relatively low, but LPeu lowest. Leptosols are shallow and often found on rocky or stony substrates, with limited soil depth. Lixisols (LXha) do not hold much organic matter and lack a well-developed structure, which could lead to moderate to low permeability. The clay content restricts water movement, while there is not much variation in texture to enhance permeability. Eutric Luvisols (LVeu) have a clay-enriched subsoil, but "Eutric" suggests a high base saturation, which can be associated with better aggregation and structure in the soil. This structure can allow for better permeability compared to the other soils, especially if the clay particles are well-flocculated and the soil has a good balance of sand, silt, and clay. Ferralic luvisols are also considered relatively permeable, but the "ferralic" qualifier indicates a high content of iron and aluminum oxides which can act as binding agent and cement to soil particles. This leads to lower permeability. Furthermore, compaction of any soil can lead to a reduction in pore space, reduced plant growth and loss of soil structure.

3.3 Riparian vegetation

The study area lies partly in the Miombo woodlands, which is a tropical bushland, wooded grassland and savanna biome of an estimated 2.7 million km2, dominated by the Brachystegia and Julbernardia tree species (Frost, 1996). The most occuring tree species in Zimbabwe, from most to least volume are Brachystegia spiciformis, Baikiaea plurijuga and Julbernardia globiflora (FRA zimbabwe, 2020). One characteristic feature of Miombo woodlands is their flush of new leaves 4-8 weeks before the first spring rains. The leaves of this flush are red, particularly those of Brachystegia spiciformis (Frost, 1996), affecting NDVI and transpiration estimates.

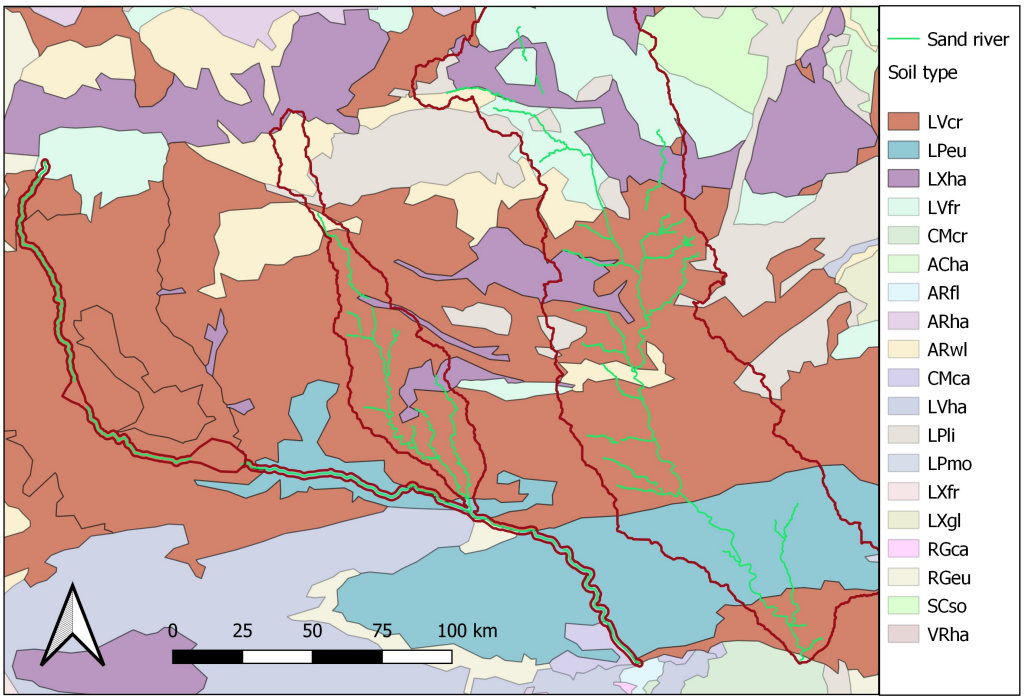


FIGURE 3.5: Soil map, data adapted from ESDAC (European Soil Data Centre). The study area mainly consists of Leptosols (chromic), Luvisols (eutric), Lixisols (haplic) and Leptosols (ferralic)

This coloration reaches a peak about 3 weeks later. Producing a functioning canopy before the rain enables the trees to begin production immediately once the rain arrives. Miombo species have horizontally and vertically extensive root systems. Frost (1996) mentions maximum recorded lateral distances of 27 m, and mentions that tap roots of these species can exceed 5m in deep soils. Roots are expected to penetrate substantially weathered regolith. Trees have evolved to preserve nutrients and water. Still, the fact that potential evaporation greatly exceeds precipitation means that dense vegetation plots have to retrieve water from sand rivers to survive.

3.4 Socio-economic situation / human activity

In 2002, around 375 thousand people lived in the Limpopo river basin part of Zimbabwe, out of the then 13.1 million (Ashton et al., 2001). The main foreign exports of Zimbabwe are minerals (gold, diamonds and platinum) and agriculture. The commercial agriculture sector initially provided exports and jobs. However, the sector has been damaged leading to Zimbabwe becoming a net importer of food products.

Because of the little and erratic rainfall, rain-fed irrigation is difficult in the southern part of Zimbabwe. Pastoral activities are predominant (Hussey, 2007) and more intensive in the communal areas than in the commercial areas. As a result of these hydrological conditions, groundwater provides the main water supply for most rural areas, while urban areas rely on large dams (Hussey, 2007). A total of 36 reservoirs were identified in close proximity to the delineated rivers. 25 of these are located in the Mzingwane sub-catchment.

Some of the sand river aquifers are being used to provide water for domestic use, livestock watering and dip tanks, commercial irrigation and market gardening. Agricultural plots are found further away from sand rivers as well as right next to

them. It is occasionally visible from satellite images that riparian vegetation has been cleared away and replaced by these plots.
The land cover cover map by the European Space Agency captures these crops very well, even small plots, as depicted in (appendix A). Approximately a third of all delineated 100m sand river segments (3581 out of 15676) lie nearby an agricultural plot (+-500m).

Methods

4.1 Overview and data requirements

In summary, two distinct sources of information were used to estimate the water storage capacity of delineated sand rivers, resulting in the production of two main maps. An overview of the methodology is illustrated in Figure **??**. The first map represents the minimum water storage capacity, derived from the accumulated evaporation of the dry season based on WaPOR v3 data. The second map contains estimates of the geometric volume of water stored within the channel, derived from depth measurements. Depth was predicted for the study area based on empirical relationships between environmental factors and the 35 suitable depth measurements. The accuracy of the resulting model was assessed with a train-test validation. The relationships were explored through a spatial, statistical analysis of 11 environmental attributes at the resolution of 100m river segments. In addition to predictions of depth, the analysis identified factors influencing riparian vegetation. Both maps required the delineation of sand rivers and their associated riparian zones. This was achieved using Sentinel-2 data. The two maps were combined by subtracting and dividing the two estimates for complementary insights about the unconsumed water availability.

Prior to these methods, the study area's environment was investigated with annual precipitation, elevation profiles and illustrations of bedrock configurations. Lastly, the seasonal behavior of vegetation and its surroundings were investigated through time series analysis of NDVI and transpiration data.

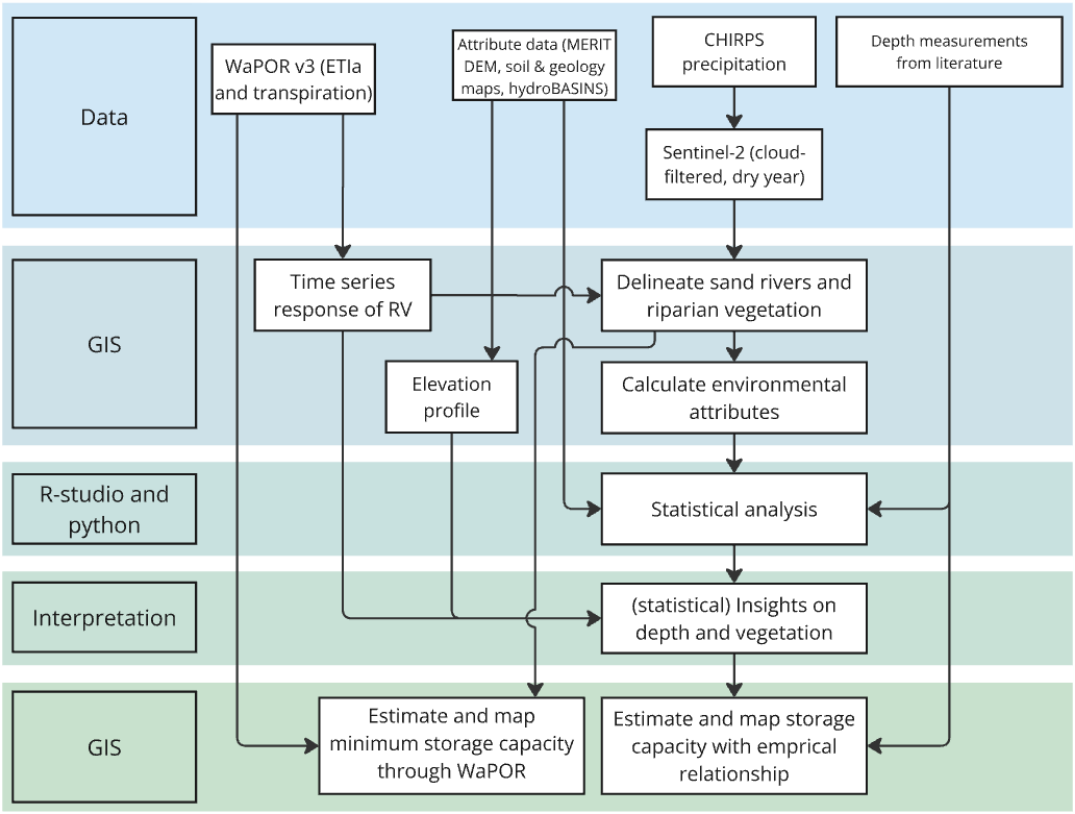


FIGURE 4.1: Flowchart of methodology

4.1.1 Data requirements

Firstly, CHIRPS precipitation data ((Climate Hazards Group InfraRed Precipitation with Station data) was downloaded mainly to choose the time period. It was downloaded from the Climate Engine (Climate Engine, 2023).

For delineation, NDESI (Normalized Difference Enhanced Sand Index) and NDVI (Normalized Difference Vegetation Index) were calculated with Sentinel-2 data in GEE. This specific sand index was chosen since sand and bare soil have very similar reflectance properties. NDESI was successfully used by Bremer (2022) and is argued to be able to discriminate sandy surfaces from soil and alluvium (Marzouki et al., 2022). For the time series analysis, cloud-filtered NDVI was downloaded from the Copernicus website. Canopy height was used as an alternative to compare with NVDI. It is a 10x10m deep-learning-derived product fused from Sentinel-2 data and sparse height data from the Global Ecosystem Dynamics Investigation LiDAR mission (Lang et al., 2023).

For estimation, WaPOR v3 data (Water Productivity through Open Access of Remotely Sensed Derived Data) was downloaded from the Food and Agriculture Organization (FAO). This was actual evaporation for the dry season evaporation estimates and transpiration for the time series analysis. The new WaPOR v3 has an increased resolution of 100x100m which is beneficial for the relatively small riparian zones, at dekadal time scale.

Channel depth was essential for the second estimate method, totaling 35 suitable measurements, compiled by (Kenyon, 2022) (see Appendix A). Depths were measured in the channel or floodplains with electrical resistivity or probing.

A DEM (Digital Elevation Model) by MERIT (Multi-Error-Removed-Improved-Terrain) (Yamazaki et al., 2023) was used to calculate the slope attribute and elevation profiles. Resolution and accuracy are important to calculate slope on smaller scale. MERIT was chosen because it incorporates other DEM's (NASA SRTM3 DEM, JAXA AW3D DEM, Viewfinder Panoramas DEM) and tries to eliminate their major error components.

A geology map provided the type of lithology and location of geological structures such as faults and dikes. The finest-resolution geology map found was georeferenced by the International Groundwater Resource Assessment Centre (IGRAC). Coarser geology was retrieved from the African Goundwater Atlas. The soil map was retrieved from the European Soil Data Centre (ESDAC).

4.2 Contextual investigation

This section consists of four types of preliminary investigation: precipitation, geohydrology, longitudinal elevation profiles and the time series analysis.

Daily CHIRPS catchment means were downloaded for both the Mzingwane and Shashani sub-catchment polygons. Precipitation was used to decide the driest year and check for homogeneously dry seasons, as local rains could disturb the vegetation analysis. Additionally, the driest year likely improves the delineation of sand rivers and riparian vegetation by reducing NDVI of non-riparian vegetation, highlighting riparian zones.

During the geohydrological investigation, relevant environmental factors and possible causes of false positives and negatives were identified with the help of visual inspection of google earth imagery and literature. A False positives is a large water storage estimate caused by large riparian zones (positive) while there is actually little water available for abstraction. False negatives occur when e.g. a small vegetated zone (negative) causes underestimations of water storage. These errors were illustrated as cross sections for intuition.

The longitudinal elevation profiles were created using QGIS. Manually edited OSM river centerlines were used to extract the MERIT DEM values.

The time-series analysis illustrates the seasonal behavior of riparian vegetation and its reliance on water from sand rivers. NDVI and transpiration were plotted of vegetation zones and its surroundings. Additionally, it illustrates the NDVI contrast of riparian zones for delineation. The data was downloaded for different plots (bare soil, riparian vegetation, farming plots, non-riparian tree cover). This was repeated in a second, spatially different region for comparison. Additionally, a temporal GIF was made with the Copernicus website to visually compare riparian vegetation with its surroundings.

4.3 Delineation of sand rivers and the riparian zone

4.3.1 Delineation of sand rivers

After choosing the year (2023) and period (Juli-September), sand rivers were delineated using sentinel-2 in GEE. Compared to the workflow of Bremer (2022) which aspires to delineate sand rivers fully-automatic, this study's method is simplified for feasible results within time-constraints. First, sentinel-2 data was downloaded for the selected dry period, and cloud filtered by taking the temporal median of cloud-filtered images. Consequently, the sand index NDESI was calculated to discriminate

sand surfaces. Two versions of NDESI were developed by Marzouki et al. (2022), of which only one is used. It uses the blue band (492 nm), the red band (664 nm), and the short wave infrared bands, SWIR1 (1613 nm) and SWIR2 (2202 nm). The NDESI equation of sentinel-2 can be written as:

$$NDESI = \frac{red(665nm) - blue(490nm)}{red(665nm) + blue(490nm)} - \frac{SWIR2(2190nm) - SWIR1(1610nm)}{SWIR2(2190nm) + SWIR1(1610nm)} \tag{4.1}$$

A static threshold was applied to filter non-sandy areas. The chosen threshold value was visually fine tuned. Only a small difference is present in spectral reflectance between sand and bare soil patches, causing a fine balance between high threshold values (resulting in underestimating sand rivers) and low threshold values (resulting in many bare soil patches). The delineation in the dry season was compared with the delineation in the rainy season.
Afterwards, the sand river was edited in Qgis. This consists of an automatic workflow model including the deletion of small bare soil areas and filling small sand river holes (see appendix B). At a handful of places where delineation failed, soil patches were manually deleted and sand river polygons were manually filled. These errors were caused by three phenomena; vegetation or mossy features in the sand river, flowing water, and outcrops. See appendix (B) for examples.

4.3.2 Delineation of the riparian zone

Similarly, a static threshold for NDVI was visually iterated towards in GEE, balancing between underestimation (patchiness) and overestimation (falsely including non-riparian vegetation). Frequency histograms were made of the distribution of NDVI and canopy height values. However, no optimal threshold was visible to discriminate between riparian zones (see appendix (B)). Furthermore, the performance of NDVI and Canopy Height were compared at a variety of sites to decide between of them.
To limit unrealistic lateral lengths of the riparian zone, occurring on few occasions in forested areas, the lateral riparian vegetation extent was limited. This limit varied with the channel width per 5 categories ("width classes"): 250m for rivers up to 20m wide, 250m for 20-40m, 650m for 40-100m, 650m for 100-300m and 1000m for 300-900m.

4.4 Spatial, statistical analysis

4.4.1 Calculating attributes

The delineated rivers were cut in 100m segments, and environmental variables were calculated and assigned per river segment. Every river segment forms a statistical point for the analysis. Only the most influential and observable environmental variables were chosen:

Attribute Name	**Why/ Influence on Riparian Vegetation or Water Storage Volume**	**How / Calculation Methodology**

Lateral Vegetation Extent	Indicates the water storage capacities as it draws water from it, in two manners: with the accumulated evaporation and with the empirical relation between vegetation extent and depth or storage volume. Additionally, vegetation can stabilize soils from erosion.	Determined by counting 10x10m delineated riparian zone pixels with 250-1000m wide "counting boxes", perpendicular to the river.
Channel width	Directly influences water storage volume, which is average depth times width. Width was also moderately correlated with depth.	Calculated as the average width of a 100m delineated river polygon segment, by dividing the area over 100m.
Slope	Controls water flow rate and sediment aggregation.	Derived from elevation data (MERIT DEM). Calculated as the change in average height over distance along the channel; 800 and 2000m. For a distance of 800m, an average height is calculated in a window of 2*400m, (forward and backward), resulting a range of 2*800m for each point.
Sinuosity	Reflects erosion and deposition patterns and slope. It can influence permeability, e.g. permeable sediment deposition at point bars.	Sinuosity is calculated as the channel length divided by the straight-line distance, for the largest observed meander wave within each of the 5 width-classes. The channel length is chosen to avoid being too long (which would lose spatial detail by averaging sinuosity) or too short (which would fail to capture full bends). For a more detailed workflow see appendix B.
Drainage Area	Influences sediment thickness. A larger drainage area increases the amount of sediment delivered to the main river by tributaries during rain and flow events. While the water-to-sediment ratio may not directly increase with drainage area, and sediment may be partially trapped in middle or upper reaches with very gentle slopes, some of it is still transported downstream during high-energy flow events. In lower reaches, this sediment can accumulate as the flow energy dissipates, especially where slopes are gentle.	Obtained from HYDROsheds polygons by HYDRObasins, calculated as the upstream area contributing to a specific point (in this case polygon) in the river or stream network.
Lithology Type	Controls soil composition and erosion rates, influencing the level of meandering and the shape of the bed of the channel and the floodplains. Rising- or heterogeneous bedrock can block water flow to riparian vegetation. Additionally, older lithology types increase the chance of leakage along fractured rock or faults.	Determined with geological maps in shapefile format.

Lithology Transition Presence	Influences erosion, meandering and sedimentation processes 2.6. Greater alluvial fill can occur upstream of a resistant rock, while thicker alluvial fill can occur downstream of resistant rock.	Mapped by buffering the boundaries of the lithologies of the georeferenced geological map (300m wide).
Faults and Dikes Presence	Influence water quantity by potentially acting as seepage lines. Values of leakage are reported especially significant at older lithologies.	Faults and dykes came as approximately 500m wide polygons of the georeferenced geological map.
Tributary Presence	Due to the method, riparian vegetation of tributaries was counted as if it belonged to the main river. Also, confluences can have increased riparian vegetation.	Tributaries were manually marked if they contributed towards counting vegetation extent of the main stem.
Logging / Woodland Removal Presence	Directly affects the amount of riparian vegetation.	Farming plots, roads and buildings were manually marked if they replaced the natural riparian zone, compared to the adjacent riparian zones. Could also be automated.
Reservoir Presence	Influence groundwater tables, trap sediment and modify the downstream water flow. Depending on the controlled release, the sediment-low water can increase downstream erosion.	Polygons were downloaded from Open Street Map. The length of marking the presence of reservoir, upstream and downstream, ranged from 1 km for small reservoirs to 10 km for the largest reservoir. .

TABLE 4.1: Table with environmental variables chosen for the analysis, including why and how they are calculated.

Table B.1 in Appendix B lists variables that were excluded and why.

Slope requires some additional explanation. For visualization, the safer 2×2000m slope was used, but the 2×800m slope was chosen for statistical analysis as it retained significantly more depth measurements. The 2×2000m slope excluded many measurements due to coarse and inaccurate DEMs requiring averaging, which could not be addressed within the available time.

For intuition, Figure (4.2) demonstrates how lateral vegetation extent was calculated in QGIS using "counting boxes". Some manual adjustments were necessary for meandering rivers, as they occasionally overlap with other river sections.

Many of the attributes were partly automated with QGIS models; see appendix (B.2.2) for the more detailed workflow.

FIGURE 4.2: The delineated sand river is cut in river sections (brown) and the counting boxes (yellow) count the number of pixels of riparian vegetation. The boxes are longer for wider width classes.

4.4.2 Statistical analysis

All river segments and their attributes were analyzed with statistical methods like (log-)linear regression and multiple linear regression (MLR).

The analysis was conducted at a 100m scale, balancing the following trade-offs. A smaller scale allows depth measurements to represent its reach better due to high variability, and to capture local phenomena like seepage or the impact of small point bars on vegetation. A bigger scale however, minimizes the influence of upstream sub-surface flows on vegetation and is less sensitive to autocorrelation and dependency.

A zoomed-in version of the scatter plots of depth versus riparian vegetation was created for depths up to 6 meters, as this is a common depth range for sand rivers and is approximately the limit of a basic suction pump.

4.5 Estimating water storage capacity

4.5.1 Dry season evaporation, a minimum storage capacity

The dry season evaporation estimates assume that all water evaporated by the riparian zone comes from the sand river's aquifer. Additionally, the evaporation from the sand in the channel is added to the estimated storage capacity. Below is a simplified workflow to create the map with dry season evaporation estimates. In summary, for every 100m river section in Qgis, the evaporated volume was calculated ($m^3/100m$) by multiplying the area of the riparian vegetation and the sand river with the averaged amount of meters evaporated by this combined area, over a 6 months dry season.

1)Download monthly raster images of WaPOR v3
2)Sum the monthly raster images to one dry season raster (april - october).
3)Filter out non-riparian and non-sand river pixels with the delineated RV zone and delineated sand river zone.
4)Calculate the mean of the summed evaporation of the sand river and the riparian zone of each 100m river section, with the help of the calculation boxes.
5)Calculate the evaporated volume by multiplying with the area.

Both the sand river area and its evaporation rates were included to ensure more complete evaporation estimates. Many 100×100m WaPOR pixels overlap riparian vegetation, sand rivers, and non-riparian areas. While pure sand river pixels have evaporation of around 56mm (April to October), riparian vegetation pixels have 400–500mm of evaporation. Excluding sand river areas from overlapping pixels would underestimate evaporation, especially in regions with small riparian zones (see example in Appendix B).

4.5.2 Water Storage capacity, Calculated with Empirical Relationships

The empirical relations that emerged from the statistical analysis were summarized in a correlation matrix plot. The most predictive environmental predictor variables for depth were combined in multiple linear regression models. The models included only a maximum of three predictor variables and one intuitive interaction factor to avoid overfitting, reduce artifacts and ensure the model remains interpretable. The interaction factor is the width times either evaporation or the extent of vegetation. The models first predict depth which is converted to storage afterward. A multiple linear regression model with three predictor variables without interaction terms looks like the following:

$$Depth = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + \epsilon \tag{4.2}$$

Here, X1, X2 and X3 represent the calculated predictor variables, while the β coefficients are parameters estimated by minimizing the sum of squared errors, between predicted and measured depth. The depth was then converted to (abstractable) water storage capacity, using width and a measured average specific yield of 0.15.

$$\text{Water storage capacity} = \text{depth} \times \text{width} \times \text{length} \times \text{specific yield} \tag{4.3}$$

To find the added performance per variable, the combinations were added in a table in the order from most to least predictive (drainage area, evaporation, RV extent, width).

All storage predictions were validated with repeated k-fold cross validation, explained below. The slope was excluded because the current calculations had gaps and did not significantly improve models. The final water storage capacity was displayed as $[m^3/100m]$.

Validation

The depth predictions were evaluated using train-test validation. Specifically, using repeated k-fold cross-validation. The chosen performance metrics were the coefficient of determination (R^2) and the mean absolute percentage error (MAPE). The

coefficient of determination was calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

And the MAPE was calculated as follows:

$$\text{Mean Absolute Percentage Error} = \frac{|\text{Actual Value} - \text{Predicted Value}|}{|\text{Actual Value}| * 100\%}$$

This relative metric helps with weighing errors across the different storage sizes, helping to prevent large storage errors from disproportionately dominating the evaluation.

K-fold cross validation splits the data into k folds and evaluates the model's performance on different splits. The data was split into 5 folds, of which four were used to train and one was used to test. This process was repeated k=5 times, so that each fold was used once for testing. To reduce variance in performance estimation, this entire process was repeated 200 times to calculate R^2 and MAPE. K-fold cross-validation ensures that all data points are used for both training and testing, increasing data efficiency and reducing the potential bias introduced by random sampling. This process was repeated for the best model with the 10 extra depth measurements available nearby but outside the study area.

The residuals were plotted to visualize the difference in scales of errors and to check whether they scatter around zero.

4.5.3 Comparing estimates

The two types of estimates were compared by dividing and subtracting them with/from each other. In theory, this reveals the "unevaporated" or "unconsumed" water availability.

In addition, the costs were compared between using a reservoir and using sand rivers to store the study area's total water storage.

Results

The maps with estimations are preceded by the time-series analysis, delineation of sand river and vegetation spatial analysis.

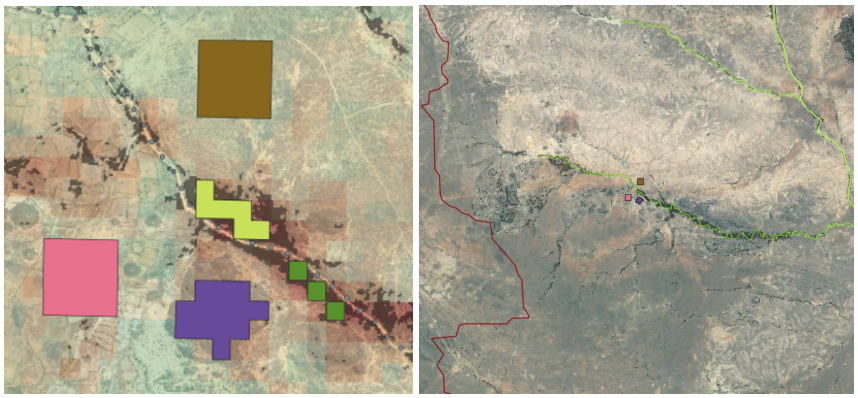
5.1 Time-series analysis on vegetation

This section presents the time-series analysis to investigate seasonal behavior. Figure 5.1 illustrated the NDVI and transpiration responses of a tributary in the Mzingwane sub-catchment. The site lies at the boundary between communal land (bare and farming plots) and commercial land (tree cover), as can be seen in Figure 5.1(b).

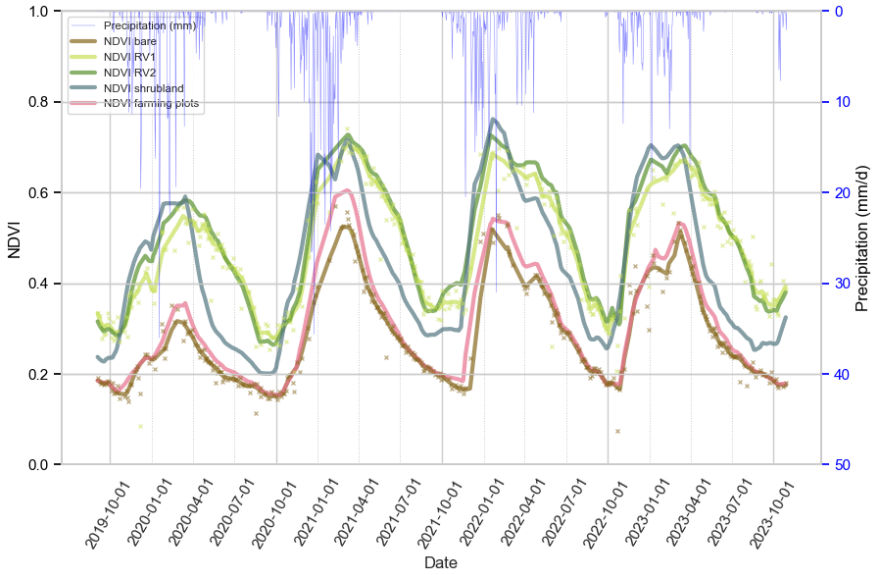
A few key observations can be made. In terms of seasonal dynamics, NDVI and transpiration roughly follow precipitation, going up in the rainy season and decrease in the dry season. Riparian (and non-riparian) tree cover maintain significantly higher NDVI values than bare soil and farming plots year-round. Although in the rainy season of 2021, the NDVI values of farming plots come closer than usual to those of riparian vegetation. Furthermore, the NDVI of non-riparian tree cover has similar values to riparian vegetation until the start of the dry season, when it starts to decrease quicker and a gap becomes visible. The difference in NDVI between riparian vegetation and its surroundings was more constant during the dry season. This information was used to choose the dry period for delineation and a corresponding threshold value of NDVI>0.42.

According to WaPOR, vegetation keeps transpiring even in the dry season, while bare soil and farming plots decrease to zero or almost zero. During the dry season, riparian vegetation plot 1 transpires more than tree cover.

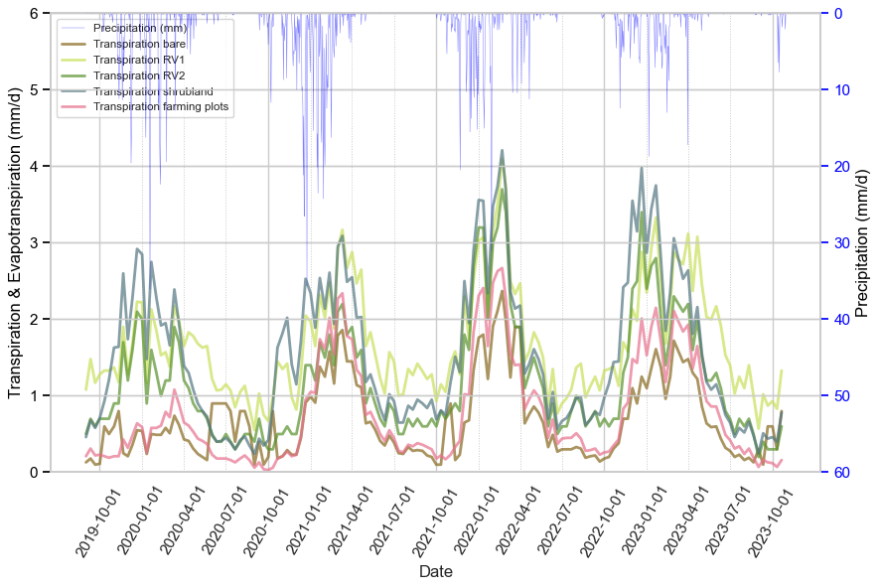
The site with a wider river in the Shashani sub-catchment can be found in appendix (C).



(a) (b)



(c)



(d)

FIGURE 5.1: Sentinel-2 NDVI and transpiration time series for five plots near a 20m wide tributary of the Mzingwane. (a) Five plot polygons with WaPOR transpiration pixels, color-coded to match plot lines. (b) Zoomed-out view of the site, between communal land (north) and commercial land (south). (c) NDVI time series (2020–2023) with daily precipitation, smoothed with LOESS, and raw data for bare and RV1 plots. (d) Transpiration time series for the same plots.

Snapshots of the temporal video of NDVI of riparian vegetation are shown in Figure (5.2. It highlights that NDVI decreases over the dry season but clearly remains greener than non-riparian surroundings all year round. For the GIF click the link in the caption.

(a) (b) (c)

FIGURE 5.2: Snapshots of a Sentinel-2 NDVI GIF of the Shashani in 2023 at (a) a few weeks after the start of the rainy season (November 20), (b) a peak in greenness (March 5), and (c) near the end of the dry season (October 11). Video link: **https://streamable.com/iutrdz**.

5.2 Delineation of sand rivers and riparian zone

The delineation of sand rivers and their adjacent riparian zones lay the foundation for understanding water storage capacity, the spatial extent of these features, and performing the spatial analysis.

Delineation examples are illustrated in Figures (5.4) and (5.3). For a large majority of the targeted sand rivers and riparian zones, a large percentage of their area was captured.

Especially in wider channels without bedrock outcrops, surface water or plant features, almost the whole channel area was captured (see figure (5.3).

(a) (b)

FIGURE 5.3: Example showing the delineation of a 370-430m wide sand river and its riparian zones. (a) Google satellite image. The tributaries and the associated confluences have increased extents of riparian vegetation. (b) Delineated pixels of the riparian zone by NDVI (>0.42), and a wide sand river which area is largely captured.

There are two occurrences where the delineation of sand rivers worked less well. The first is the disturbance of features with different colors. These are grey bedrock outcrops, blue water and green tree cover or features that look like moss or weed, illustrated in Figure (5.4). Other examples can be found in appendix (C). Secondly, delineation failed for rivers smaller than 30m. For channels between 30m and 40m, the delineation is very patchy. Tributaries with widths of 5-40m were therefore given shape by buffering the centerline with their (manually inspected) average width.

(a) (b)

FIGURE 5.4: Example showing the delineation of a sand rivers and the riparian zone. (a) Google satellite image of a sand river of the 100-300m width class, showing some surface flow and mossy features. (b) The delineated results of the sand river and riparian pixels. Water flows and mossy features cause a small amount of underestimation of the sand river area.

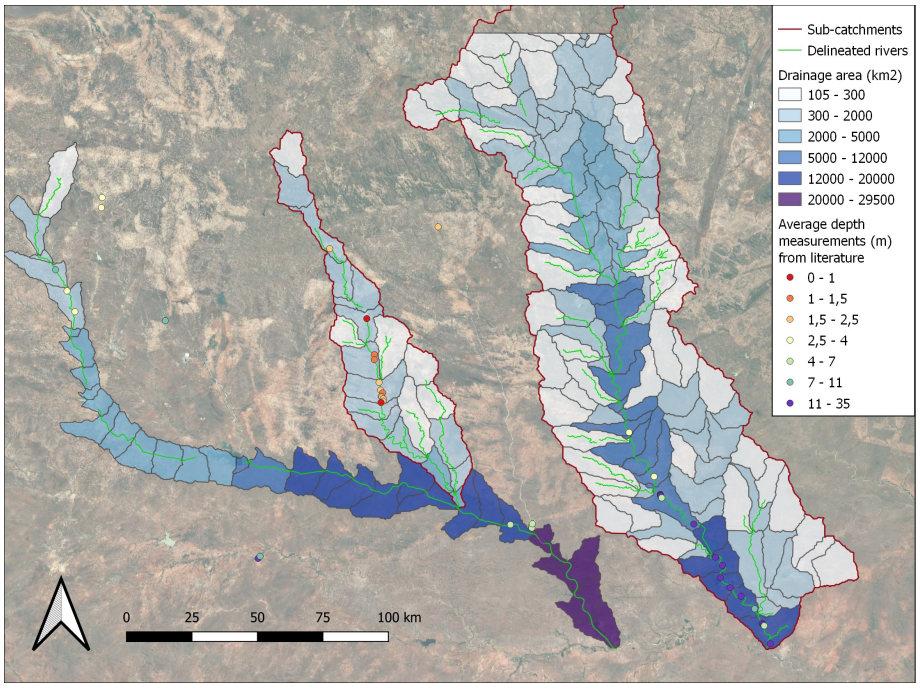
The delineation of riparian zones worked better for large zones and for places where its surroundings have browner, more bare-ish features like in communal lands. Smaller riparian zones are harder to delineate as they often appear patchy and with ambiguous boundaries; a challenge that persists also to the naked eye. The NDVI

threshold (>0.42) was able to capture smaller zones than the Canopy Height threshold (see Figure C.4 in Appendix (C). In commercial land, tree cover can sometimes be so dense that some pixels surpass the NDVI threshold. This can be seen a little bit in figure (5.4). Large vegetation zones often come with the occurrence of tributaries, particularly in the confluences, as can be seen in figure (5.3). It is not uncommon for tributaries to have greater riparian extents than main stems.

5.3 Spatial, statistical analysis

5.3.1 Calculating attributes

The values and resolution of the most dominant predictor variable, drainage area, were visualized for the study area in Figure (5.5). The total drainage area of the Shashani sub-catchment is smallest: 2834 km^2. It is followed by the Mzingwane sub-cachment with 15990 km^2. The Shashe river had the largest drainage area of 29000 km^2. Depth measurements were included in the figure. They range from 0.9m in the middle of the (small) Shashani to 35 meters in the lower (large) Mzingwane sub-catchments.



(a)

FIGURE 5.5: Visualisation of the data used for the drainage area attribute (km^2), showing the lvl12 HydroBASINS polygons. The rivers' main stems are included in green. The legend's distribution is by equal count (quantile).

Other attributes, such as slope and sinuosity, are visualized in appendix C. Slope ranges from mostly 0.005 m/m (1:200) in the upstream hills/headwaters to mostly 0.0012 m/m (1:833) in the lower sections of the sub-catchments (see Figure C.5) in C. The gentlest slopes are present in the Shashe river and the lower reaches of the Mzingwane.

High sinuosity values typically match with the level of meandering. However, the sinuosity values match less well for river widths smaller than 25m.

5.3.2 Statistical analysis

Relationships with depth and geometric storage

The relationships between the observed environmental variables were summarized in a correlation matrix (Figure 5.6) and Table 5.1. Key relationships are further illustrated through scatter plots. They provided the basis for estimating water storage capacity and mapping, which is presented in section 5.4.

The most significant correlated observable variables with the depth are drainage area(r=0.66), dry season evaporation (r=0.64), riparian vegetation extent (r=0.59), slope (r=0.41, p=0.01) and width (r=0.31) illustrated in Figure 5.6. All p-values were below 0.05 except for links with sinuosity and slope. Depth explains a large percentage of the variation of the geometric storage volumes (r=0.96).

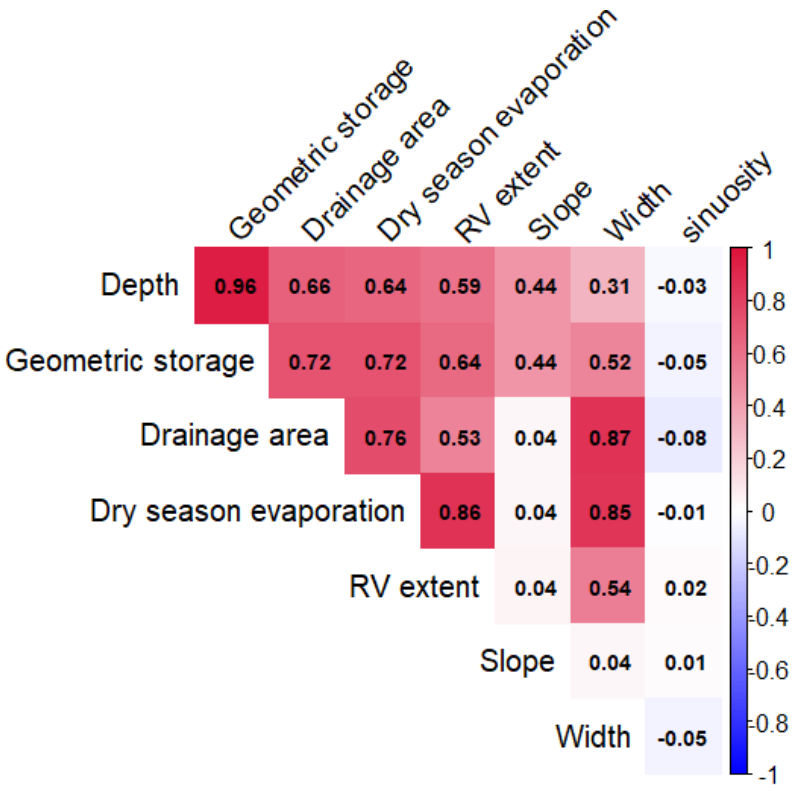


FIGURE 5.6: Pearson correlation coefficient matrix of the observed environmental variables. Only the links with sinuosity had p-values higher than 0.05.

Dry season evaporation is moderately correlated with measured geometric storage volume (r=0.72, assuming a specific yield of 0.15), shown in Figure 5.7. Almost all data points lie below the grey 1:1 line, where evaporated volume is smaller than the channel's geometric storage volume. As evaporation rates increase, geometric water storage increases even more. This results in a gap of potentially unconsumed water availability. The three exceptions have extensively vegetated point bars, further discussed in the discussion chapter.

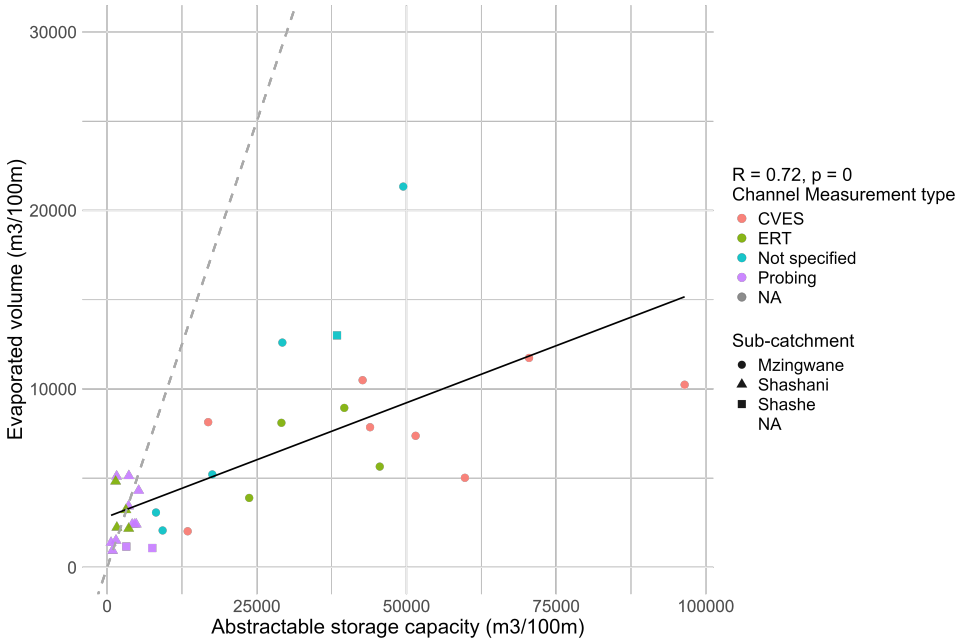
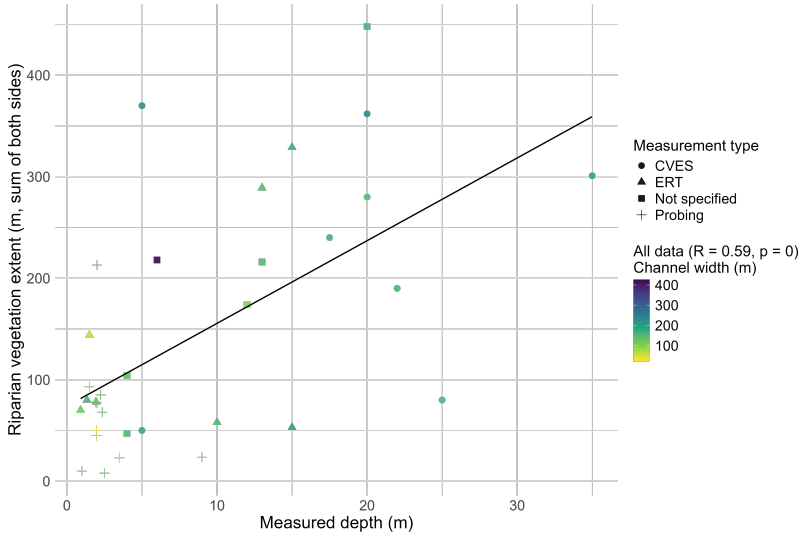


FIGURE 5.7: Scatter plot of dry season evaporation and geometric water storage capacity of the channel (assuming sy=0.15) derived from measured depths. The pearson correlation coefficient is 0.72.

Depth and riparian vegetation were moderately correlated, as shown in Figure 5.8, with an overall correlation of r=0.59. This decreases to r=0.44 at smaller rivers for depths up to 7m (see Appendix C. Nonetheless, some useful data gaps can be observed: channel segments with more than 210m of riparian vegetation are at least 2m deep, and those with more than 100m of riparian vegetation are at least 1.5m deep. The colored width classes indicate that the deepest rivers are consistently wide and extensively vegetated, while the shallowest rivers are small with relatively sparse vegetation (<100m).



(a)

FIGURE 5.8: Riparian vegetation versus measured channel depth, including width and measurement type. A moderate Pearson correlation coefficient is present (r=0.59).

Drainage area had the strongest relationship with depth, with a Pearson correlation coefficient of 0.66 (see Figure (5.9). The shown by the colored depths, all measurement locations with shallow depths (<2.4m) and small drainage areas lie in the Shashani sub-catchment, while the deeper depths (>4m, averaging 15m) with large drainage areas lie mostly in the Mzingwane. The depth measurements were incomplete; there were no depth measurements of the Mzingwane in upstream regions, and no measurements of the Shashani in the downstream regions.

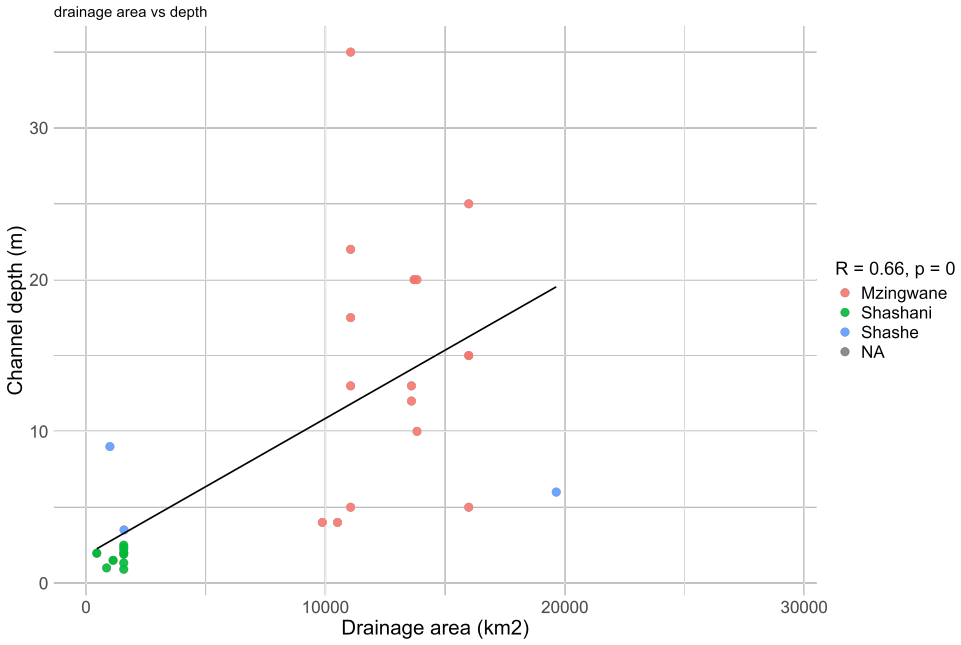


FIGURE 5.9: Scatter plot of depth versus drainage area, colored by sub-catchment.

Other interesting relations were found with depth, like slope (r=0.44, see Figure C.8 in Appendix C). While the correlelation coefficient is small, on average gentle slopes had deeper channels (17 m) and steeper slopes had shallower channels (3 m). However, this relationship was influenced by two notable exceptions: a shallow measurement on a gentle slope in the Shashani and a deep measurement on a steep slope in the Mzingwane. Visual inspection of steep slopes consistently revealed many bedrock outcrops while gentle slopes rarely had outcrops (see appendix C).

Furthermore, the abstractable storage capacity was moderately correlated with vegetation extent (r=0.64), illustrated in Figure C.11) in Appendix C. Coloring by lithology shows most small storages are underlain with older gneiss (Shashani). The storages underlain with basalt are all greater than 15000 $m^3/100m$.

See appendix (C) for more combinations of attribute coloring and multi linear regression plots.

Influences on riparian vegetation

Understanding the influences on riparian vegetation can help explain their variation and uncertainty of estimates.

In addition to water storage, correlated influences were width, drainage area and slope. Width was moderately correlated with riparian vegetation (r=0.54), shown in Figure (C.16) of appendix C. The predictive power does not seem large, but clear

data gaps are present at the upper left and lower right quadrant. The drainage area is positively correlated with riparian vegetation (r=0.44) and plotted in Figure (C.12) in Appendix C. Lastly slope was not significantly correlated but showed a higher average vegetation extent at gentle slopes.

The remaining influences on riparian vegetation are summarized in table 5.1, which compares various variables with the overall mean vegetation extent (76m). The largest observed influences on the extent of riparian vegetation were confluences/tributaries (299% of average), dikes (16%), Basalt (305%) and the presence of a reservoir upstream (156%). Especially the Zovhe dam - also underlain by Basalt - had increased vegetation upstream and downstream. As discussed in chapter discussion, collinearity should be kept in mind; other factors than reservoirs or Basalt could be the actual cause of increase in riparian vegetation.

Variables that showed less impact were sinuosity, soil type and finer-scale georeferenced lithology. Human activity/logging had a small but frequently occurring influence. For bar plots and illustrations of observed influences, see Appendix C.

TABLE 5.1: Mean Riparian Vegetation Extent (of both sides together) Compared to the Overall Mean for Each Environmental Variable

Environmental Variable	**Mean Riparian Vegetation Extent (m)**	**Extent as Percentage of Overall Mean (%)**
Overall Mean Riparian Vegetation Extent	**76 m**	**100%**
Tributary / confluence	227 m	299%
Human activity / logging	68 m	90%
Down of Reservoir	116 m	156%
Upstream of Reservoir	89 m	119%
Dike	12 m	16%
Fault	51 m	67%
Lithology Transition	102 m	135%
Basalt	232 m	305%
Basement Complex	54 m	71%
Point Bars (Sinuosity > 1.5)	67 m	88%

5.4 Estimating and mapping storage capacity

5.4.1 Method 1: dry season evaporation

The final map with dry season evaporation estimates is shown in Figure (5.10). The total estimated water storage capacity within the delineated rivers is 49.1 × 10^6 m^3. The values range from approximately 0 to 65000 m^3/100m. This could irrigate 6.5 ha/100m if all water would be abstracted. These units were converted to irrigation potential by assuming a crop needs 500mm over a dry season with 50% irrigation efficiency, resulting in a water requirement of 10000 m^3 per hectare.

Most tributary segments have a smaller storage capacity below 2500 m^3/100m (potentially irrigating 0.25 ha/100m). The lower reaches of the Shashe and the main stem of the Mzingwane rivers contain between 5000 and 50000 m^3/100m. As discussed in the chapter (**??**), this conservative estimate does not equal the amount of water suitable for abstraction when the needs of riparian vegetation are considered.

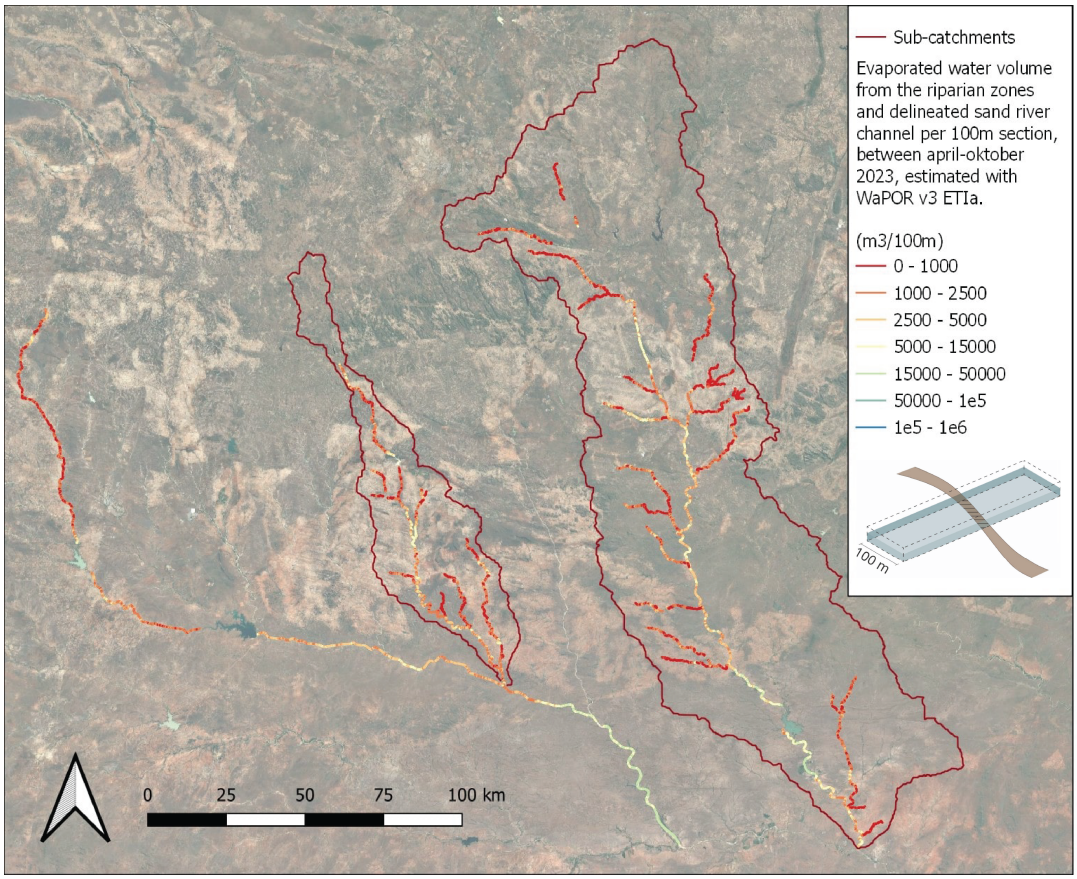


FIGURE 5.10: Map of sand rivers and the amount of volume (m^3) the riparian vegetation evaporated in the dry season (april-oktober) 2023 of each 100m river section in the study area, as estimated by WaPOR v3. Assuming all riparian vegetation retrieve their water from the sand river's aquifer, this is the minimum water storage capacity.

5.4.2 Method 2: geometric water storage

Going beyond this conservative estimate, the geometric water storage capacities were modeled and mapped. The models with different combinations of predictor variables were compared in table (**??**).

Models that included drainage area performed consistently better, with relatively little difference in performance when other predictors were added. However, evaporation stood out slightly among the options.

Riparian vegetation extent and evaporated volume alone can provide estimates of water storage capacity (as shown in Figure C.11 in Appendix C). However, when including additional variables such as upstream area, width, and slope in a multiple linear regression model, the contribution of vegetation extent diminishes, as shown in statistic model summaries in Appendix C.

The best-performing combination, with a relatively limited amount of predictor variables, was Model 8. It included drainage area, evaporation, width and the interaction between evaporation and width, achieving an R^2 of 0.64 and a MAPE of 61%. This model was chosen to compute the map of water storage capacity. The final equation that predicts the depth goes as follows:

$$\begin{aligned} \text{Depth} = 0.8205 + 0.000796 \cdot X_1 - 0.026 \cdot X_2 + 0.0011 \cdot X_3 \\ -0.00000224 \cdot (X_3 \cdot X_2) \end{aligned} \tag{5.1}$$

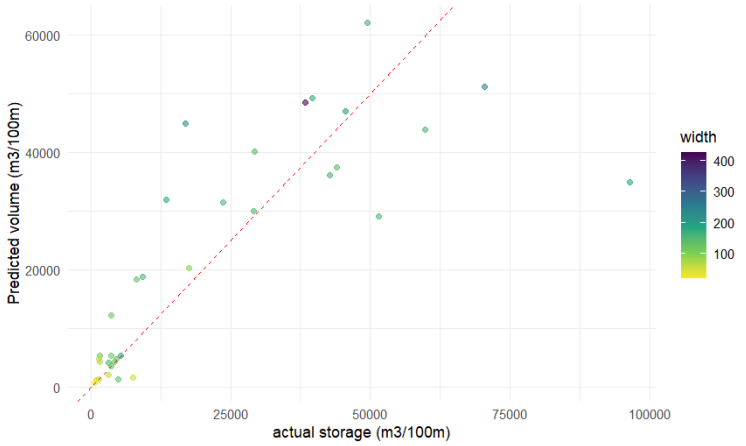
Where X1 is the drainage (upstream) area (km^2), X2 the width (m) and X3 the dry season evaporation ($m^3/100m$). The intercept term 0.8205 provides the base depth estimate when all predictors are zero. The negative coefficients for width and the width-evaporation interaction likely adjust for the inter-correlation with drainage area, reducing the estimated depth for wider rivers. As discussed later, these corrections may reflect either an underlying pattern or an artifact tailored to this specific dataset.

All multiple linear regression models demonstrated better predictive accuracy when trained and tested on smaller storage capacities (below 30,000 m^3 per 100m), while single regression models using vegetation did not.

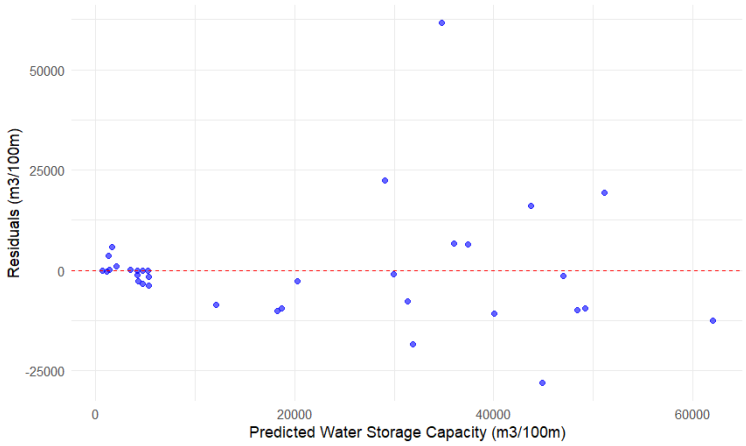
TABLE 5.2: Model performance on predicting depth, using repeated k-fold cross validation, using different predictor variables and interaction terms. The interaction factor is width x RV extent or width x evaporation for models using RV extent or evaporation, respectively. "*" Means the model was trained and tested on smaller storage capacities (below 30,000 m^3 per 100m).

Model	**Upstream Area**	**Evaporation**	**RV Extent**	**Width**	**Interaction factor**	**R^2**	**MAPE (%)**
Model 1	x	-	-	-	-	0.19	75
Model 2	x	x	-	-	-	0.23	77
Model 3	x	-	x	-	-	0.3	72
Model 4	x	-	-	x	-	0.46	72
Model 5	x	x	-	x	-	0.60	69
Model 6	x	-	x	x	-	0.60	73
Model 7	x	-	x	x	x	0.62	69
Model 7*	x	-	x	x	x	0.76*	59*
Model 8	x	x	-	x	x	**0.64**	**61**
Model 8*	x	x	-	x	x	0.81*	61*
RV Extent Linear	-	-	x	-	-	0.40	132
RV Extent Linear*	-	-	x	-	-	0.34*	96*
Evaporation	-	-	x	-	-	0.32	113
Evaporation*	-	-	x	-	-	0.37*	92*

Figure 5.11 presents the actual versus predicted water storage capacities using MLR model 8 and the residual plot using MLR model 8, with drainage area color-coded for intuition. The residuals show no clear trend, although the model tends to underestimate at larger actual values above 50,000 m^3/100m and overestimates in the smaller range of 10,000 to 35,000 m^3/100m. This pattern closely resembles the predicted depth results (see Appendix C), except that one small actual storage value is also underestimated there. Predicted depths reach up to 25m, compared to the measured maximum of 35m.



(a)



(b)

FIGURE 5.11: (a) Predicted versus actual water storage capacity (sy=0.15), calculated with multiple linear regression (MLR) model 8. (b) Residuals of the predicted storage volumes calculated with MLR model 8, averaging zero.

The resulting map is depicted in Figure 5.12. The total estimated water storage capacity within the delineated river is 132 × 10^6 m^3 (which could irrigate 13,200 ha). This is 0.0010% of total annual precipitation (as calculated by a total area of 45,000 km^2 times a yearly average of 286mm). The map excludes water stored in floodplains. Storage is primarily concentrated in the main stems, more so than dry season evaporation. The values reach up to 220,000 m^3/100m.

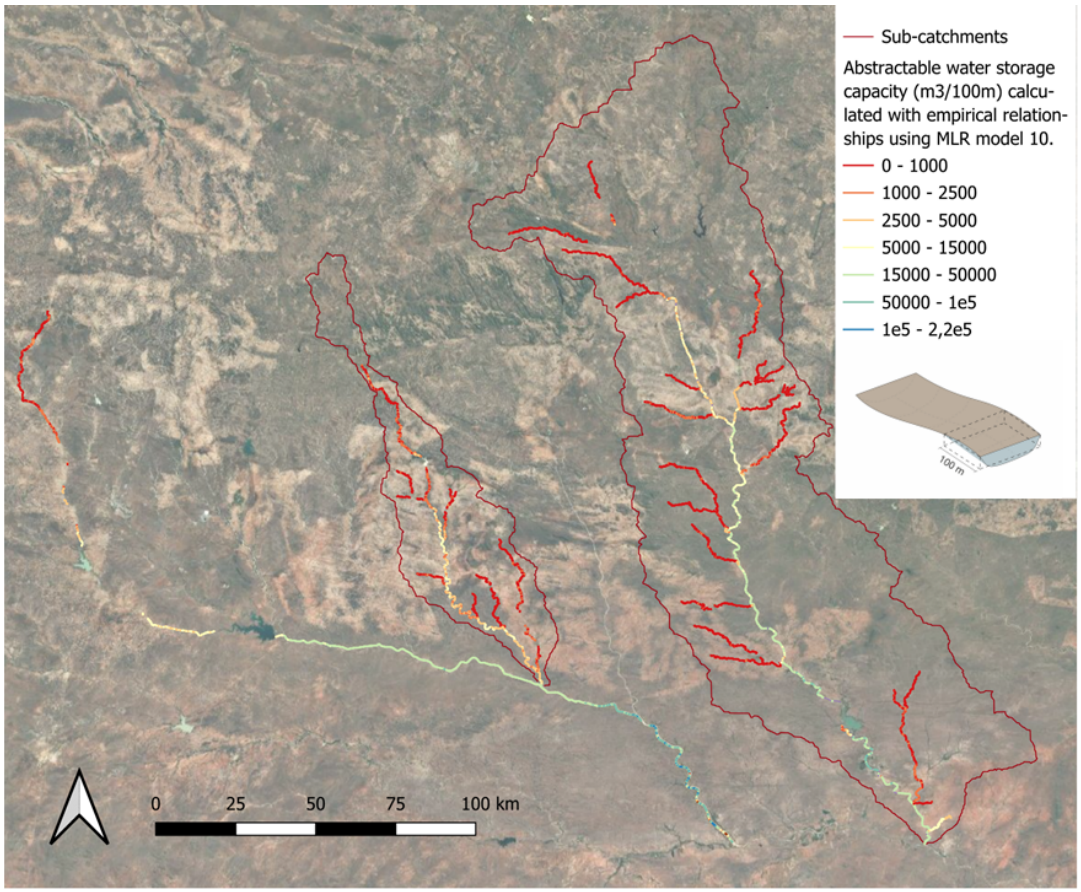


FIGURE 5.12: Map of sand rivers and the abstractable storage capacity for every 100m river section, calculated with MLR model 8, assuming a specific yield of 0.15.

5.4.3 Combining the two estimate methods

The maps can be combined to reveal insights about the unconsumed or unevaporated water availability.

The total estimate by the MLR model 8 is nearly 3 times higher than the total dry season evaporation. This is mainly due to the difference in the larger volumes, which hold most water availability. For the MLR model, the top 10% contains approximately 55% of water, illustrated in the exceedance plot Figure C.31) in Appendix C). Especially for the MLR model estimates, water is concentrated more in medium and larger rivers. The smallest 60% of rivers contain only 5% of the water, compared to 17% for evaporation.

The fraction between the two estimates was calculated and plotted in Figure (5.13). It shows the percentage of water that is evaporated from the estimated geometric storage volume. The most unconsumed water availability is located in the main stems. This is also illustrated by the difference between the two estimates, depicted in Figure C.32 in Appendix C. At tributaries, geometric storage estimates are frequently exceeded by evaporation.

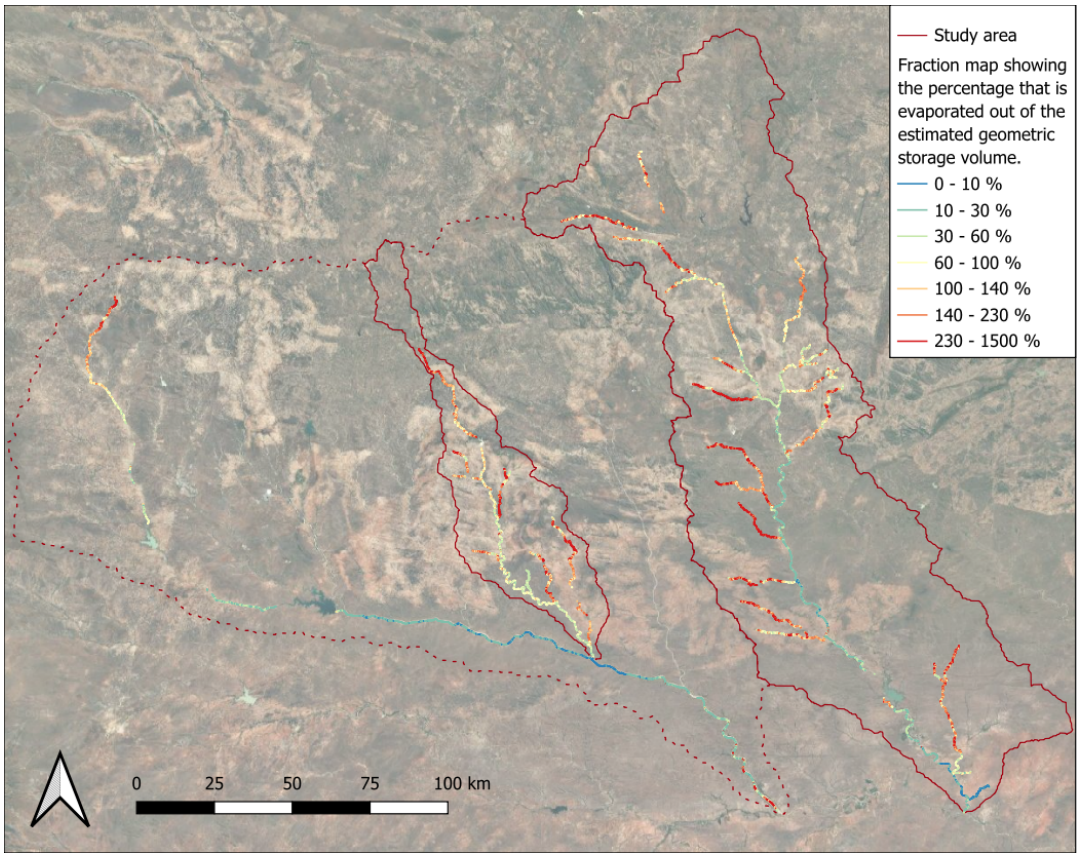


FIGURE 5.13: Map of delineated sand rivers and the fraction of water that is consumed. Calculated by dividing dry season evaporation with the geometric storage volume, assuming a specific yield of 0.15.

Lastly, is map with depths capped at 7m, considering the maximum depth of standard suction pumps (see Figure C.30 in Appendix C). It totaled 75.2 x 10^6 m^3, with an unconsumed total of 26 x 10^6 m^3. This could sustainably irrigate 2600ha, which is approximately 0.16% of the area of the Shashani and Mzingwane subcatchments (4.5 x10^6 ha).

5.4.4 Cost comparison

The material costs of a well-point system up to 5 meters deep are $31.50 (Hussey, 2007), resulting in a water cost of approximately $0.0032 per m^3. Assuming one well-point can irrigate one hectare or 10,000 m^3, this is about 300 times cheaper than the median final reservoir costs of $1 per m^3, as reported by Petheram et al. (2019).

Using reservoirs to store the total geometric estimate of 130 × 10^6 m^3 would cost around $130 million, whereas sand river well-points would cost just $0.4 million. This comparison excludes water transport costs and labor costs of setting up the well-point.

TABLE 5.3: Cost comparison of sand rivers and reservoirs

Description	**Hussey (2007)**	**Petheram & McMahon (2019)**
Installation/final cost	$31.50	$86 x$10^6$
Cost per m^3 (well-point)	$0.0032	$1
Cost for 130x10^6 m^3	$0.4 x$10^6$	$130 x$10^6$

Discussion and recommendations

This study aimed to enhance understanding of the extent of sand rivers and their nature-based capacity to store water through a remote sensing approach. The results demonstrate that sand rivers and a rough indication of their water storage potential can be mapped remotely with minimal field data. These findings could have significant implications for water management in semi-arid regions, although they also raise questions about sustainability and their accuracy.

This chapter discusses the results, focusing on the vegetation's behavior, the delineation method, and an assessment of the predicted water storage capacities and their uncertainties.

6.1 Seasonal Behavior of Riparian Vegetation

Prior to the mapping, the seasonal behavior of riparian vegetation was investigated to help assess vegetation as indicator of water storage. In summary, the results showed that riparian vegetation remained greener than its surroundings, continued to transpire during the dry season (more so than non-riparian tree cover), and showed a drop in NDVI only several months after the dry season began.

These findings highlight the reliance of riparian vegetation on sand river aquifers. The observed NDVI and transpiration declines could indicate water limitations, although other possibilities exist. One possibility is that vegetation fully depletes the available water in sand rivers during the dry season. Alternatively, water may remain in the channel but becomes inaccessible to plants, potentially due to limited water connectivity, such as low water gradients from declining water tables or obstruction by bedrock. This aligns with some models that found vegetation used significant amounts of water during the dry season, but water generally remained available for abstraction (Moulahoum, 2018, Benito et al., 2010, Love, 2013).

Secondly, these declines may not solely reflect water scarcity. Plant species have adapted to cycle their leaves annually, a form of phenology, shedding leaves in the dry season to reduce water loss. This adaptation could explain reduced transpiration and greenness, even in the presence of some accessible water. Additional factors include lower dry-season temperatures which may further reduce evaporation rates, and adaptations such as deep rooting and internal water storage may enable plants to conserve water early in the season, further complicating interpretations.

Further investigation into the role of leaf cycling and other adaptations is necessary to better understand riparian vegetation dynamics alongside sand rivers and assess the water availability. Future research could extend this analysis to other locations to confirm whether riparian vegetation fully depletes available water. The timing of NDVI and transpiration declines could possibly serve as an indicator of water availability and inform abstraction practices. Additionally, assessing the recovery of vegetation after extremely dry years could provide insights into resilience and the limits of sand river storage capacity.

6.2 Delineation Method and Automation Potential

6.2.1 Sand Rivers

Delineation lays the foundation of indicating water storage. For sand rivers, the results showed that the semi-automatic method was able to delineate large areas (percentage wise) of sand rivers wider than +- 40m. Compared to Bremer (2022), the only other automatic study, this method was simpler and less patchy, but did require the manual editing of river centerlines.

Mostly small underestimations arose due to color differences (blue water, green vegetation and grey bedrock). These could be manually corrected, though with some ambiguity. The delineation process for connected rivers is especially fast for wider rivers when OpenStreetMap (OSM) centerlines are available; however, it is slower where these are missing or need manual adjustments, which is often the case for rivers narrower than 50m.

For future studies in other areas, more complex delineation approaches might be necessary if the materials of the channel - sand or otherwise - have a variety of colors. Alternatively detecting flow in the rainy season, as done by Walker et al. (2019). Alternatives like machine learning classifications have potential but would need to be sophisticated, as performing supervised classification in GEE had poor results. For example, the global land cover classification by ESA (European Space Agency, 2021) successfully delineates medium to large sand rivers at 10x10m resolution. This product also successfully classifies small farming plots. It opens up the possibility to automate the marking of disturbance or logging of riparian vegetation, or identify untouched potential of sand rivers.

6.2.2 Riparian zones

While thresholds had to be visually fine-tuned to balance overestimation and underestimation, both smaller and larger riparian zones were captured at least approximately. While NDVI was more effective than Canopy Height for smaller zones, Canopy Height may serve as a more conservative, less patchy predictor for larger zones. Other studies that use (un)supervised classification (Nguyen et al., 2019) or LiDAR and Object Based Image analysis (OBIA) (Johansen et al., 2011, Blaschke et al., 2011) seem more powerful, but are not globally available or computationally more intensive.

One notable limitation can occur at forest-like surroundings where non-riparian tree-cover can occasionally pass the NDVI threshold. This was seen at a reach in the upper/middle Mzingwane where the riparian zone had to be limited.

In future studies in other regions, whether for riparian zones or sand river channels, threshold values should be recalibrated, or dynamic thresholding methods could be considered.

6.3 Maps: Dry Season Evaporation and Geometric Water Storage Capacity

The two main estimate results differ in their levels of uncertainty and potential applications for water resource management. Together, they provide the most valuable insights to support the development of climate-resilient livelihoods, such as farmer-led irrigation systems.

6.3.1 Applications and study comparison of Dry season Evaporation

This remote sensing approach offers an initial indication of water storage without the need for field data. Estimated results totaled a 49x10^{6}m3 (enough to irrigate 4,900 ha). Most water was located at the main stems, with additional local hotspots.

With this method only, however, uncertainty remains about how much water is available for abstraction without disrupting riparian vegetation and its ecosystem services. This uncertainty is addressed by incorporating depth measurements and geometric storage estimates.

Firstly, the depth measurements - converted to geometric storage - were almost always bigger than dry season evaporation (see Figure 5.7). This reaffirms that dry season evaporation serves as a conservative estimate. The four exceptions are mostly small channels with large floodplains, where vegetated point bars caused higher evaporation (Appendix C). Also, the moderate regression below the 1:1 line which resulted in a gap - with sustainable irrigation potential - quickly becomes significant. For example, a dry season evaporation estimate of 7000 m3/100m has on average around 25,000 m3/100m, leaving 22,000 of untapped water. The gap is more consistent at higher values (>5000 m3/100m). The depth measurement studies validate this result to some extent, reporting leftover water at the end of the dry season in successful irrigation areas (e.g. Owen et al., 2005, Love, 2013, Blok et al., 2017). For example, piezometer data from Love et al. (2011) in the lower Mzingwane and Shashe regions showed water levels never dropped below 0.9m during the dry season, which also indicates minimal seepage losses. These findings roughly align with the maps showing 10-30% water loss to evaporation. For an exemplary site in the Shashani, this was 72% water loss at the end of the dry season, compared to the maps showing 30-60% (Blok et al., 2017).

Future research could further validate the evaporation-storage gap, especially in areas unaffected by suitability bias. As discussed in section 6.4, WaPOR's uncertainty - with a MAPE of around 40% - is notable but does not significantly undermine its ability to capture local variations of dry season evaporation.

6.3.2 Applications and Study Comparison including Geometric Storage Capacity

The geometric storage estimates complement dry season evaporation by enabling estimates of unconsumed water storage. While these estimates require depth measurements and come with significant uncertainties (MAPE of 61% plus a bias in depth measurements, see section 6.4), they provide a useful first approximation in a semi-automatic way.

The unconsumed difference between dry season evaporation and geometric storage estimates highlighted significant sustainable potential. With depths modeled to 25m, this was 75 x 10^6 m^3 (potentially 7500ha, with 60% of total unevaporated). With a depth capped the 7m reach of suction pumps, it totaled a 2600ha irrigation potential. While this represents only 0.16% of the total area, it could double the 2,600 ha irrigated in 2011 in the lower Mzingwane valley (Love et al., 2011). Even a small increase in irrigated area could meaningfully impact many smallholders in water-stressed regions.

Estimates at tributaries suggests that many of them were emptied by evaporation. Here, many people in decentralized rural areas could benefit. However, estimates exclude floodplain storage and still show occasional hotspots with untapped

water availability. More field data is needed in tributaries about unconsumed water availabilities.

This approach is far cheaper than reservoirs (see Table 5.3) and offers a distributed water supply without negative ecological consequences.

Two other studies (Moyce et al., 2006, Love, 2013) have estimated geometric storage by extrapolating limited depth measurements. These methods, which do not account for local depth variations, are less precise. While comparisons are challenging due to differing study areas, Love (2013) estimated an unconsumed alluvial aquifer potential for the Mzingwane Catchment that could support 1,250–2,800 ha, which aligns with this study's 2,600 ha (limited to 7m). Moyce et al. (2006) reported geometric storage of 175,000–5.4 million m^3 for the lower Mzingwane river, compared to this study's cap-less estimate of 40 x 10^6 m^3 for entire Mzingwane. The discrepancy largely stems from their use of a conservative depth, averaging approximately 3.5 m, compared to 15 m.

These studies also incorporate floodplain storage (delineated with riparian zones and false color composites) and regional seepage risks (based on geology and field data). Future research could integrate these elements for more complete estimates.

6.4 Estimate Limitations

A key limitation of the maps is their accuracy. In summary, both estimation methods are affected mainly by data uncertainties: WaPOR has a 40% error margin, and the geometric storage model has a 61% MAPE plus depth measurement bias. While these uncertainties are significant on a local scale, they still allow the identification of broad targets, such as main river stems and reaches with high evaporation.

6.4.1 Limitations of Dry Season Evaporation

The largest source of uncertainty for dry season evaporation is WaPOR data. Blatchford et al. (2020) reported a bias of 0.1 mm/day and a MAPE of 36% in a savanna/woodland biome, along with an overall mean bias of 0.6 mm/day and MAPE of 40% based on data from 14 eddy covariance (EC) stations across Africa. Further validation is limited due to a lack of eddy-covariance (EC) towers and due to variations in evaluation methods (Blatchford et al., 2020, Tran et al., 2023). Despite these uncertainties, the 40% error is unlikely to hinder the identification of key targets such as main stems and high-evaporation reaches. The studies also suggest slight underestimation during the dry season in Miombo woodlands (Zimba et al., 2023) and in areas of relatively high local evapotranspiration (Blatchford et al., 2020).

Another notable uncertainty is the unconsumed or available water when the method is not complemented by geometric storage estimates or measurements. Especially at smaller rivers with less evaporation, the unconsumed water availability is uncertain. However, in this study area, larger rivers consistently have unconsumed water. Whether this is the case for other areas, however, needs further validation.

Evaporation losses from the channel's top layer (0.8m) early in the dry season were not excluded from the total estimates. This water could still be used if abstracted shortly after the last flow event.

Finally, a minor shared uncertainty is the imperfect delineation of areas, including manual editing and threshold fine-tuning.

6.4.2 Limitations of the geometric (MLR) storage estimates

In summary, the primary uncertainty is a local MAPE of 61%, along with potential biases in depth measurements.

Data Uncertainties

Uncertainties stem largely from the depth measurements and the limited explanatory power of predictor variables.
Specifically, in addition to possible errors in the depth measurements themselves, a sampling bias, dependency and insufficient coverage are present.

While predictors like vegetation may be less location-dependent than drainage area, even a few depth measurements in these unmeasured regions could substantially improve accuracy.

Depth measurements were unevenly distributed, with depth measurements completely missing in the middle/upper Mzingwane sub-catchment, lower Shashani, and middle Shashe (Figure 5.5). As a result, the MLR model relied on data from other areas, making the true depth and the depth-predictor relationships uncertain at the sub-catchment level (Figure 5.9). The relationship between geometric water storage and vegetation could be less location dependent than drainage area.

A sampling bias was likely introduced, as sites were selected based on irrigation potential, possibly leading to overestimation. Additionally, measurements were often clustered, which could skew regressions.

Secondly, the predictor variables cannot fully explain depth. Although the sum of r^2 of all individual predictors is large (>1), not all variety was captured due to multicollinearity between predictors. Uncaught errors are illustrated in section 6.4.3.

Multicollinearity can confuse MLR models. While it still allows for extracting more information than a single predictor could, it makes the individual contributions of predictors harder to interpret. Despite multicollinearity, highly variable depth data, and possible measurement errors, the negative terms in model 8 have some rationale. For instance, they could compensate for the non-linear overemphasis of drainage area in the largest rivers. Since at the largest drainage areas - on average - depth even decreases again.

Method

The residuals of the model average zero, possibly helping to reduce total estimates with large amounts of calculations. However, the depth measurements may be biased, making predictions wrong regardless if the residuals cancel out. Additionally, the model slightly overestimates medium-sized storages and underestimates large ones.

Furthermore, slope calculations were incomplete, limiting the insights on its potential influence.

6.4.3 Understanding (Invisible) Errors

Understanding errors or biases in the maps can help to reduce them.

False positives (vegetation but very little storage) and false negatives (underestimations) can occur. They could explain why dry season evaporation is exceeded by geometric water storage. The main false negative is likely lateral confinement (see Figure 6.1). Multiple studies report confinement by clay, rising bedrock or

bedrock outcrops, such as (Moulahoum, 2018). The second is water flow limitations induced by small gradients. Other causes of false negatives include limited fertile bank/floodplain extent (restricting vegetation growth) and logging/clearing riparian vegetation (regardless of water storage present).

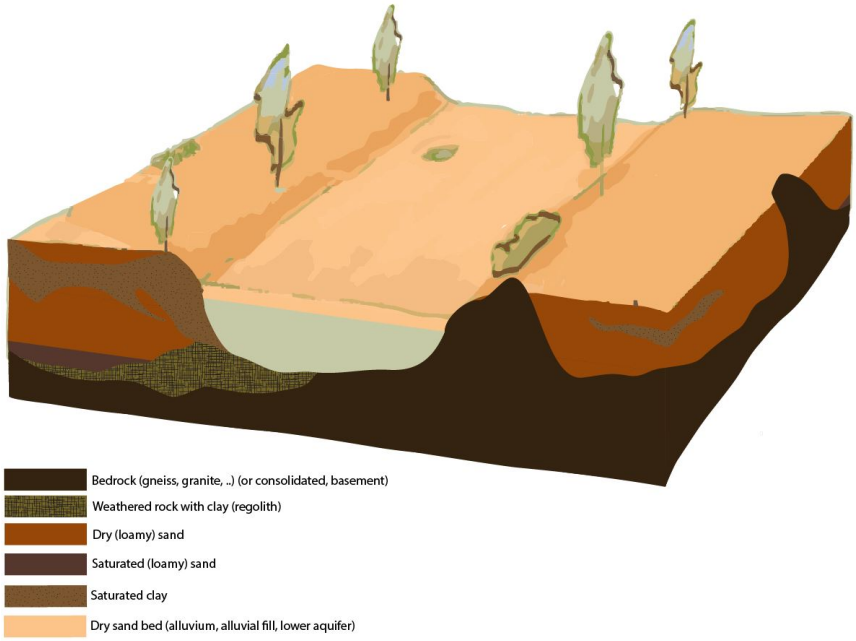


FIGURE 6.1: The main suspected false negative (no RV but still storage) is lateral confinement by rising bedrock or clay. While the channel aquifer can have large water storage, it cannot reach and sustain riparian vegetation.

A main expected false positive is lateral seepage into regolith (see Figure C.36 in Appendix C). Other causes of false positives include local gaining conditions (Lateral groundwater flow sustains vegetation but never reaches the channel) and riparian vegetation consumes all available water.

Furthermore, water leakage, caused by faults, fractured rock or dikes is inaccessible to both vegetation and well-points. For geometric storage estimates this could cause overestimation. Additional inaccuracies can result from recent sand mining activities.

Some of the above errors might be deduced with remote sensing, though field investigations might be necessary. Possibly remotely deducible examples are surfaced confining bedrock and clay, faults inducing seepage, logging of riparian vegetation, floodplain materials indicating hydraulic conductivity (tried out by Moyce et al. (2006)) and (Mpala et al., 2016).

6.5 Insights from the Spatial Analysis

Drainage area and riparian vegetation are new (regression) predictors of water storage. Available literature mostly mention slope, seepage lines, hard-soft boundaries and erosion of soil (e.g. Hussey, 2007, Moyce et al., 2006.

The influence of slope is in reality greater than the results showed. Zooming in on steep slopes consistently reveals bedrock outcrops (possibly dikes, C.39 in Appendix C). A large amount of bedrock outcrops leaves little room for sand. Future research

could improve the slope calculation method, by averaging over longer distances and filling the data gaps by using adjacent values.

Unexpectedly, vegetated point bars were not statistically significant. This was likely due to their inconsistent vegetation cover, their cover often balanced out by the sparse vegetation typically found on the opposing cut banks, and imperfect calculation of sinuosity (particularly for smaller rivers).

6.5.1 Limitations of the Statistical Analysis

The many possible (inter-correlated) causes of vegetation and depth complicate drawing conclusions about the statistical links. A main example is the large storage volumes and sudden increase in riparian zones at the lower Mzingwane. They can be caused by (a combination of) large drainage areas, gentle slopes, Basalt, presence of reservoirs and the presence of faults and dikes. Moyce et al. (2006) and Kenyon (2022) partially account the increased sediment fill to hard-soft rock boundaries at the many faults and dikes. While these may occur more at field scale, they were not observed well with the coarse geological maps.

Similarly, soil maps were too coarse to assess their impact. Also, their applicability is low as floodplains fall outside FAO's category and they do not directly represent erodability. Land cover is expected to better represent soil erodability and is therefore recommended for future modeling.

6.6 Final Recommendations

The two current remote sensing methods are first trials that need further validation and could still be improved.

First and foremost, when using the model(s) for targeted irrigation planning, it is necessary to verify whether water is actually available. Ecosystem services should be considered during site selection. This could imply recording water levels at the end of the dry season and probing for depths.

To improve both estimates and understanding, fieldwork could investigate causes of local variations in riparian vegetation, as well as other false positives and negatives. This can include factors like soil quality or water connectivity (impermeable clay, water gradients). Furthermore, automated adjustments could be made to help correct for overestimated evaporation rates at confluences/tributaries and for decreased rates at human logging. On the other hand, increased vegetation and evaporation could potentially improve predictions by extrapolating them to adjacent river sections, provided that the river reach is known to limitedly vary in aquifer volume. Just looking at slope over a sattelite image can indicate depth.

To improve the empirical (MLR) method, fieldwork should include simple probing to address data gaps and assess the universality of the empirical relationships. Depth measurements are particularly needed in all smaller rivers (with smaller drainage area), gentle slopes, and at downstream Shashani. This would help to assess the predictor variables at catchment-level, in particular drainage area. Future studies could (better) incorporate slope, land cover (for erodability) and lastly faults and dikes. The impact of faults and dikes could be better captured with remote detection, and include the possible compartmentalizing effect they have on aquifers.

6.6.1 Extrapolating to Other Areas

The method could be applied to other areas. For doing so, here is a summary with some guidelines and key recommendations.

The delineation process is semi-automatic but requires specific conditions. Select imagery from a homogeneous dry season to avoid local peaks in vegetation caused by rain events, but avoid extreme dry years where vegetation may be completely absent. Varying colors in sand river channels might require more complex methods such as machine learning or dynamic thresholding. Optionally, manually draw river centerlines.

Limited constraints are expected for estimating dry season evaporation, but inspections remain necessary about unconsumed water availability.

For estimating geometric water storage capacity in new areas, it is essential to have some depth measurements to add to the training and validation of a new (MLR) model, as the universality of the empirical relationships are not yet certain. They are expected to vary with for instance soil erodability, and the influence of slope along with interactions with other variables. It is recommended to spread these measurements across different drainage areas, widths and vegetation extents and slopes, also at catchment-level, so that the model represents the whole area.

Conclusion

To obtain the objectives, the following research question and sub-questions were set up:

Can remotely observed spatial patterns like riparian vegetation adequately indicate the capacity of sand rivers to naturally store water, for irrigation and other human uses?

1. What are key climatological, hydrological and geo-morphological factors that influence riparian vegetation?

Riparian vegetation is influenced by a combination of water quantity and connectivity, assuming the most scarce resource for (adapted) riparian plants is water.

Water quantity depends on amount and duration of rainfall, dry season duration, influencing recharge of floodplain aquifers (Hussey, 2007). Geomorphological factors such as drainage area, slope (controlling flow velocity), sediment supply (land use and erodability of soil), geology type and channel meandering affect vegetation by shaping aquifer dimensions. Losses and inputs also affect water quantity like vertical seepage (more likely at older lithology types) and subsurface flow (determined by slope and subsurface material). Water connectivity depends on permeability (reduced by clay and rising bedrock) and water gradients, with limited connectivity in shallow aquifers reducing vegetation extent.

2. How does the riparian vegetation respond to rainfall patterns?

The year-round higher NDVI and delayed decrease in NDVI during the dry season, suggest the vegetation's dependence on the sand river's aquifer. This is likely true whether the drops in NDVI and transpiration mean the aquifer is fully depleted or not, supplemented by riparian vegetation transpiring even at the end of the dry season. However, influences such as the plants' adaptation to cycle leaves annually, preserve water early, store water internally and the winter's lower temperatures complicate interpretation.

3. How can sand rivers and riparian vegetation be delineated?

The simple sand index threshold method effectively delineated most river channels wider than 40m, except where channels were densely vegetated or contained outcrops. While manual editing was necessary to obtain river centerlines and connect sand river polygons, the method can be largely automated, especially in regions with existing OpenStreetMap (OSM) river centerlines. NDVI thresholds were able to capture small riparian zones (<20m), and bigger zones better.

4. How can a minimum water storage capacity be estimated with evaporation?
Using global WaPOR v3 evaporation data and the area of the sand river channel and riparian zone, the evaporation can be summed over the dry season. This assumes

all evaporated water comes from the sand river aquifer.

5. How can geometric water storage capacities be predicted with empirical relationships?

By using 35 depth measurements along with remotely sensed environmental attributes — such as dry season evaporation, drainage area, and channel width — empirical relationships can be developed. These relationships were integrated into a multiple linear regression model, allowing for the prediction of depth across the entire study area.

6. What can the estimates teach about total water availability and suitable locations for water abstraction?

There is a significant nature based capacity to store water in the study area's sand river channels (totaling 123 × 10^6 m^3 potentially irrigating 12,300 ha). The two estimates can be combined for an unconsumed water availability, showing a significant sustainable potential (totaling 83 x10^6 m^3 potentially irrigating 8300ha). Depth measurements showed that dry season evaporation was consistently exceeded by geometric storage, suggesting that rivers with medium to large evaporation rates have most unconsumed storage potential. The maps suggest that water storage is concentrated in the main river stems (the largest 10% hold 45-55% of water). Especially reaches are promising with large vegetated areas and gentle slopes. All geometric storage estimates at tributaries were relatively small. Most of them were exceeded by dry season evaporation, suggesting that using their water would be in direct competition with riparian vegetation. However, some local hotspots point towards some water availability. This could potentially benefit many people living closeby in decentralized rural areas. These estimates are uncertain since there is a shortage of data at smaller rivers. While local uncertainties are still significant, they provide a useful first approximation.

7. How uncertain are the estimates and can understanding them help us to reduce them?

Both methods had large local uncertainties stemming mostly from the data. WaPOR has a MAPE of around 40%, but the dry season evaporation indicated limited information about unconsumed water availability without the help of geometric estimates in smaller rivers. The MLR model's depth measurements were not very representative and the predictor variables were limited. The model's uncertainty is significant (R^2 = 0.64; MAPE = 61%). However, the estimates are still very useful for total estimates and first approximations. Identifying theoretically errors could help us to reduce them, such as remotely sensing confining clay or bedrock, faults indicating seepage and the extent and connectivity of floodplains

Although uncertainties remain and field validation is still needed, the models show promise for wide-scale planning and development. The findings highlight that sand rivers, with rough estimates of their decentralized, cost-effective, and sustainable water storage, can be mapped remotely with minimal effort or field data. This could open up possibilities for water management in semi-arid regions, offering support for farmers and paving the way for farmer-led irrigation initiatives.

Appendix A, Study area

The geology map at a part of the Limpopo river basin is depicted in figure A.1. G: Older gneiss (high grade metamorphic, formed by metamorphosis of granite) of the Craton of the post African and Pliocene Surfaces (pre-cambrian). Bv: Mainly mafic metavolcanics with subordinate metasediments. (Bulawayan)
Yg: Younger intrusive (igneous) granites. Gn: Various paragneisses (gneiss derived from sedimentary rock). Ba: Basalt. Fine-grained. extrusive igneous rock formed by rapid near-surface cooling of (mafic) lava. Another map mentions this region is sediments overlain by volcanics. St: Sandstone. S: Recent and older alluvial deposits (gravel, sand, silt and clay). Mb: Various paragneisses of the Limpopo mobile belt. Other, coarser, geology maps summarize G, Bv, Yg, Gn together as "basement", but still in distinguish basalt and unconsolidated.

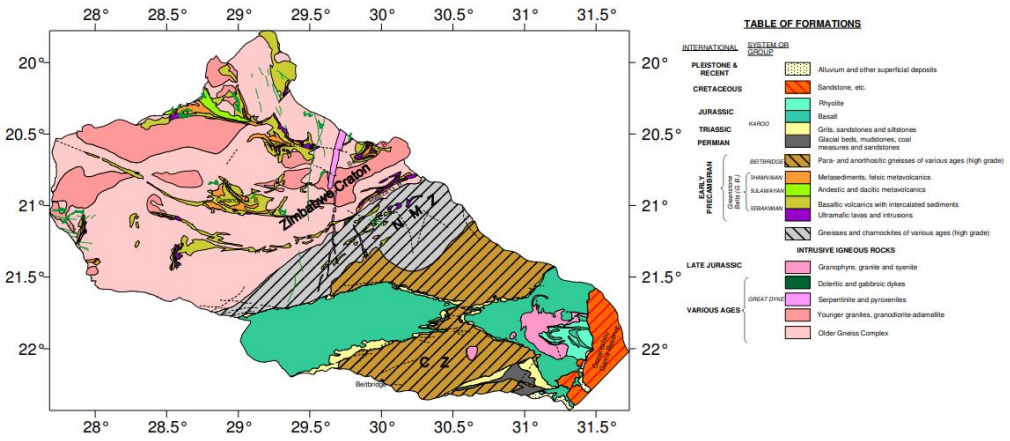


FIGURE A.1: Geology map with different rock types in the Zimbabwe Craton and the location of the Northern Marginal zone (NMZ) and Central Zone (CZ) of the LMB

Daily precipitation means of the Mzingwane sub-catchment and its dry dry season in the hydrological year 2023 are shown in Figure A.2

Depth measurements are depicted in figure A.3

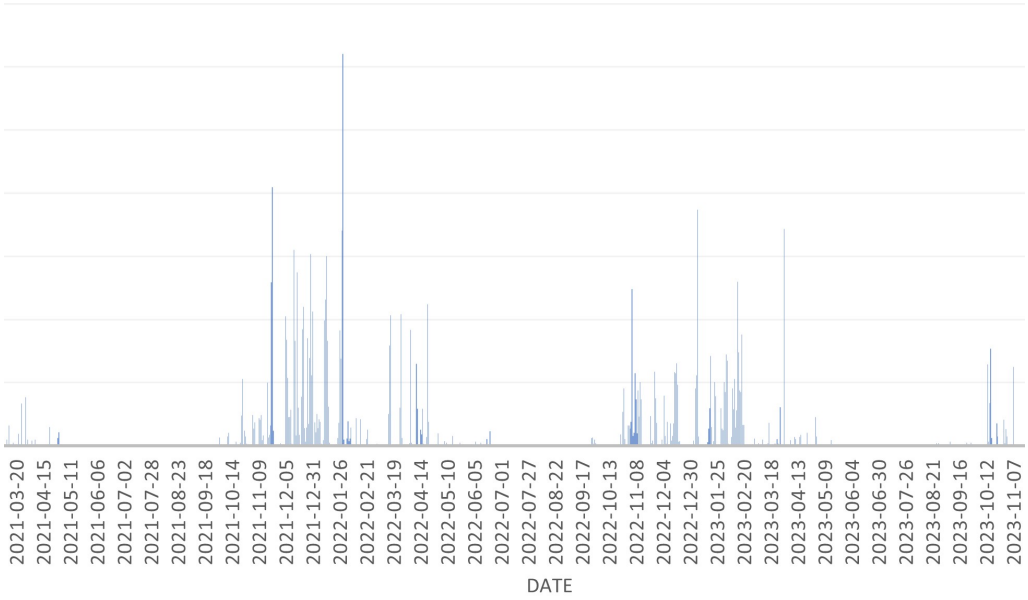


FIGURE A.2: Daily precipitation means of the Mzingwane sub-catchment of the hydrological years 2021-2023

Name	River	Country	AvgThick_m	MinThick_m	MaxThick_m	DataType	Sylold	Author	Loc
Gulubane	Shashe	Zimbabwe	9,00	9,00	10,00	Probing		Davies et al	Riverbed
Mathangwane	Shashe	Botswana	3,50	3,50	5,00	Probing		Davies et al	Riverbed
Chadibe	Shashe	Botswana	3,00	3,00	5,00	Probing		Davies et al	Riverbed
Malala	Mzingwane	Zimbabwe	25,00			Field Test	0,0540	Masvopo et al	Riverbank
Malala	Mzingwane	Zimbabwe	5,00			Field Test	0,0540	Masvopo et al	Riverbed
C2a	Mzingwane	Zimbabwe	15,00			CVES	0,2000	Owen & Dahlin	Floodplain
C2b	Mzingwane	Zimbabwe	30,00			CVES	0,2000	Owen & Dahlin	Riverbed
C1b	Mzingwane	Zimbabwe	25,00			CVES	0,2000	Owen & Dahlin	Riverbed
L1a	Mzingwane	Zimbabwe	40,00			CVES	0,2000	Owen & Dahlin	Riverbed
L1c	Mzingwane	Zimbabwe	5,00			CVES	0,2000	Owen & Dahlin	Riverbed
L5a	Mzingwane	Zimbabwe	40,00			CVES	0,2000	Owen & Dahlin	Riverbed
L5c	Mzingwane	Zimbabwe	5,00			CVES	0,2000	Owen & Dahlin	Riverbed
L1b	Mzingwane	Zimbabwe	20,00			CVES	0,2000	Owen & Dahlin	Riverbed
L5b	Mzingwane	Zimbabwe	35,00			CVES	0,2000	Owen & Dahlin	Riverbank
DOBRN	Mzingwane	Zimbabwe	4,00			Field observations	0,1070	Love et al	Riverbed
ZEZANI	Mzingwane	Zimbabwe	4,00			Resisitivity, assumed as closeby	0,0770	Love et al	Riverbed
KWALU	Mzingwane	Zimbabwe	13,00			Resisitivity, assumed as closeby	0,0770	Love et al	Riverbed
FERG	Mzingwane	Zimbabwe	12,00			Interview	0,1070	Love et al	Riverbed
BISHOP	Mzingwane	Zimbabwe	20,00			Interview	0,0900	Love et al	Riverbed
MZNGA	Mzingwane	Zimbabwe	20,00			Interview	0,0900	Love et al	Riverbed
BENFER	Mzingwane	Zimbabwe	20,00			Interview	0,0900	Love et al	Riverbed
MASSASS	Mzingwane	Zimbabwe	10,00			ERT	0,0430	Love et al	Riverbed
RRANCH	Mzingwane	Zimbabwe	15,00			Resisitivity, assumed as closeby	0,0430	Love et al	Riverbed
MALALA	Mzingwane	Zimbabwe	15,00			ERT	0,0430	Love et al	Riverbed
C1a	Mzingwane	Zimbabwe	20,00					Owen & Dahlin	Floodplain
C1c	Mzingwane	Zimbabwe	5,00			CVES		Owen & Dahlin	Floodplain
L1c	Mzingwane	Zimbabwe	15,00			CVES		Owen & Dahlin	Riverbed
Mabuza	Shashani	Zimbabwe	1,97			Probing		Mpala	Riverbed
Tshelanyemba	Shashani	Zimbabwe	1,90			Probing		Mansell & Hussey	Riverbed
a	Shashani	Zimbabwe	1,00			Probing Cross Section		Mpala et al	Riverbed
b	Shashani	Zimbabwe	1,50			Probing Cross Section		Mpala et al	Riverbed
c	Shashani	Zimbabwe	1,50			Probing Cross Section		Mpala et al	Riverbed
d	Shashani	Zimbabwe	2,00			Probing Cross Section		Mpala et al	Riverbed
e	Shashani	Zimbabwe	2,00			Probing Cross Section		Mpala et al	Riverbed
f	Shashani	Zimbabwe	2,50			Probing Cross Section		Mpala et al	Riverbed
Pb5	Shashani	Zimbabwe	1,94	0,50	3,00	ERT	0,1500	Blok	Riverbed
Pb6A/B	Shashani	Zimbabwe	0,91	0,50	3,00	ERT	0,1500	Blok	Riverbed
Pb3A/B	Shashani	Zimbabwe	0,90	0,50	3,00	ERT	0,1500	Blok	Riverbed
Pb2	Shashani	Zimbabwe	1,50	0,50	3,00	ERT	0,1500	Blok	Riverbank
Pb1	Shashani	Zimbabwe	1,33	0,50	3,00	ERT	0,1500	Blok	Riverbed
CrossSectionA	Shashani	Zimbabwe	2,75			Probing		Moulahoum	Riverbed
CrossSectionB	Shashani	Zimbabwe	2,25			Probing	0,1500	Moulahoum	Riverbed
CrossSectionC	Shashani	Zimbabwe	2,30			Probing	0,1500	Moulahoum	Riverbed
CrossSectionD	Shashani	Zimbabwe	2,50			Probing	0,1500	Moulahoum	Riverbed
CrossSectionE	Shashani	Zimbabwe	2,35			Probing	0,1500	Moulahoum	Riverbed
Pb4	Shashani	Zimbabwe	2,13	0,50	3,00	ERT	0,1500	Blok	Riverbed

FIGURE A.3: Summarized depth measurements georeferenced from literature.

Appendix B, method

B.1 Delineation

B.1.1 Delineating sand rivers

sand river workflow steps for one widthclass in Qgis:

-1 Draw or edit OSM centerlines manually (so that the automatic centerline problem is skipped)
-2 Buffer the centerline of interest to the biggest river width you can find
-3 Clip raster to buffer to exclude bare soil patches
-4 raster calculater where band=0 to exclude nodata
-5 Polygonize
-6 Buffer the centerline to the smallest river width you can find
-7 Merge polygonised raster with buffer to connect river reaches
-8 Dissolve
-9 Multipart to singleparts, so that bare soil patches don't merge to sand river in next step
-10 Buffer then debuffer (e.g. 20m) to fill river gaps
-11 Delete smaller singleparts with area < 10000 m^2 (tool extract by expression)
-12 Manually add and delete sand river areas (vegetation and water makes it patchy)
-13 Dissolve

Repeat for all width classes. Steps 2 – 11 could be automated in the python console or model builder of Qgis.

B.1.2 Delineating riparian zones

Below are the frequency histograms of NDVI and canopy height of a site at the Shashani and a site at the Shashe.

B.2 Statistical analysis

B.2.1 Choosing attributes

Other excluded attributes include Topographic Wetness Index (TWI), surface roughness, Multi-Resolution Valley Bottom Flatness (MRVBF) and change in slope. These were excluded because they are all derived from the same DEM and similar to slope, thus likely decreasing parsimony. Planform curvature, profile curvature, change in width and specific yield were proposed by Kenyon (2022). The increased information adding one of these was not considered worth the added complexity of adding another predictor value. Most of these are derived from DEM's, the same source of information as slope.

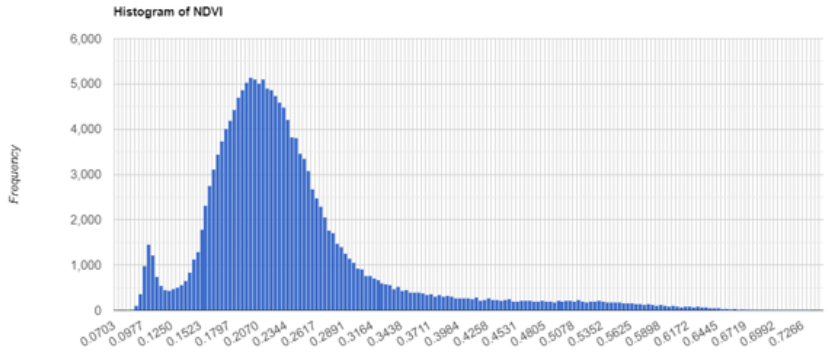


FIGURE B.1: NDVI frequency histogram of southern Shashani. There is no obvious minimum where riparian vegetation is seperated from other areas. NDVI of 0.42 was finally chosen.

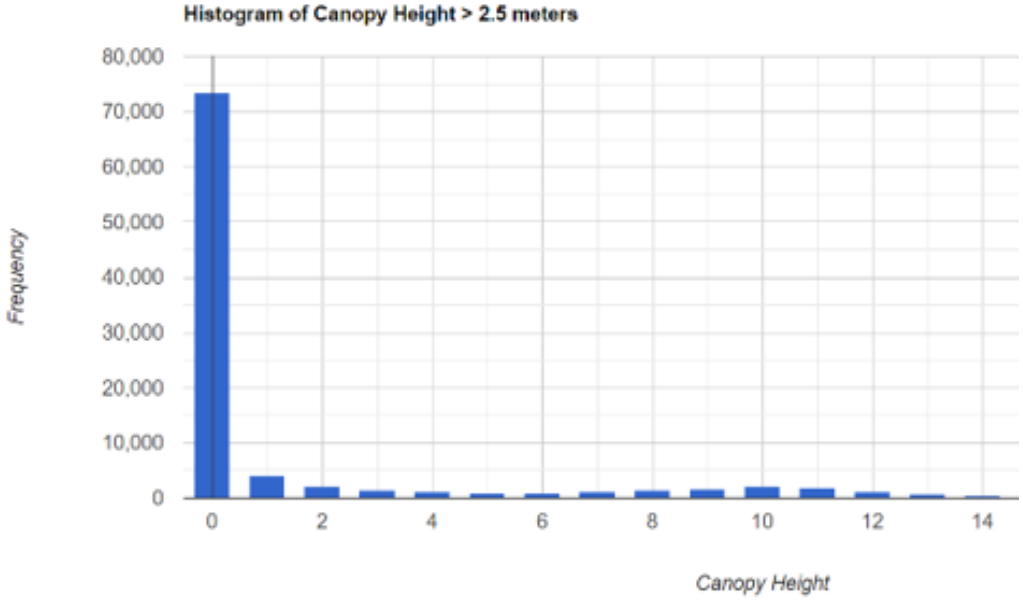


FIGURE B.2: Canopy Height frequency histogram of southern Shashani. The seperation between riparian vegetation and other areas seems clearer as a lot of non-riparian areas have 0 height. Visually, however, a threshold of >(3-6m) best captured riparian zones. So also at canopy height, there was an ambiguous transition and no clear optimum or threshold.

Attribute Name	**Why excluded**	**What/relevance**
Topographic Wetness Index (TWI)	All attributes, other than slope, coming from a DEM likely have limited unique added value compared to slope. Left out considering parsimony and time constraints.	Is commonly used to identify areas with higher water accumulation potential, therefore possibly sand too.

Multi-Resolution Valley Bottom Flatness (MRVBF)	All attributes, other than slope, coming from a DEM likely have limited unique added value compared to slope. Left out considering parsimony and time constraints.	Flatness could indicate the accumulation of water and possibly extent of floodplains, in addition to gentle slopes.
surface roughness	All attributes, other than slope, coming from a DEM likely have limited unique added value compared to slope. Left out considering parsimony and time constraints.	Retrieved from Kenyon, 2022. Could indicate objects causing surface roughness like bedrock outcrops and vegetation in the channel.
profile curvature	All attributes, other than slope, coming from a DEM likely have limited unique added value compared to slope. Left out considering parsimony and time constraints. Additionally, requires high resolution DEM data.	Is the longitudonal rate of change in slope. Indicates shape (convex, concave) and might indicate uniform bed materials.
Planform curvature	No known method to calculate well with remote sensing. Even a wide river (+-500m) would only have 5 measurement points due to a DEM's resolution (90x90m).	Retrieved from Kenyon, 2022. It is the lateral rate of change in slope, positive for ridges and negative for valleys. Might inform about subsurface faults and sediment load.
change in width	Considering channels only sporadically significantly change in width, it was excluded due to time constraints.	).
specific yield	Not observable from space.	Is somewhat correlated to slope as calculated by Kenyon (2022).

TABLE B.1: Table with excluded environmental variables, including why

B.2.2 Calculating attributes

B.2.3 Cutting sand rivers and calculating widths

1. smoothen centerline once
2. Points along geometry (100m interval)
3. create perpendicular lines with geometry by expression:
extend(make_line($geometry, project($geometry,40, radians("angle"-90))),40,0)
4. Insert id's with rownumber in field calculator
5. clip sand river polygon with lines
6. Delete the first polygon reach of every river. (They will disturb the area calculation since they will have the same ID as the second polygon)
7. Delete the last polygon reach of every river. (they are misformed)
8. Buffer perpendicular lines by 2.5m
9. Assign ID's to cut river sections with tool join attributes by location (overlap, one-to-one. Join ID's of buffered perpendicular lines)
10. Dissolve by section ID to merge braided rivers to one area (dissolve field to ID)
11. Do a manual check and check if features are (chrono)logical
12. in attribute table: add width by area/100 (for 100m interval)

Manual editing is required after cutting sand rivers, namely merging braided rivers and sections that are cut an extra time (see figures below). The solution was to

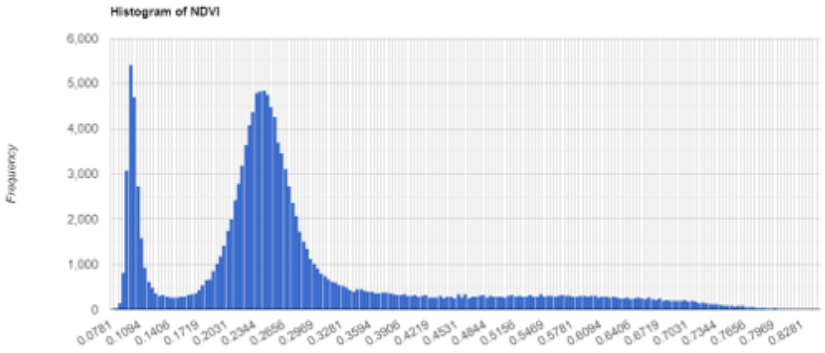


FIGURE B.3: NDVI frequency histogram of south-eastern Shashe. There is no obvious minimum where riparian vegetation is seperated from other areas. NDVI of 0.42 was finally chosen.

join attributes by a (shifted) line. In complete words, make shifted lines, give them a row number, join attributes (row number of the lines) by location to the river sections, and then merge the river sections per row number.

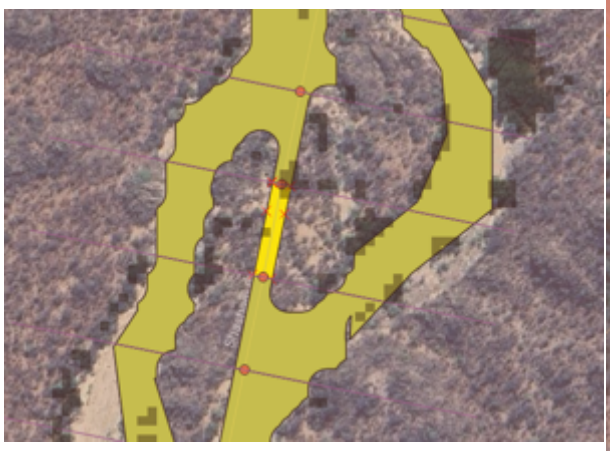


FIGURE B.4: Braided river results in separate polygons per transect that need to be merged

FIGURE B.5: Cut section gets separated and forms an extra polygon

B.2.4 Shifting points to hold information

All attributes are stored in points that are shifted by 50m, intersecting, lying in both the river section and the corresponding calculation box (see below). This avoids the problem that the "join attributes by location" tool mixes up points that touch two river sections or calculation boxes.

B.2.5 Cutting and calculating riparian vegetation extent

1.Clip NDVI raster result with the dynamic buffer (farthest RV extent for every river width of 20-40m, 40-100m, 100-300m, 300-900m)

FIGURE B.6: Outcrops and vegetation cause underestimation in the delineation of the sand river. Here it is clear what to fix.

FIGURE B.7: The sand river gaps are manually filled ar what to fix it is clear what to fix

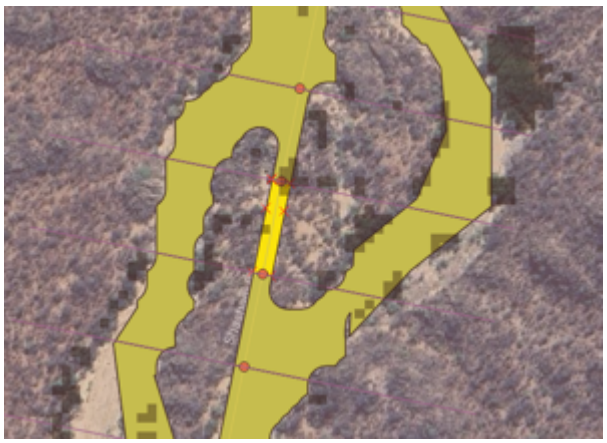


FIGURE B.8: Vegetation causes a large underestimation in the delineation of the sand river. This is more ambiguous as one can choose to include the vegetation as sand river or not. Automatic delineation often skips small of these vegetation plots, so this sand river is filled except for the small vegetation plot.

2.Polygonize
3.Create longer perpendicular lines
4.Buffer perpendicular lines (40m)
5.Delete the lást buffered line of each river to later make sure vegetation is calculated

downstream a river reach.
6.Manually clip buffered lines that fall into another river reach or its riparian vegetation, because of sharp bends.
7.Zonal statistics tool (count) to count the amount of 10x10m pixels of vegetation in the buffer 8.If any “invalid geometries” occur, use the check geometries tool to identify and delete random invalid geometries.
9.Convert count to area by multiplying with 100 in attribute table.
10.Create interpolated points to store data in with offset 75m (To dodge a very annoying problem of not being able to match the right calculation box to the right river section).
11.Combining attribute tables: join attributes by location (contain) 12.Select odd id1 features by expression: “id1" % 2 = 1
13.Delete odd numbered rows so that vegetation area’s are matched to the intersected river section upstream

14. The result needs a few manual edits. The calculation boxes occasionally counted riparian vegetation of another section. They were manually shortened.

FIGURE B.9: These are the cur river sections and the counting boxes that count the number of pixels of riparian vegetation. The counting boxes are longer for wider width classes and, when the river meanders, can occasionally intersect other sections and count too many pixels. This was manually edited.

Then, below is a fix used for when calc boxes overlap multiple statpoints. The average is taken of multiple rows in excel and are reduced to one row. Conditions used are row number and width. =IF(AND(B2=B1, D2=D1), "", AVERAGEIFS(F:F, B:B, B2, D:D, D2))

B.2.6 Automatic workflow overview

The following workflow with automatic Qgis models were used to calculate all attributes and add them to the information-holding that represent river sections. Some steps could be combined to one model in the future. Per river width class:
1. Run the model that cuts the river into 100m sections, calculates river width and creates RV counting boxes. These are combined because the river sections and counting boxes need a manual intermezzo.
2. Manually edit incorrectly cut river sections.

3. Manually cut RV counting boxes that count RV of other river sections. 4. Run the final model that:
4.1. Runs the model that counts riparian vegetation pixels.
4.2. Runs the model that copies elevation, averaged per river section.
4.3. Runs the model that copies drainage area, to the river section points.
5. Run the model that runs multiple models which calculate sinuosity, corresponding to the current width class.
6. Add soil type, geology type, lithology transitions, geologic structures with the join attributes by location tool. 7. Add depth data to the river sections and then the river section points with the join by location tool.
8. Add anthropogenic interference and tributary interference using the join by location tool.
9. Add widthclass identifier, riverID identifier and sub-catchment identifier using the join by location tool.
10. Merge the information-holding river section points of all width classes. 11. Export to xlsx. Custom sort the data: RiverID lowest to highest, widthclass lowest to highest and finally rownumber from lowest to highest. 11.1 Shift sinuosity since the data is assigned to each beginning point instead of the middle. Do this with an SQL function in DB manager (in qgis data manager) instead field calculator, because field calculator randomly shifts the values:
SELECT row_nr, sinuosity_,
LAG(sinuosity_, 15) OVER (ORDER BY row_nr) AS shifted_sinuosity
FROM "40_100_Statpoints";
10.2 Fill sinuosity gaps in excell with helper columns:
sinu_fill1 fill value above: B=IF(A3=0, A2, A3)
sinu_fill2 fill value above: C=IF(B3=0, B4, B3)
12. Calculate slope in excel.
12.1 Calculate the moving average from elevation (column A) for 1200m slope for a cell in row 9, with IF statements that check whether the river (ID) is the same (B3=B15):
C=IFERROR(IF(B3=B15, AVERAGE(OFFSET(A9, -6, 0, 13)), ""), "")
12.2 Calculate the 1200m slope in column D for a cell in row 2:
D=IFERROR(IF(ABS(C14-C2)/1200<=0.1, (C14-C2)/1200, ""), "")

B.2.7 Dry season evaporation

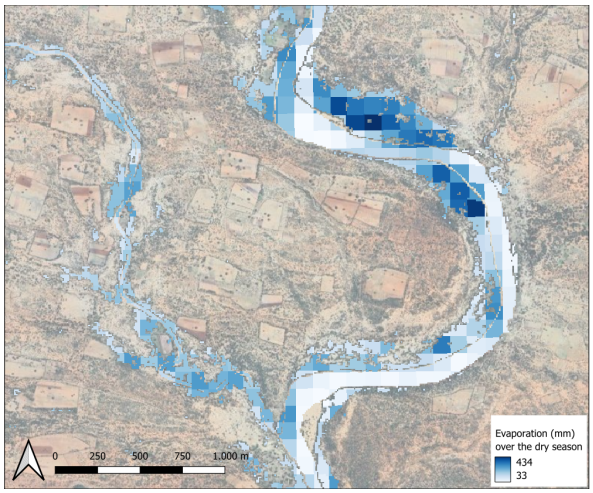


FIGURE B.10: Visualization example of dry season evaporation at the Shashani river. Many pixels overlap sand river, riparian zones and non-riparian zones.

The calculation of the evaporated volume by the sand river (area in m), the riparian zone (area count in 10x10m) and their mean evaporation (in mm):

_mean/1000_2*(area+("_count" /100))

Appendix C, results

C.1 Contextual/preliminary investigation

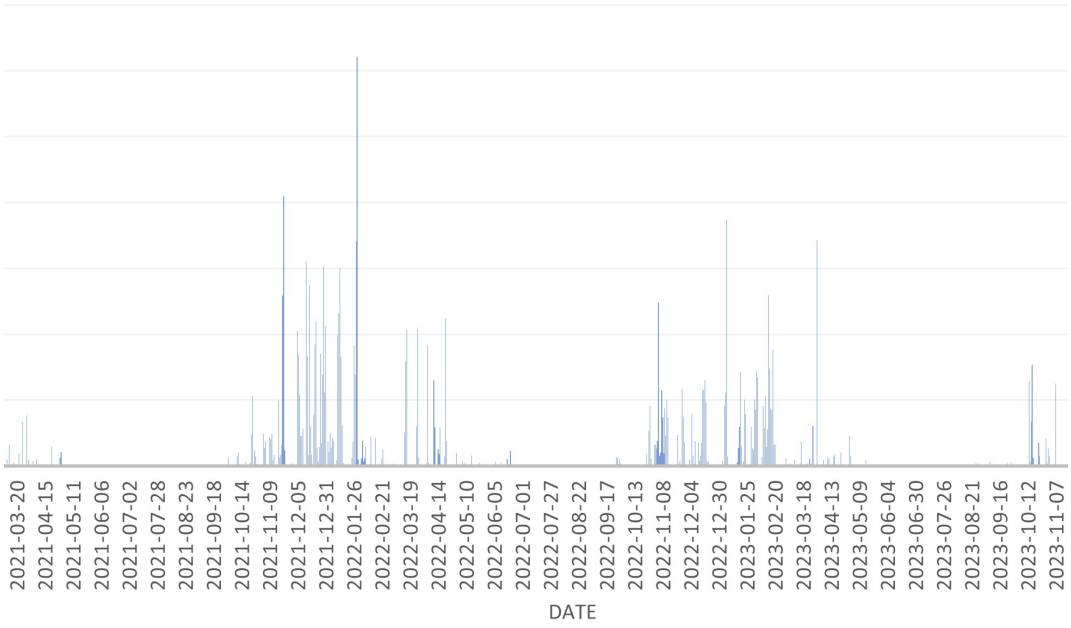
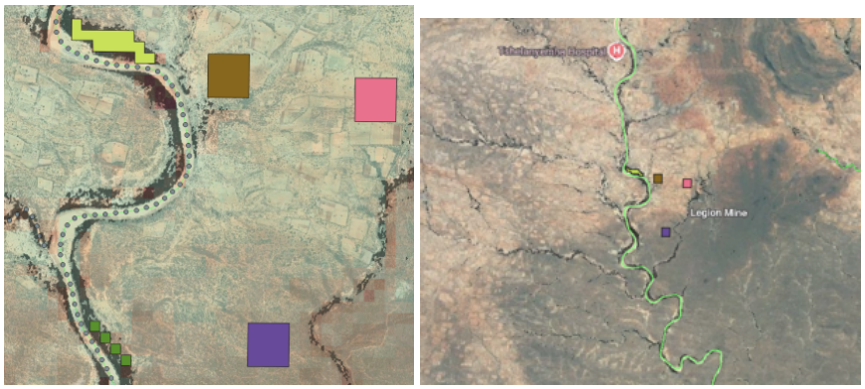
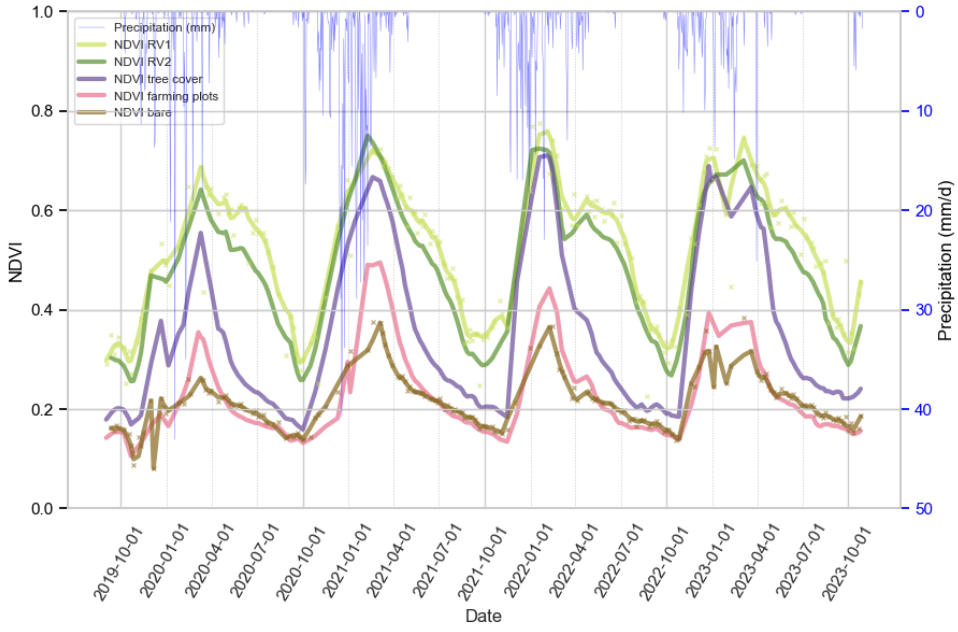


FIGURE C.1: The CHIRPS daily precipitation means of the Mzingwane sub-catchment. April till September of the hydrological year 2023 are completely dry, therefore hosting a meteorological homogeneous setting



(a) (b)



(c)

FIGURE C.2: Sentinel-2 NDVI time series of five plots in the Shashani sub-catchment. (a) The polygons of the five plots, e.g. bare is brown and riparian vegetation is green. (b) Zoomed out picture of the environment of the site; a boundary can be seen between communal land in the north (bare and farming plots) and commercial land in the south (tree cover) (c) NDVI time series of the five plots of the hydrological years 2020-2023, with daily precipitation in the background. The NDVI data was smoothed with the LOESS method. Original data can be seen of the bare and RV1 plots.

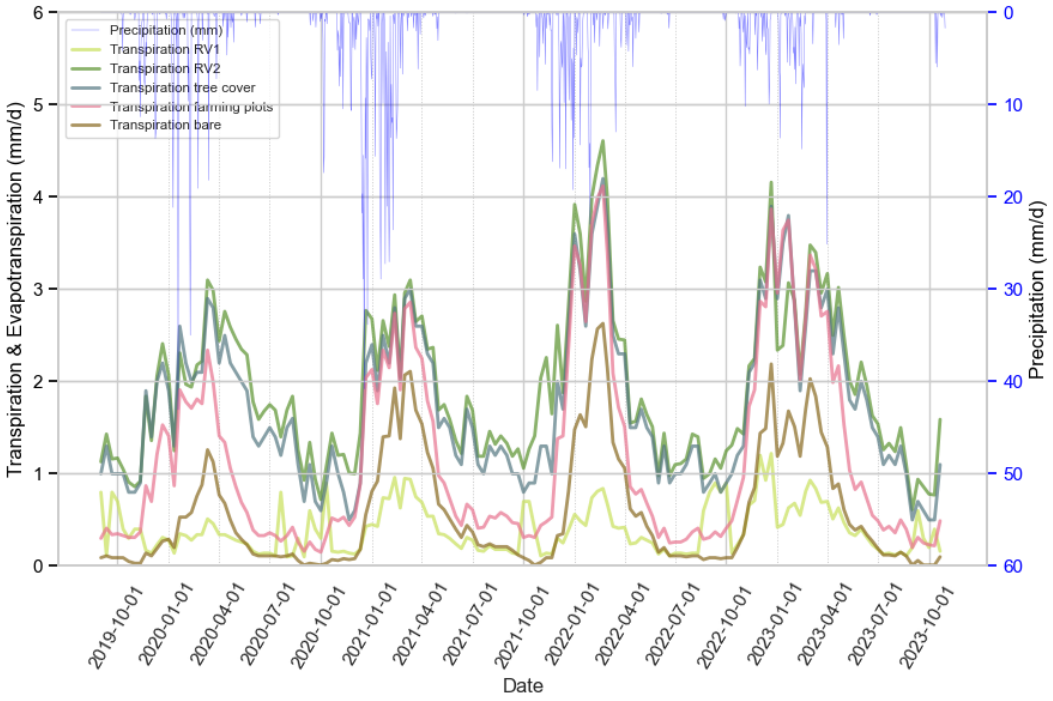


FIGURE C.3: The transpiration time series of the same five plots in the Shashani sub-catchment, with daily precipitation in the background. E.g. bare is brown and riparian vegetation is green.

C.2 Delineation

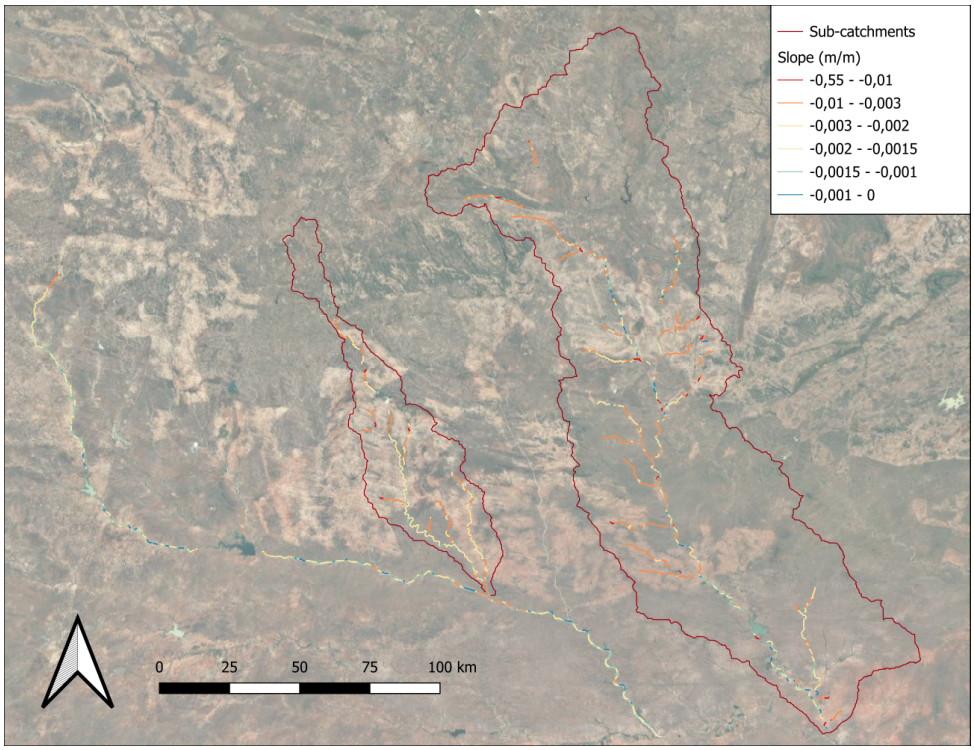
(a) (b)

FIGURE C.4: Example showing the delineation of a 100m wide sand river and its riparian zones. (a) Delineated sand river and its riparian zone by sentinel-2 NDVI (>0.42). A small riparian zone is harder and more ambiguous to capture. It looks like some farming plots have replaced riparian vegetation. (b) Google sattelite image and the delineated riparian zone by Canopy Height (>4m). Canopy Height results in more inconsistent blocks that sometimes overlap the sand river.

C.3 Spatial, statistical analysis

C.3.1 Calculating attributes

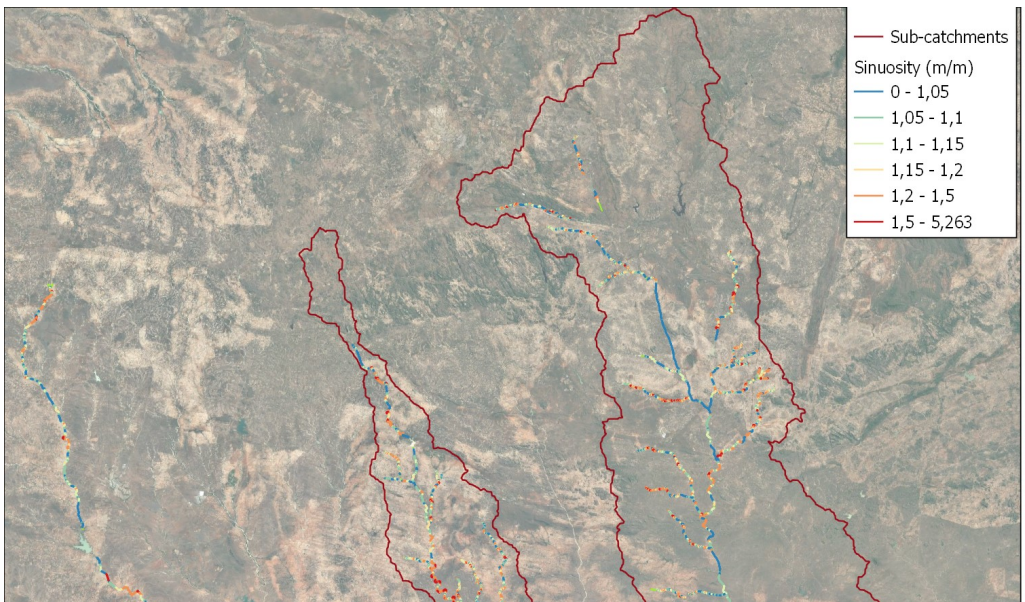
The calculated sinuosity attribute typically matches the level of meandering, especially for river reaches with a width above approximately 25m. This can be seen at the meandering river in the lower left corner of figure (C.6 (b)). For widths smaller than 25m, the sinuosity values match less well, as can be seen in the smaller attributes in the right of Figure (C.6 (b)).



(a)

(b)

FIGURE C.5: (a) Visualisation of the slope attribute (m/m). Each point is averaged over 4000m. (b) Zoomed in on the Shashani sub-catchment, including depth measurements.

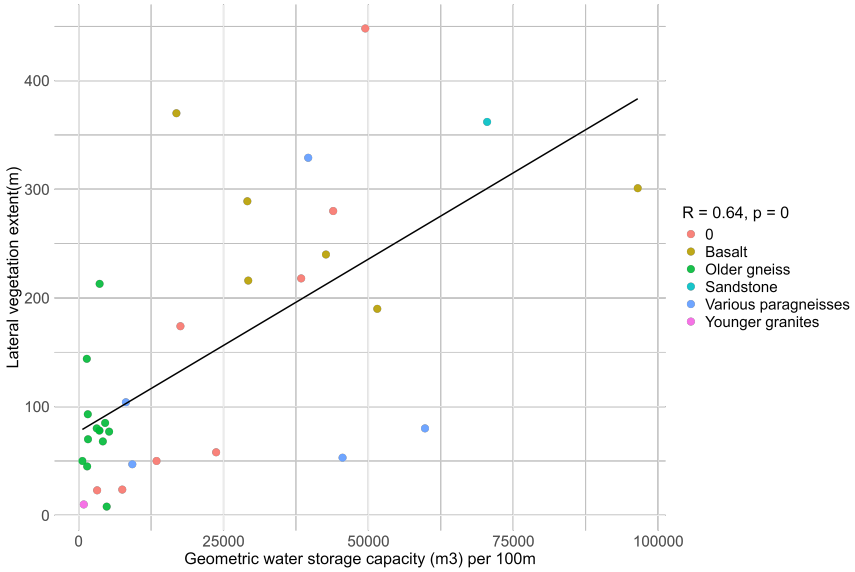


C.3.2 Statistical results

Relationships with depth and storage

Extra scatter plots are plotted here such as width, slope and versus riparian vegetation, as well as other coloring combinations and the bar plots of influences on riparian vegetation.

The relationship between riparian vegetation extent and water storage capacity is likely not completely linear. At large vegetation extents, the storage plateaued. Even larger vegetation extents occurred near depth measurements (up to 700m) which would lead to large overestimations. Furthermore, a triangular data gap can be observed, more clearly in the log-linear plot.



(a)

FIGURE C.7: Riparian vegetation versus geometric water storage capacity per 100m, assuming a specific yield of 0.15.

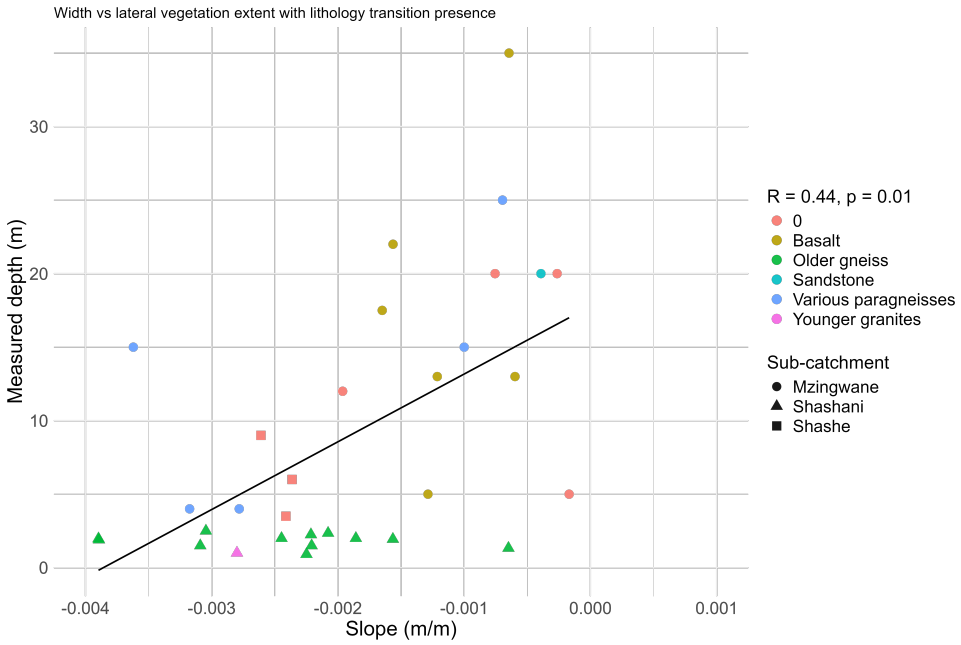


FIGURE C.8: Scatter plot of depth (m) versus slope (m/m, 2x800m resolution), colored by lithology type, shaped by sub-catchment. Pearson correlation coefficient is 0.51. The "0" lithology type means no data.

Zooming in on depths up to 7m (C.9), the correlation coefficient decreases (to r=0.44), suggesting high variation of vegetation for shallower rivers.

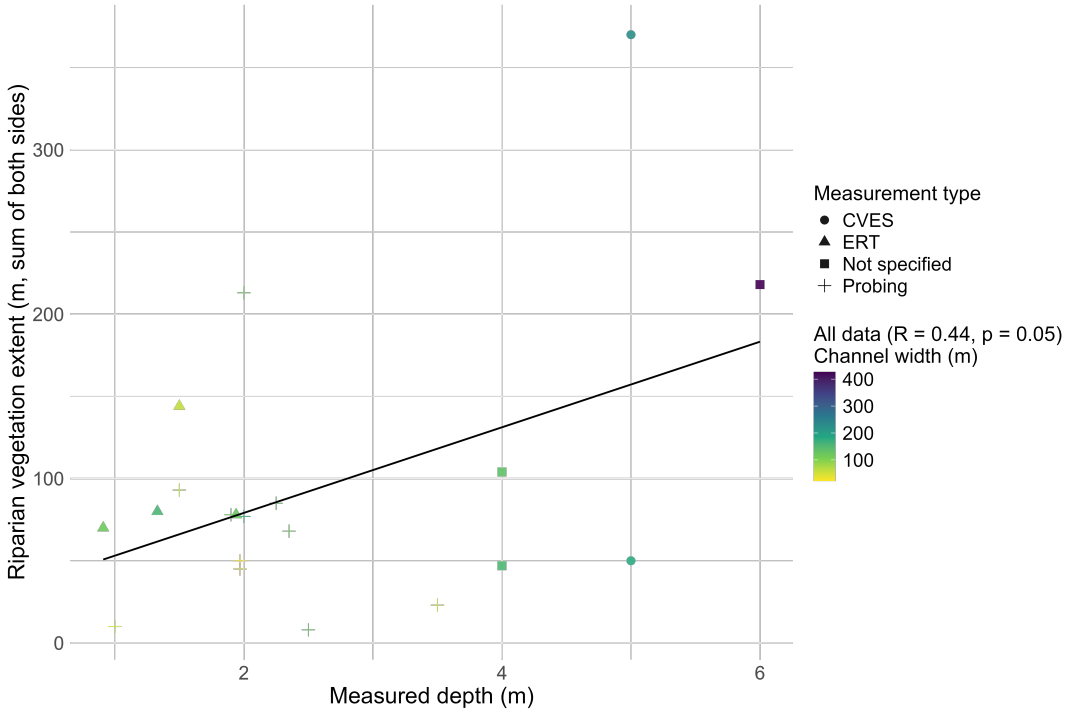


FIGURE C.9: Depth versus riparian vegetation extent for shallower depths up to 7m only

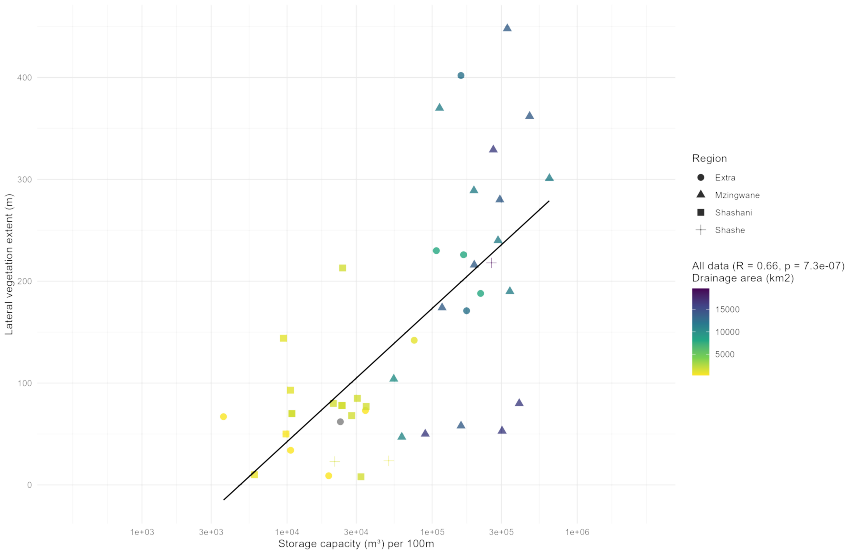
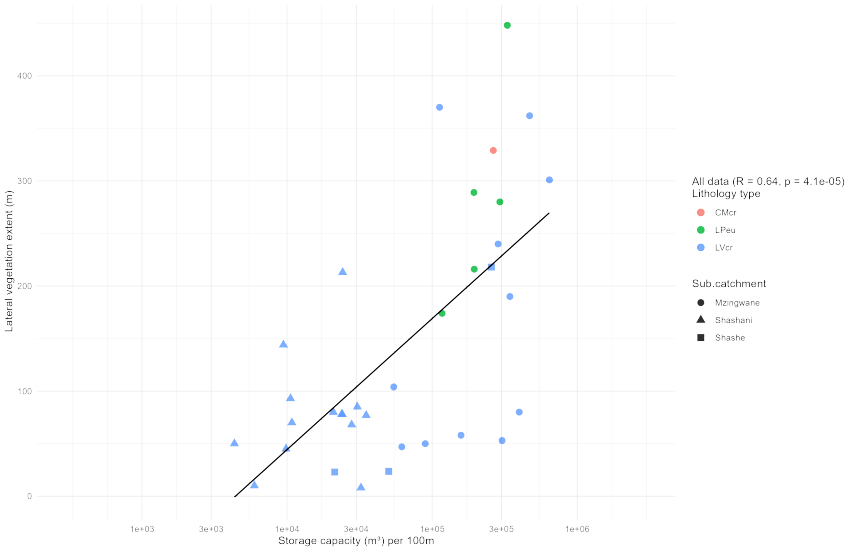
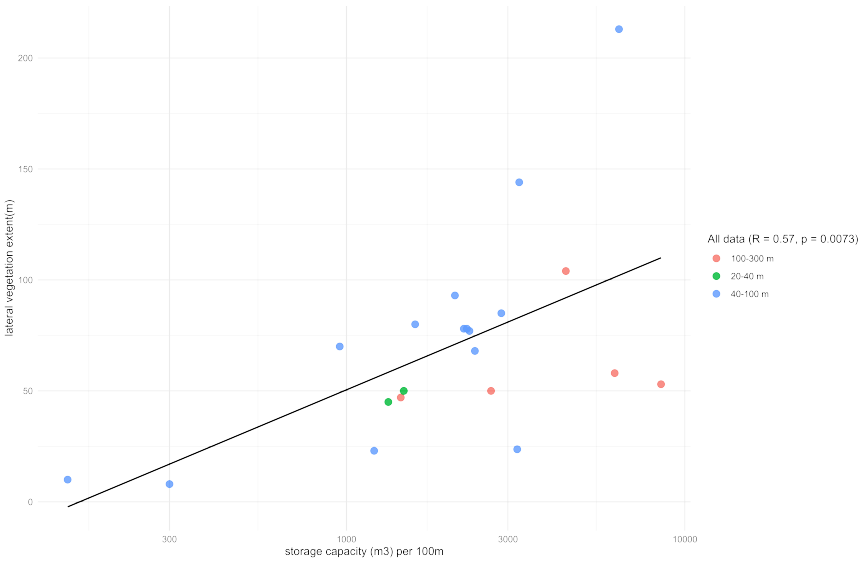


FIGURE C.10: Riparian vegetation versus storage capacity (width*depth*length) per 100m, with extra data points, colored with drainage area.



(a)



(b)

FIGURE C.11: Riparian vegetation versus water storage capacity (excluding specific yield) per 100m. (a) With logarithmic x-axis, colored with soil type. (b) Storage data limited to 10,000 cubic meter.

The drainage area versus riparian vegetation is plotted with a logarithmic x-axis in figure (C.12), colored by width class. Small jumps occur whenever a new drainage polygon starts. Big jumps occur whenever a tributary joined with significant upstream area. A data gap is visible at the 300-900m width class; rivers at 29000 km ^{2} have at least 50m riparian vegetation.

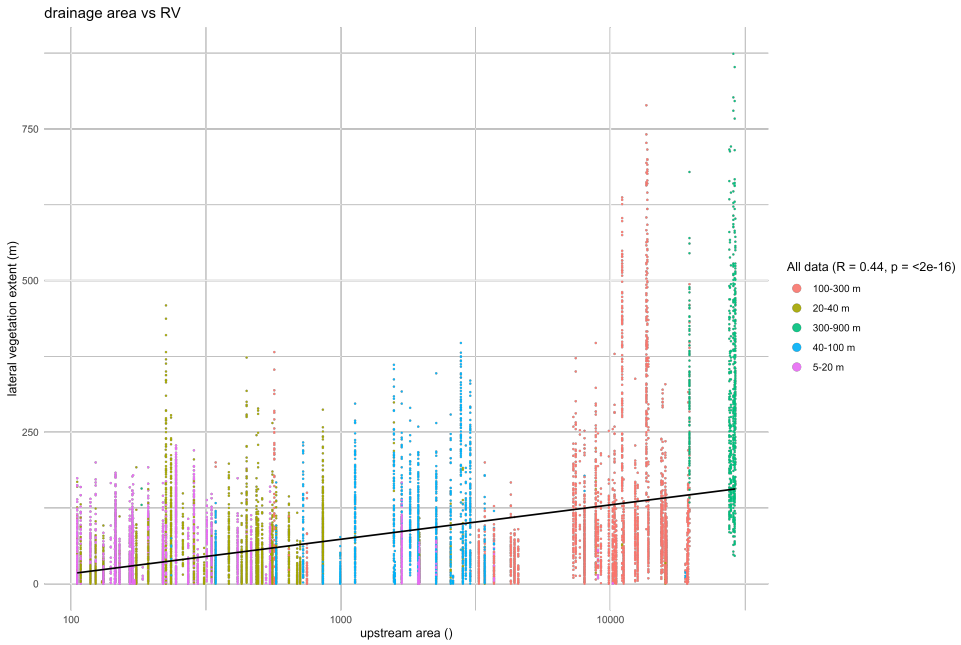


FIGURE C.12: Riparian vegetation versus drainage area in km ^{2}

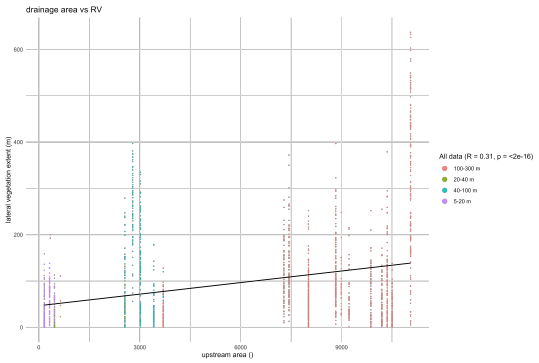


FIGURE C.13: Scatter plot of lateral vegetation extent versus drainage area (km ^{2})of the main stem of the Mzingwane

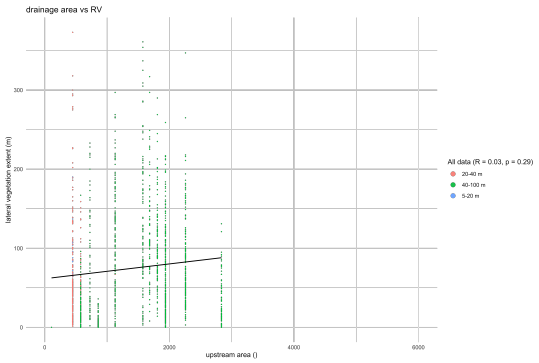
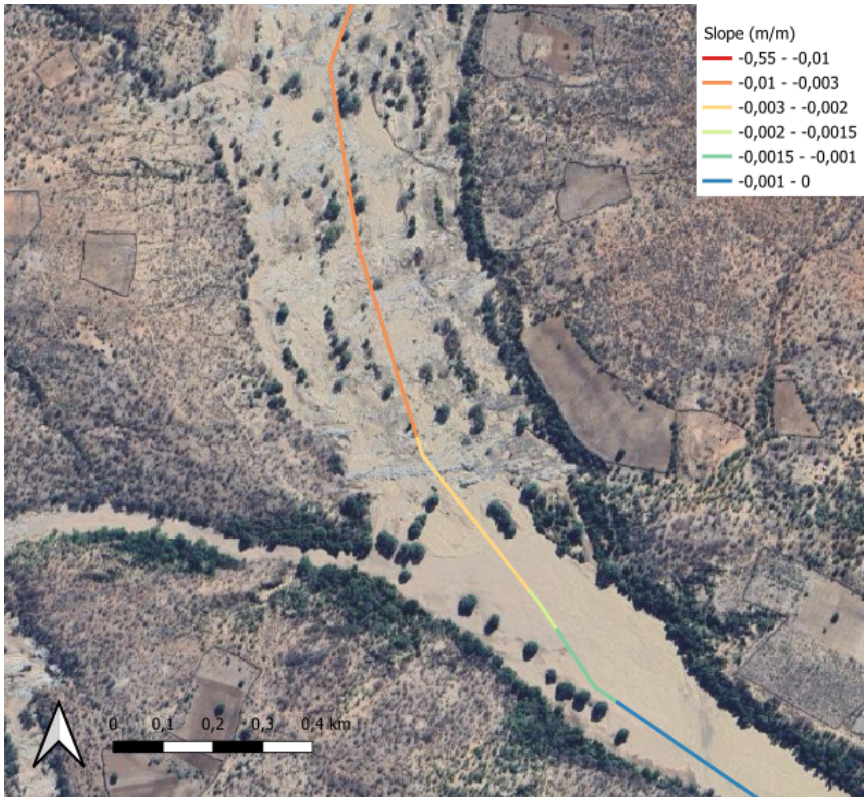


FIGURE C.14: Scatter plot of lateral vegetation extent versus drainage area (km ^{2}) of the main stem of the Shashani

Slope versus vegetation extent is depicted in figure (**??**. While slope was not significantly correlated with riparian vegetation extent, the average vegetation extent

is higher at gentle slopes (approximately 250m, >-0.03 m/m) than at steeper slopes (approximately 70m, <-0.03 m/m).



(a)

FIGURE C.15: One of many examples where steep slope comes with bedrock outcrops. When the slope becomes gentle, only smooth sand is visible.

Influences on vegetation

Width versus riparian vegetation is moderately correlated (r=0.54), shown in Figure (C.16). The predictive power does not seem large, but clear data gaps are present at the upper left and lower right quadrant. River sections with underlying basalt have the highest widths and highest lateral vegetation extents.

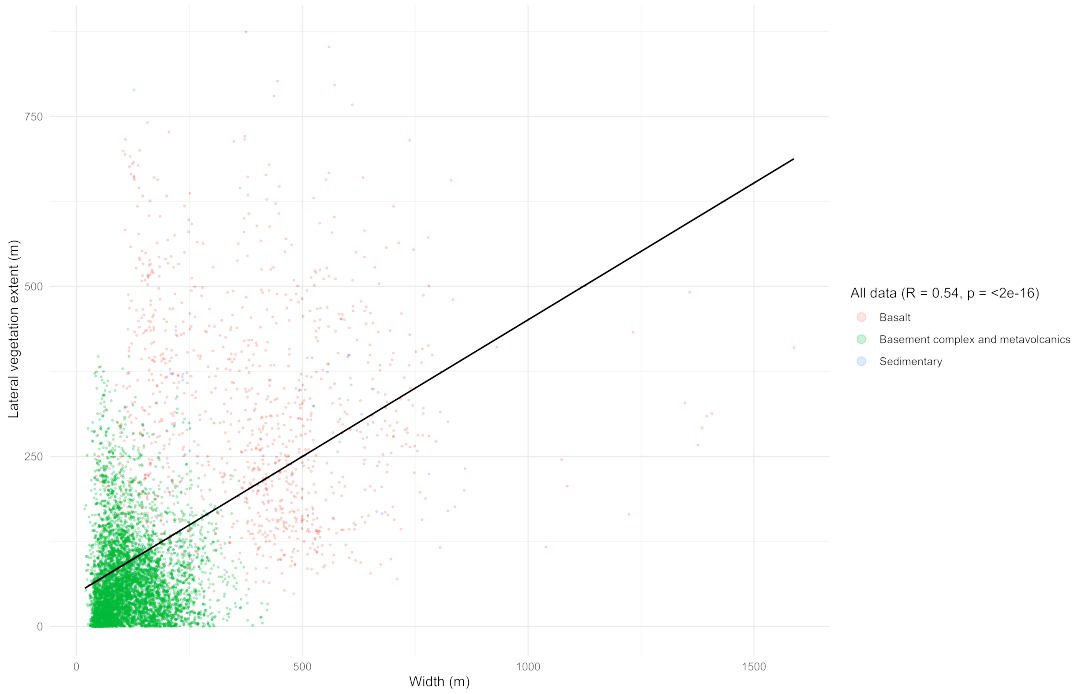


FIGURE C.16: Riparian vegetation versus width of the channel segment.

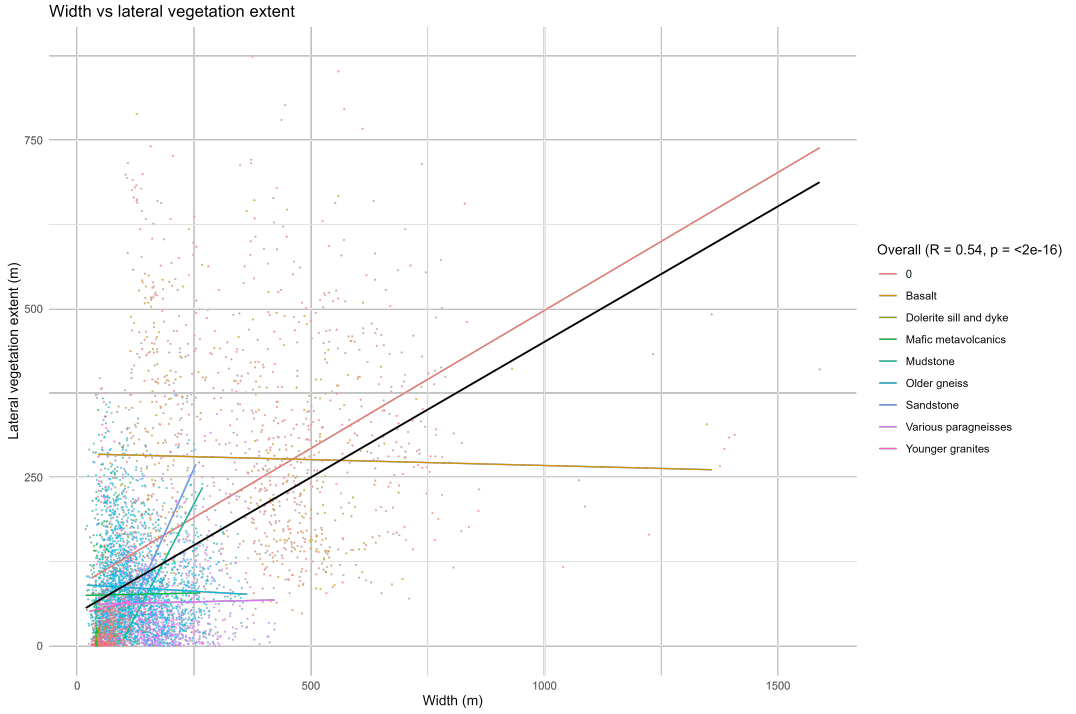


FIGURE C.17: Width versus riparian vegetation, colored by lithology, without river widths smaller than 40m

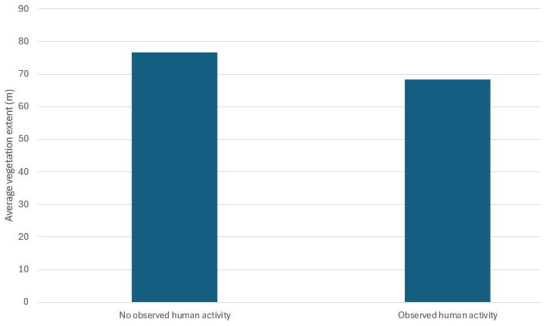


FIGURE C.18: Bar plot of the average vegetation extent (m) of river sections that are disturbed by human activities or not. All river sections were marked where it looked like human features had replaced riparian vegetation or medium to large scale irrigation was present. Features include mainly farming plots and roads.

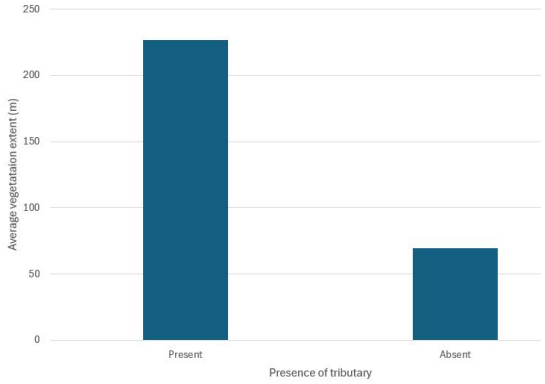


FIGURE C.19: Bar plot of the average vegetation extent (m) of river sections that are influenced by tributaries or not. All river sections were marked whose counting box was including a significant amount of a tributary's riparian vegetation.

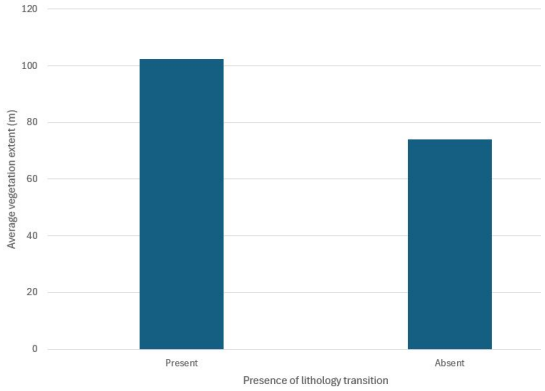


FIGURE C.20: Bar plot of the average vegetation extent (m) of river sections that are upstream or downstream near a lithology transition. Boundaries between lithologies were buffered to a width of 300m

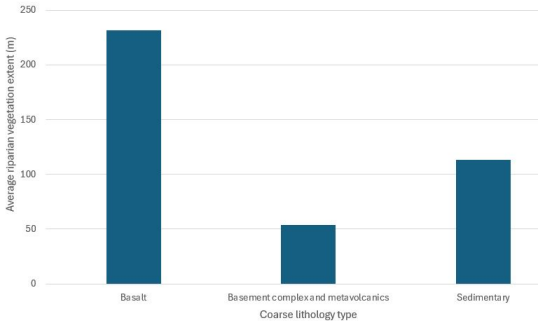


FIGURE C.21: Bar plot of the average vegetation extent (m) of river sections with different (coarse) lithology types. Basalt has the most riparian vegetation extent, the basement complex the least. Mind collinearity; all of the sedimentary region lies in the lower Mzingwane, and Basalt is either lower Shashe or lower Mzingwane.

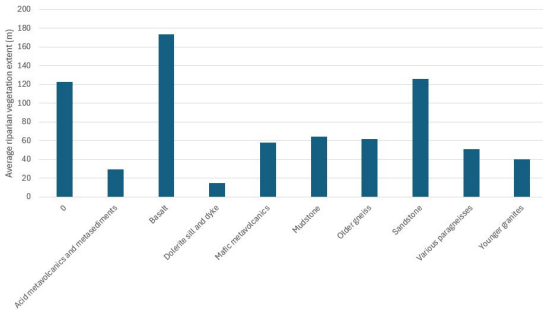


FIGURE C.22: Bar plot of the average vegetation extent (m) of river sections with different (georeferenced) lithology types. The "0" lithology type is no data.

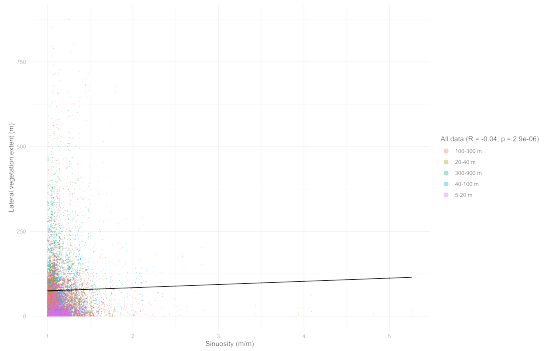


FIGURE C.23: Scatter plot of sinuosity versus riparian vegetation. The highest values of riparian vegetation, as well as the average, do not seem to occur at meandering rivers.

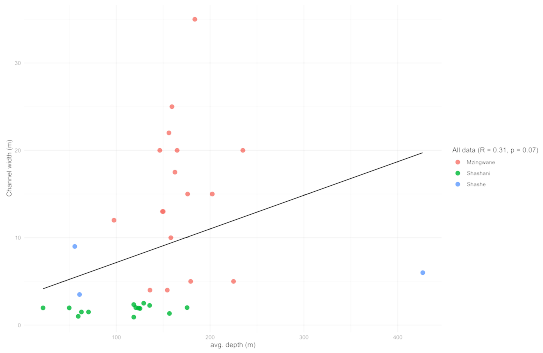


FIGURE C.24: Scatter plot of channel width (m) versus channel depth (m), colored by sub-catchment. A weak Pearson correlation coefficient is present of 0.31

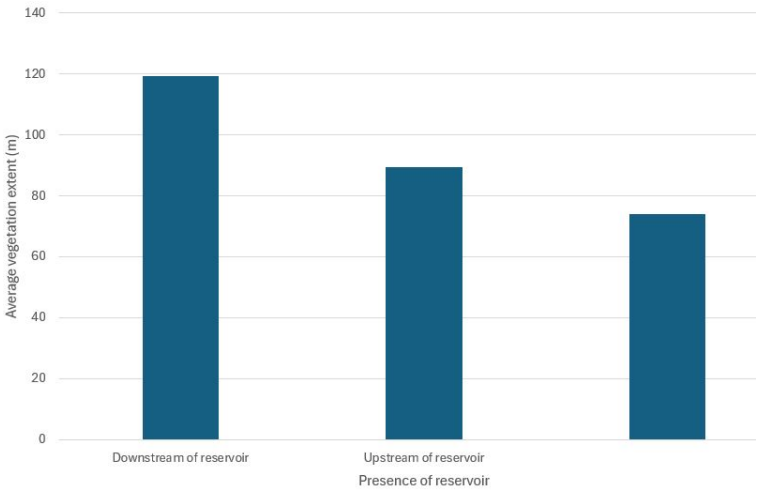


FIGURE C.25: Bar plot of the average vegetation extent (m) of river sections that are upstream, downstream of a reservoir or not near one. Distances taken range from 1000m from small reservoirs to 10000m at big reservoirs.

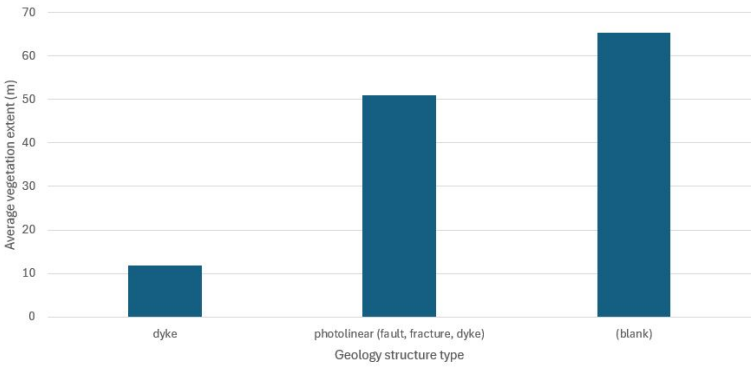


FIGURE C.26: Bar plot of the average vegetation extent (m) of river sections with and without geological structures. Dikes have less riparian vegetation, fault lines and other fractures slightly less.

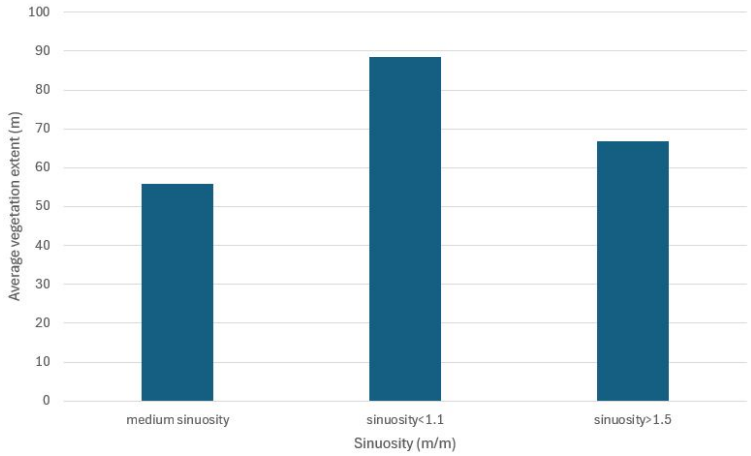


FIGURE C.27: Bar plot of the average vegetation extent (m) of straight rivers (sinuosity<1.1) or meandering rivers(sinuosity>1.5), or anything in between. Riparian vegetation of both river sides are included in the calculation.

C.3.3 Estimating and mapping water storage capacity.

Predicted vs actual depth by model 10 (C.28), shows one difference compared to the predicted water storage capacity, namely an underestimation at a smaller actual storage.

Figure C.29 displays the percentage of evaporation over geometric storage, showing the (un)consumed difference in another manner. In contrary to the unconsumed map, it shows that some yellow and green tributaries have unconsumed water availability.

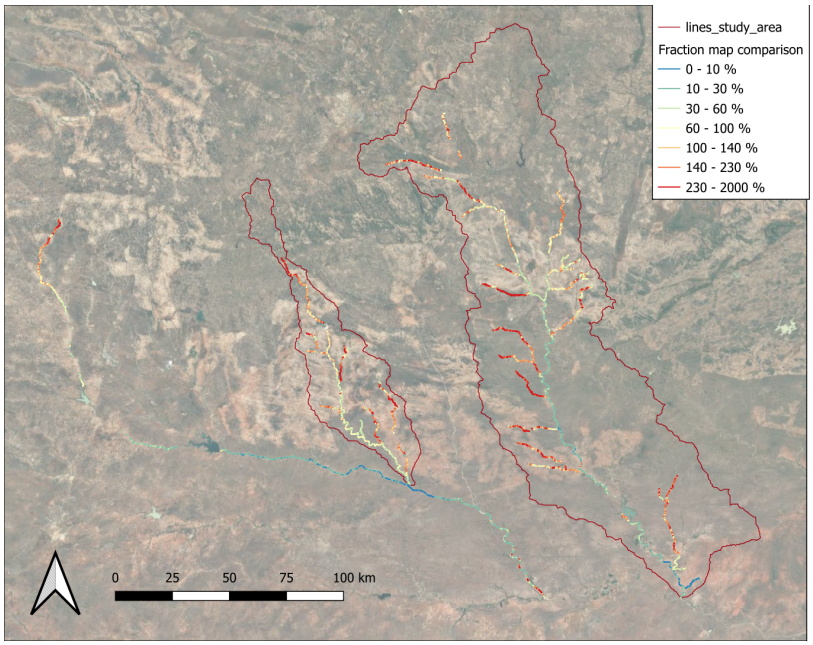


FIGURE C.29: Fraction linemap, displaying evaporation / geometric storage x 100%.

Figure C.30 shows the map with MLR estimated storage capacities with depths capped to 7 meters, similar to the maximum of suction pumps.

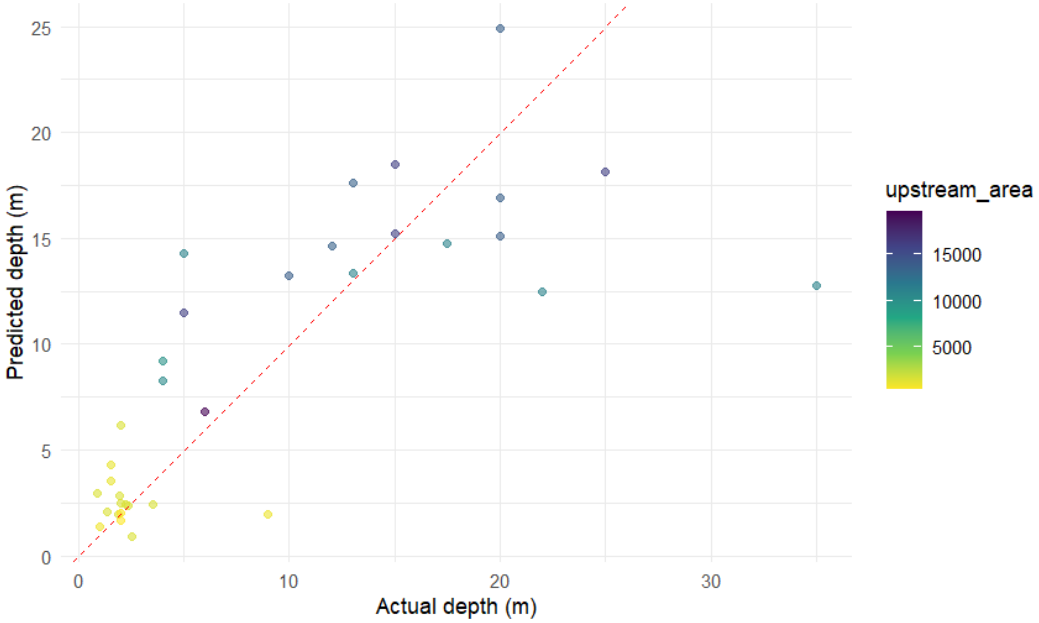


FIGURE C.28: Predicted vs actual depth

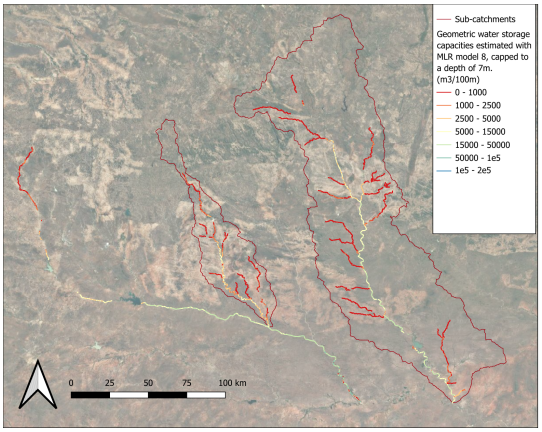


FIGURE C.30: Estimated geometric water storage capacities with MLR model 8, capped to 7 meters, assuming a specific yield of 0.15.

The model summary in Table C.1 provides information on the significance of each predictor variable in a multi-linear regression model with slope and interaction terms. The table can be read as follows. The Estimate column shows the regression coefficient for each predictor variable, representing average change in the response variable (depth) for a one-unit increase in the predictor variable, assuming all other variables are held constant. The t value is the estimate divided by the standard error. The p-value indicates the statistic significance (by comparing a t-statistic to a t-distribution, depending on the sample size). Upstream area and with*upstream area are most significant. Riparian vegetation is the most insignificant both with and without slope (p=0.99 and p=0.88), but completely removing it would increase the R-squared with 0.3 and the MAPE from 65% to 75%. When zooming in on smaller storages, however, all p values decrease and slope becomes the most insignificant with a high p value of 0.43.

TABLE C.1: Statistical Summary of Regression Model With Slope and Interaction Terms, Predicting Storage from Environmental Factors

Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	5.169e+00	6.356e+00	0.813	0.4235
width	-2.631e-02	4.309e-02	-0.611	0.5468
upstream_area	1.298e-03	5.434e-04	2.388	0.0245 *
RV_extent2	4.342e-04	4.786e-02	0.009	0.9928
slope_800m	9.125e+02	1.068e+03	0.854	0.4009
width:upstream_area	-2.438e-06	3.170e-06	-0.769	0.4487
width:RV_extent2	1.068e-04	2.649e-04	0.403	0.6901

TABLE C.2: Statistical Summary of Regression Model Without Slope, Predicting Storage from Environmental Factors

Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	1.853e+00	5.008e+00	0.370	0.7142
width	-2.076e-02	4.238e-02	-0.490	0.6282
upstream_area	1.336e-03	5.388e-04	2.479	0.0197 *
RV_extent2	6.937e-03	4.701e-02	0.148	0.8838
width:upstream_area	-2.785e-06	3.128e-06	-0.890	0.3811
width:RV_extent2	1.019e-04	2.635e-04	0.387	0.7020

TABLE C.3: Statistical Summary of Regression Model Trained On Storages Smaller Than 25000 m3/100m, Predicting Storage from Environmental Factors

Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	3.043e+00	2.278e+00	1.336	0.20287
width	-2.001e-02	1.508e-02	-1.327	0.20569
upstream_area	1.429e-03	4.389e-04	3.255	0.00575 **
RV_extent2	-2.818e-02	2.110e-02	-1.335	0.20305
slope_800m	-3.184e+02	3.942e+02	-0.808	0.43271
width:upstream_area	-6.310e-06	2.884e-06	-2.188	0.04616 *
width:RV_extent2	1.974e-04	1.197e-04	1.649	0.12134

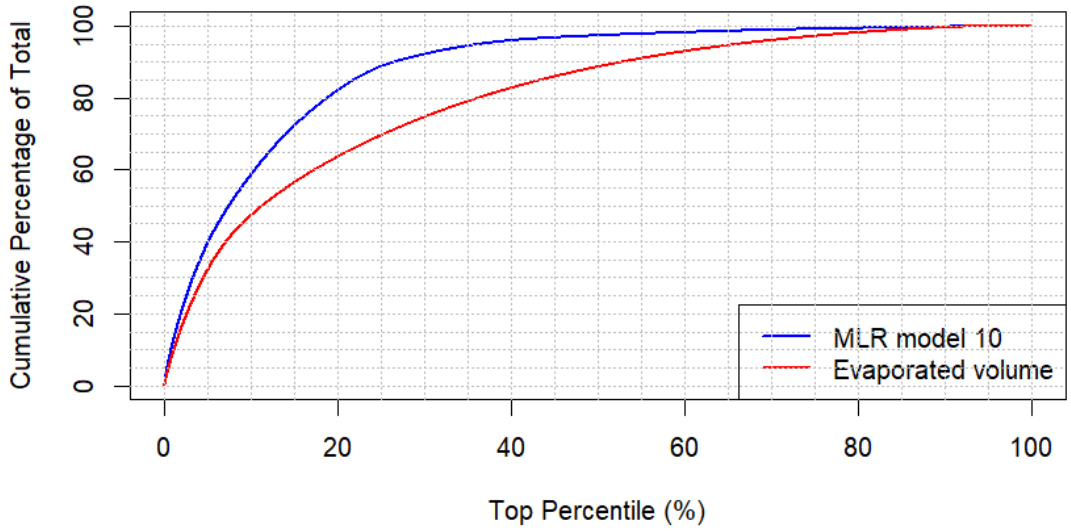


FIGURE C.31: Exceedence plot showing the distribution of the storage estimates. For example, the biggest 10% of rivers contain 55% of all water storage.

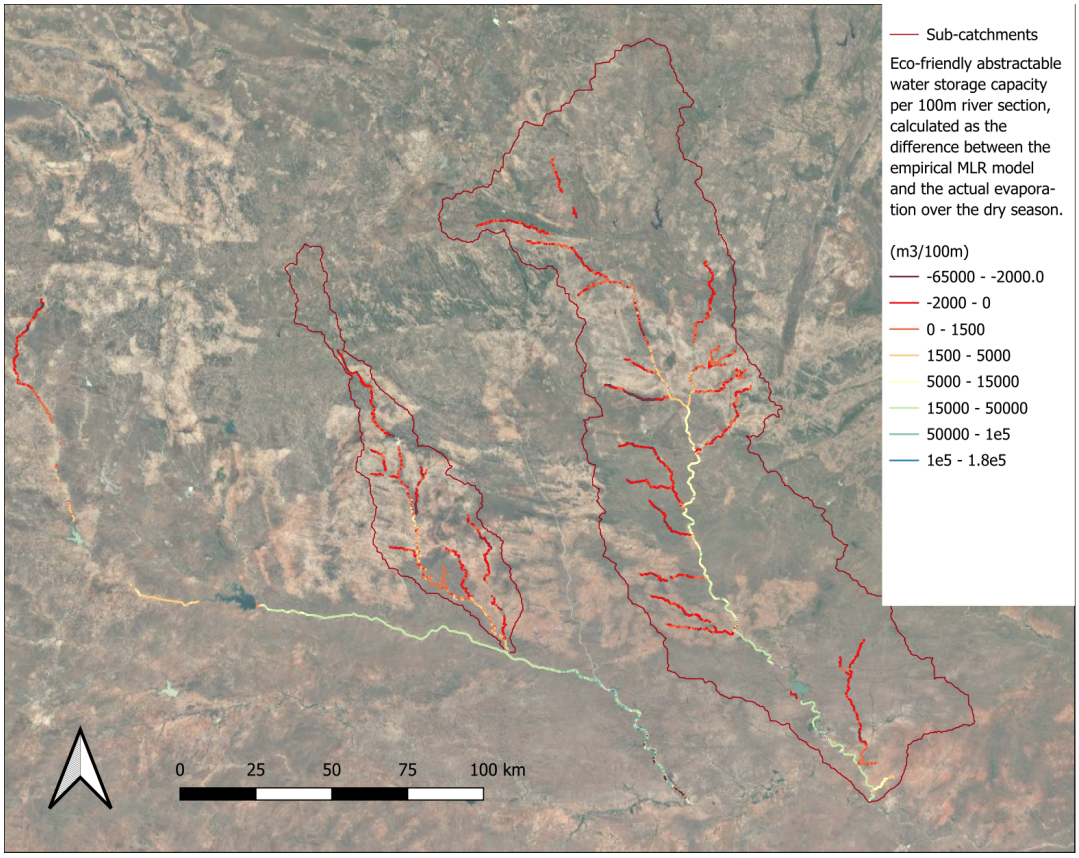


FIGURE C.32: Map of delineated sand rivers and the unconsumed abstractable water storage capacity for every 100m river section, calculated as the difference between the dry season evaporation and the empirically calculated storage volume with MLR model 8, assuming a specific yield of 0.15.

FIGURE C.33: Exception 1 with point bar and adjacent tree cover where dry season evaporation exceeds geometric water storage

FIGURE C.34: Exception 2 with point bar and adjacent tree cover where dry season evaporation exceeds geometric water storage

C.4 Exceptions where Dry season evaporation exceeds geometric water storage

These exceptions can be explained by mainly the inclusion of floodplain water in the dry season evaporation estimates, while the geometric storage calculation considers only the channel's area and depth. Other possible factors include subsurface flows feeding downstream riparian vegetation, gaining conditions in slightly hilly locations, increased evaporation due to subcrops raising water levels near the surface, data inaccuracies and including the evaporation of non-riparian vegetation.

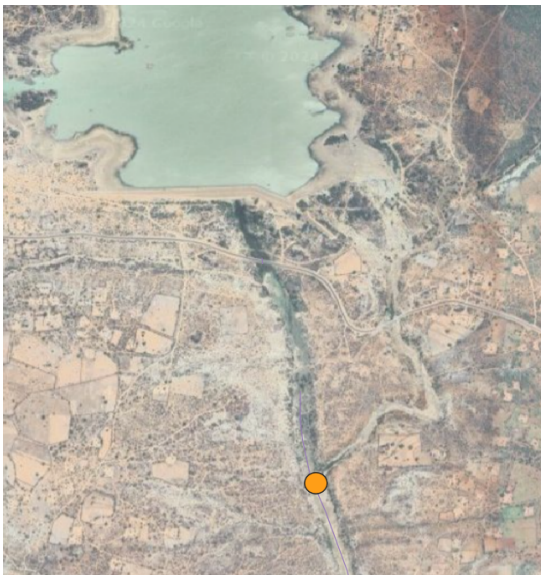


FIGURE C.35: Exception 3 with upstream reservoir where dry season evaporation exceeds geometric water storage

C.5 False positives and negatives

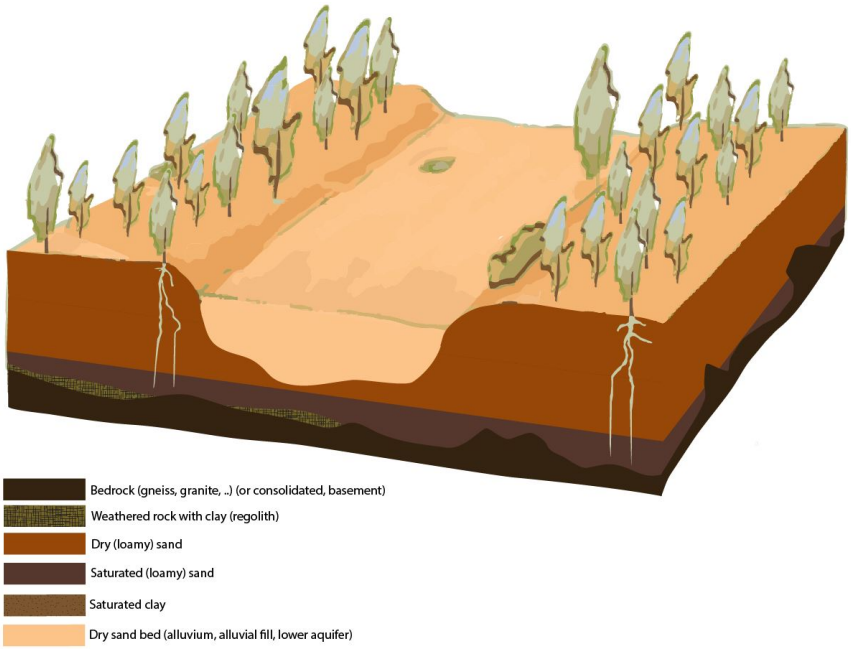


FIGURE C.36: A main suspected false positive is fast lateral seepage/recharge. While a lot of riparian vegetation may be present, water spreads laterally into the regolith of banks after the rainy season, accessible to only tap roots of vegetation. The water gradient line in blue shows that while tap roots can penetrate regolith and reach the water, a wellpoint may not. (Replace with a version that has regolith and water gradient)

C.6 Slope illustration

FIGURE C.37: Multi-resolution valley bottom flatness (MRVBF)

FIGURE C.38: Example of slope revealing many bedrock outcrops)

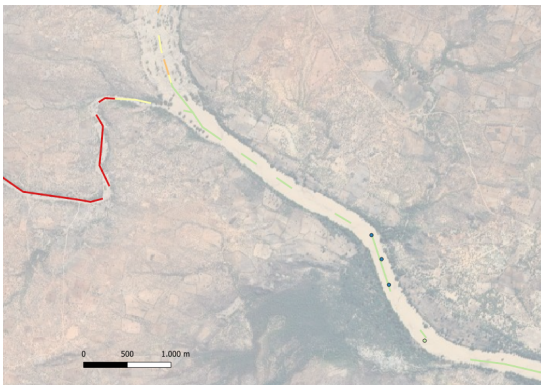


FIGURE C.39: Slightly zoomed out example showing the MLR model captured shallow depths at the bedrock outcrops. Both estimation methods would need validation for the transition in the middle)

Bibliography

A4store (2024). *Smallholder farming families Adapt African Alluvial Aquifers to Strengthen Their Own Resilience*. https://sandrivers.org/a4store/. Accessed: 2024-11-19.

Aerts, Jeroen et al. (2007). "Robustness of sand storage dams under climate change". In: *Vadose Zone Journal* 6.3, pp. 572–580.

Anderson, I. P. et al. (1993). "Physical Resource Inventory of the Communal Lands of Zimbabwe. An Overview. (+ maps) | ISRIC Library and map collection". https://library.wur.nl/WebQuery/isric/2277211.

Ashton, Peter J et al. (2001). "An overview of the impact of mining and mineral processing operations on water resources and water quality in the Zambezi, Limpopo and Olifants Catchments in Southern Africa". In: *Contract Report to the Mining, Minerals and Sustainable Development (Southern Africa) Project, by CSIR-Environmentek, Pretoria and Geology Department, University of Zimbabwe-Harare. Report No. ENV-PC* 42, pp. 1–362.

Benito, Gerardo et al. (2010). "Management of alluvial aquifers in two southern African ephemeral rivers: implications for IWRM". In: *Water Resources Management* 24, pp. 641–667.

Blaschke, Thomas et al. (2011). "Object-based image analysis for vegetation mapping and monitoring". In: *Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications* 2010, pp. 241–272.

Blatchford, Megan L et al. (2020). "Evaluation of WaPOR V2 evapotranspiration products across Africa". In: *Hydrological processes* 34.15, pp. 3200–3221.

Blok, T. et al. (2017). "Groundwater storage in sand rivers". Unpublished internship report. Acacia Water.

Bremer, K. (2022). "Detection of Sand Rivers using Remote Sensing, an Earth Observation–based identification method for ephemeral rivers In Sub-Saharan Africa". Unpublished report. IHE, Delft.

Chauruka, M. (2022). "A livelihoods approach to assessing the role of sand river aquifers for irrigation purposes". Unpublished thesis report. IHE, Delft.

Climate Engine (2023). *CHIRPS: Climate Hazards Group InfraRed Precipitation with Station Data*. https://support.climateengine.org/article/53-chirps.

Cuthbert, MO et al. (2016). "Understanding and quantifying focused, indirect groundwater recharge from ephemeral streams using water table fluctuations". In: *Water Resources Research* 52.2, pp. 827–840.

De Hamer, W et al. (2008). "Potential water supply of a small reservoir and alluvial aquifer system in southern Zimbabwe". In: *Physics and Chemistry of the Earth, Parts A/B/C* 33.8-13, pp. 633–639.

Dufour, Simon et al. (2019). "Tracing the scientific trajectory of riparian vegetation studies: Main topics, approaches and needs in a globally changing world". In: *Science of the total environment* 653, pp. 1168–1185.

Duker, A et al. (2020a). "Using nature-based water storage for smallholder irrigated agriculture in African drylands: Lessons from frugal innovation pilots in Mozambique and Zimbabwe". In: *Environmental Science & Policy* 107, pp. 1–6.

Duker, AEC et al. (2020b). "Shifting or drifting? The crisis-driven advancement and failure of private smallholder irrigation from sand river aquifers in southern arid Zimbabwe". In: *Agricultural Water Management* 241, p. 106342.

European Commission, Joint Research Centre (2013). *Soil Atlas of Africa*. Accessed: September 30, 2024. Luxembourg: Publications Office of the European Union. URL: https://esdac.jrc.ec.europa.eu/content/soil-map-soil-atlas-africa.

European Space Agency (2021). *WorldCover 10 m Land Cover Map for 2020 (v1.0)*. https://esa-worldcover.org. Accessed: 2024-11-27.

Frost, Peter (1996). "The ecology of miombo woodlands". In: *The miombo in transition: Woodlands and welfare in Africa* 266.

Hussey, Stephen W (2007). *Water from sand rivers. Guidelines for abstraction.*

Johansen, Kasper et al. (2011). "Automatic geographic object based mapping of streambed and riparian zone extent from LiDAR data in a temperate rural urban environment, Australia". In: *Remote Sensing* 3.6, pp. 1139–1156.

Kenyon, L. (2023) (2022). "Assessment of Sand River Morphology as a Function of Environmental Variables". Unpublished master's thesis.

Knight, Jasper (2022). "Geomorphology and landscapes of the Limpopo River system". In: *Landscapes and Landforms of Botswana*. Springer, pp. 287–298.

Lane, Emory W (1955). "Design of stable channels". In: *Transactions of the American society of Civil Engineers* 120.1, pp. 1234–1260.

Lang, Nico et al. (2023). "A high-resolution canopy height model of the Earth". In: *Nature Ecology & Evolution* 7.11, pp. 1778–1789.

Lasage, R et al. (2008). "Potential for community based adaptation to droughts: Sand dams in Kitui, Kenya". In: *Physics and Chemistry of the Earth, Parts A/B/C* 33.1-2, pp. 67–73.

Love, D. (2013). "Targeting the under-valued resource: an evaluation of the water supply potential of small sand rivers in the northern Limpopo Basin". Unpublished PhD report.

Love, D et al. (2011). "A water balance modelling approach to optimising the use of water resources in ephemeral sand rivers". In: *River Research and Applications* 27.7, pp. 908–925.

Love, David et al. (2010). "Changing hydroclimatic and discharge patterns in the northern Limpopo Basin, Zimbabwe". In: *Water sa* 36.3, pp. 335–350.

Mansell, MG et al. (2005). "An investigation of flows and losses within the alluvial sands of ephemeral rivers in Zimbabwe". In: *Journal of Hydrology* 314.1-4, pp. 192–203.

Marzouki, Anass et al. (2022). "Normalized Difference Enhanced Sand Index for desert sand dunes detection using Sentinel-2 and Landsat 8 OLI data, application to the north of Figuig, Morocco". In: *Journal of Arid Environments* 198, p. 104693.

Morin, Efrat et al. (2009). "Flood routing and alluvial aquifer recharge along the ephemeral arid Kuiseb River, Namibia". In: *Journal of Hydrology* 368.1-4, pp. 262–275.

Moulahoum, A. W. (2018). "Using Field Assessment and Numerical Modelling Tools to Optimize a Water Abstraction System in the Shashane Sand River Aquifer (Zimbabwe). Delft: UNESCO-IHE Institute for Water Education." Unpublished MSc Thesis.

Moyce, William et al. (2006). "Alluvial aquifers in the Mzingwane Catchment: their distribution, properties, current usage and potential expansion". In: *Physics and Chemistry of the Earth, Parts A/B/C* 31.15-16, pp. 988–994.

Mpala, SC et al. (2016). "The hydrology of sand rivers in Zimbabwe and the use of remote sensing to assess their level of saturation". In: *Physics and Chemistry of the Earth, Parts A/B/C* 93, pp. 24–36.

Mpala, Sibonakaliso C et al. (2020). "Modelling the water level of the alluvial aquifer of an ephemeral river in south-western Zimbabwe". In: *Hydrological Sciences Journal* 65.8, pp. 1399–1415.

Nguyen, Uyen et al. (2019). "Mapping vegetation types in semi-arid riparian regions using random forest and object-based image approach: A case study of the Colorado River Ecosystem, Grand Canyon, Arizona". In: *Ecological Informatics* 50, pp. 43–50.

Nissen-Petersen, Erik (1998). "Water from sand-rivers". In.

Owen, Richard et al. (2005). "Alluvial aquifers at geological boundaries: geophysical investigations and groundwater resources". In: *Groundwater and human development*. Balkema, pp. 233–246.

Owen, RJ (1989). "The use of shallow alluvial aquifers for small scale irrigation with reference to Zimbabwe." In.

Owen, R.J.S. (1994). "Water Resources for Small-scale Irrigation from Shallow Alluvial Aquifers in the Communal Lands of Zimbabwe". Unpublished MPhil Thesis.

— (2000). "Conceptual Models for the Evolution of Groundwater Flow Paths in Shallow Aquifers in Zimbabwe." Unpublished MPhil Thesis.

Pace, G et al. (2021). "Remote sensing depicts riparian vegetation responses to water stress in a humid Atlantic region". In: *Science of The Total Environment* 772, p. 145526.

Pastora Estrada, A.G. (2023). "Estimation of water consumption by riparian vegetation in a sand river aquifer; a remote sensing and open data approach in the Shashani river, Zimbabwe." Unpublished MSc thesis.

Petheram, C et al. (2019). "Dams, dam costs and damnable cost overruns". In: *Journal of Hydrology X* 3, p. 100026.

Rietkerk, Max et al. (2002). "Self-organization of vegetation in arid ecosystems". In: *The American Naturalist* 160.4, pp. 524–530.

Rwasoka, DT et al. (2011). "Estimation of actual evapotranspiration using the Surface Energy Balance System (SEBS) algorithm in the Upper Manyame catchment in Zimbabwe". In: *Physics and Chemistry of the Earth, Parts A/B/C* 36.14-15, pp. 736–746.

Saveca, Paulo Sérgio Lourenço et al. (2022). "Assessing groundwater dynamics and hydrological processes in the sand river deposits of the Limpopo River, Mozambique". In: *Frontiers in Water* 3, p. 731642.

Scanlon, Bridget R et al. (2006). "Global synthesis of groundwater recharge in semiarid and arid regions". In: *Hydrological Processes: An International Journal* 20.15, pp. 3335–3370.

Stella, John C et al. (2013). "Riparian vegetation research in Mediterranean-climate regions: common patterns, ecological processes, and considerations for management". In: *Hydrobiologia* 719, pp. 291–315.

Tran, Bich Ngoc et al. (2023). "Uncertainty assessment of satellite remote-sensing-based evapotranspiration estimates: a systematic review of methods and gaps". In: *Hydrology and Earth System Sciences* 27.24, pp. 4505–4528.

Twidale, CR (2004). "River patterns and their meaning". In: *Earth-Science Reviews* 67.3-4, pp. 159–218.

Vaezi, Ali Reza et al. (2017). "Contribution of raindrop impact to the change of soil physical properties and water erosion under semi-arid rainfalls". In: *Science of the Total Environment* 583, pp. 382–392.

Walker, David et al. (2018). "Alluvial aquifer characterisation and resource assessment of the Molototsi sand river, Limpopo, South Africa". In: *Journal of Hydrology: Regional Studies* 19, pp. 177–192.

Walker, David et al. (2019). "Ephemeral sand river flow detection using satellite optical remote sensing". In: *Journal of Arid Environments* 168, pp. 17–25.

Yamazaki, Dai et al. (2023). *MERIT DEM: Multi-Error-Removed Improved-Terrain Digital Elevation Model*. https://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/. Accessed: YYYY-MM-DD.

Zaimes, George et al. (2007). *Understanding Arizona's riparian areas*. College of Agriculture and Life Sciences, University of Arizona (Tucson, AZ).

Zimba, Henry et al. (2023). "Phenophase-based comparison of field observations to satellite-based actual evaporation estimates of a natural woodland: miombo woodland, southern Africa". In: *Hydrology and Earth System Sciences* 27.8, pp. 1695–1722.