GROUNDWATER MANAGEMENT INSTITUTE

ƵUTARI
IMPACT. ENGINEERED.

PEGASYS
CHANGING LIVES CHANGING WORLDS

L2K2 Consultants

SADC-GMI – Groundwater Drought Risk Interventions (GMI-GDRI)

Assessment of surface water availability

August 2020

Boxes, Tables, Figures, Maps, Photos and Illustrations as specified

This report is accessible online through SADC-GMI website: www.sadc-gmi.org
Citation: SADC-GMI, (2020). *Assessment of Surface Water Availability.* SADC GMI report: Bloemfontein, South Africa.

EXECUTIVE SUMMARY

To determine population hotspots in SADC that are most in need of domestic water supply interventions, and to assess the viability of these proposed interventions, it is necessary to first quantify the surface water availability and risk on a regional scale. Given the project's restraints, it was agreed upon to follow a GIS-based methodology. An initial assessment pinpointed various **global precipitation, streamflow** and **catchment runoff raster datasets** that are both reliable and freely available (see *Draft Summary Review Report (SADC-GMI, 2020))*. As part of this report, these datasets were validated against global gauge datasets of discharge, runoff and rainfall. This validation process showed **WaterGAP v2.2** to be the most reliable dataset for discharge and runoff, and **WorldClim v2.1** the most reliable for rainfall. Statistical indices based on hydro-meteorological data are commonly used to quantify droughts and their severity, and as such, statistical analyses of the datasets was undertaken. For runoff, discharge and rainfall, the following indices were calculated through time series analyses: **Mean annual values, Seasonality, Index of Seasonal Variation, Coefficient of Variation** and **Runoff Coefficient**. Following this, the indices were **normalised** and **weighted**, and a **sensitivity analysis** was performed to determine the impact of different indices on the combined **surface water risk index**, and the final surface water risk map. A **qualitative validation** process showed that the surface water risk map correlated well with existing drought maps and reports throughout SADC. The final surface water risk map and accompanying index maps will be used to identify the surface water interventions for the priority areas.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

Acronym	Definition
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station
CRU	Climate Research Unit
CoV	Coefficient of Variation
DRC	Democratic Republic of the Congo
FAO	Food and Agriculture Organization
GDR	Groundwater Drought Risk
GDRI	Groundwater Drought Risk Intervention
GIP	Groundwater Information Portal
GIS	Geographic Information System
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GRDC	Global Runoff Data Centre
GSIM	Global Streamflow Indices and Metadata Archive
GSWP	Global Soil Wetness Project
MCA	Multi-Criteria Analysis
NASA	National Aeronautics and Space Administration
SADC-GIP	SADC-Groundwater Information Portal
SADC-GMI	Southern African Development Community Groundwater Management Institute
SADC	Southern African Development Community
SPI	Standard Precipitation Index
WMO	World Meteorological Organisation
WWF	World Wide Fund for Nature

1 INTRODUCTION

1.1 Background

The Southern African Development Community Groundwater Management Institute (SADC-GMI) is implementing the project: *Assessment of Groundwater Resources Development Priority Intervention Areas in the Southern African Development Community (SADC) Region (SADC GMI-GDRI),* which seeks to bring the role of groundwater in securing water supply during periods of droughts to the forefront and to provide for proactive planning, recommendations and management of groundwater and surface water systems. The project aims to identify areas that are prone to drought in the SADC region and provides information on groundwater and surface water resource availability.

The project makes use of existing geospatial, hydro-meteorological and hydrogeological datasets and entails a practical assessment of the groundwater and surface water resources which can be quickly mobilised to support sustainable domestic water supply investments in areas with high groundwater drought risk and limited access to safe domestic water supply. The study will eventually identify the most adequate and cost-effective infrastructure interventions in the areas in most need.

1.2 Purpose of this Report

To identify areas of priority water supply interventions, a Geographic Information System (GIS)-based approach is being followed. This approach essentially consists of three distinct components: a multi-criteria analysis (MCA) to determine population vulnerability hotspots, a revised groundwater drought risk (GDR) analysis and a surface water availability assessment.

This report focuses on the surface water availability assessment. It discusses the datasets used in the assessment, describes the methodology which was followed to generate a surface water risk map and presents the outcome of the analysis. This entailed the following key tasks:

- data collection of time series raster data;

- validation of raster data using point data;

- development of surface water indices;

- normalization of surface water indices;

- weighting of normalized surface water indices to generate a surface water risk map

2 ASSESSING SURFACE WATER AVAILABILITY

Droughts can arise from a range of hydrometeorological drivers which suppress precipitation and/or limit surface water and groundwater availability, causing significantly drier conditions than normal, and leading to water shortage (Svoboda & Fuchs, 2016). Droughts can be characterized in terms of location, severity and duration. Drought indices are typically used to quantify hydrometeorological information and to ultimately identify locations, severity and duration of droughts (Nagarajan, 2009).

Figure 2-1 illustrates the methodology that was followed in assessing the availability of surface water and deriving the surface water drought risk map, and the following chapters in this report will follow the same methodology.

Global precipitation, streamflow and catchment runoff time series datasets were collected and validated (Section 0). Global GIS delineated catchment data was also collected. These global datasets were selected based on the following criteria:

- No financial contributions required (freely available)
- Validated and/or calibrated with observed data (not only using satellite data)
- Covering all or the majority of SADC countries
- Data extending over a period of at least 30 years
- References in peer reviewed journals
- Credible data custodians

From this data, relevant statistical indices were calculated to quantify hydro-meteorological characteristics at appropriate scales (Section 4). These indices included Mean Annual Values, Seasonality, Index of Seasonal Variability, Coefficient of Variation and Runoff Coefficient. These indices were calculated per catchment unit, at a scale that was agreed upon in Section 3.

The indices were subsequently normalized (Section 5) and weighted (Section 6) to produce a combined surface water drought risk index.

Finally, the combined surface water drought index was used to produce a surface water risk map, that was validated against other drought risk maps (Section 7).
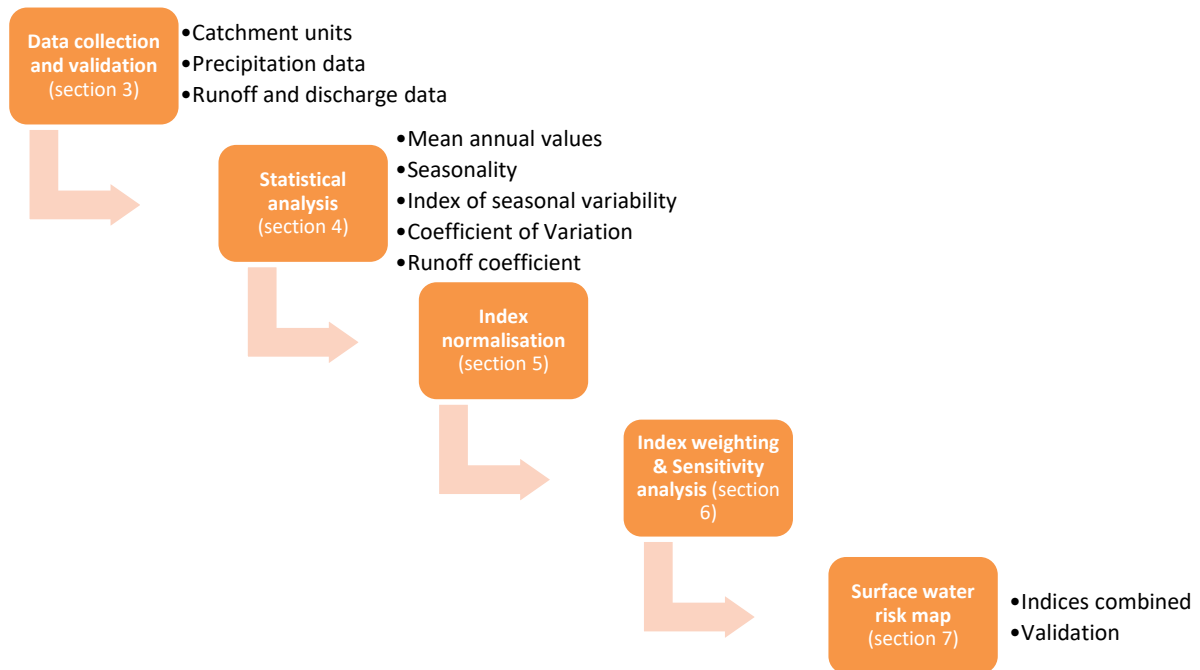
Figure 2-1: Methodology followed to derive the surface water drought risk map

3 DATA COLLECTION AND VALIDATION

Taking into consideration the limited timeframe of the project, applicable and freely-available global datasets at an appropriate quality, scale, recency/date, format and projection were used. Details regarding the various datasets investigated and evaluated based on the above criteria are provided in the "*Draft Summary Review Report*" (SADC, 2020) and summarized in the following sections.

Note: Some datasets, including spatial data, are used across all three analysis components as described in Section 1.2 - i.e. some data layers are relevant to vulnerability mapping, the revised GDR analysis and the surface water assessment. The use of the data layers is thus not exclusive to any one component of the project.

3.1 Catchment Units

Precipitation, runoff and discharge global data comes in the form of raster datasets, at varying scales. In order to combine these datasets to produce a risk map, these raster datasets must be processed to a uniform scale. Given the nature of surface water and catchment hydrology, 'catchment unit' polygons are used to create uniformity for the statistical analysis. Thus, the statistical analysis presented in Section 4 will be done per catchment unit.

Hydrological data and maps based on **SH**uttle **E**levation **D**erivatives at multiple **S**cale (HydroSHEDS) is a mapping product that provides hydrographic information for regional and global-scale applications. HydroSHEDS has been developed by the Conservation Science Program of World Wildlife Fund (WWF), in partnership and collaboration with the U.S. Geological Survey (USGS); the International Centre for Tropical Agriculture (CIAT); The Nature Conservancy (TNC) and others. HydroSHEDS is based on high-resolution elevation data obtained from the Shuttle Radar Topography Mission (SRTM) (Linke, et al., 2019).

Underpinning the HydroSHEDS database are amongst others the HydroATLAS compendium, the HydroBASINS watershed shapefiles and the HydroRIVERS river network.

HydroATLAS provides a fully-global data compendium that gathers and presents a wide range of hydro-environmentally relevant characteristics at both sub-basin and river scale.

HydroRIVERS provides a global river network delineation derived from HydroSHEDS data at 15 arc-second resolution.

HydroBASINS presents a series of polygon layers that were derived from HydroSHEDS data at 15 arc-second resolution and that depict watershed boundaries and sub-basin delineations at a global scale (Lehner, 2014). These sub-basins provide a global coverage of consistently sized and hierarchically nested catchment areas at different scales (from tens to millions of square kilometers), supported by a coding scheme that allows for analysis of watershed topology such as up- and downstream connectivity. A level 1 catchment distinguishes the continent, level 2 splits the continents into 9 sub-units and at level 3 the largest river basins of each continent start to break out. From level 4 onwards the largest river basins are broken down into the tributaries using high resolution elevation data (Lehner, 2014) up to level 12. From the HydroBASINS dataset, catchments were extracted based on level 7 and level 8 sub-basins respectively.

Table 3-1 shows the number of catchments per SADC country. The level 8 catchments were considered more appropriate for this analysis, due to its higher resolution, and will ensure data quality without compromising on computation time (Figure 3-1). The level 8 sub-basins are referred to as "catchment units" in this report. A typical catchment unit is illustrated in Figure 3-2.

Table 3-1: HydroBASINS level 7 and level 8 sub-basins per country

Country	Number of level 8 sub-basins	Number of level 7 sub-basins
Angola	1840	599
Botswana	902	318
Comoros	3	3
Democratic Republic of the Congo	3392	873
Lesotho	89	24
Madagascar	902	264
Malawi	224	90
Mozambique	1330	427
Namibia	1238	441
South Africa	1829	618
Swaziland	34	15
Tanzania	1498	374
Zambia	1264	419
Zimbabwe	657	187
Total	**15202**	**4652**

Figure 3-1: Catchment Level 8 HydroBASIN sub-basin
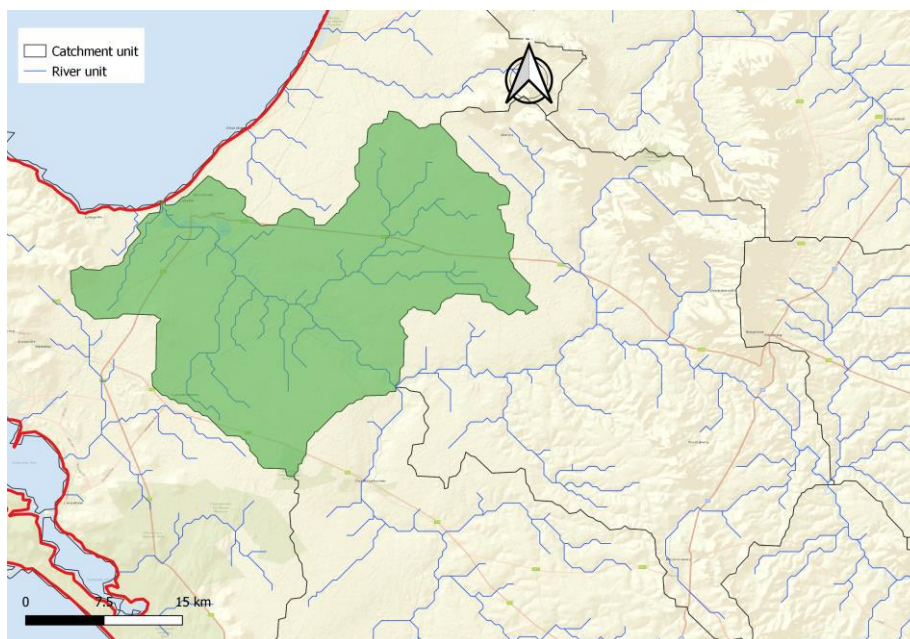
Figure 3-2: Catchment unit (Level 8 HydroBASIN sub-basin)

3.2   Precipitation data

A number of global precipitation datasets ranging from observed rainfall records at point locations to gridded estimates of rainfall from satellite-derived or advanced meteorological estimation methods are readily available. The main datasets which were evaluated are briefly summarized in Table 3-2. A more extensive table is attached as Appendix A, where references and related comments are also provided.

In the following sections, the datasets are briefly discussed - specifically their relevance for use in the assessment of surface water availability. A distinction is made between primary datasets and validation datasets.

Table 3-2: Summary of global precipitation datasets which were considered

Data Type	Dataset	Years available	Temporal resolution	Spatial resolution	Use in this project
Gauge-based	GPCC	1901-2010	Daily, Monthly	0.5°x 0.5°	Possible validation
	CRU	1901- near present	Monthly	0.5°x 0.5°	Possible validation
Satellite-based	GPCP	1979 – 2010	Daily, Monthly	2.5°x 2.5°	Possible primary rainfall dataset
	CHIRPS	1981-2018	Daily, monthly	0.05° x 0.05°	Possible primary rainfall dataset
	WorldClim	1960-2018	Monthly	0.5°x 0.5°	Possible primary rainfall dataset

3.2.1 Primary datasets

3.2.1.1 WorldClim-time series raster

The WorldClim database (Fick & Hijmans, 2017) is a database of interpolated gridded global climate surfaces at a spatial resolution of 0.5°. It is considered one of the most popular global datasets providing invaluable data for data-sparse areas (Wango, et al., 2018; Fick & Hijmans, 2017). WorldClim v1.4 contains average monthly climatic gridded data for the period between 1960 to 1990, while historical monthly data from 1960 to 2018 are available from the updated WorldClim v2.1 dataset.

The WorldClim v2.1 model used data from the most recent Climate Research Unit gridded Time Series (CRU TS-4.03) dataset from the Climate Research Unit (CRU) at the University of East Anglia for bias correction. The CRU is widely recognised as one of the world's leading institutions concerned with the study of natural and anthropogenic climate change (Harris, et al., 2020).

WorldClim employs satellite-derived (such as elevation and vegetation cover) and gauge-based data. (observation station data is interpolated using thin-plate smoothing spline algorithms and combined with the satellite-derived database). The WorldClim database includes information from 47,554 precipitation stations, which were used for validation (Fick & Hijmans, 2017) and interpolation with satellite data to create a complete dataset. According to Wangi et al (2018), the WorldClim datasets offers acceptable correlation to station data including temporal and seasonal variation. Data uncertainties mainly occurred in areas with sparse station data as well as in areas with high variation in elevation (Hijmans, et al., 2005).

3.2.1.2 CHIRPS

CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) is a satellite-derived dataset supported by funding from the USAID, NASA and NOAA. The CHIRPS dataset used interpolation techniques along with long periods of precipitation estimates based on infrared Cold Cloud Duration observations (Funk, et al., 2015). The CHIRPS algorithm applied blending methods between satellite-derived information, gauge information and the infrared Cold Cloud Duration observations to create a 35+ year quasi-global rainfall dataset which spans between 50°S to 50°N including all longitudes. The dataset has a high spatial resolution of 0.05° and presents a daily, pentadal and monthly rainfall time series from 1981 to 2018. CHIRPS data have been applied to support drought monitoring as well as to analyse shifts in precipitation in numerous African countries, including data sparse areas such as the Sahel (Dinku, et al., 2018; Badr, et al., 2016; Funk, et al., 2015).

3.2.1.3 GPCP

The most widely recognised global merged dataset (Sun et al., 2017) is the Global Precipitation Climatology Project (GPCP) dataset, first released in 1997. The GPCP is based on sequential combination of microwave, infrared as well as gauge data. Satellite data is obtained from the National Oceanic and Atmospheric Administration (NOAA). The algorithm entails that various satellite precipitation datasets are merged e.g. the Geostationary Operational Environmental Satellites Precipitation Index (GPI), the Outgoing long-wave radiation precipitation index (OPI) and the Special Sensor Microwave/Imager (SSM/I). The derived dataset

merges satellite data with rain gauge data and adjusts the satellite estimates to the gauge bias (Sun et al., 2017) The GPCP dataset has a spatial resolution of 2.5° and contains monthly data from 1979 to the near present. According to Wang (2020), the GPCP is useful in model validation as well as global precipitation analysis. It has been used and referenced in a number of studies and journals, and used extensively for studies in SADC (Driver, 2014; Masih, et al., 2014; Malisawa & Rautenbach, 2012).

3.2.2   Validation datasets

3.2.2.1   GPCC

The GPCC (Global Precipitation Climatology Centre) is one of the most utilised and referenced gridded gauge-based precipitation dataset in academic studies and journals (Sun et al., 2017). The GPCC was established in 1989 on request of the World Meteorological Organization (WMO) and is currently operated by the Deutscher Wetterdienst (DWD). The GPCC has the largest dataset and spans the greatest time period, between 1901 to 2013, with monthly data from over 85,000 stations worldwide. Various institutions, such as the WMO, FAO and UNESCO make use of different data products from GPCC for water and climate-related research (Deutscher Wetterdienst, 2018). The calculation of the gridded precipitation datasets consists of three main steps (Rudolf & Schneider, 2005): interpolation from stations to regular 0.5° grid points; calculation of area-mean precipitation for the grid cells; as well as the assessment of area-mean precipitation for larger grid cells or other areas (e.g. river basins). An empirical interpolation weighting method is followed to extrapolate the gauge data to gridpoints. While this form of measurement is relatively accurate and trusted, and the large temporal data extent is useful for deriving mean annual precipitation and predicting climate impacts, the poor station coverage over Equatorial Africa implies poor data accuracy in some areas (Schneider, et al., 2016). Figure 3-3 presents the GPCC global gauge monitoring stations.
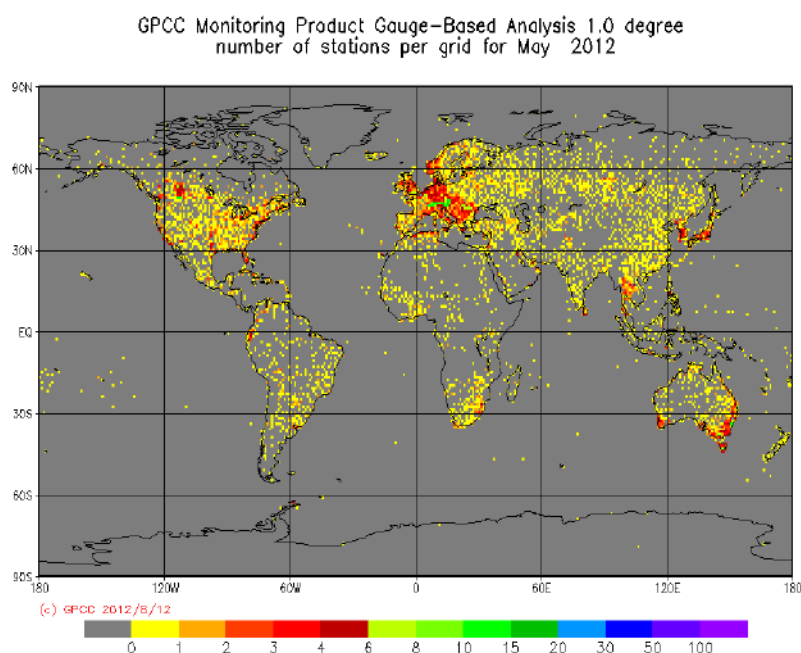

Figure 3-3: GPCC global gauge monitoring stations. Data retrieved from (https://climatedataguide.ucar.edu/climate-data/gpcc-global-precipitation-climatology-centre)

3.2.2.2 CRU

The Climate Research Unit (CRU) at the University of East Anglia is widely recognised as one of the world's leading institutions concerned with the study of natural and anthropogenic climate change (Harris, et al., 2020)The Climate Research Unit gridded Time Series (CRU TS) dataset is derived by interpolation (angular-distance weighting method) of monthly climate anomalies from station observation data. The angular-distance weighting method provides improved traceability between each gridded value and the input observation data. The CRU provides monthly data at a spatial resolution of 0.5° and ranges between 1901 and 2018 (Harris, et al., 2020). The CRU monthly precipitation data were obtained through the auspices of national meteorological agencies (NMAs), the WMO, the CRU, the Centro International de Agricultura Tropical, the Food and Agriculture Organization (FAO), and others (Sun et al., 2017). The overriding objective of CRU was to present complete global coverage. This is achieved by filling in missing station values by a) anomalising the series with the corresponding station data between 1961 and 1990; b) applying the angular distance weighting method to interpolate the values into grid points; and then c) converting the anomaly grid into actual values. Although this process might cause decreasing variance in climate data, the CRU database can still be used for global and regional trend analysis (Harris, et al., 2020). The decreased variance will have minimal impact on the annual averages to be used in this project.

3.2.2.3 NOAA-observed data points

The National Oceanic and Atmospheric Administration (NOAA) previously had three data centres including the National Climatic Data Centre, the National Geophysical Data Centre and the National Oceanographic Data Centre. These three data centres have merged into the National Centres for Environmental Information (NCEI) making the NCEI the world's largest provider of weather and climate data. Land-based observations are collected from instruments sited at locations on every continent (NOAA, 2020). NCEI provides a broad level of service associated with land-based observations. These include data collection, quality control, archive, and removal of biases associated with factors such as urbanization and changes in instrumentation through time. Data on sub-hourly, hourly, daily, monthly, annual, and multiyear timescales are available. However, in data sparse SADC areas, especially in Angola, the Democratic Republic of Congo, Madagascar and Mozambique, the observation stations can present inaccurate data and should therefore be used with caution. Figure 3-4 presents the NOAA observation stations within the SADC region.
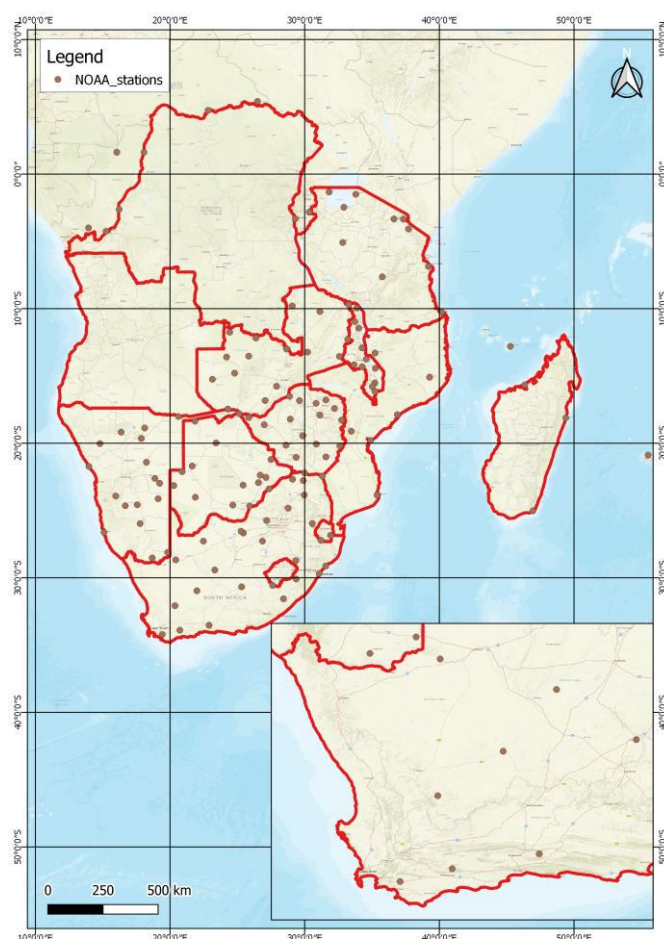
Figure 3-4: NOAA observation stations within the SADC region

3.2.3   Data validation

As described in the "*Draft Summary Review Report*" (SADC-GMI, 2020) the above datasets were assessed and evaluated in terms of various criteria and ultimately, WorldClim was selected as the preferred precipitation dataset for the assessment of surface water availability under this project.

Validation of the WorldClim v2.1 gridded data in SADC was performed by comparing WorldClim data with observed precipitation data from the NOAA dataset at observed data points based on Mean Annual Precipitation. A total of 126 NOAA stations were selected across the SADC region, with at least 3 stations in each country, subject to available stations. Using a raster point sample method, annual average precipitation raster values were extracted at each observation data point. The comparison between the WorldClim v2.1 annual average precipitation and the NOAA annual average precipitation is presented in Figure 3-5. At 66 locations, NOAA and WorldClim v2.1 mean annual precipitation values are within 10%. Stations where the comparison was less good, occur in central African countries such as Angola, Tanzania and the Democratic Republic of Congo and Madagascar. However, it is was assumed that the WorldClim v2.1 dataset is more reliable than the NOAA station data in the above-mentioned countries.
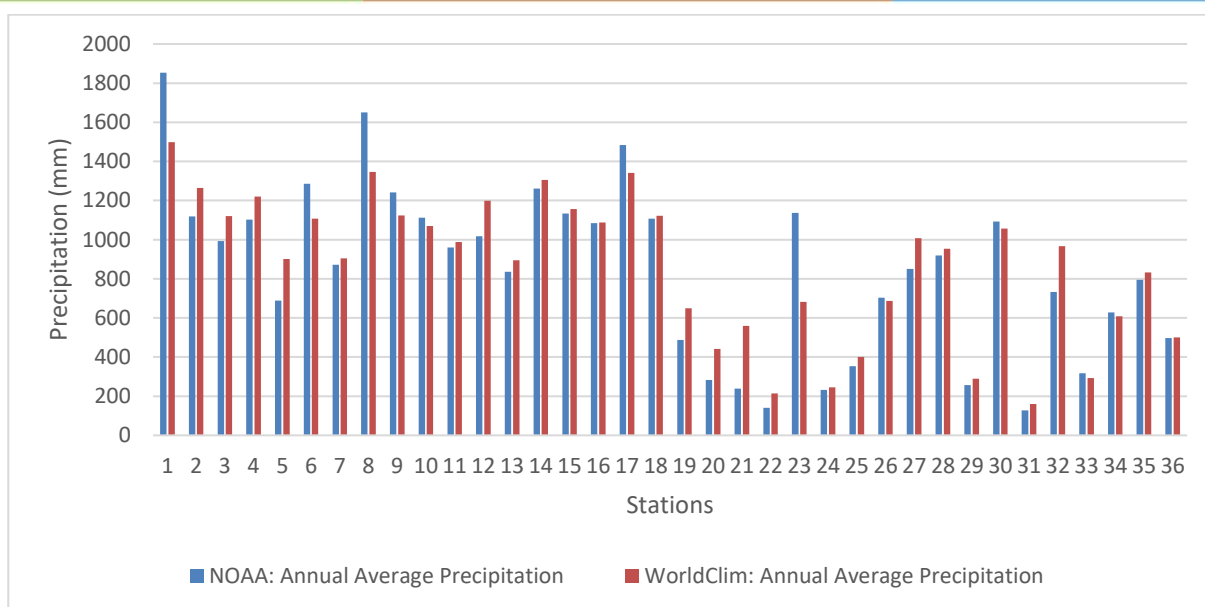
Figure 3-5: Comparison of NOAA and WorldClim rainfall data (sample of 36 stations)

3.3   Runoff data

A number of global datasets of catchment runoff and streamflow data are readily available. These can be categorised into: gauge-based, model- and/or simulation-based as well as reanalysis data-sets. The main datasets which were evaluated are briefly summarized in Table 3-3. An extensive table is attached as Appendix A where references and related comments are also provided.

In the following sections, the datasets are briefly discussed - specifically their relevance for use in the assessment of surface water availability. A distinction is made between primary datasets and validation datasets.

Table 3-3: Summary of global runoff datasets which were considered

Data Type	Dataset	Years available	Temporal resolution	Spatial resolution	Use in this project
Gauge-based	GRDC	1901 -near current	Daily, Monthly	Point data	Possible validation and bias-correct
Model-based	WaterGAP	1901-2016	Monthly average	0.25°x0.25°	Possible primary runoff dataset
	GRUN	1901-2014	3-hourly	0.5°x0.5°	Possible primary runoff dataset
	GeoSFM	1998 -2005	Monthly	0.25°x0.25°	Possible primary runoff dataset

3.3.1 Primary datasets

3.3.1.1 WaterGAP

WaterGAP v2.2 (Müller Schmied, et al., 2014) is a global water assessment model consisting of two main components, namely: the Global Water Use model and the Global Hydrology model. The Water Use model considers basic socio-economic factors to estimate domestic, industrial and agricultural water use, while the Hydrology model incorporates physical and climate factors to simulate runoff and groundwater recharge based on the computation of daily water balances of the soil and canopy. Both components of the model have been calibrated and tested against data on water use and runoff from river basins throughout the world (Alcamo, et al., 2003). The data collection period generally expands from 1901 to 2016. The spatial resolution of the output data is 0.5° and is presented in monthly time series raster bands – as part of a "netCDF" file. Approximately 3,000 global observation stations were used to validate the model (Alcamo, et al., 2003). According to the custodian of WaterGAP v2.2, there is a likelihood that overestimations of flow might occur in data sparse areas. However, during the analysis with WaterGAP, no such overestimations where observed. .

3.3.1.2 GRUN

The GRUN dataset contains a gridded global reconstruction of monthly runoff timeseries data. Runoff within the context of the GRUN model, is defined by Ghiggi et al. (2019) as "the amount of water drained from a given land unit (i.e. grid cell) which eventually enters the river system, including groundwater flow and snowmelt". In-situ streamflow observations from the Global Streamflow Indices and Metadata Archive (GSIM) and the GRDC dataset were used to train a machine learning algorithm that predicts monthly runoff rates based on antecedent precipitation and temperature from the Global Soil Wetness Project Phase 3 (GSWP3) meteorological dataset (Ghiggi, et al., 2019) The runoff data has a monthly resolution with a spatial resolution of 0.5°, covering the period from 1901 to 2014. The model tends to overestimate runoff in arid regions such as areas in southern Africa (Ghiggi, et al., 2019).

3.3.1.3 GeoSFM

A further model-based runoff dataset is generated with the Geospatial Streamflow Model (GeoSFM). The geospatial streamflow modeling system is parameterized with global terrain, soils and land cover data and runs with satellite-derived precipitation and evapotranspiration datasets (Asante, et al., 2008). The dataset is created by using simple linear methods to transfer water through subsurface, overland and river flow phases. The resulting monthly flows are expressed in terms of standard deviations from mean annual flow and presented at a spatial resolution of 0.25°. In sample applications, the modeling system was used to simulate flow variations in the Congo, Niger, Nile, Zambezi, Orange and Lake Chad basins between 1998 and 2005, and the resulting flows were compared with mean monthly values from the open-access Global River Discharge Database. The main limitations of GeoSFM include its inability to predict absolute flow magnitude and difficulties in characterizing flow travel time in basins with significant wetlands or reservoir systems. However, the model can provide independent monitoring information to water managers working in river systems with limited in-situ data (Asante, et al., 2008).

3.3.2 Validation datasets

3.3.2.1 GRDC

The Global Runoff Data Centre (GRDC) is an international data centre operating under the auspices of the World Meteorological Organization (WMO). Their dataset is a collection of quality controlled historical mean daily and monthly discharge data. Time series data on river discharge is available at more than 9 900 stations in 159 countries. The Southern Africa Flow Database of SA FRIEND constitutes a sub-dataset under the GRDC and is also obtainable from the GRDC website. The Southern Africa Flow Database was established between 1992 and 1997 to support rainfall-runoff modelling and it contains flow time series data from about 850 stations across southern Africa. The data have to be requested from grdc@bafg.de. In general, however, stations are limited in developing countries.

For this project, data at 881 stations across SADC were sourced from GRDC (Figure 3-6). The average record length is 44 years and most runoff stations have daily and monthly time-series. The GRDC has been used extensively in research papers and projects across the SADC region, including rainfall-runoff modelling in data scarce areas such as the DRC (Tshimanga & Hughes, 2014) and basin-wide research, such as in the Zambezi River Basin (McCartney, et al., 2013).
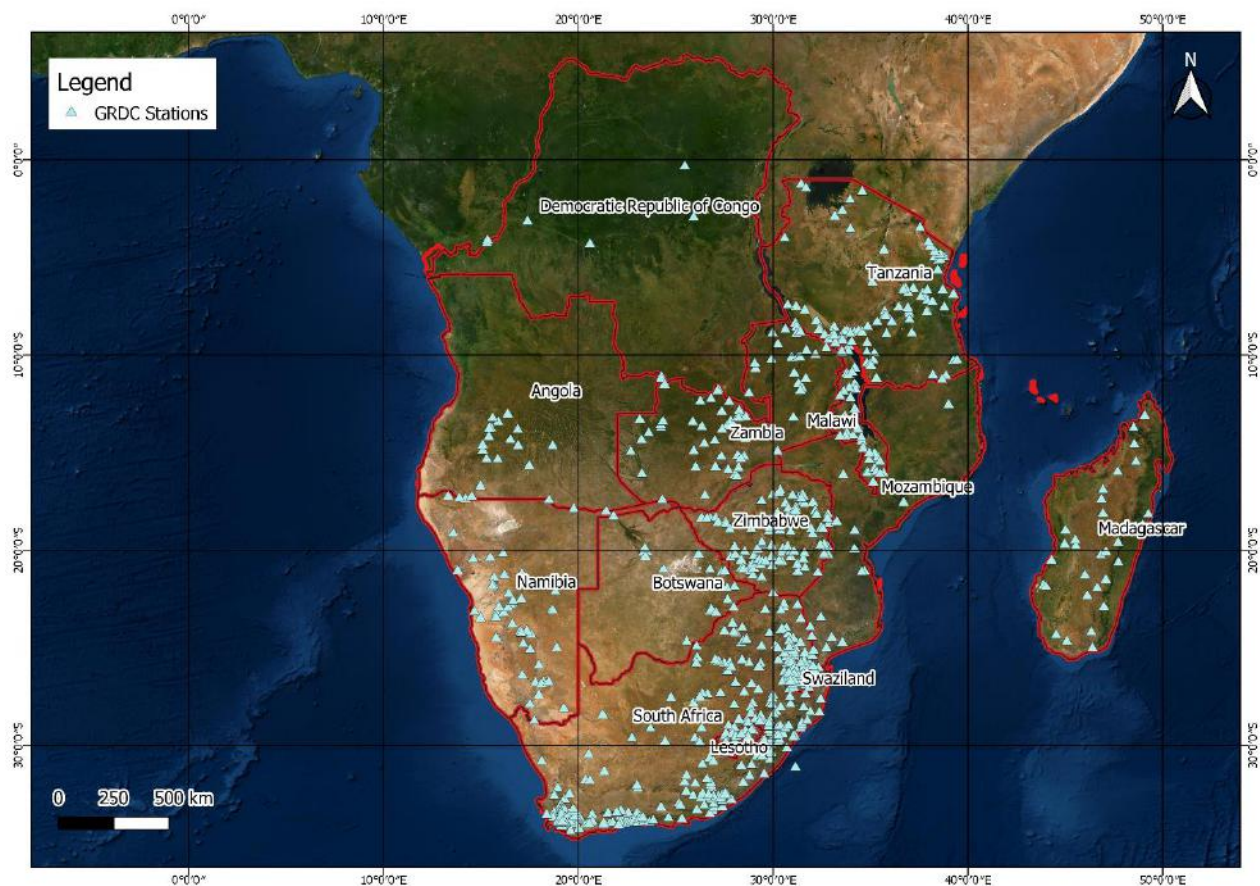

Figure 3-6: Observation Stations of the GRDC within the SADC region

3.3.3   Data validation

As described in the *Draft Summary Review Report (SADC-GMI, 2020)*", the above datasets were assessed and evaluated in terms of various criteria. Although every data set evaluated has shortfalls especially in data sparse areas, WaterGap v2.2 was selected as the preferred runoff dataset for the assessment of surface water availability under this project. Although the custodian of WaterGap v2.2 mentioned that overestimations might occur in data sparse areas, no such overestimations where observed when analysing the data. WaterGap v2.2 provides a combined dataset with both satellite-based as well as gauge-derived aspects which provide a versatile and validated dataset.

The WaterGAP v2.2 discharge raster dataset was verified against the GRDC station data, specifically catchment areas upstream of gauges and mean annual discharge. The GRDC stations in countries such as South Africa and Namibia, where the GRDC station catchment area is within 10% of the area of the catchment unit, the runoff and the discharge values were also within 10%. However, stations where the catchment area of the GRDC station did not compare with the area of the catchment unit were not used in the validation, as the streamflow and runoff would also not be comparable. With regard to all the stations, catchment areas and mean annual discharge at only 20% of the validation locations corresponded closely (within 10%). Only the stations with catchment areas within 20% of the areas of the catchment units where used for validation. It was found that GRDC stations had gaps in their monitoring data especially in data sparse countries Furthermore, the catchment areas documented in the GRDC were not always accurately demarcated. Based on the validation of the comparable stations, and various successful applications of the WaterGAP dataset in many river basins across the world, which included accurate validation, it was decided to use WaterGAP for this analysis.

4   STATISTICAL ANALYSIS

Statistical indices based on hydro-meteorological data are commonly used to quantify droughts on the landscape for any given time period (Svoboda & Fuchs, 2016) and provide numerical representations of drought severity. Statistical analyses to quantify precipitation, streamflow and runoff characteristics were thus undertaken, and surface water indices were calculated across SADC, at catchment unit scale, based on WorldClim and WaterGAP timeseries data (1960 – 2018) - averaged per catchment unit.   The methodology which was followed to determine the statistical indices as well as the motivation for using specific indices as drought indicators are discussed in the following sections.

4.1   Mean Annual Values

Mean annual precipitation, discharge (streamflow) and runoff values, averaged over a catchment unit, provide an indication of average long-term precipitation, streamflow, runoff and recharge. Figure B1, Figure B2, and Figure B3 presents the MAP, mean annual discharge and mean annual runoff per catchment unit over the SADC region, respectively. Refer to Appendix B

4.2   Seasonality

The seasonal index represents the extent to which precipitation and discharge (streamflow) vary between seasons in any hydrological year (starting in October). It was calculated as the difference between values during the wettest season (three wettest / highest flow consecutive months), expressed as a percentage of the corresponding annual precipitation or streamflow value, and the value during the driest season (three driest / lowest flow consecutive months), expressed as percentage of the corresponding annual values. A high seasonal variability index indicates that the bulk of the rainfall or flow occurs in the wet season, while the rest of the year experiences relatively low rainfall or flow, therefore suggesting a higher drought risk. Figure B4 and Figure B5 present the seasonality of precipitation and discharge respectively. Refer to Appendix B

4.3   Index of seasonal variability

The index of seasonal variability indicates the extent of intra-annual (month-to-month) fluctuation of rainfall and streamflow over a single year (Pitman, et al., 2008). It is calculated by using a mass curve method, i.e. the cumulative departure of mean calendar monthly rainfall or flow from the mean monthly rainfall or flow (expressed as percentage MAP). It was calculated by using a mass curve method as illustrated in Figure 4-1. The higher the index of seasonal variability, the greater the drought risk. Figure B6 and Figure B7 present the index of seasonality variability for precipitation and discharge respectively. Refer to Appendix B.
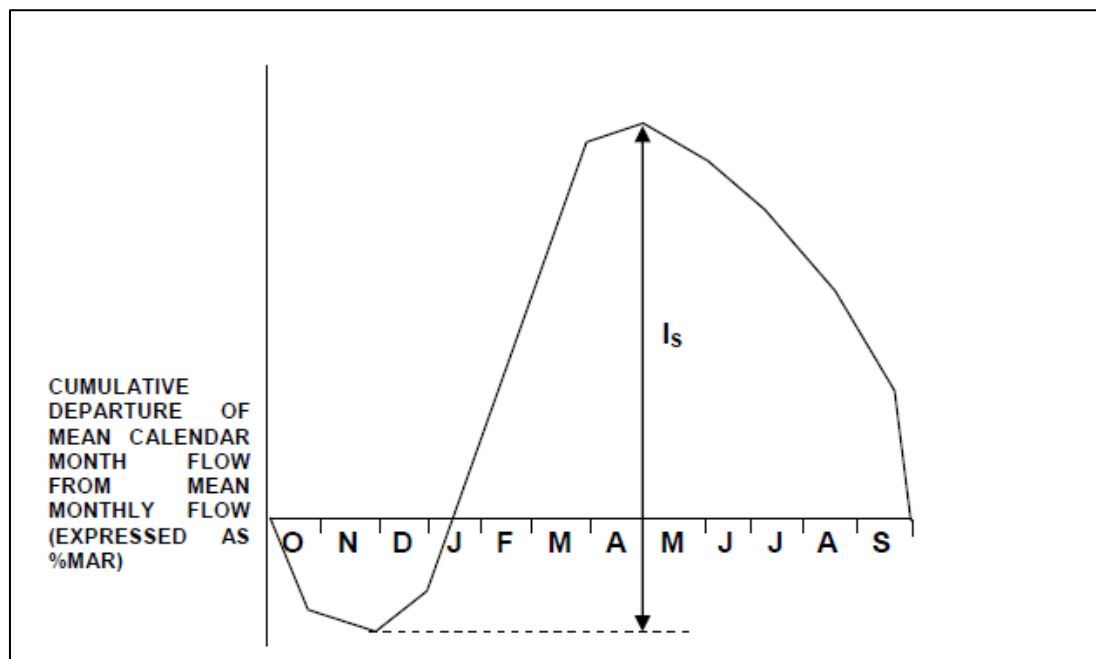
Figure 4-1: The Index of Seasonal Variability (Is)

4.4 Coefficient of Variation

The coefficient of variation of mean annual precipitation or discharge provides an index of climatic risk, indicating the likelihood of fluctuations from year to year (inter-annually). The higher the coefficient of variation, the more variable is the inter-annual variability and the greater the drought risk. Figure B8 and Figure B9 present the coefficient of variation for precipitation and discharge over the SADC region, respectively. Refer to Appendix B.

4.5 Runoff Coefficient

The runoff coefficient is a dimensionless factor that relates the amount of surface water runoff from a catchment to the amount of precipitation received. It represents the integrated effect of catchment losses and hence depends upon the nature of land surface, slope, degree of saturation, and rainfall intensity. A high runoff coefficient may indicate flash flooding areas during storms as water moves fast overland on its way to a river channel or a valley floor. The runoff coefficient per catchment unit was determined as the mean annual runoff (WaterGAP) expressed as a percentage of the mean annual precipitation (WorldClim). Figure B10 presents the runoff coefficients over the SADC region. Refer to Appendix B.

5   INDEX NORMALIZATION

Index normalization was undertaken to standardize the different index values to values between 0 and 1, and to allow comparison and integration of a number of indices.

5.1   Normalization Methods

The normalization techniques which were considered are defined below:

Percentage of Maximum

$$v_i = \frac{a_i}{\max(a_i)} \tag{1}$$

Percentage of Range

$$v_i = \frac{a_i - \min(a_i)}{\max(a_i) - \min(a_i)} \tag{2}$$

Unit Vector

$$v_i = \frac{a_i}{\sqrt{(\sum_i a_i^2)}} \tag{3}$$

where: a_i : the criterion measurement for any given Scenario; and

      v_i : normalized value of a_i.

There is no one single method that can prove itself to be the globally acceptable approach for normalization. Rather, characteristics of various indicators and parameters have to be evaluated and a normalisation process has to be selected that can support comparison of various parameters at a comparable scale.

As a general guidance the following recommendations are provided:

- If the normalized values are expected to range between 0 and 1, use 'percentage of range'.

- If the values of the indicators considered should remain constant in the interval [0; 1], the 'unit vector' technique should be used.

- If there is no basis for favouring one over the other, use 'percentage of maximum' - it is the most commonly used technique.

5.2   Normalisation of statistical indices

The surface water indices as determined in Section 4 are summarized in Table 5-1, Table 5-2 and Table 5-3 respectively. The absolute value range presents the minimum value and the maximum value of the specific index relating to the SADC catchment units. A direction for each index was selected based on how the index impacts the drought risk, such that the drought risk is maximized. The maximum drought risk is represented by 1. The normalization method used to normalize each index is also indicated.

Table 5-1: Precipitation indices normalization

Index	Absolute Value Range	Direction: Drought Risk	Normalization Method
Mean Annual Rainfall	9 mm – 3284 mm	Max as 0; Min as 1 The higher the rainfall, the lower the drought risk	Percentage of Max
Seasonality	11% MAP - 86% MAP	Max as 1, Min as 0 High inter-seasonal percentage, the higher drought risk.	Percentage of Range
Index of Seasonal Variability	3% MAP - 63% MAP	Max as 1, Min as 0 The higher the seasonal variation, the higher drought risk	Percentage of Range
Coefficient of Variation	1% - 51%	Max as 1, Min as 0 The higher the CoV the larger the range of data with respect to the mean, the higher the year-to-year fluctuation, the higher the drought risk	Percentage of Range

Table 5-2: Discharge indices normalization

Index	Absolute Value Range	Direction: Drought Risk	Normalization Method
Mean Annual Discharge	0.2 MCM/yr – 1 453 639 MCM/yr	Max as 0; Min as 1 The higher the discharge, the lower the drought risk	Percentage of Max
Seasonality	5 % - 95 %	Max as 1, Min as 0 High inter-seasonal percentage indicates non-perennial rivers, thus also higher drought risk.	Percentage of Range
Index of Seasonal Variability	1% MAR – 65% MAR	Max as 1, Min as 0 The higher the seasonal variation, the higher drought risk	Percentage of Range
Coefficient of Variation	13 % - 582 %	Max as 1, Min as 0 The higher the CoV the larger the range of data with respect to the mean, the higher the year-to-year fluctuation, the greater the drought risk	Percentage of Range

Table 5-3: Runoff indices normalization

Index	Absolute Value Range	Direction: Drought Risk	Normalization Method
Mean Annual Runoff	0 mm – 2262 mm	Max as 0; Min as 1 The higher the runoff, the lower the drought risk	Percentage of Max
Runoff Coefficient	0% - 7 %	Max as 0; Min as 1 The higher the runoff coefficient, the lower the drought risk	Percentage of Max

6   INDEX WEIGHTING AND SENSITIVITY ANALYSIS

A surface water risk map was produced by superimposing/combining the surface water indices determined for rainfall, discharge and runoff. The different indices were combined through a simple linear algorithm and associated weighting scheme based on the relative importance of various indices to derive a spatially distributed surface water risk map across the SADC region.

A sensitivity analysis was performed on the weightings to determine the impact of different indices on the combined surface water risk index. The sensitivity analysis involved the investigation of five different scenarios. Scenario 1 was the control scenario in which all the indices are equally weighted. Scenarios 2, 3, 4 and 5 were set up so that all indices remain constant, while the index under investigation was varied such that the impact of each index on the overall surface water risk index can be gauged. The scenarios are summarized in Table 6-1.

Table 6-1: Sensitivity Analysis

		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
		Equal Weighting	Average	Inter-seasonality	Index of Seasonality	Coefficient of Variation
Rainfall	Average rainfall (mm)	Constant	**Varying**	Constant	Constant	Constant
	Inter-seasonality	Constant	Constant	**Varying**	Constant	Constant
	Index of seasonality	Constant	Constant	Constant	**Varying**	Constant
	Coefficient of variation (%)	Constant	Constant	Constant	Constant	**Varying**
Discharge	Average discharge (mm)	Constant	**Varying**	Constant	Constant	Constant
	Inter-seasonality	Constant	Constant	**Varying**	Constant	Constant
	Index of seasonality	Constant	Constant	Constant	**Varying**	Constant
	Coefficient of variation (%)	Constant	Constant	Constant	Constant	**Varying**
Runoff	Mean annual runoff (mm)	Constant	**Varying**	Constant	Constant	Constant
	Runoff coefficient (%)	Constant	Constant	Constant	Constant	**Varying**

The results of the sensitivity analysis are presented in Table 6-1. From Figure 6-1 it is evident that the largest change in combined risk coefficient is as result of the Scenario 2, where the average indices are varied.
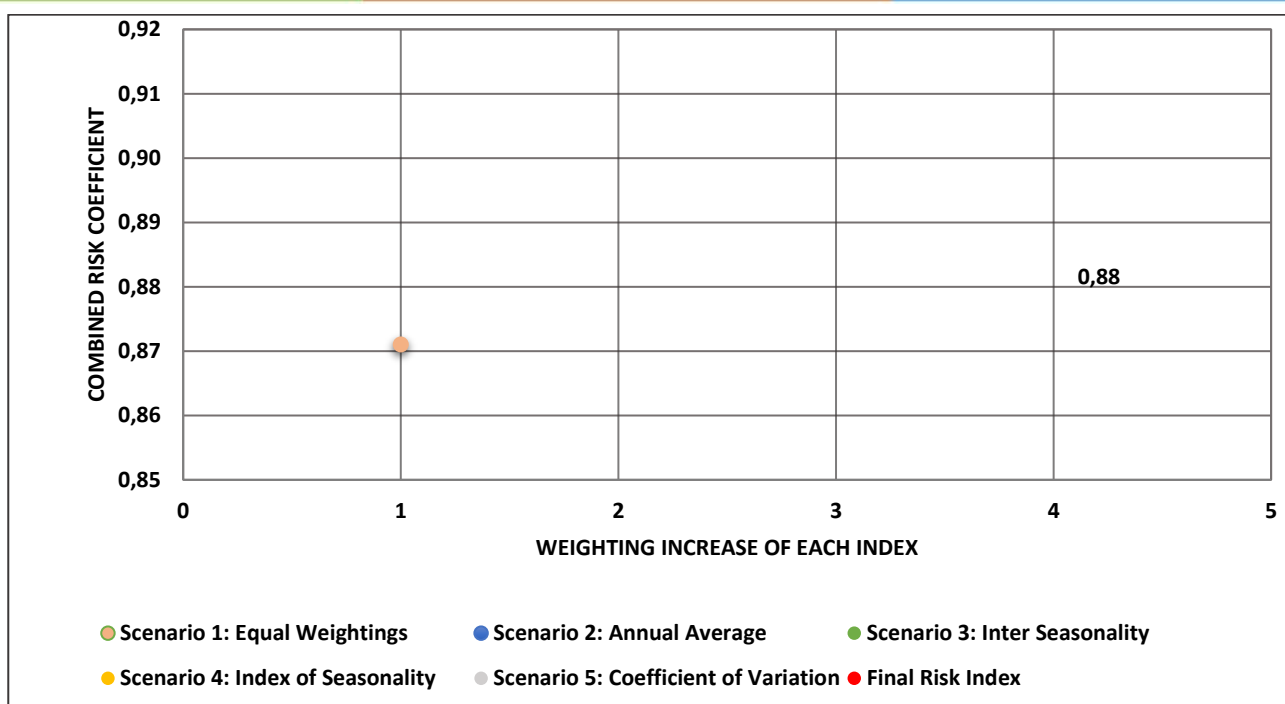
Figure 6-1: Impact of weighting for each index on the combined surface water risk index

The average values represent absolute precipitation, discharge or runoff, while seasonality, index of seasonal variability, coefficient of variation and runoff coefficient represent dimensionless indices. The sensitivity analysis confirmed the importance of not assigning too great a weight to absolute values as this could potentially skew the resulting surface water risk index. Weights were therefore assigned to indices which measure inter- and intra-annual variability such that their combined weight significantly exceeds that of the "average" value indices. Coefficient of variation was assigned the highest weight as it measures inter-annual variability - an important factor when considering drought risk (Svoboda & Fuchs, 2016). The final weightings of the different surface water indices are shown in Table 6-2 below.

Table 6-2: Final weighting of surface water indices

	Surface water indices	Final Weightings
Rainfall	**Average rainfall (mm)**	0.11
	Seasonality	0.06
	Index of Seasonality	0.06
	Coefficient of variation (%)	0.15
Discharge	**Average discharge (mm)**	0.11
	Seasonality	0.06
	Index of Seasonality	0.06
	Coefficient of variation (%)	0.15
Runoff	**Mean annual runoff (mm)**	0.11
	Runoff coefficient (%)	0.15

7 FINAL RISK MAP

7.1 Final surface water risk map

The final surface water risk map is presented in Figure 7-1. Discharge and runoff data was not available for some of the island states. The available rainfall data was used to calculate the rainfall indices and derive a surface water risk index for the island states.
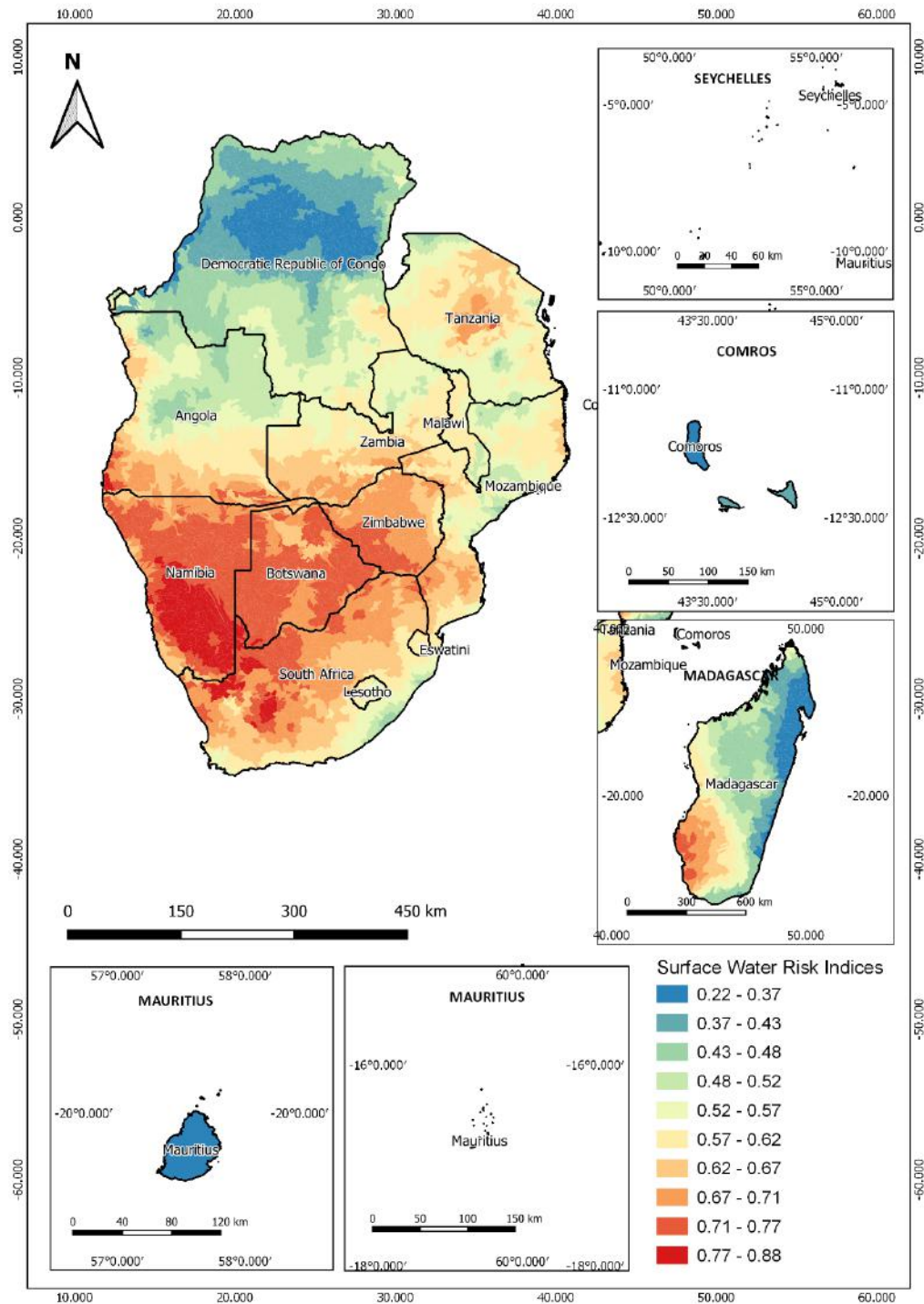

Figure 7-1: Surface water risk map

7.2 Validation of surface water risk map

The surface water risk map (Figure 7-1) highlights areas within the SADC region where surface water drought conditions are prevalent. A qualitative validation process is followed where geographical areas exposed to more frequent droughts are identified using a number of existing derived drought maps as well as reports on droughts throughout southern Africa.

According to Figure 7-1, areas of severe surface water drought include: south-western and central Namibia, most of Botswana, south-western Zimbabwe, northern South Africa as well as southern Angola, southern Zambia, southern Mozambique as well as central Tanzania and south-western Madagascar.

According to the SADC Climate Services Centre (2018/2019), extreme drought was declared over most of the south-western parts of Southern African Development Community due to below average rainfall during the 2018/2019 rainfall season. According to Figure 7-2, extreme drought conditions are indicated mainly over southern Angola, southern Zambia, northern Zimbabwe, northern Botswana, north-western South Africa and most of central-northern Namibia. Moderate to severe drought is also affecting most of Angola, Namibia, Botswana, Zimbabwe, South Africa, Lesotho and Zambia. Pockets of dryness are indicated over most of Tanzania, western and eastern DRC, Eswatini, southern Mozambique and western Madagascar. The areas identified as drought areas by the SADC Climate Services Centre (2018/2019), concur with the identified drought areas of the surface water drought risk map (Figure 7-1).
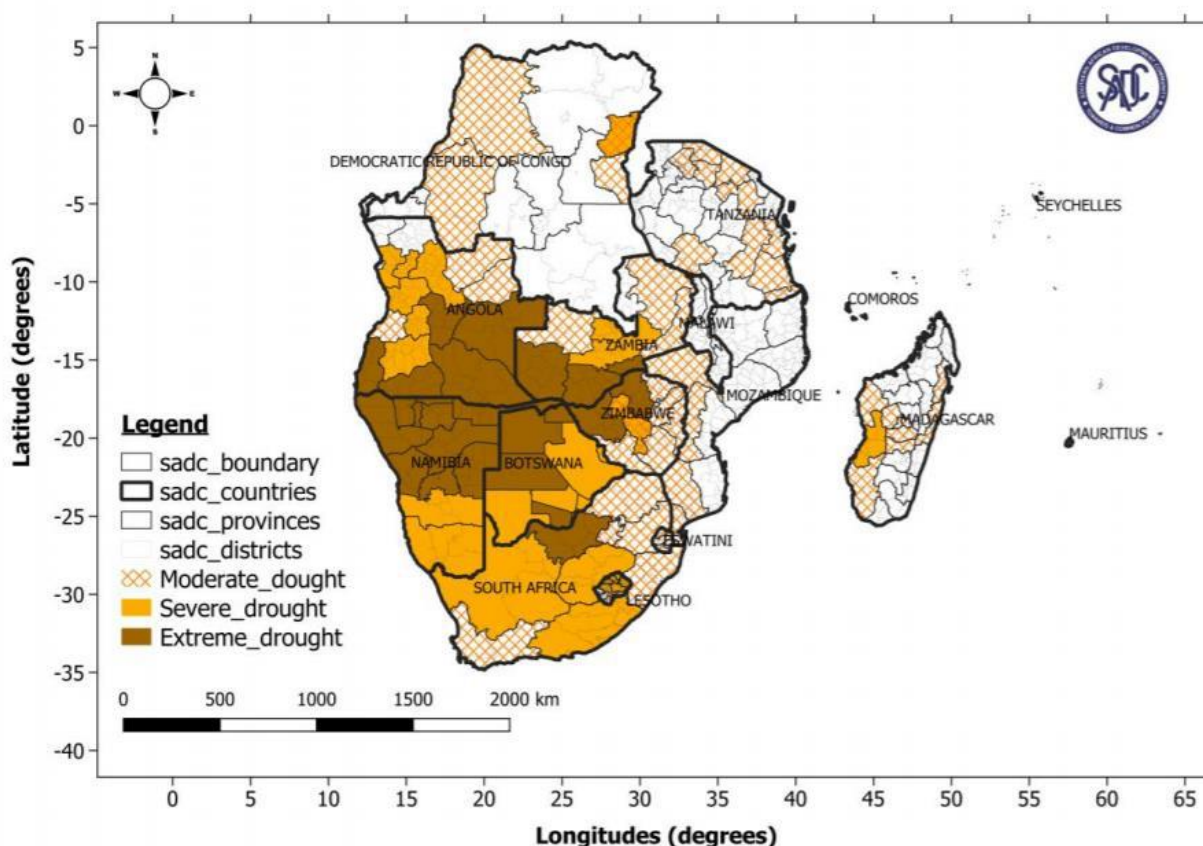

Figure 7-2: Drought situation over the SADC Region for 2018/2019 rainfall season
(https://www.sadc.int/files/5615/5991/5186/SADC_DROUGHT_MONITOR_2018-19_SEASON_JUNE_2019.pdf)

The Famine Early Warning System Network as well as the U.S. Geological Survey was used to identify and monitor drought risk areas in February 2019. Figure 7-3 depicts soil moisture anomalies in February 2019. Areas with more (green) or less (red) water in the upper layers of the ground than the norm for the month (Stevens & Hansen, 2019). Namibia and southern Angola and southern Zambia, northern Botswana and norther Zimbabwe as well as western Madagascar show especially dry soils. The areas highlighted by dry soil moisture highlight the same regions as identified through the surface water risk map (Figure 7-1).
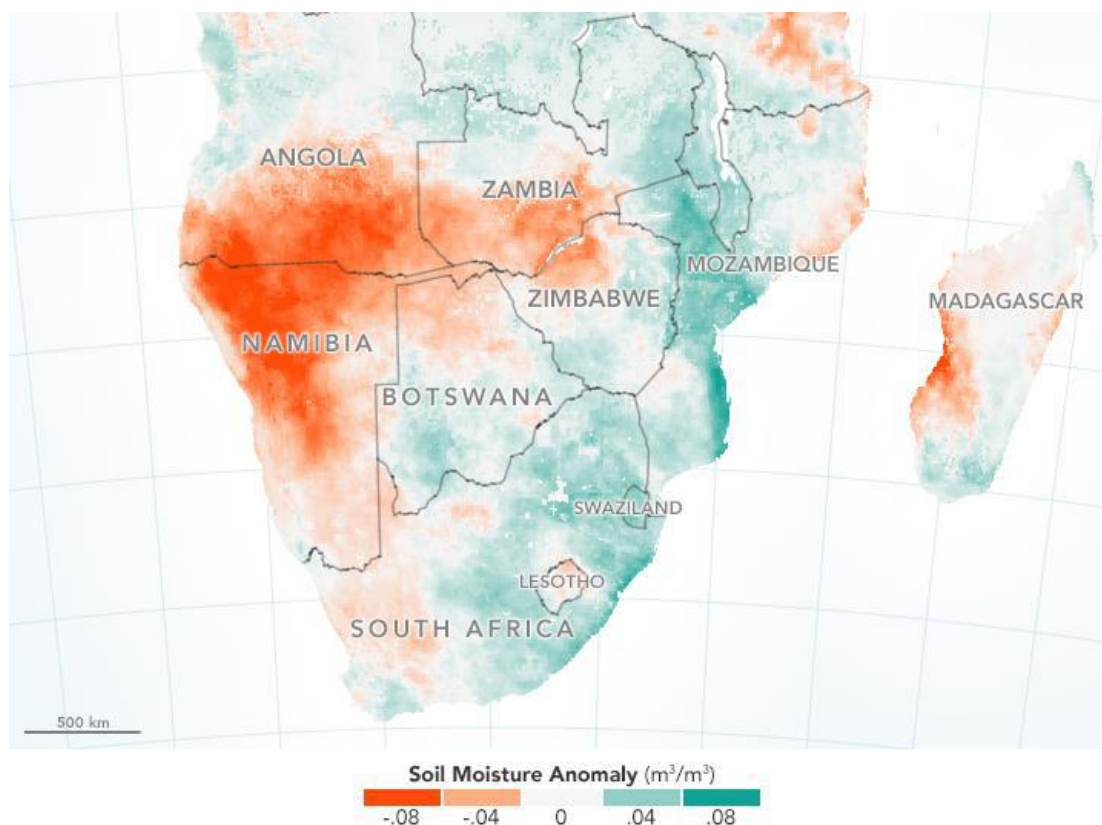

Figure 7-3: Soil moisture anomaly February 2019
(https://earthobservatory.nasa.gov/images/144704/drought-harms-corn-crops-in-southern-africa)

The United Nations Convention to Combat Desertification (UNCCD) held the African Drought Conference in 2016 in Windhoek where the history of droughts within Africa where discussed and documented (Tadesse, et al., 2018). According to Tadesse et al., (2018) frequent droughts have been recorded across southern Africa within the 20th and 21st century including: Namibia, South Africa, northern Botswana, northern Zimbabwe, southern Zambia, southern Angola as well as parts of Mozambique, Madagascar and Tanzania. The regions highlighted by Tadesse et al., (2018) show consistency with the surface water risk map (Figure 7-1).

8 CONCLUSION AND WAY FORWARD

The assessment of surface water resources makes use of freely available and accessible global hydrological datasets, specifically discharge and runoff from WaterGAP v2.2 (Döll et al. 2003), and rainfall from WorldClim v2.1 (WorldClim, 2020). These global datasets form part of the WWF HydroATLAS compendium. These global datasets were validated against point discharge, runoff and rainfall data from the GRDC and NOAA respectively, and show a good level of data integrity. A set of statistical indices were then calculated using the time series data available at each unit catchment, including MAP, index of seasonality, coefficient of variation and drought index for rainfall and MAR, index of seasonality, coefficient of variation as well as the drought index for runoff. Finally, these indices were normalised and combined to develop an integrated surface water risk map. This surface water risk map will be overlayed with the revised groundwater drought risk map (deliverable 3 of this project) and the population vulnerability priority areas map (deliverable 4 of this project) to produce a final hotspot map. This hotspot map will be used to pinpoint areas for high level water supply interventions, from both ground and surface water (deliverable 5 of this project). The surface water and precipitation maps that have been produced in this report will form the basis for identifying the most appropriate surface water interventions for the hotspot areas.

9 REFERENCES

Alcamo, J. et al., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. *Hydrological Sciences Journal,* 48(3), pp. 317-337.

Asante, K., Artan, G., Prevez, S. & Rowland, J., 2008. A linear geospatial streamflow modeling system for data sparse environments. *International J.River Basin Management ,* 6(3), pp. 233-241.

Badr, H., Dezfuli, A., Zaitchik, B. & Peters-Lidard, C., 2016. Regionalizing Africa: Patterns of Precipitation Variability in Observations and Global Climate Models. *Journal of Climate,* Volume 29.

Deutscher Wetterdienst, 2018. *Global Precipitation Analysis Products of the GPCC.* [Online] Available at: https://opendata.dwd.de/climate_environment/GPCC/PDF/GPCC_intro_products_v2018.pdf [Accessed 26 May 2020].

Dinku, T. et al., 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society,* 144(S1), pp. 292-312.

Driver, P., 2014. *Rainfall variability over southern Africa,* s.l.: University of Cape Town.

Fick, S. & Hijmans, R., 2017. WorldClim 2: new 1-km spatial resolution climate surface for global land areas. *International Journal of Climatatology,* 10(1002).

Funk, C. et al., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data,* 2(150066).

Ghiggi, G., Humphrey, V., Seneviratne, S. & Gudmunddon, L., 2019. GRUN: an observation-based global gridded runoff dataset from 1902 to 2014. *Earth Systems Scientific Data,* Volume 11, pp. 1655-1674.

Harris, I., Osborn, T. J., Jones, P. & Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data,* 7(109).

Harris, I., Osborn, T., Jones, P. & Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data,* 7(109).

Hijmans, R. et al., 2005. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology ,* Volume 25, pp. 1965-1978.

Lehner, B., 2014. *HydroBASINS: Technical Documentation Version 1.c,* s.l.: WWF.

Linke, S. et al., 2019. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Scientific Data,* 6(283).

Malisawa, M. S. & Rautenbach, C. J., 2012. Evaluating water scarcity in the Southern African Development Community (SADC) region using a climate moisture index (CMI) indicator. *Water Science & Technology: Water Supply,* 12(1), pp. 45-54.

Masih, I., Maskey, S., Mussa, F. E. F. & Trambauer, P., 2014. A review of droughts on the African continent: a geospatial and long-term perspective. *Hydrology and Earth System Sciences,* Volume 18, pp. 3635-3649.

McCartney, M., Cai, X. & Smakhtin, V., 2013. *Evaluating the Flow Regulating Functions of Natural Ecosystems in the Zambezi River Basin,* Colombo, Sri Lanka: International Water Management Institute (IWMI).

Müller Schmied, H. et al., 2014. Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration. *Hydrology and Earth System Sciences,* Volume 18, pp. 3511-3538.

Nagarajan, R., 2009. Drought Assessnebt. In: 1, ed. *Drought Indices.* India: Springer, pp. 164-204.

NOAA, 2020. *National Centres of Environmental Information.* [Online] Available at: https://www.ncdc.noaa.gov/ [Accessed 25 June 2020].

Pitman, W., Kakebeeke, J. & Bailey, A., 2008. *Water Resources Simulation Model for Windows,* Pretoria: Water Resource Commission.

Rudolf, B. & Schneider, U., 2005. *Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations.* s.l., s.n., pp. 231-247.

Schneider, U. et al., 2016. The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources. *Proceedings of the International Association of Hydrological Sciences,* Volume 374, pp. 29-34.

Stevens, J. & Hansen, K., 2019. *Earth Observatory-NASA: Drought Harms Corn Crops in Southern Africa.* [Online]
Available at: https://earthobservatory.nasa.gov/images/144704/drought-harms-corn-crops-in-southern-africa
[Accessed 28 July 2020].

Svoboda, M. & Fuchs, B., 2016. *Handbook of Drought Indicators and Indices.* Integrated Drought Management Programme, Integrated Drougjt Manaagement Tools and Guidelines Series 2 ed. Geneva: World Meteorological Organization and Global Water Partnership.

Tadesse, T., Tsagai, D., Stefanski, R. & Moreno, P., 2018. *Strategic Framework for Drought Risk Management and Enhancing Resilience in Africa,* Germany: United Nations Convention to Combat Desertification (UNCCD).

Tshimanga, R. M. & Hughes, D. A., 2014. Basin-scale performance of a semidistributed rainfall-runoff model for hydrological predictions and water resources assessment of large rivers: The Congo River. *Water Resources Research,* 50(2), pp. 1174-1188.

Wango, T., Musiega, D. & Mundia, C., 2018. Assessing the Suitability of the WorldCLim Dataset for Ecological Studies in Southern Kenya. *Journal of Geographic Infromation System,* 10(6), p. 16.

APPENDIX A: METADATA FOR SOURCES WITHOUT MAP

Table A1: Precipitation Datasets and Sources

Dataset	Source	Source URL	Original Metadata URL	Licence	Data type (raster or vector)	Scale	Frequency or Time Series	Date Created	Data Collection Period	Geographic-al Coverage	Other Comments About Data Use, Limitations and Processing Done
GPCC	GPCC	https://climatedataguide.ucar.edu/climate-data/gpcc-global-precipitation-climatology-centre	https://psl.noaa.gov/data/gridded/data.gpcc.html	None	Rastar	0.5x0.5	Monthly	1989	1901-2013	Worldwide	Accurate interpolated gauge data, but not good data coverage in central Africa
CHIRPS		https://data.chc.ucsb.edu/products/CHIRPS-2.0/	https://data.chc.ucsb.edu/products/CHIRPS-2.0/	None	Rastar	0.05 x 0.05	Daily, pentadal, and monthly		1981-2018	quasi-global (50°S-50°N)	Previously used in SADC regions
GPCP	GPCP	https://climatedataguide.ucar.edu/climate-data/gpcp-monthly-global-precipitation-climatology-project	https://psl.noaa.gov/data/gridded/data.gpcp.html	None	Raster	2.5x2.5	Monthly	2012	1997-2020	Worldwide	Satellite data and gauge data is combined
CRU	The CRU of the University of East Anglia	http://www.cru.uea.ac.uk/	http://www.cru.uea.ac.uk/data	None	Raster	0.5x0.5	Monthly		1901-2015	Worldwide	Used for global and regional trend analysis
WorldClim	WorldClim v2.1	https://www.hydrosheds.org/images/inpages/BasinATLAS_Catalog_v10.pdf	https://www.worldclim.org/data/monthlywth.html	Creative Commons CC-BY 4.0	Rastar	2.5x2.5	Monthly		1960-2018	Worldwide	Used for global and regional trend analysis

Table A2: Runoff and Discharge Datasets and Sources

Dataset	Source	Source URL	Original Metadata URL	Licence	Data type (raster or vector)	Scale	Frequency or Time Series	Date Created	Data Collection Period	Geographic-al Coverage	Other Comments About Data Use, Limitations and Processing Done
GRDC	GRDC	https://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html	Request from grdc@bafg.de	None	Vector/ point		Daily, Monthly	1988	1901 to near current	Worldwide	Sparse in Central and Northern Africa
SA FRIEND	GRDC	https://www.bafg.de/GRDC/EN/04_spcldtbss/45_SAFL/saflow_node.html	Request from grdc@bafg.de	None	Vector/ point		Daily, Monthly	1988	1901 to near current	Part of the SADC region	Sparse in Central and Northern Africa
GRUN	GRDC	https://figshare.com/articles/GRUN_Global_Runoff_Reconstruction/9228176	https://figshare.com/articles/GRUN_Global_Runoff_Reconstruction/9228176	None	Raster grid	0.5	3-hourly		1901-2014	Worldwide	Relatively good data coverage over SADC
GeoSFM		https://www.researchgate.net/figure/The-GeoSFM-software-is-a-semi-distributed-hydrologic-model-developed-as-an-extension-of_fig3_228779249		None	Raster	0.25	Monthly		1998 - 2005	Worldwide	Useful for modelling water resources in data sparse
WaterGAP v2.2	HydroS WaterGAP v2.2	https://www.hydrosheds.org/page/hydroatlas	https://www.hydrosheds.org/images/inpages/BasinATLAS_Catalog_v10.pdf	Creative Commons CC-BY 4.0	Raster grid	0.25	Monthly average		1971-2000	Worldwide	Limited data period available and not recently updated

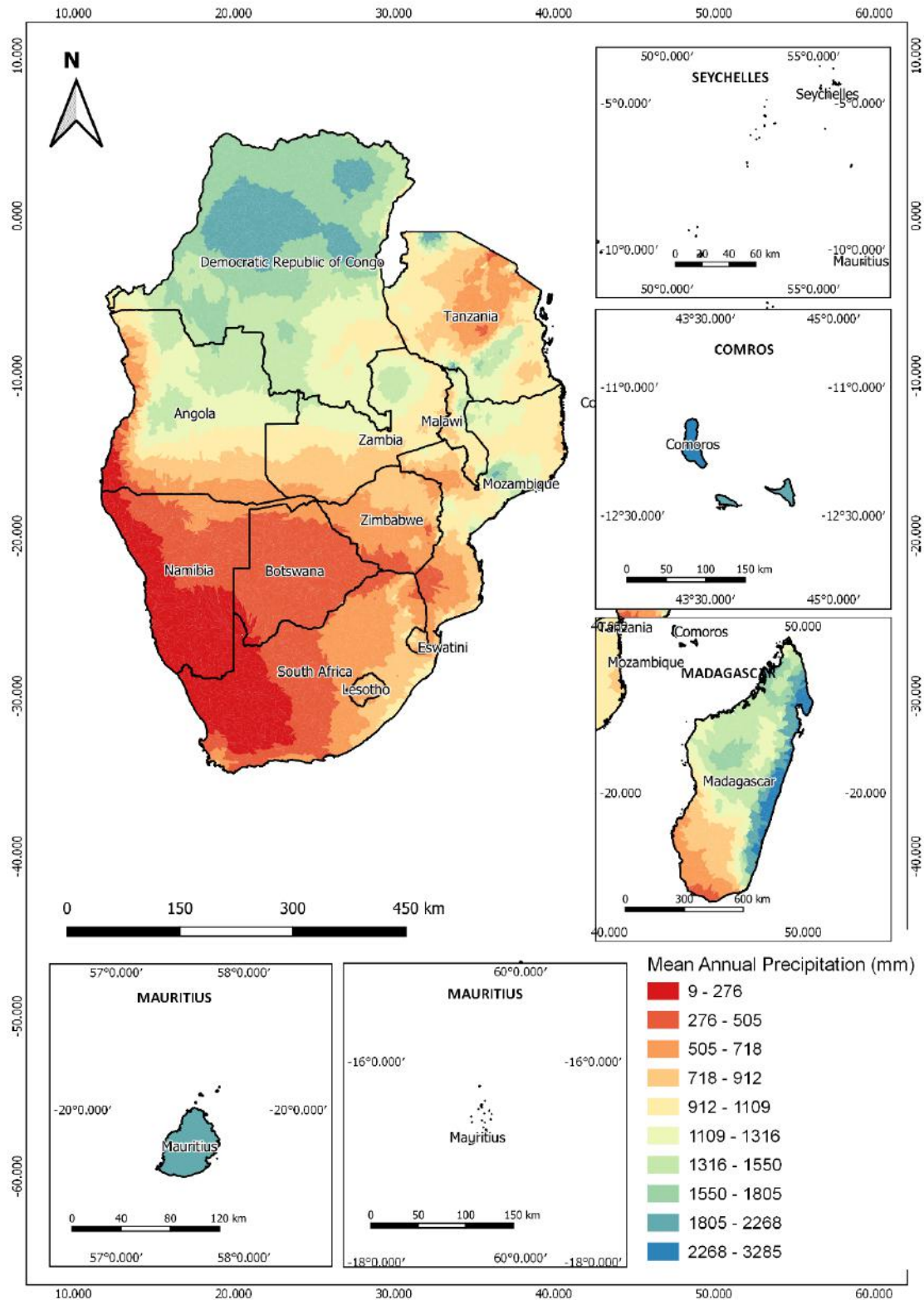# APPENDIX B: STATISTICAL INDICES

B1: Mean Annual Values

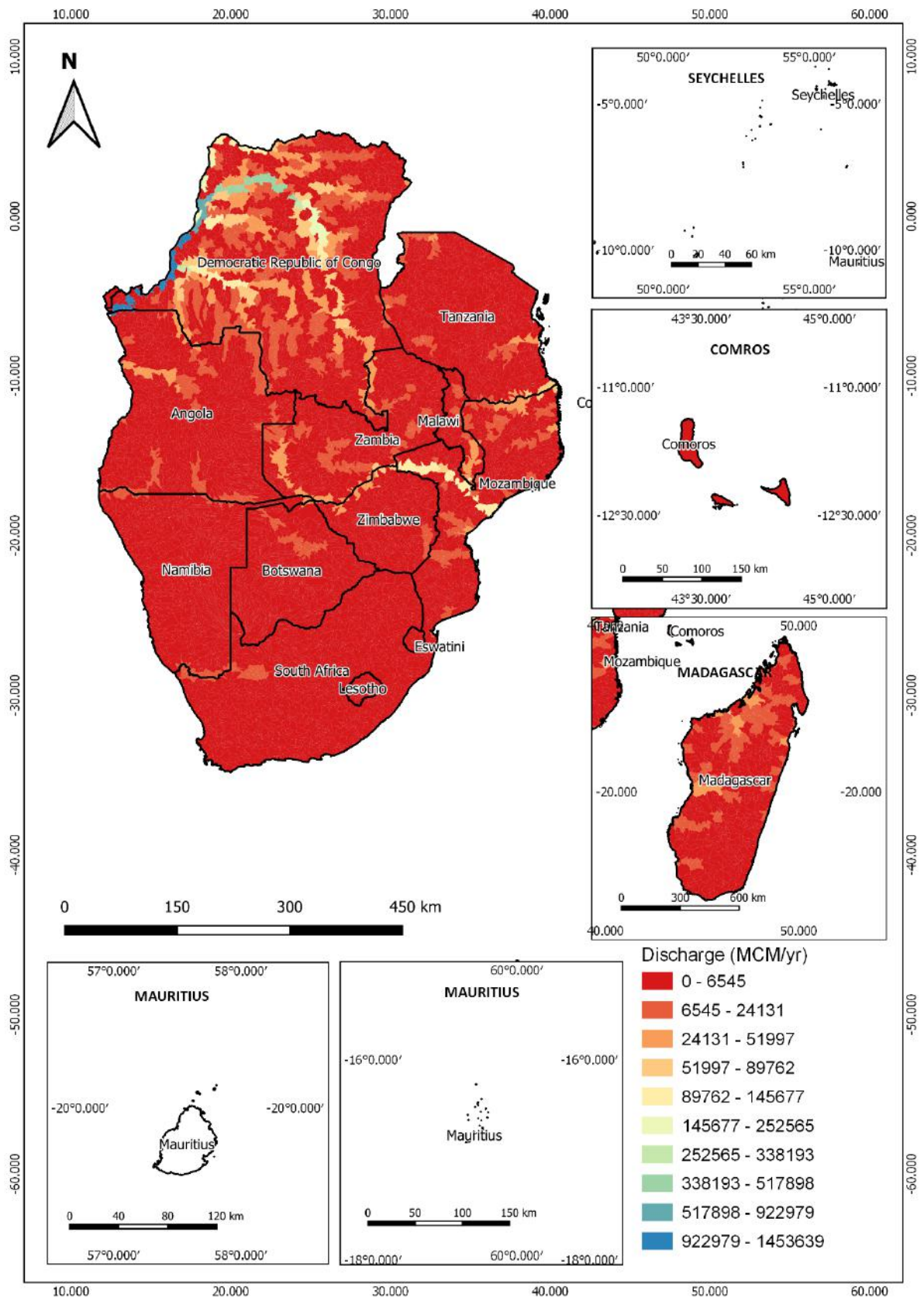
Figure B1: Mean annual precipitation (mm)

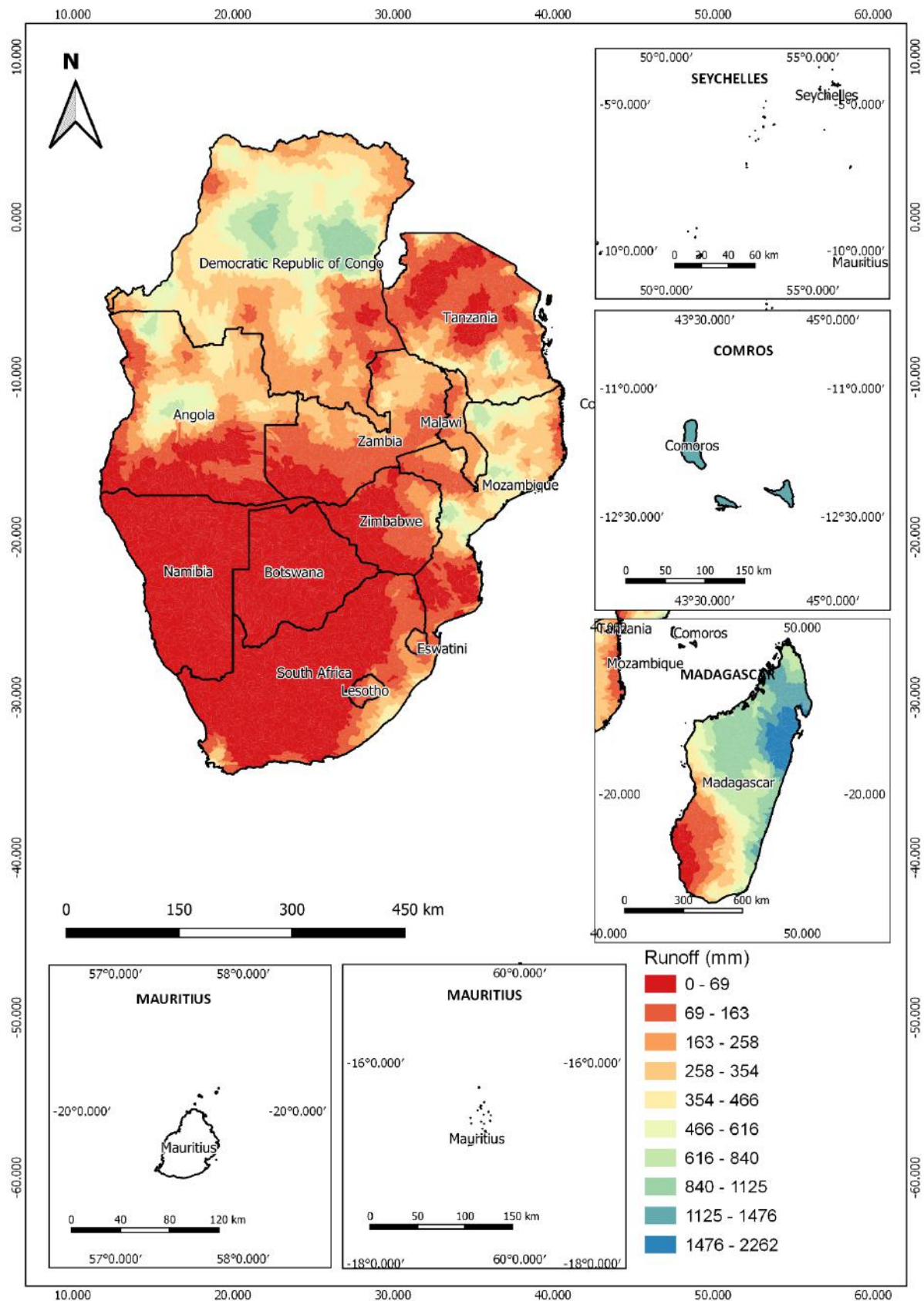**Figure B2: Mean Annual Discharge (million cubic meters per year)**

Figure B3: Mean Annual Runoff (mm)
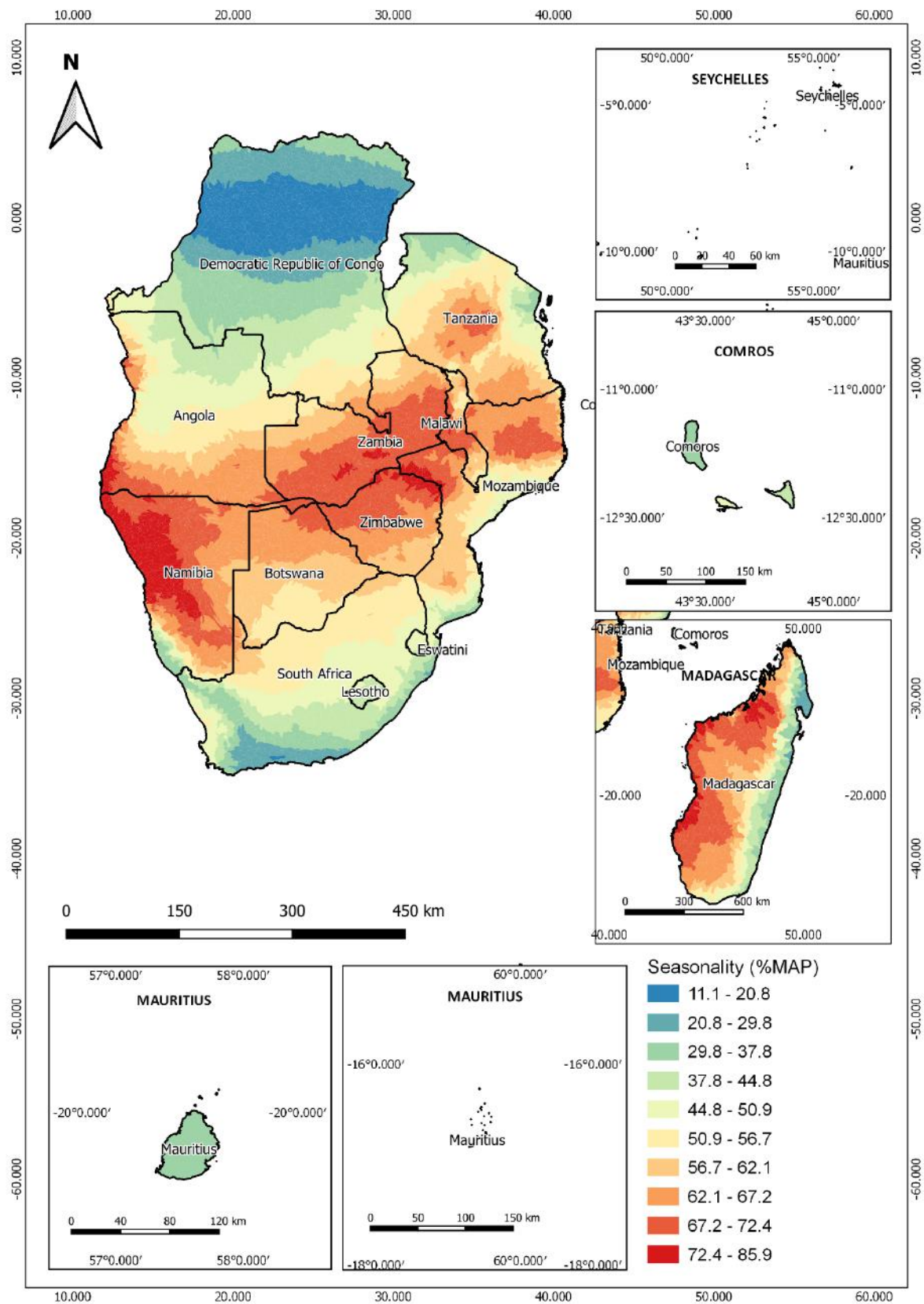
B2: Seasonality

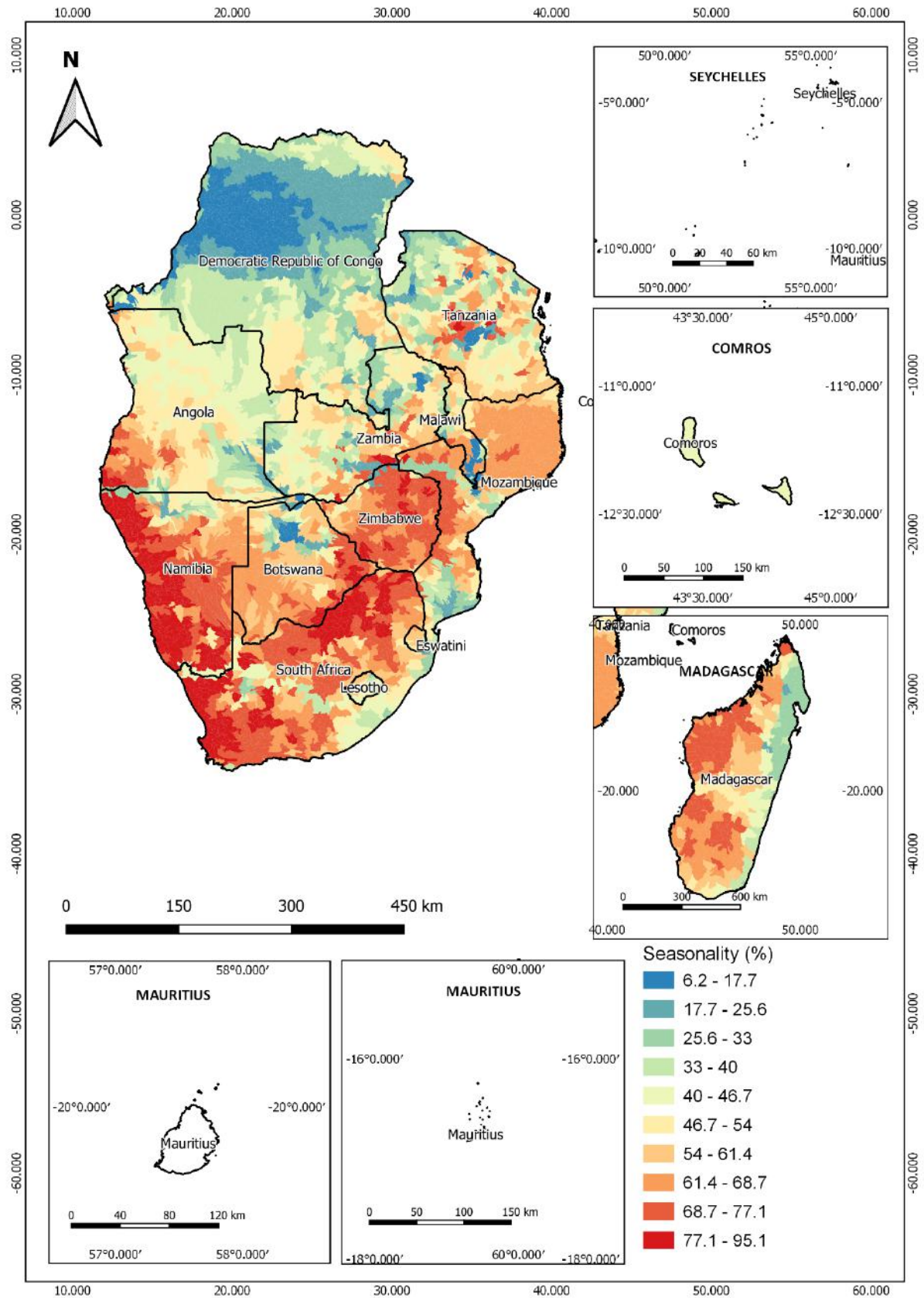**Figure B4: Seasonality of Precipitation**

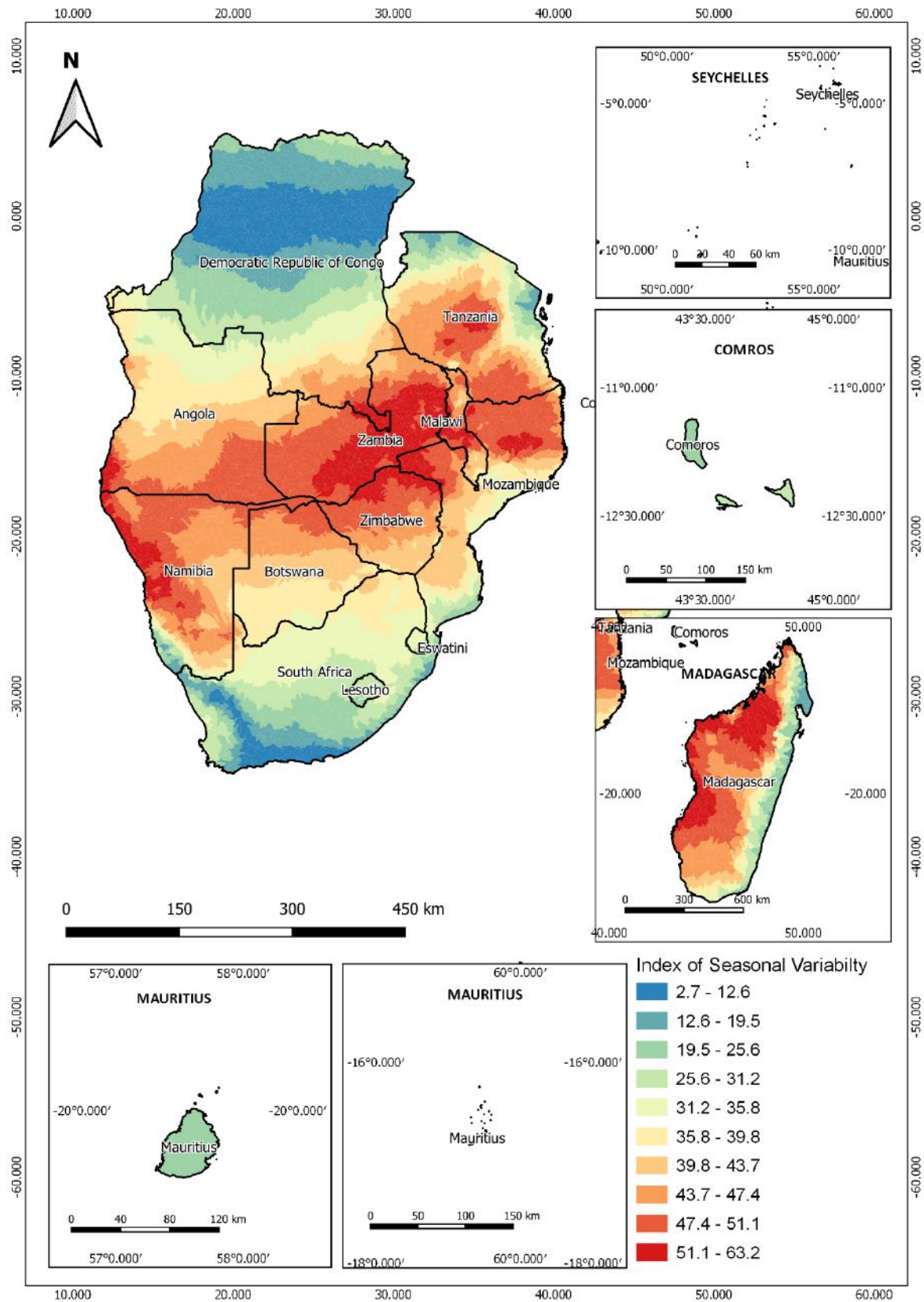**Figure B5: Seasonality of Discharge**

B3: Index of Seasonal Variability
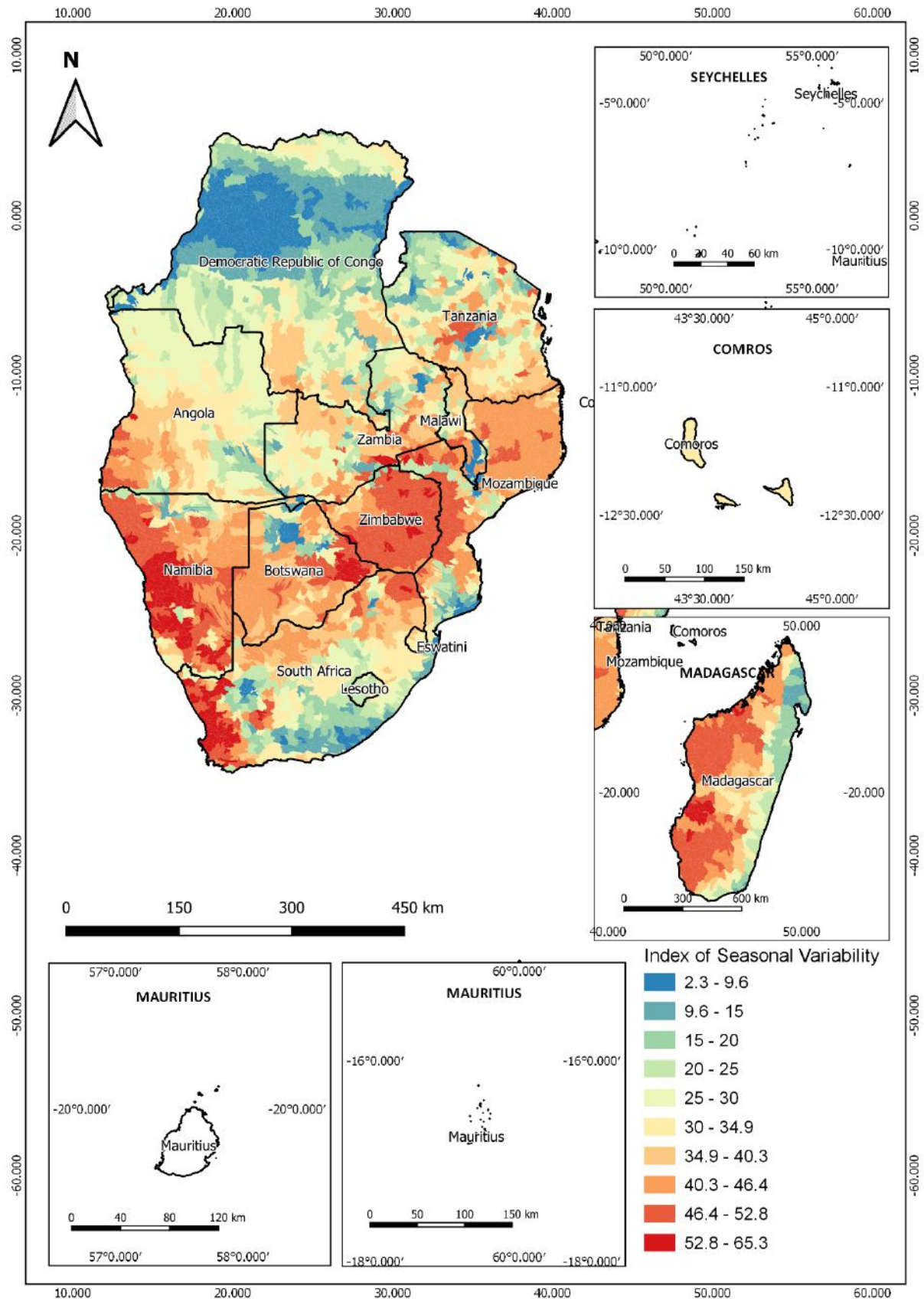

Figure B6: Index of Seasonal Variability Precipitation

Figure B7: Index of Seasonality of Discharge
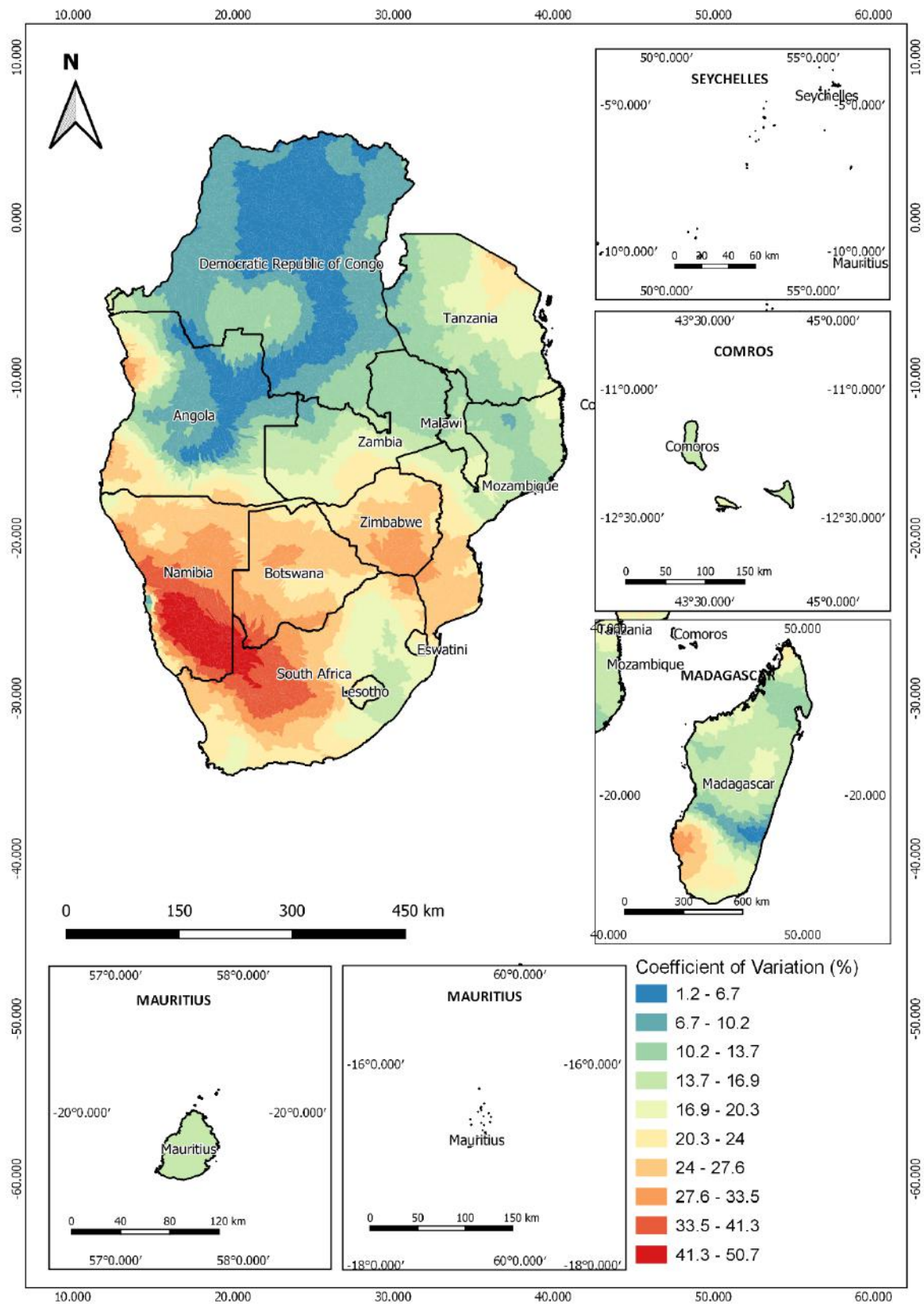
B4: Coefficient of Variation

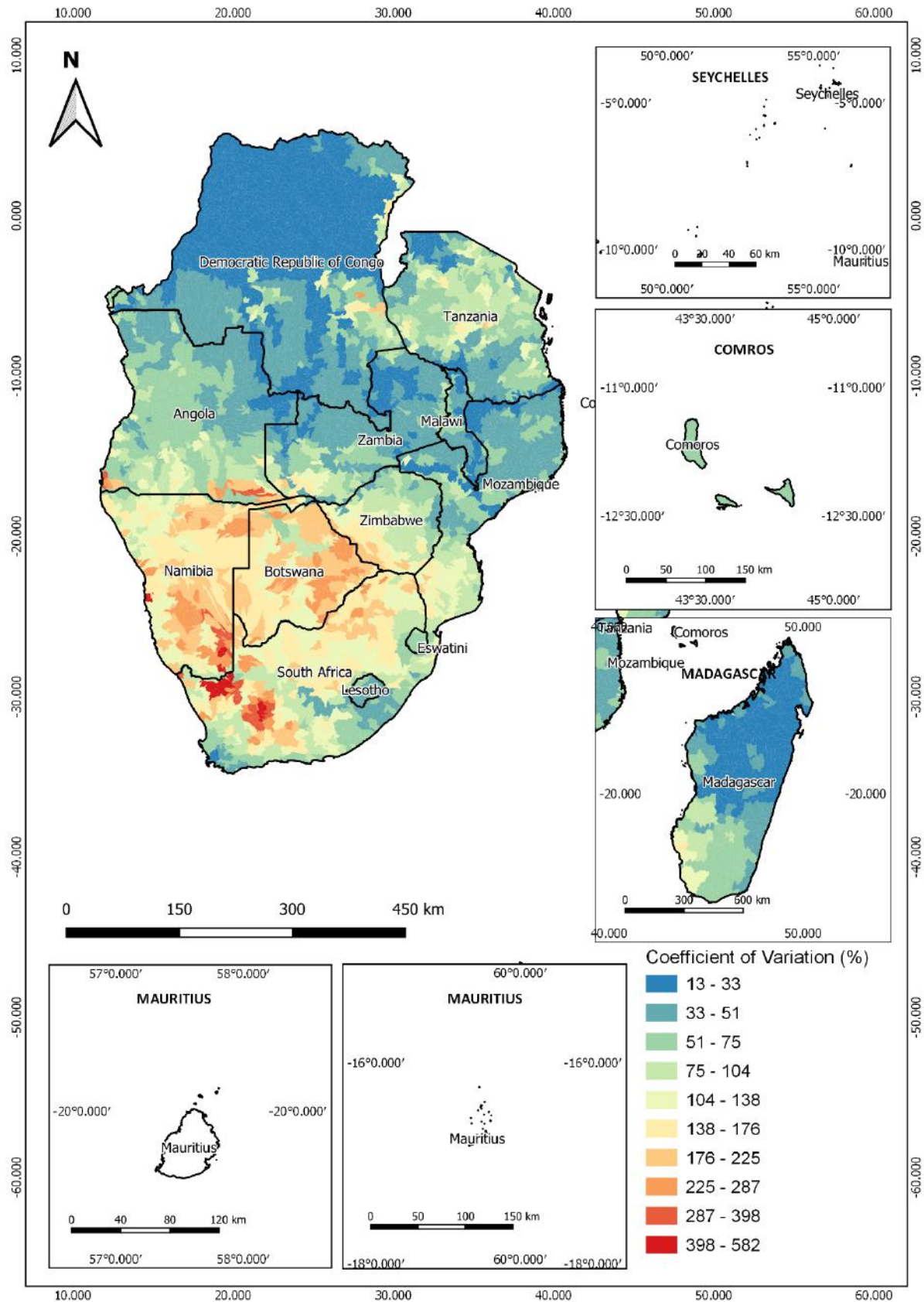**Figure B8: Coefficient of Variation Precipitation**

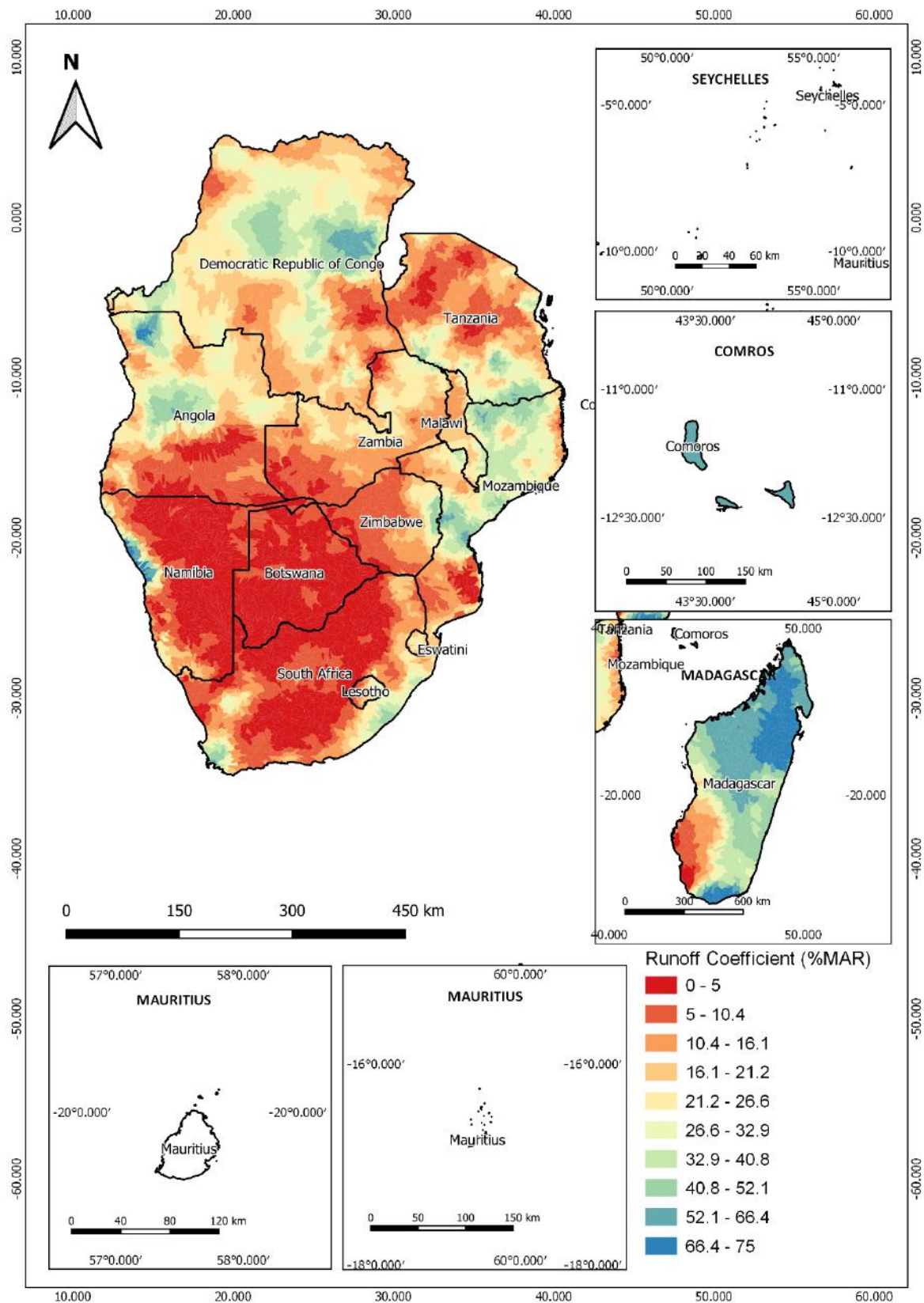**Figure B9: Coefficient of Variability of Discharge**

B5: Runoff Coefficient

Figure B10: Runoff Coefficient as percentage of MAP