Regionalization of Baseline Climatologies and Time Series for the Okavango Catchment

Thomas Weinzierl, Olaf Conrad, Jürgen Böhner & Jan Wehberg

Abstract: The resolution of modern Global and Regional Climate Models is sufficient for many purposes. Nevertheless, the downscaling of model output is necessary when a higher spatial resolution is needed for smaller scale applications. This article introduces intermediate results of an approach for climate regionalization using the free and open source GIS software SAGA, applied on historical and future REMO regional climate model scenarios for the Okavango River Basin of southwestern Africa. The approach makes use of an SRTM digital elevation model and regionalizes temperature via a linear elevation correction/polynomial fit and precipitation with a geographically weighted regression. The results are then compared to the WorldClim interpolated station dataset. For four locations of interest without historical climate observations, time series are generated.

Keywords: downscaling; elevation correction; geographically weighted regression; regional climate.

Introduction

The Okavango River is called the ‘fifeline’ of southwestern Africa (Mendelsohn & Obed 2004) as it provides water in an otherwise largely arid environment and a considerable number of people in the three riparian states directly or indirectly derive their livelihoods from it. Originating in the humid Bié Plateau of central Angola, the river crosses Namibia’s Caprivi Strip and disperses into the largest inland delta in the world, a complex series of heavily vegetated channels and shallow basins in Botswana. The climate in the Okavango Catchment is strongly influenced by the Subtropical High Pressure Zone bringing in dry air from the south and the Intertropical Convergence Zone bringing in moist air from the north (Mendelsohn et al. 2002). Rainfall in the region is highly seasonal with a wet season in the (south) summer, typically from around October to March and a dry season from about April to September. Temperatures are generally high with an extreme daily variability and high evaporation rates. Namibia is one of the driest countries in the world and has no perennial rivers at all originating within its borders. It benefits from the summer rainfalls mostly in its northeastern parts while the southwest has average precipitation rates of only 50 mm and less. The reason for this gradient is the cold Benguela Current in the South Atlantic with its upwellings at Namibia’s coasts. Southern Africa in general and the Okavango basin in particular are subject to processes such as landuse, socioeconomic and climatic changes (Mbawa 2003 & 2011, Weinzierl & Schilling 2013) and a hotspot for biodiversity loss (Biggs et al. 2008). These changes require appropriate adaptation measures as they directly and indirectly affect many sectors like environment, agriculture, water resources, energy, wildlife and diseases. Therefore, policy makers are in need of suitable spatial and nonspatial information such as maps and time series that allows them to identify vulnerable areas, estimate the range and magnitude of climate change impact and thus make reasonable decisions. Assessments of future climate development in the catchment are in high demand because water is a key ecosystem service and the local economies depend on its availability.

The study region can be characterized as a poor data region, as climate observations for southwestern Africa, especially Angola, are relatively scarce and often low quality with only very short measured periods. This is mostly due to historic reasons such as the decades of civil war in parts of the catchment. In their 4th assessment report, the Intergovernmental Panel on Climate Change IPCC indicates a temperature increase above the global annual mean warming for southern Africa toward the end of the century, accompanied with a very likely decrease in precipitation. Especially the dry subtropical regions are expected to be a hotspot of global
warming (IPCC 2007). In the recent years, Global Climate Models (GCM) have emerged as an adequate tool to predict large-scale climatic conditions and assess the impacts of anthropogenic loading of the atmosphere with greenhouse gases (von Storch et al. 1993). These models undergo constant improvement and today reach horizontal resolutions between 100 and 300 km. However, many critical impacts of climate change will occur at sub-regional or local scales, for example at drainage basins and agricultural areas (Wetterhall et al. 2005) but it will not be possible to accurately compute scales of 50 km or less with a Global Climate Model in the near future (Spekat et al., 2008). Therefore, the regionalization of GCM output is required for analyses on smaller scales. A common practice is to use nested dynamical models (RCM - Regional Climate Models) such as REMO that simulate climatic variables of a certain region through the laws of physics and achieve resolutions of 25 km and higher. For even finer resolutions, further downscaling is required. While there are multiple ways of resolving sub-synoptic features using physical or statistical relationships with corresponding upsides and downsides (considering e.g. technical complexity, computing time and physical consistency), the possibilities for the application of small scale climate data are also “endless”, ranging from impact studies to water supply, economy, tourism and the identification of bioclimatic envelopes. The purpose of this article is to present preliminary results of the authors’ regionalization scheme for RCM output for the Okavango region. The work was conducted within the interdisciplinary project “The Future Okavango” (TFO), aiming at the improvement of land use and resource management in the Okavango Catchment as part of the BMBF (German Federal Ministry of Education and Research) funding measure “Sustainable Land Management”.

Data

Four local study sites have been chosen by the project participants: Chitembo (Angola), Caïundo (Angola), Mashare (Namibia) and Seronga (Botswana). At these locations, measurement of climate variables has only recently begun and there is a need for historical and future model data (time series) while spatially explicit data is of interest at the catchment level, for example to determine hotspots of climate change within the domain. Unless otherwise stated, the authors worked with monthly data, though the presented approach is transferable to any other temporal resolution without difficulty assuming the necessary data is available.

The climate data to be regionalized had been computed in a horizontal resolution of 25x25 km with the Regional Climate Model REMO (Jacob 2001) forced with the Global Circulation Model ECHAM for the domain of longitude 8° - 32°, latitude 10° - 27°. Haensler (2011) showed that REMO well reproduces the main temporal and spatial climatic features of the southern African region including the strong seasonality and large interannual variability of rainfall patterns and therefore concludes that the model is suitable for climate studies in the region despite some coastal warm bias. Included is a historical control run spanning the period 1950 – 2005 and two future emission scenarios from 2006 – 2100 in daily temporal resolution. The dataset comprises 18 climate variables at surface level as well as seven tropospheric variables at the 500, 700, 850, 925, 1000, 1005, and 1015 hPa levels.

Table 1: Overview of the TFO core sites

<table>
<thead>
<tr>
<th>TFO Core Site</th>
<th>Country</th>
<th>Coordinates</th>
<th>Altitude in m (REMO)</th>
<th>Altitude in m (SRTM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chitembo</td>
<td>Angola</td>
<td>16.925 E, -13.630 S</td>
<td>1542</td>
<td>1560</td>
</tr>
<tr>
<td>Catundo</td>
<td>Angola</td>
<td>17.805 E, -16.495 S</td>
<td>1181</td>
<td>1155</td>
</tr>
<tr>
<td>Mashare</td>
<td>Namibia</td>
<td>20.213 E, -17.898 S</td>
<td>1076</td>
<td>1068</td>
</tr>
<tr>
<td>Seronga</td>
<td>Botswana</td>
<td>22.415 E, -18.816 S</td>
<td>985</td>
<td>982</td>
</tr>
</tbody>
</table>
1000 hPa level. For this approach, the most important variables are temperature, precipitation and the speed and direction (u- and v-components) of the wind. In the TFO context, two IPCC scenarios or RCP (Representative Concentration Pathways) are used, RCP4.5 as a medium-low and RCP8.5 as a high scenario, in which the values 4.5 and 8.5 indicate the projected increase in radiative forcing (W/m²) until the year 2100 (Meinshausen et al. 2011). The Digital Elevation Model (Fig. 1) used for the elevation correction and as a predictor of the Geographically Weighted Regression of model output was originally derived by the first Shuttle Radar Topography Mission (SRTM) in the year 2000 and upgraded 2005. It was published by the US Geological Survey (USGS) and is freely available in resolutions of up to 90m worldwide. For further use, the DEM was projected to geographic coordinates and resampled to fit the target resolution of 1000m. It is quite common for this kind of DEM to contain error sinks and nodata-gaps known as voids even in the newest versions. These voids were closed using spline interpolation as recommended for small and medium-sized gaps in higher altitudes and heterogeneous terrain by Reuter et al. (2007) in their evaluation of different interpolation methods for SRTM voids.

As mentioned above, climate observations for the Okavango Catchment, especially the Angolan part, are scarce and often heavily fragmented. For comparison with model data, two sources of observational data are used. First, the renowned Global Historical Climate Network (GHCN) monthly dataset (Lawrimore et al. 2011), which contains point-source data for a number of stations in the domain but not necessarily inside the catchment. Wherever possible, the climate stations nearest to the catchment were used. Second, the gridded Worldclim dataset (Hijmans et al. 2005) of temperatures and precipitation, compiled and interpolated from average monthly station data to a resolution of 30 arc-seconds or 1km².

Methods

This publication is focused on the regionalization of temperature and precipitation with a polynomial fit of 3D temperature data and a geographically weighted regression of 2D precipitation data respectively. As a comparison, temperature is also regionalized by correcting the altitude with a constant lapse rate.

The handling of large data quantities such as climate data makes certain demands concerning the choice of tools used in the process. There is no perfect approach and tools may differ from researcher to researcher and from institution to institution, however, certain useful instruments emerged over time. An advantage of the presented approach is that, besides an up-to-date Personal Computer (using any of the established operating systems) and the discussed software and data, nothing else is needed. For easy reproduction, an effort was made to utilize freely available open-source software rather than commercial products.

The storage of climate data in netCDF file format has become widely established in climate sciences over the recent years. A metadata standard for netCDF files has been introduced with the so-called "climate and forecast (CF) metadata conventions". Based on this standard a number of applications have been developed that allow relatively easy handling of such data. One of them is CDO (Climate Data Operators), a command-line-based set of tools for the basic handling of netCDF files, developed by the Max Planck Institute for Meteorology in Hamburg for processing and analysis of data produced by a variety of climate and numerical weather prediction models. Supported basic tasks include file operations, simple statistics, arithmetic, interpolation or the calculation of climate indices (Kaspar et al. 2010). Visualization of spatial explicit data, geographical projection, resampling, calculations and further tasks are carried out in SAGA (System for Automated Geoscientific Analyses), a cross-platform open source Geographic Information System (GIS) software developed by the Departments of Physical Geography in Göttingen and Hamburg (Conrad 2006). Its modular structure enables the exchange and addition of libraries for a wide range of geoscientific analyses. As geoprocessing tasks are often performed on a number of different datasets or on large datasets with numerous records, they can be time consuming. The SAGA distribution therefore contains a command-line version (SAGA CMD) for the automation of simple to relatively complex work flows by scripting in Python or other languages. The scripts are recyclable and can be used again when written data nonspecific (Tucker 2004).

Fig. 2: Difference between 25x25km REMO elevation and 1x1 km SRTM. Red areas are underestimated and blue areas overestimated by the regional climate model. In green areas (predominantly flat areas) the assumed elevation approximately coincides.
Temperature

Temperature controls a large variety of environmental processes and is an essential input parameter for land surface models, for example in hydrology, ecology and climatology (Gao et al. 2012). Throughout the troposphere (up to about 8km above the poles and about 18km over the equator), with an increase in elevation and therefore an increasing distance from the heating earth surface, air temperature decreases (Mohanakumar 2008). The vertical lapse rate averages 6.5°C per kilometer or 0.0065°C per meter (Root 2011). GCM/RCM temperature data can be regionalized by performing an elevation correction, based on the assumption that elevation is the main factor determining the distribution of surface temperature (Feulner et al. 2013). Because of the “coarse” resolution (2.5 x 25km) of the RCM data, the altitude assumed by the model is relatively inaccurate (Fig. 2), which (especially in mountainous areas) has a distorting effect on variables that change with elevation. Therefore, an elevation correction is required, for which there are two established methods, the linear and polynomial. In a first step, DEM and model data are projected to the same coordinate system and resampled to fit the target resolution of 1km using spline interpolation (Conrad 2006). The resampling process for itself adds no information to the data. While the linear correction method uses a constant vertical lapse rate to correct 2m temperature, a polynomial trend can be calculated and fitted through the free atmosphere temperatures (FAT) at certain pressure levels, provided that three-dimensional data is available. When added a high resolution DEM, a value for the corrected surface height can be derived from the polynomial equation. As shown by Kawohl (2013), the differences between corrected elevation with linear and polynomial method are marginal, except for very mountainous regions. The southern part of the Okavango Catchment is relatively plain, so the temperature values of the RCM are already quite exact in this area because the smoother the topography, the more realistic the elevation assumed by REMO (Fig. 2).

Precipitation

Precipitation is a crucial factor in this mainly semiarid region and the most important driving variable in hydrological modeling and analysis (Wetterhall et al. 2005). Precipitation downscaling is performed using Geographically Weighted Regression (GWR) as proposed by Brunsdon et al. (1998). Instead of fitting one global regression model GWR creates a set of local regression models, for which the influence of each observation is weighted according to its distance to the

Fig. 3: Average temperature 1950-2000 for March in a) original REMO 25x25km resolution b) Worldclim 1 x1 km interpolated station data c) 1 x1 km altitude adjusted T2m and d) 1 x1 km interpolated Free Atmospheric Temperatures.
Fig. 4: Average precipitation 1950-2000 for March in a) original REMO 25x25km, b) Worldclim interpolated station data 1x1 km, c) downscaled with GWR and d) GWR including residual correction.

Fig. 5: Average precipitation 1950-2005 for March in a) original 25x25km and b) downscaled 1 km resolution. The active Okavango Catchment and four TFO core sites are marked in red.
model location. This way GWR is able to reflect in dependence of the location changing relations between observed and predictor variables Brunsdon et al. (2001) showed the advantages of GWR compared to global regression for the regionalization of precipitation using elevation as single predictor. In this study the modeled precipitation is taken as observation. Besides elevation a windward-leeward index has been used as additional predictor, assuming that areas exposed to wind force the moving air masses to rise and thus yield a higher precipitation than areas in the lee side. This index is calculated for each month separately based on the DEM and the modeled wind field.

To generate time series for the four TFO core sites and to be able to compare the model output with climate station observations, nearest-neighbor interpolated values for the 9 grid cells around each location of interest were extracted from the regionalized data. This technique was used to compensate potential deviations. For a better overview and to exclude the extreme intra-year variations, the data was combined to a yearly temporal resolution. An 11-year symmetric moving mean was calculated and added to better illustrate long-term trends.

**Results and Discussion**

The regionalized monthly temperature and precipitation values were compared with available station observations from the GHCN network. In doing so, the height-corrected temperatures were found to represent the observed data quite good with comparatively low Root Mean Square Errors (Fig. 6b). In the (south) winter months, the model overestimated the temperatures a little at most stations (figure 6a). As the winter months are also the dry season in the region, both mean deviation and RMSE at the precipitation stations are very low for that time period because there is practically no precipitation. In the summer months, the model tends to overestimate rainfall by around 20-60 mm each month (Fig. 6c & d).

A comparison of the regionalized monthly long-term values from 1950 to 2000 with the gridded WorldClim dataset interpolated from stations, also in 1km resolution, shows that the model reproduces the regional climate in terms of temperature and precipitation very well (Fig. 3 and 4), except the warm and wet bias mentioned in chapter 2, especially in the northwestern coastal regions. Precipitation is best represented by the GWR including residual correction. Time series generated from the regionalized data (see Fig. 7 to 14) suggest that all four core sites will experience rising temperatures in the future. The estimated change ranges from 2 to 2.5 °C in scenario rcp4.5 to almost 6 °C in scenario rcp8.5. Most of the core sites with the exception of Chitembo as the northernmost site also show a decreasing precipitation of around 20% in the scenarios, most distinctive in the second half of the century. This development might put further pressure on the already stressed water situation of the present.

The added value of the performed elevation correction increases with the topography of the targeted area. In this case study, the linear 12m correction performs better than the polynomial interpolation of FAT. A likely reason is that in the available dataset, only two to three of the available pressure levels are situated above the surface for most of the domain.

Note that the results presented have not yet been bias-corrected. For a bias-correction, an in-depth residual analysis of temporal and spatial regularities as well as observed data with substantial spatial and temporal resolution for the region is required. However, working with monthly data, an adjustment according to Skaggen et al. (2002) using the mean monthly ratio between model and observed precipitation in the control period and a monthly regression equation of temperature might lead to satisfactory results.

**Conclusions**

The article presented preliminary results of the climate regionalization scheme with a target resolution of 1km for the Okavango region. With the presented methods, a dataset of high resolution monthly mean surface temperatures and precipitation was generated for the historical timeframe of 1950-2005 and two scenarios from 2006-2100. An elevation correction of temperature data added value especially in the more mountainous areas of the northern catchment. For the downscaling of precipitation data, a Geographically Weighted Regression algorithm with DEM based orographic parameters as predictors so far proved to be a reasonable approach. A possibility to enhance this
method is to find additional suitable predictors such as remote sensing data and its derivatives (e.g. NDVI, NPP). The presented methods are implemented in the open source software SAGA-GIS and therefore freely available to the scientific community for reproduction and improvement. The downscaled results can, for example, be used to improve spatially explicit assessments of the projected temperature and precipitation changes, by comparing the scenarios with a historic reference period. Also, a comparison between the presented results and those using different GCM/RCM combinations (with EC-EARTH, WRF) as well as a deduction of parameters such as evaporation is currently in progress. Naturally, regionalized climate data can only be as good as the model data that is used for input. With the rapid technical development in recent years, the underlying global and regional climate models undergo constant improvement.

Acknowledgement

We would like to thank our colleagues in the TFO project for their support and many stimulating discussions. Also, we owe special thanks to our project partners from the Climate Service Center Hamburg, who run the REMO simulations this work is based on.

References


Affiliations

Thomas Weinzierl* (thomas.weinzierl@uni-hamburg.de), Olaf Conrad (conrad@geowiss.uni-hamburg.de), Jürgen Böhner (boehner@geowiss.uni-hamburg.de), Jan Webber (webberg@geowiss.uni-hamburg.de) Department of Geography, University of Hamburg Bundesstr. 55 20146 Hamburg, GERMANY
ANNEX

ECHAM/REMO: Chitembo

**temperature scenarios (elevation-adjusted) with 11-year moving mean**

![Temperature graph]

*Corresponding author

**Fig. 7:** Time series of regionalized ECHAM/REMO temperature scenarios for Chitembo (Angola).

**Fig. 8:** Time series of regionalized ECHAM/REMO precipitation scenarios for Chitembo (Angola).
Fig. 9: Time series of regionalized ECHAM/REMO temperature scenarios for Caiundo (Angola).

Fig. 10: Time series of regionalized ECHAM/REMO precipitation scenarios for Caiundo (Angola).
**Fig. 11:** Time series of regionalized ECHAM/REMO temperature scenarios for Mashare (Namibia).

**Fig. 12:** Time series of regionalized ECHAM/REMO precipitation scenarios for Mashare (Namibia).
**Fig. 13:** Time series of regionalized ECHAM/REMO temperature scenarios for Seronga (Botswana).

**Fig. 14:** Time series of regionalized ECHAM/REMO precipitation scenarios for Seronga (Botswana).