Downscaling climate projections over large and data sparse regions: Methodological application in the Zambezi River Basin

Nadav Peleg¹ | Scott Sinclair¹ | Simone Fatichi¹,² | Paolo Burlando¹

¹Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland
²Department of Civil and Environmental Engineering, National University of Singapore, Singapore

Abstract
Climate impact studies often require climate data at a higher space–time resolution than is available from global and regional climate models. Weather generator (WG) models, generally designed for mesoscale applications (e.g., 10³–10⁵ km²), are popular and widely used tools to downscale climate data to finer resolution. One advantage of using WGs is their ability to generate the necessary climate variables for impact studies in data sparse regions. In this study, we evaluate the ability of a previously established state of the art WG (the AWE-GEN-2d model) to perform in data sparse regions that are beyond the mesoscale, using the Zambezi River basin (10⁶ km²) in southeast Africa as a case study. The AWE-GEN-2d model was calibrated using data from satellite retrievals and climate re-analysis products in place of the absent observational data. An 8-km climate ensemble at hourly resolution, covering the period of 1976–2099 (present climate and RCP4.5 emission scenario from 2020), was then simulated. Using the simulated 30-member ensemble, climate indices for both present and future climates were computed. The high-resolution climate indices allow detailed analysis of the effects of climate change on different areas within the basin. For example, the southwestern area of the basin is predicted to experience the greatest change due to increased temperature, while the southeastern area was found to be already so hot that is less affected (e.g., the number of ‘very hot days’ per year increase by 18 and 9 days, respectively). Rainfall intensities are found to increase most in the eastern areas of the basin (1 mm·d⁻¹) in comparison to the western region (0.3 mm·d⁻¹). As demonstrated in this study, AWE-GEN-2d can be calibrated successfully using data from satellite reanalysis products in the absence of ground station data and can be applied at larger scales than the mesoscale.

Keywords
climate change, climate indices, extreme climate indices, rainfall, stochastic downscaling, temperature, weather generator, Zambezi River
1 | INTRODUCTION

Anthropogenic activities are altering climate conditions (Huber and Knutti, 2012), triggering consequences such as changes of the hydrological cycle and hence water resources at global, regional, and local scales (Kundzewicz et al., 2008; Wu et al., 2013). Models of different complexity can be used to explore and quantify the hydrological alterations; for example future changes in the streamflow of small catchments (Peleg et al., 2015) or large river systems (Fatichi et al., 2015), flow in urban environments (Willems et al., 2012), modification of floods (Hirabayashi et al., 2013), changes in water availability (Immerzeel et al., 2010) and changes in hydropower operation decisions (Anghileri et al., 2018) have all been analysed using a range of approaches to simulate climate change scenarios.

In general, hydrological models require climatic inputs at a range of spatial and temporal resolutions, from $10^{-3}$ to $10^2$ km and minutes to years, depending on the structure of the hydrological model that is used (e.g., lumped or distributed model, process based or conceptual model), as well on size and complexity of the catchment (e.g., topographic variability, rural vs. urban catchment) (Bloschl and Sivapalan, 1995; Cristiano et al., 2017). In absence of better information, large catchments, in the order of $10^4$–$10^6$ km$^2$, can be modelled using climate inputs at $10^2$ km and daily to monthly resolution to estimate hydrological fluxes and storages, such as runoff (Christensen and Lettenmaier, 2007). The recent improvements of distributed hydrological models (Fatichi et al., 2016b) and access to computational resources, however, raised the demand for climate input data at higher space–time resolutions. Recent studies have demonstrated the improvement in hydrological simulations when using climate data at the fine resolution of $10^3$–$10^4$ km and hourly to daily scale to reproduce the hydrology over large catchments (Chen et al., 2017; Mateo et al., 2017).

Concurrently, the use of hydrological models to explore how the water cycle is responding to climate change requires climate variables that are derived from climate models, such as General Circulation Models (GCMs) or Regional Climate Models (RCMs) (Xu, 1999a; Christensen et al., 2004). Due to the typically coarse spatial resolution and temporal archiving of the output from GCMs (10$^2$ km and daily/monthly) and RCMs (10$^1$ km and daily), a downscaling procedure is often applied to increase the resolution of the data and match the requirements for the inputs of hydrological model used for impact analyses (e.g., Burlando and Rosso, 2002). Downscaling approaches can be generally divided into three categories: dynamical, statistical, and stochastic (Xu, 1999b; Fowler et al., 2007; Maraun et al., 2010; Peleg et al., 2019). The stochastic approach has been popular among hydrologists, as it is relatively efficient in terms of computational demands and does not require advanced statistical knowledge or computational capabilities to allow a direct computation of the stochastic climate uncertainty (Burlando and Rosso, 1991; Fatichi et al., 2013; Bordoy and Burlando, 2014; Glenis et al., 2015; Peleg et al., 2019). Stochastic uncertainty, also called internal or natural climate variability, originates from the chaotic nature of the climate system (Deser et al., 2012). Its relative contribution to the total uncertainty in climate change projections can be significant and it is often the largest source of uncertainty for precipitation. However, its importance depends both on the spatial and temporal scales and on the period of interest (Hawkins and Sutton, 2011; Fatichi et al., 2016a; Giuntoli et al., 2018).

Stochastic downscaling is often performed using weather generator (WG) models in hydrological impact studies (e.g., Fatichi et al., 2015; Peleg et al., 2015; Chen et al., 2016; Camici et al., 2017; Singer and Michaelides, 2017; Sorup et al., 2018). WGs simulate time series of multiple climate variables either for a given location (Semonev and Barrow, 1997; Fowler et al., 2005; Kilsby et al., 2007; Fatichi et al., 2011), multiple locations (Keller et al., 2015) or over gridded domains (Peleg et al., 2017) while preserving the observed statistics and internal space–time correlations between the variables. WGs can be used to simulate time series of climate variables for a future period by re-parameterizing the model to follow the changes in climate statistics that are derived from GCMs or RCMs (e.g., Burlando and Rosso, 1991; Fatichi et al., 2011, 2013; Peleg et al., 2019). In such a way, climate model statistics are implicitly downscaled locally.

In addition to the uses of WGs described above, WG models are also useful to infill missing climatic data and generate coherent climate time series for data sparse regions (Wilby and Yu, 2013; Camberlin et al., 2014; Wilby et al., 2014). One powerful advantage in applying WG in ungauged basins is that the necessary climate variables for hydrological studies can be generated for the period of interest and at the required space–time resolution starting from an originally coarser (in space and time) information. These are often taken from remote sensing and climate re-analysis products, which are assumed to be the best (and often only) alternative to ground station measurements when the latter are not available.
The scope of this study is to evaluate how a previously established stochastic downscaling methodology that makes use of state-of-the-art, distributed WGs performs in large and data-sparse regions. This is particularly challenging because gridded WGs are designed for mesoscale applications (e.g., $10^1$–$10^5 \text{ km}^2$) and are typically trained using a wealth of ground and radar observations. In this study, we used the recently developed Advanced WEather GENerator for 2-dimensional grid model (AWE-GEN-2d, Peleg et al., 2017, 2019). Its applicability to a large and data-poor region is demonstrated by calibrating it for the Zambezi River Basin (ZRB), a large catchment in Africa with some sparse climate station data at the daily scale (e.g., Gumindoga et al., 2019), yet with poor coverage of climate data at the sub-daily scale (hourly). Climate indices of rainfall and temperature for the present (1976–2005) and future (2020–2099) climates, as well as the stochastic uncertainty of the predictions, were computed using the AWE-GEN-2d model outputs. A successful outcome of using the WG model for downscaling climate variables over large and data-sparse regions can increase the benefits and enlarge the usefulness of WG applications. This is the case because reanalysis products and climate models are increasingly available at the global scale and can be used to downscale climate change signals from large scales to fine space–time resolutions, that is, at $10^3 \text{ km}$ and sub-daily scales for large regions around the world, where detailed climate observations—at these scales—are lacking.

2 STUDY AREA

The ZRB (Figure 1) is the fourth largest basin of Africa with an area of about 1.4 million km$^2$, and elevation that ranges from approximately 2,500 m to sea level. The basin is located in southeast Africa and its waters are shared by eight riparian states (Angola, Botswana, Malawi, Mozambique, Namibia, Tanzania, Zambia, and Zimbabwe), with almost 40 million inhabitants living within the catchment.

The Zambezi River covers roughly 2,600 km from its source at the heart of the mainland before reaching its delta at the coast of the Indian Ocean. Seasonal flood-plains and swamps are dominant in the western part of the basin, inundating large areas ($10^3 \text{ km}^2$). The basin includes several large lakes and artificial reservoirs for hydropower generation, including Lake Malawi (northeastern part of the basin, 30,000 km$^2$), Lake Kariba (5,500 km$^2$) and Lake Cahora Bassa (2,700 km$^2$). Mean annual rainfall varies between 1,500 mm in the north-eastern part of the basin to 500 mm in the south (Figure 2). Average annual temperatures range between 26°C in the southeastern areas of the basin and 18°C in the north-east and northwest areas of the basin. About 8% of the rainfall contributes to the river runoff and the remaining 92% is lost via evapotranspiration (Kling et al., 2014). The dry season (July to August) is mild and dry, while the wet season (November to March) is wet and hot.

![Digital elevation model and river network of the Zambezi River basin](image)
A long record of observations of climate variables, in the order of 30 years, is recommended to calibrate and validate weather generator models (Fatichi et al., 2013; Kim and Ivanov, 2015; Peleg et al., 2017). In this case study, AWE-GEN-2d was set to simulate the climate variables at sub-daily resolution; hence, sub-daily observations are required. In addition, AWE-GEN-2d is used to simulate the climate variables over a 10 km grid. In such a case, multiple climate ground stations, with inter-distance in the order of 10–100 km, would be the ideal data for a satisfactory validation of the model. While multiple climate stations are present in the ZRB (e.g., Cohen Liechti et al., 2012; Thiemig et al., 2012; Gumindoga et al., 2019), only a few of them are archiving climate variables at sub-daily resolution. Moreover, the data is often not accessible, and has variable quality and record length. Therefore, we consider the ZRB to be a data sparse region.

To overcome this issue of lack of densely distributed sub-daily station observations, alternative climate products were used. We targeted freely accessible climate data to calibrate and validate the weather generator. Data were obtained from climate reanalysis and remote sensing products that fulfilled the required space–time resolution of the model. The weather generator was validated using independent sets of climate data, obtained from gridded climate products, station measurements, and information extracted from the scientific literature. Projections of the future climate were based on regional climate models. The data sources for both the present and future climates are described in the following sections.

### 3.1 Present climate data

Gridded rainfall data were obtained from the CMORPH (Joyce et al., 2004) product. CMORPH rainfall estimates were investigated in the literature and compared against rain-gauge measurements and another reconstructed product, and then were assumed to represent various rainfall characteristics over the ZRB (Cohen Liechti et al., 2012; Thiemig et al., 2012; Awange et al., 2016; Gumindoga et al., 2019) and surrounding areas (Sahlu et al., 2017). CMORPH was preferred when compared to other remote sensing and climate reanalysis products due to its relatively high space–time resolution (8-km and 30 min) that is required to calibrate the weather generator. We used a recently released bias-corrected version of CMORPH (Xie et al., 2017), for which data is available for the period 1998–2017. To the best of our knowledge, this bias-corrected version of CMORPH was not yet evaluated over Africa. A simple comparison between annual rainfall interpolated based on rain-gauges (Figure 2a, following Gumindoga et al., 2019) and the newly CMORPH product (Figure 2b) showed that the bias-corrected CMORPH product underestimates the annual rainfall by about 60 mm/y when averaged over the catchment. This is an improvement in comparison to the precedent
uncorrected CMORPH product for which Thiemig et al. (2012) reported an overestimation of 235 mm·y$^{-1}$.

We note, however, that: (a) the interpolated annual rainfall corresponds to a different period (gauged data span between 1998 and 2010, with many gaps, see Guminzoga et al., 2019) than the CMORPH data, thus possibly affecting the above comparison; and (b) interpolating ground measurements using different techniques may lead to different rainfall maps compared to the annual rainfall map presented by Thiemig et al. (2012).

Data for all the other climate variables that are required for the calibration of the weather generator, namely near-surface (at 2 m) air temperature, near-surface relative humidity, wind speed at 850 hPa, shortwave global radiation and total cloud cover, were obtained from the MERRA-2 climate reanalysis product (Gelaro et al., 2017). This product was chosen due to its frequent temporal resolution and relatively high spatial cover (grid size is approximately 50 km × 66 km and temporal resolution is 1 hr). The lack of ground data challenges the evaluation of the accuracy of the MERRA-2 products. Draper et al. (2018) compared the 2-m air temperature from the MERRA-2 with CRU data at the global scale. For the stations thereby included, they reported, over most of the ZRB area, no bias between the two products for maximum, minimum and diurnal temperature range, and a negative bias of up to 2°C for the eastern area of the basin. The accuracy of the shortwave radiation variable from the MERRA-2 product over Africa was recently evaluated against four ground stations in Ethiopia by Stettz et al. (2019), who reported a good agreement between the climate reanalysis data and the observed data. Miao et al. (2019) compared the cloud cover of MERRA-2 with Cloud-Sat/CALIPSO data, showing that the MERRA-2 underestimates the cloud cover for the ZRB region by a maximum of 10%. Torralba et al. (2017) compared the wind speed trends from MERRA-2 and JRA-55 climate reanalysis products, concluding that the two products yield similar trends, thus suggesting MERRA-2 to be a good data source to evaluate the wind speed at 850 hPa. Overall, the MERRA-2 reanalysis product is thus likely to provide robust data for calibration of the WG, compensating the lack of ground measurements over the ZRB domain. MERRA-2 is available on the same period of the CMORPH data (1998–2017).

Data from multiple independent sources were used for validating the model. Near-surface air temperature (simply temperature from hereafter), incoming global radiation and relative humidity were validated using gridded climate products from the WorldClim (Fick and Hijmans, 2017), Global Solar Atlas (World Bank, 2017) and Biosphere Atlas (New et al., 1999). Additionally, at the point scale, meteorological data were obtained from ground stations reported in the Global Historical Climatology Network (GHCN, Menne et al., 2012), scientific reports (Dörner, 2013; Suri et al., 2014; Kapumpu, 2015) and published articles (McGregor and Nieuwolt, 1998; Shahin, 2006).

### 3.2 Regional climate models

Future climate trajectories were obtained using an ensemble of 22 RCMs (Table 1) from the Africa-CORDEX initiative (Giorgi et al., 2009; Mariotti et al., 2011; Nikulin et al., 2012). The chosen RCMs have daily temporal resolution and spatial resolution of about 50 km (0.44°). Past studies demonstrated the ability of using an ensemble multi-model mean approach (as done in this study, see Section 4.2) to reliably project future changes in rainfall

| RCM     | GCM          | #1 | #2 | #3 | #4 | #5 |
|---------|--------------|----|----|----|----|----|
| CCLM4-8-17 | CNRM-CM5 | x  | x  | x  | x  |    |
|         | EC-earth     | x  | x  | x  |    |    |
|         | HadGEM2-ES  | x  | x  |    |    |    |
|         | MPI-ESM-LR  | x  | x  | x  |    |    |
| CRCM5   | CanESM2     | x  | x  |    |    |    |
|         | MPI-ESM-LR  | x  |    |    |    |    |
| HIRHAM5 | EC-earth     | x  | x  |    |    |    |
|         | NorESM1-M   | x  |    |    |    |    |
| RACMO22T | EC-earth    | x  | x  | x  |    |    |
|         | HadGEM2-ES  | x  | x  |    |    |    |
| RCA4    | CanESM2     | x  |    |    |    |    |
|         | CNRM-CM5    | x  | x  | x  | x  |    |
|         | CSIRO-Mk3-6-0 | x |    |    |    |    |
|         | EC-earth     | x  | x  | x  |    |    |
|         | GFDL-ESM2M  | x  |    |    |    |    |
|         | HadGEM2-ES  | x  |    |    |    |    |
|         | IPSL-CM5A-MR |    |    |    |    |    |
|         | MIROC5      | x  |    |    |    |    |
|         | MPI-ESM-LR  | x  | x  | x  |    |    |
|         | NorESM1-M   |    |    |    |    |    |
| REMO2009 | EC-earth    | x  | x  |    |    |    |
|         | MPI-ESM-LR  | x  | x  |    |    |    |

**Note:** Climate models used in previous studies, which are discussed in Section 3.2, are marked with an ‘x’: #1 Onyutha et al. (2019); #2 Gibba et al. (2019); #3 Warnatzsch and Reay (2019); #4 Luhunga et al. (2016); #5 Haensler et al. (2013).
and temperature over central and east Africa domains. Nikulin et al. (2012) compared the annual cycle from 10 RCMs (overlapping RCMs to the ones used in this work are indicated in Table 1) with rainfall data emerging from various satellites and climate reanalysis products, concluding that there is a good agreement between the observed and the simulated multi-model mean rainfall. Endris et al. (2013) used the same RCMs as Nikulin et al. (2012) to show that the simulated rainfall multi-model mean adequately represents the rainfall climatology over east Africa. Haensler et al. (2013) used a multi-model mean approach to estimate the changes in annual rainfall over central Africa based on an ensemble composed of RCMs from the Africa-CORDEX and GCMs. They found a low bias (5%) between the ensemble estimates and the observed rainfall over most of the domain, with a small area (over north Zambia) exhibiting higher bias up to a maximum of 20%. The RCMs were evaluated to minimum and maximum temperature and rainfall over Tanzania (Luhungu et al., 2016) using data from ground stations. It was found that the RCMs capture the annual cycle of minimum and maximum temperature as well as mean annual rainfall well. Van Vooren et al. (2019) reported that the simulated multi-model mean rainfall overestimates ground observations and climate reanalysis data over Ethiopia. They concluded that the overestimation is a result of orography effect over high mountainous region that is not properly simulated by the CORDEX models, however, we note that the topography of the ZRB is less significant in comparison to the area studied by Van Vooren et al. (2019). The ability of the RCMs to reproduce extreme rainfall was examined by Gibba et al. (2019), who concluded that the multi-model mean produces rainfall magnitude within the range of the observations over Africa domain. Warnatzsch and Reay (2019) evaluated the ability of the RCMs to reproduce the climate in Malawi. They concluded that while the multi-model mean can be used to estimate future changes in temperature, there is a need to improve the representation of rainfall in the climate models, as RCMs simulated outputs were considerably divergent from the observations. Onyutha et al. (2019) also indicated that rainfall emerging from Africa-CORDEX multi-model mean should be treated with care, as they found biases in the rainfall magnitudes over the Lake Victoria basin. However, they concluded that the RCMs had enough skill to be used in the context of studying changes in climate.

In this demonstrative study, only the RCP4.5 emission scenario, characterized by a moderate increase in global temperature and greenhouse gas concentration (Meinshausen et al., 2011), is considered. RCP4.5 is chosen as an intermediate case between the extreme emission scenario (RCP8.5) and the desirable emission scenario (RCP2.6) targeted by the ‘Paris agreement’ in 2015 (Sanderson et al., 2016), which seems, however, unlikely to be achieved (Millar et al., 2017; Raftery et al., 2017).

4 | METHODS

4.1 | The AWE-GEN-2d model

The AWE-GEN-2d model (Peleg et al., 2017) is a stochastic weather generator that can be used to downscale ensembles of gridded climate variables at high spatial (sub-kilometre) and temporal (sub-daily) resolutions. The model is parameterized to simulate the key climate variables that are needed for hydrological, agricultural, and ecological applications for the present climate using a combination of stochastic and physically-based methods and can be re-parameterized using information obtained from climate models to simulate climate variables representative of future periods. A detailed description of the model and the calibration process is presented by Peleg et al. (2017), whereas the methodology to re-parameterise it for future climate simulation is described by Peleg et al. (2019).

AWE-GEN-2d was used in this study to simulate a downscaled ensemble of rainfall, temperature, cloud cover, shortwave radiation, relative humidity and near-surface wind speed for the ZRB. The model was first calibrated using CMORPH (for rainfall) and MERRA-2 (for all other climate variables) products and subsequently used to simulate climate variables at hourly and 8-km resolution for the period of 1976–2005, as presented and discussed in Section 5.1. The AWE-GEN-2d model was then re-parameterized, as explained in the following sections, to simulate the period of 2020–2099. To assess the natural (stochastic) climate uncertainty in the present and future periods (Peleg et al., 2019), an ensemble consisting of 30 realizations was simulated. For the present climate, for example, the control period, the ensemble consists of 30 realizations of 30 years each, corresponding to the period of 1976–2005. For the future, each realization is representative of a stochastic climate trajectory (2020–2099) originated from the multi-model mean of climate change signals derived from the RCM ensemble.

4.2 | Factors of change

AWE-GEN-2d was re-parameterized using the “factors of change” (FC) approach according to a quite established methodology (Burlando and Rosso, 1991; Kilby et al., 2007;
Anandhi et al., 2011; Faticchi et al., 2011, 2013; Bordoy and Burlando, 2014; Peres and Cancelliere, 2018; Peleg et al., 2019). This approach considers the changes in the long-term mean of climate variables and accounts for their seasonality. FC are computed as the ratio or arithmetic difference (depending on the climate variable) between statistics of climate variables from a realization corresponding to the future and one of the historical period, both obtained from a climate model (RCM or GCM). FC can be applied to modify statistics at finer temporal scales than the scale of the original climate model with considerations about scale dependency, which may be non-trivial for higher order statistics (Faticchi et al., 2011; Bordoy and Burlando, 2014). Here, FC are computed at the daily scale and applied to the weather generator at an hourly scale following Peleg et al. (2019).

FC were computed for the mean change in rainfall and temperature for 7 overlapping periods of 30 years each, covering the future climate from 2010 until the end of the century (i.e., 2010–2039, 2020–2049, ..., 2070–2099). The FC were computed on a seasonal basis, considering the months December–January–February (DJF), March–April–May (MAM), June–July–August (JJA) and September–October–November (SON).
(DJF), March–April–May (MAM), June–July–August (JJA) and September–October–November (SON). Note that FC for rainfall means were not derived for the dry season (JJA) because rainfall is consistently close to zero during this period and even large percentage change would reflect in only few mm of differences. The FC were computed first for each of the RCMs individually (examples for the FC of mean rainfall and temperature are given in Figures S1 and S2, respectively). The climate model uncertainty (Deser et al., 2012), emerged from the spread of FC computed with the different climate models, and it seems to contribute significantly to the total uncertainty in climate for this region (Figure S3). AWE-GEN-2d can be used to estimate both the stochastic (natural) uncertainty and climate model uncertainty (see Peleg et al., 2019) by running multiple simulations of the future climate. In this case, accounting for the climate model variability and the natural (stochastic) uncertainty would require performing 22 climate models × 80 years × 30 realizations. However, we limited the number of simulations to 80 years × 30 realizations by using a multi-model mean simulation approach, because this allows to fit the purpose of the paper, while limiting the data output to a reasonable amount but still delivering meaningful results for the mean change in climate. This approach required computing the multi-model mean FC over the ZRB for each of the periods and seasons (Figure 3). A bi-linear interpolation was used to rescale the FC from the original 50-km resolution of the climate models to 8-km resolution of the weather generator, similarly to Peleg et al. (2019). The FC show an increase in the spatial heterogeneities of both rainfall (Figure 3a) and temperature (Figure 3b) toward the end of the century and during the warmer season (SON), with no notable shifts in rainfall seasonality predicted by the climate models.

### 4.3 Changes to the rainfall spatial structure

Changes in climate are expected to affect the rainfall spatial patterns and structure, especially in a domain as large as the ZRB. The changes in statistics describing the rainfall field at a given time cannot be directly computed from the RCMs due to the relatively coarse spatial and temporal resolution of the climate models. In recent studies, Wasko et al. (2016), Lochbihler et al. (2017) and Peleg et al. (2018) have demonstrated that there is a relation between increasing temperature and changes in rainfall area properties. Here, we computed the scaling relationship between temperature and two key spatial properties of the rainfall: wet area ratio (the number of wet grid cells divided by the total number of grid cells for each time step) and the rainfall spatial coefficient of variation (the rainfall standard deviation divided by the rainfall mean for each time step). Paschalis et al. (2013) and Peleg et al. (2020) explain in details the effects of these two parameters on the rainfall spatial structure. These rainfall properties, obtained from the CMORPH data, were regressed with temperature data, obtained from the MERRA-2 reanalysis, using a linear fit on a logarithmic scale as suggested by Peleg et al. (2018). The results of the rainfall properties scaling with temperature are presented in Figure 4. After excluding the grid cells above or close to surface water bodies, which are known to be biased in the CMORPH product (Tian and Peters-Lidard, 2010), the computed regression coefficients of each grid cell were averaged over the basin. We found that the wet area ratio increases at a rate of 1.06% °C⁻¹ and that the rainfall spatial coefficient of variation increases at a rate of 1% °C⁻¹. These parameters were modified in the AWE-GEN-2d simulations assuming that the scaling laws derived in the present climate will hold true in the future.

**FIGURE 4** Scaling of rainfall wet area ratio (a) and rainfall spatial coefficient of variation (b) with temperature in each 8 × 8 km² grid cell in the present climate.
climate with higher mean temperature as computed with the FC (Figure 3b).

4.4 | Climate indices

A number of climate indices for the rainfall and temperature were analysed, as listed in Table 2. We focus on two 30-year periods representing the present climate (1976–2005) and the end of the century climate (2070–2099). First, we computed the mean value for each of the analysed climate indices from each realization (consisting of a 30-year period). Then, we extracted statistics out of ensemble, that is, the mean of the mean was computed for the 30 realizations of 30 years both for the present and future periods (see also Peleg et al., 2019). Additionally, the 5-95th percentile range, representing the stochastic climate uncertainty was computed from the ensemble (i.e., the 5-95th percentile of the 30 mean values).

5 | RESULTS

5.1 | Present climate

The simulated climate variables were first evaluated for their spatial distribution at the annual scale. The

### Table 2  List of climate indices

| Abbreviation | Definition | Unit | Figure |
|--------------|------------|------|--------|
| TXx          | Hottest day of the year: Annual maximum of daily $T_{\max}$ | $^\circ$C | 7 |
| TNn          | Coldest night of the year: Annual minimum of daily $T_{\min}$ | $^\circ$C | 7 |
| TXx          | Warmest night of the year: Annual maximum of daily $T_{\min}$ | $^\circ$C | 7 |
| TXn          | Coldest day of the year: Annual minimum of daily $T_{\max}$ | $^\circ$C | 7 |
| TXx7d        | The hottest week of the year: Maximum 7-day mean of daily $T_{\max}$ | $^\circ$C | 7 |
| TX > 25°C    | Summer days: Annual number of days with $T_{\max}$ > 25°C | Days | 7 |
| TX > 30°C    | Hot days: Annual number of days with $T_{\max}$ > 30°C | Days | 7 |
| TN > 20°C    | Tropical nights: Annual number of days with $T_{\min}$ > 20°C | Days | 7 |
| TXp95%       | 95th percentile of daily $T_{\max}$ | $^\circ$C | S2 |
| TXp99%       | 99th percentile of daily $T_{\max}$ | $^\circ$C | S2 |
| TNp5%        | 5th percentile of daily $T_{\min}$ | $^\circ$C | S2 |
| TX95P        | Hot days: Annual number of days with $T_{\max}$ > TXp95% of the present climate | Days | 8 |
| TX99P        | Very hot days: Annual number of days with $T_{\max}$ > TXp99% of the present climate | Days | 8 |
| MEA          | Mean daily rainfall | Mm·d$^{-1}$ | S3 |
| INT          | Wet-day intensity: Mean rainfall amount on a wet day | Mm·d$^{-1}$ | 9 |
| Rx1h         | Maximum of 1-hr rainfall | Mm·h$^{-1}$ | 10 |
| Rx1d         | Maximum of 1-day rainfall | Mm·d$^{-1}$ | 10 |
| Rx5d         | Maximum of 5-day rainfall | Mm·5d$^{-1}$ | 10 |
| Rp90%        | 90th percentile of daily rainfall | Mm·d$^{-1}$ | 11 |
| Rp95%        | 95th percentile of daily rainfall | Mm·d$^{-1}$ | 11 |
| Rp99%        | 99th percentile of daily rainfall | Mm·d$^{-1}$ | 11 |
observed long-term mean areal annual rainfall over the ZRB, as estimated from the CMORPH data, is 862 mm, which compares well with the simulated long-term mean of 812 mm. The simulated rainfall was found to underestimate the CMORPH rainfall over 61% of the ZRB and overestimate the rainfall over 4.5% of the area when only differences larger than 20 mm y \(^{-1}\) between the annual rainfall maps are considered (Figure 2e). Yet, only few of the grid cells have under- and over-estimations that exceed the 10% threshold when compared to CMORPH (mean underestimation is 5% and mean overestimation is 4%), meaning that the simulated data is well within the natural climate variability, which is around 6% (as estimated from stochastic simulations) for the ZRB. Also in term of rainfall occurrence, it was found that the simulated rainfall tend to underestimate the length of the wet spells (Figure S4), however, similarly to the rainfall amount, this underestimation is affecting only a fraction of the basin and with a bias that is mostly lower than 10%.

The inter-annual variation in rainfall is reproduced correctly. The coefficients of inter-annual variation are

---

**FIGURE 5** Annual maps for temperature (a–c), shortwave radiation (d–f) and relative humidity (g–i), of observations (different products a, d, g), modelled with AWE-GEN-2d (b, e, h) and respective differences (c, f, and i)
0.1 and 0.08 for the observed (20 years) and simulated (30 years) data, respectively. The observed annual lag-1 autocorrelation of the process is very low (−0.03) and this property is also preserved in the simulated time series (0.08).

The model outputs were validated at the annual scales also for temperature, shortwave radiation and relative humidity (Figure 5). Most of the simulated grid cells (79%) have less than ±1°C difference in comparison to the observed data (Figure 5c). 14% of the area is overestimated by more than 1°C in the model (average of 1.54°C per grid cell) and 7% is underestimated (−1.41°C on average). Similarly, most of the simulated shortwave radiation grid cells (64%) have differences below ±5 W·m⁻² in comparison to the observed data (Figure 5f) and only a few of them (3%) exceed a ±15 W·m⁻² difference. The simulated relative humidity underestimates the observed data over the entire domain. The mean underestimation is of 5.7% and maximum underestimation of 11.6% at the annual scale.

In addition to the spatial evaluation carried out on reanalysis data, the simulated climate variables were evaluated for different locations at monthly and sub-daily scales, following at-station data availability. For Lusaka

**FIGURE 6** Observed (red) and simulated (blue) climate variables for Lusaka airport station, computed from 30 realizations of 30-year each. Solid lines represent the median and blue areas represent the 5-95th percentile range of the natural (stochastic) climate variability. Values are reported for monthly rainfall (a), rainfall extremes at daily and hourly scales (b), monthly air temperature (c), air temperature daily cycle (d), monthly shortwave radiation (e), shortwave radiation daily cycle (f), monthly relative humidity (g), and monthly wind speed (h)
station (Figure 6), the model represents well the monthly statistics of the rainfall, with a maximum underestimation of 38 mm·m⁻¹, which corresponds to a 18% bias, whereas other months have zero bias; near-surface wind speed shows up to 0.2 m·s⁻¹ difference, relative humidity up to 9% underestimation; and temperature has a maximum of 1.7°C difference with most months having zero bias. Yet, the model underestimates the global radiation in almost all months. This is likely due to an overestimation of the cloud cover (although no observed data are available for additional evaluation and confirmation of this possible cause) and the inability to represent the local topography in a given location, as a result of the relatively coarse resolution of the grid scale (8-km). The extreme rainfall intensities at the daily scale are reproduced well (less than 5% bias), but no observed data is available at the hourly scale for comparison. The diurnal cycle can be evaluated only for temperature (Figure 6d), and the model reproduces well the daily dynamics, with a small overestimation of the maximum (−1°C) and underestimation of the minimum (−2°C) daily temperatures. No observed data is available for direct comparison with sub-daily dynamics of incoming shortwave radiation, which are, however, following sun position expectations (Figure 6f). The bias of the model changes somehow between locations, but unfortunately, the lack of sub-daily observations does not allow a complete evaluation of the model performance across the entire basin.

The at-station validation carried out at other locations within the ZRB, where ground observations are available, produced similar results. More specifically, in addition to the Lusaka station, the weather generator performance was evaluated for the stations of Bulawayo, Harare, Lilongwe, and Kitwe (Figure S5–S8). The climate variables (rainfall, temperature, relative humidity and wind speed) are well simulated for Lilongwe station, which is located east to Lusaka. In Kitwe, a station located north of Lusaka in a wetter region, the seasonal dynamic and magnitude of temperature and wind speed are properly reproduced by the weather generator, yet the model underestimates the rainfall during the wet months (underestimation of 50 mm·m⁻¹) of January and December. Harare and Bulawayo are located south to Lusaka. While the climate in Harare is a bit colder but generally similar to Lusaka, the climate in Bulawayo is both colder and considerably drier. For Harare, the model reproduces the climate variables (rainfall, temperature, relative humidity, and wind speed) with very high skill, with only a small overestimation of the temperature during May to July (up to 2°C). For Bulawayo, the weather generator overestimates the rainfall for the wet period of December to February (on average by 66 mm·m⁻¹), overestimate the temperatures between April to December (by an average of 1.3°C), but successfully reproduces the relative humidity and wind speed.

5.2 | Future temperature indices and extremes

Temperature is generally projected to increase toward the end of the century, yet the magnitude of change differs spatially over the ZRB (see Figure 3b). As a result, climate indices and their changes have different spatial signature over the basin (Figure 7). A west–east trend is found for the changes in the hottest day of the year, where an increase of 2.7°C is projected on the west of the basin, and a lower increase of 1.8°C is projected at the eastern border of the catchment along the coast (TXx, in Figure 7). A similar trend was found for the changes in the coldest night of the year (TNn), where temperatures are projected to increase from 2.4°C in the west to 1.8°C in the east. Because of the increase in temperature, the number of summer days (TX > 25°C) significantly increases over the basin—by 144 days in the west to 48 days in the east. The number of tropical nights (TN > 20°C) also increases, but with a trend that follows the current distribution in tropical nights, that is along the coast in the east a moderate increase (1–4 days) is found because in the current climate almost the entire year is already characterized by tropical nights; in the center of the basin and along the major lakes the increase is in the range of 50–80 days; and in the west of the basin the increase in the number of tropical nights is as high as 150 days in comparison to present climate. The opposite spatial trend is observed when examining increase in the number of hot days (TX > 30°C), where the largest increase is detected along the coast at the eastern part of the basin (86 days), while in the west of the basin the increase in number of hot days is in the order of 19 days. No overlaps are found between the 5-95th percentile of the indices representing the future climate and the indices representing the mean of the present climate (Figure 7), implying that the changes to the indices exceed the natural climate variability for this region, that is, the changes are beyond what has ever been experienced in the past.

The increase in the number of hot days and very hot days was also explored by examining the changes to a given percentile threshold (95th and 99th, see Figure S9), instead of a fixed temperature threshold (25 and 30°C). This helps avoiding false perceptions when such thresholds are already exceeded very often in the present climate (Figure 7). Similar spatial trends are found for both
FIGURE 7  Indices of temperature for the present and future climate (RCP4.5) computed from 30 realizations of 30-year each that were simulated using the weather generator. The mean of the indices is presented for both present and future climates where q05 and q95 represent the 5-95th percentile range of the indices for the future (see Section 4.3 for details). The abbreviations of the indices are given in Table 2.
indices (Figure 8a,b), where the lowest increase in the number of hot and very hot days are found along the coast to the east (12 and 4.5 days, respectively) and the largest increase is found at the south-western region of the basin (49 and 21 days, respectively). The uncertainty in the increase of the number of hot and very-hot days (the 5-95th percentile difference, Figure 8) reaches 10% of the change for the hot days (up to 5 days of uncertainty)

**FIGURE 8** Number of days with maximum temperature larger than a given percentile, the 95th (a) and 99th (b) for the future climate (RCP4.5, end of the century). The percentile of the maximum daily temperature are computed using the present climate from 30 realizations of 30-year each that were simulated with the weather generator. The q05 and q95 represent the 5-95th percentile range of the indices (see Section 4.3 for details)

**FIGURE 9** Rainfall intensity during wet-days computed from 30 realizations of 30-year each that were simulated using the weather generator for the present and future (RCP4.5) climates. The q05 and q95 represent the 5-95th percentiles for the future climate (see Section 4.3 for details)
and 20% of the change for the very hot days (up to 3 days of uncertainty).

5.3 | Future rainfall indices and extremes

Rainfall mean is projected to change toward the end of the century, but the magnitude of the change varies in space and across seasons over the ZRB (see Figure 3a). As a result, changes in rainfall intensities are also heterogeneous over the basin (Figure 9 and Figure S10). In general, rainfall in wet days is projected to intensify over the basin, with a lower increase in the western regions of the basin (average increase of 0.3 mm·d\(^{-1}\)) in comparison to the eastern regions (average increase of 1 mm·d\(^{-1}\)). The highest increase in rainfall intensity is projected north and west of Lake Malawi (an increase of 2 mm·d\(^{-1}\)). The increase in rainfall intensity is not statistically significant when averaged over the entire basin; future rainfall intensity in some parts of the basin, especially in its western areas, fall well within the natural variability of the rainfall, that is, the mean rainfall intensity of the present is within the 5-95th percentile range of rainfall intensity of the future (Figure 9).

Changes to rainfall extreme at the hourly scale (Figure 10) are within the range of natural variability. Hourly maximum rainfall intensity tends to weaken over 70% of the basin, without a clear spatial trend. However, the changes in maximum rainfall intensity at the hourly scale are relatively small, in the order of \(\pm 2\) mm·h\(^{-1}\). At the daily scale, a clear increase in maximum rainfall is detected over the basin, affecting 98% of the area. The largest increase is found in the eastern areas of the basin (>10 mm·d\(^{-1}\)), the smaller increase is detected for the north-western areas (2–4 mm·d\(^{-1}\)). In the south-western areas, a mix of positive (around 1 mm·d\(^{-1}\)) and negative changes can be detected. All areas of the basin show an increase for the maximum 5-day cumulated rainfall, with a similar spatial trend as

![Figure 10](image-url)
found for the daily maximum rainfall, that is, higher increase in the eastern parts of the basin. Note that while for the hourly and daily scales the changes in maximum rainfall intensity are relatively low (maximum of 5 and 10% in comparison to present climate, respectively), the changes in the 5-day maximum cumulated rainfall are as high as 30% in some part of the basin. In other words, changes for the hourly and daily scales are within the range of the natural (stochastic) variability for the present climate (12 and 11.5% on average across the basin, respectively), and only the changes in maximum rainfall intensity at the 5-day scale exceed the natural variability, which is 10% on average.

Examining the increase of daily rainfall for different percentiles (Figure 11) confirms that although rainfall is, in general, intensifying in the future climate in comparison to the present, the intensification is less pronounced for the higher percentiles. For example, intensification is found over the entire basin both for the 90th and 99th percentile of daily rainfall, with a stronger intensification in the eastern part of the basin in comparison to the western areas, as reported above. However, while the 90th percentile of the rainfall is 13% larger, on average, over the basin, the 99th percentile of rainfall is larger, on average, by only 8%.

### 6 SUMMARY OF CHANGES AND HYDROLOGICAL IMPLICATIONS

Climate indices computed for temperature (Section 5.2) point to an increase in temperature projected toward the end of the century over the ZRB that is well beyond the natural climate variability, with temperature changes of different magnitudes between the eastern and western areas of the basin. Most of the climate indices computed for rainfall (Section 5.3) show a moderate increase in rainfall intensity that falls within the natural climate variability, with magnitudes of the change that are heterogeneous in space.

---

**Figure 11** The 90th (Rp90%), 95th (Rp95%) and 99th (Rp99%) percentiles of daily rainfall (computed using all days) for the present and future (RCP4.5) climates computed as average of 30 realizations of 30-year each that were simulated using the weather generator
Generalizing these findings, a reduction in the water availability can be hypothesized over the ZRB for the end of the century and for RCP4.5 emission scenario because of warmer temperatures and likely higher evapotranspiration rates in the areas of the basins, which are not water limited. However, because of: (a) the large spatial variability in the changes in temperature and rainfall mean; (b) the fact that the lower rainfall intensities are increasing at a higher rate than the heavy rainfalls; and (c) the finding that hourly extremes are only moderately changing, while the 5-day rainfall extremes are considerably increasing—the impact on median and extreme flows is difficult to predict. This will require the application of a hydrological model to explore how these climatic changes will affect the components of the hydrological cycle in detail.

7 | BENEFITS OF AWE-GEN-2D COMPUTED CLIMATE INDICES

Climate indices for present and future climate conditions can be computed directly from regional climate models. However, there are several benefits of using gridded stochastic weather generator models, such as the AWE-GEN-2d model, to perform this task. First, climate variables obtained from RCMs require adjustments to the local climate, in the form of bias-correction (e.g., quantile mapping), before they can be used to compute the climate indices. Dosio (2016), for example, showed that climatic absolute-threshold indices (e.g., TX > 25°C) are strongly influenced by the bias-correction, as they depend mainly on the present mean climate value. The parameterization of AWE-GEN-2d, and the re-parameterization for future climate conditions based on the FC method is accounting implicitly for the bias in climate models and makes the step of bias correction unnecessary. Comparing the change in rainfall intensity on wet days, INT (future minus present climates), over the ZRB computed directly using a multi-model mean of 22 RCMs (Figure 12a) and computed using the AWE-GEN-2d outputs (Figure 12b) reveals considerable details that do not emerge when raw RCMs data are used. Moreover, the AWE-GEN-2d model structure explicitly enables preserving the cross-correlation between climate variables (Peleg et al., 2017), which is an important feature in climate downscaling (Maraun, 2016). We note that the re-parameterization of AWE-GEN-2d that is shown here is limited to changes in the first moments of rainfall and temperature and to the spatial structure of rainfall. However, AWE-GEN-2d can be theoretically re-parameterized also for the higher order statistics (e.g., standard deviation or skewness) and for other climate statistics, (e.g., mean cloud cover, dry–wet period transitions, mean storm duration) as long as reliable reference data are available. For a data scarce region, reparametrizing first order statistics is deemed most appropriate and it is also coherent with the available outputs saved by current RCMs for Africa domain. In addition, we note that the parameterization of the AWE-GEN-2d model could not be better than the reference datasets that are used (e.g., CMORPH and MERRA-2), even if previous studies showed that this uncertainty is likely lower than the uncertainties that one would encounter using RCMs directly (e.g., Gampe et al., 2019).

A noteworthy feature of using AWE-GEN-2d is that the downscaling in time performed makes it possible to compute hourly climate indices (e.g., all the temperature related indices presented in this study and RX1h for rainfall), which cannot be computed directly from RCMs in the Africa-CORDEX, because the climate variables are only available at daily resolution. The downscaled values to hourly resolution are shown to be reliable as AWE-GEN-2d model reproduces fairly well the diurnal cycle of

![Figure 12](image-url)
many climate variables (Peleg et al., 2017) and, therefore, can be used to compute climate indices that are based on sub-daily dynamics, such as the diurnal temperature range (e.g., Lindvall and Svensson, 2015). The downscaling in space further allows analysis of the climate indices at fine spatial resolution (8 km), which otherwise would have been remained unknown as the regional climate models for Africa region simulate the climate variables at a coarser resolution of 50 km. For example, changes to the temperature in the complex topographical region surrounding the lakes to the north-east of the basin (Figure 7), which is poorly represented using a 50 km resolution elevation map, would remain unresolved. The high spatial resolution of the weather generator model permits the production of smoother climate index maps that better capture the local topography and land cover effects (Figure 12a) in comparison to the information obtained from RCMs (Figure 12b), thus allowing a more differentiated assessment of the impacts. With a careful parameterization of the AWE-GEN-2d model, based on better observations, climate variables can be simulated at even finer resolutions, for example, sub-kilometre and sub-hourly, as shown to be possible by previous studies (Peleg et al., 2017, 2019), thus supplying inputs for climate impact predictions that would otherwise be impossible to obtain and that are conversely useful to predict the effects on specific processes for which small scale variability can be of importance (e.g., crop yield).

Last, using a stochastic weather generator model, the uncertainty about climate indices due to the intrinsic climate variability can be explicitly quantified, thus helping decision makers to identify adaptation solutions that are more resilient than those computed on the basis of single deterministic trajectories. We acknowledge that, in this study, for the reasons explained further above, only the stochastic component of total climate uncertainty was calculated (e.g., in Figure 7) and that, in addition to the stochastic uncertainty, the uncertainties emerged from different climate models and emission scenarios contribute to the total uncertainty of climate change projections (Deser et al., 2012; Fatichi et al., 2016a). For a better understanding of the relative contribution of the different climate uncertainty components to the total uncertainty over the ZRB, we provided a general estimate of the magnitude of the climate model uncertainty by inferring it from the FC of mean temperature and rainfall. This was done by computing the difference between the 5th and 95th percentile emerging from the climate models for the locations indicated in Figure 1, for the period 2070–2099. The values so obtained were compared to the stochastic uncertainty derived from the weather generator model outputs for the same period, quantile range, and locations (Table 3). The stochastic climate uncertainty accounts for one third of the uncertainty in rainfall mean, which is 10% on average compared to a 27% climate model uncertainty. For mean temperature, stochastic uncertainty is much less important (0.1°C on average), when compared with that due to climate models (1.8°C).

While the purpose of this study was not to quantify the total uncertainty associated with predictions of future climate in the ZRB, but to demonstrate the considerable increase of knowledge about them that the use of an advanced stochastic weather generator allows, it is possible, as shown by Peleg et al. (2019) to properly compute the total climate uncertainty over the ZRB, emerging from different climate models and emission scenarios. To achieve this, one can extend the methodology presented in this study, by accounting for multiple climate models to calculate a single median distribution of factors of change. This can be done by stochastically simulating multiple climate realizations per each individual climate model and emission scenario as in Peleg et al. (2019) or reproducing multiple climate trajectories combining the 22 RCMs by means of a Bayesian approach (see Fatichi et al., 2013, 2016a).

### TABLE 3
Climate model uncertainty computed as the 5–95th percentile range of the mean temperature and mean rainfall FC obtained from the 22 RCMs used in this study, and the stochastic uncertainty computed for the same percentile range from AWE-GEN-2d outputs

| Location   | Climate model uncertainty | Stochastic uncertainty |
|------------|--------------------------|------------------------|
|            | Temperature (°C) | Rainfall (%) | Temperature (°C) | Rainfall (%) |
| Lusaka     | 1.7                     | 20                 | 0.1             | 9            |
| Kitwe      | 1.7                     | 28                 | 0.1             | 6            |
| Harare     | 1.8                     | 24                 | 0.1             | 12           |
| Lolongwe   | 1.7                     | 25                 | 0.1             | 9            |
| Bulawayo   | 2.2                     | 36                 | 0.1             | 13           |

**Note:** Uncertainties are computed for the locations indicated in Figure 1 and for the period 2070–2099 using the RCP4.5 emission scenario.

### 8 | CONCLUSIONS

The ZRB was used as an illustrative case study to demonstrate what state-of-the-art WGs, such as AWE-GEN-
2d, could produce in terms of climatic forcing for present and future conditions over large and data sparse regions, also overcoming the limitations in spatial and temporal scales of climate models. As demonstrated in this study, stochastic WGs can be successfully calibrated using data from climate reanalysis products to compensate for the absence of dense ground station data and produce results that exceed the level of detail achievable from climate models only or from interpolation of very sparse ground data. The gridded stochastic WGs can account for spatially heterogeneous changes in climate variables along with the natural climate variability (stochastic uncertainty), which is a critical approach of impact studies, as it allows a better quantification of the significance of a given change. This is probably the major advantage of using a gridded stochastic WG to downscale climate variables because stochastic climate uncertainty can be accounted for in a seamless way in the subsequent impact studies (e.g., hydrological predictions) at spatial and temporal resolution that, so far, considerably exceed the capabilities of climate models.

In summary, we scrutinized the use of an advanced gridded stochastic WG in a real world situation, which is far from common literature applications, where plentiful observational data are available to calibrate and test the model. Weather generators have been designed with mesoscale applications in mind, and as such they have not been considered for application at larger spatial scales. This study demonstrates that this limitation does not constrain their use at larger scales, and that, even though large-scale applications violate some assumptions used in the development of WGs, the results are generally satisfactory and the value of this approach for impact studies opens new possibilities of impact assessment in regions characterized by sparse data and coarse model climate simulation outputs. AWE-GEN-2d proved to be a very valuable tool to downscale to fine spatial and temporal scales ($10^0$–$10^1$ km and hourly) the key climate variables that are needed for hydrological, ecological, and agricultural impact studies and to estimate the changes of climate indices over large regions.

ACKNOWLEDGEMENTS
The work presented here is funded by the Decision Analytic Framework to explore the water-energy-food Nexus in complex transboundary water resource systems of fast developing countries (DAFNE project, https://dafne-project.eu, Horizon 2020 programme WATER 2015 of the European Union, GA no. 690268). We thank the Africa-CORDEX initiative for supplying the regional climate models.

Data availability
The climate ensembles generated using the AWE-GEN-2d models for the Zambezi River Basin (present and future periods) are available upon request (www.hyd.ifu.ethz.ch/research/water-resources/dafne). Any element of AWE-GEN-2d is free to use, modify, copy or distribute provided it is exclusively for academic use and source code developers are properly acknowledged and cited (https://hyd.ifu.ethz.ch/research-data-models/awe-gen-2d.html).

ORCID
Nadav Peleg https://orcid.org/0000-0001-6863-2934

REFERENCES
Anandhi, A., Frei, A., Pierson, D.C., Schneiderman, E.M., Zion, M. S., Lounsbury, D. and Matonse, A.H. (2011) Examination of change factor methodologies for climate change impact assessment. *Water Resources Research*, 47, W03501. https://doi.org/10.1029/2010wr009104.

Anghileri, D., Botter, M., Castelletti, A., Weigt, H. and Burlando, P. (2018) A comparative assessment of the impact of climate change and energy policies on alpine hydropower. *Water Resources Research*, 54(11), 9144–9161. https://doi.org/10.1029/2017wr022289.

Awange, J.L., Ferreira, V.G., Forootan, E., Khandu, Andam-Akorful, S.A., Agutu, N.O. and He, X.F. (2016) Uncertainties in remotely sensed precipitation data over Africa. *International Journal of Climatology*, 36(1), 303–323. https://doi.org/10.1002/joc.4346.

Bloeschl, G. and Sivapalan, M. (1995) Scale issues in hydrological modeling - a review. *Hydrological Processes*, 9(3–4), 251–290. https://doi.org/10.1002/hyp.3360090305.

Bordoy, R. and Burlando, P. (2014) Stochastic downscaling of climate model precipitation outputs in orographically complex regions: 2. Downscaling methodology. *Water Resources Research*, 50(1), 562–579. https://doi.org/10.1002/wrcr.20443.

Burlando, P. and Rosso, R. (1991) Extreme storm rainfall and climatic change. *Atmospheric Research*, 27(1), 169–189. https://doi.org/10.1016/0169-8095(91)90017-Q.

Burlando, P. and Rosso, R. (2002) Effects of transient climate change on basin hydrology. 1. Precipitation scenarios for the Arno River, Central Italy. *Hydrological Processes*, 16(6), 1151–1175. https://doi.org/10.1002/hyp.1055.

Cambrin, P., Gitau, W., Oetitti, P., Ogullo, L. and Bois, B. (2014) Spatial interpolation of daily rainfall stochastic generation parameters over East Africa. *Climate Research*, 59(1), 39–60. https://doi.org/10.3354/cr01198.

Camici, S., Brocca, L. and Moramarco, T. (2017) Accuracy versus variability of climate projections for flood assessment in Central Italy. *Climatic Change*, 141(2), 273–286. https://doi.org/10.1007/s10584-016-1876-x.

Chen, J., Brissette, F.P. and Zhang, X.C.J. (2016) Hydrological modeling using a multisite stochastic weather generator. *Journal of Hydrologic Engineering*, 21(2), 04015060. https://doi.org/10.1061/(asce)he.1943-5584.0001288.
Chen, Y.B., Li, J., Wang, H.Y., Qin, J.M. and Dong, L.M. (2017) Large-watershed flood forecasting with high-resolution distributed hydrological model. *Hydrology and Earth System Sciences*, 21(2), 735–749. https://doi.org/10.5194/hess-21-735-2017.

Christensen, N.S. and Lettenmaier, D.P. (2007) A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River basin. *Hydrology and Earth System Sciences*, 11(4), 1417–1434. https://doi.org/10.5194/hess-11-1417-2007.

Christensen, N.S., Wood, A.W., Voisin, N., Lettenmaier, D.P. and Palmer, R.N. (2004) The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, 62(1–3), 337–363. https://doi.org/10.1002/BL CL.0000013684.13621.1f.

Cohen Liechti, T., Matos, J.P., Boillat, J.L. and Schlies, A.J. (2012) Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi River basin. *Hydrology and Earth System Sciences*, 16(2), 489–500. https://doi.org/10.5194/hess-16-489-2012.

Cristiano, E., ten Veldhuis, M.C. and van De Giesen, N. (2017) Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas - a review. *Hydrology and Earth System Sciences*, 21(7), 3859–3878. https://doi.org/10.5194/hess-21-3859-2017.

Deser, C., Phillips, A., Bourdette, V. and Teng, H.Y. (2012) Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics*, 38(3–4), 527–546. https://doi.org/10.1007/s00382-010-0977-x.

Dörner, S. (2013) *Assessment of Flood Risk for Mazoya Settlement Arising from Improvements Targeted for Bombay and Lumumba Drain*. KG, Germany: H.P. GAUFF INGENIEUR GmbH & Co.

Dosio, A. (2016) Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models. *Journal of Geophysical Research-Atmospheres*, 121(10), 5488–5511. https://doi.org/10.1002/2015JD024411.

Draper, C.S., Reichle, R.H. and Koster, R.D. (2018) Assessment of MERRA-2 land surface energy flux estimates. *Journal of Climate*, 31(2), 671–691. https://doi.org/10.1175/JCLI-d-17-0121.1.

Endris, H.S., Omondi, P., Jain, S., Lennard, C., Hewitson, B., Chang’a, L., Awange, J.J., Dosio, A., Ketiem, P., Nikulin, G., Panitz, H.J., Buchner, M., Stordal, F. and Tazalika, L. (2013) Assessment of the performance of CORDEX regional climate models in simulating east African rainfall. *Journal of Climate*, 26(21), 8453–8475. https://doi.org/10.1175/jcli-d-12-00708.1.

Fatichi, S., Ivanov, V.Y. and Caporali, E. (2011) Simulation of future climate scenarios with a weather generator. *Advances in Water Resources*, 34(4), 448–467. https://doi.org/10.1016/j.advwatres.2010.12.013.

Fatichi, S., Ivanov, V.Y. and Caporali, E. (2013) Assessment of a stochastic downscaling methodology in generating an ensemble of hourly future climate time series. *Climate Dynamics*, 40(7), 1841–1861. https://doi.org/10.1007/s00382-012-1627-2.

Fatichi, S., Rimkus, S., Burlando, P., Bordoy, R. and Molnar, P. (2015) High-resolution distributed analysis of climate and anthropogenic changes on the hydrology of an alpine catchment. *Journal of Hydrology*, 525, 362–382. https://doi.org/10.1016/j.jhydrol.2015.03.036.

Fatichi, S., Ivanov, V.Y., Paschalis, A., Peleg, N., Molnar, P., Rimkus, S., Kim, J., Burlando, P. and Caporali, E. (2016a) Uncertainty partition challenges the predictability of vital details of climate change. *Earth Future*, 4(5), 240–251. https://doi.org/10.1002/2015ef000336.

Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B.A., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M. and Tarboton, D. (2016b) An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, 537, 45–60. https://doi.org/10.1016/j.jhydrol.2016.03.026.

Fick, S.E. and Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. https://doi.org/10.1002/joc.5086.

Fowler, H.J., Kilsby, C.G., O’Connell, P.E. and Burton, A. (2005) A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change. *Journal of Hydrology*, 308(1–4), 50–66. https://doi.org/10.1016/j.jhydrol.2004.10.021.

Fowler, H.J., Blenkinsop, S. and Tebaldi, C. (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12), 1547–1578. https://doi.org/10.1002/joc.1556.

Gampe, D., Schmid, J. and Ludwig, R. (2019) Impact of reference dataset selection on ECMWV evaluation, bias correction, and resulting climate change signals of precipitation. *Journal of Hydro meteorology*, 20(9), 1813–1828. https://doi.org/10.1175/JHM-d-18-0108.1.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Silva, A.M.D., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M. and Zhao, B. (2017) The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419–5454. https://doi.org/10.1175/JCLI-d-16-0758.1.

Gibba, P., Sylla, M.B., Okogbue, E.C., Gaye, A.T., Nikkiema, M. and Kebe, I. (2019) State-of-the-art climate modeling of extreme precipitation over Africa: analysis of CORDEX added-value over CMIP5. *Theoretical and Applied Climatology*, 137(1–2), 1041–1057. https://doi.org/10.1007/s00704-018-2650-y.

Giorgi, F., Jones, C. and Asrar, G.R. (2009) Addressing climate information needs at the regional level: the CORDEX framework. *Bulletin - World Meteorological Organization*, 58(3), 175–183.

Giontoli, I., Villarini, G., Prudhomme, C. and Hannah, D.M. (2018) Uncertainties in projected runoff over the conterminous United States. *Climatic Change*, 150(3–4), 149–162. https://doi.org/10.1007/s10584-018-2280-5.

Glenis, V., Pinamonti, V., Hall, J.W. and Kilsby, C.G. (2015) A transient stochastic weather generator incorporating climate model uncertainty. *Advances in Water Resources*, 85, 14–26. https://doi.org/10.1016/j.advwatres.2015.08.002.
Kundzewicz, Z.W., Haile, A.T., Makurira, H. and Reggiani, P. (2019) Performance of bias correction schemes for CMORPH rainfall estimates in the Zambezi River basin. *Hydrology and Earth System Sciences*, 23, 2915–2938. https://doi.org/10.5194/hess-23-2915-2019.

Haenler, A., Saeed, F. and Jacob, D. (2013) Assessing the robustness of projected precipitation changes over Central Africa on the basis of a multitude of global and regional climate projections. *Climatic Change*, 121(2), 349–363. https://doi.org/10.1007/s10584-013-0863-8.

Hawkins, E. and Sutton, R. (2011) The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, 37(1–2), 407–418. https://doi.org/10.1007/s00382-010-0810-6.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S. (2013) Global flood risk under climate change. *Nature Climate Change*, 3(9), 816–821. https://doi.org/10.1038/nclimate1911.

Huber, M. and Knutti, R. (2012) Anthropogenic and natural warming inferred from changes in Earth’s energy balance. *Nature Geoscience*, 5(1), 31–36. https://doi.org/10.1038/ngeo1327.

Immerzeel, W.W., van Beek, L.P.H. and Bierkens, M.F.P. (2010) Climate change will affect the Asian water towers. *Science*, 328 (5984), 1382–1385. https://doi.org/10.1126/science.1183188.

Joyce, R.J., Janowiak, J.E., Arkin, P.A. and Xie, P. (2004) CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3), 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:Camtpg>2.0.CO;2.

Kapumpu, C. (2015) *Predicting the Global Solar Radiation on a Horizontal Surface: A Case Study for Zambia*. Lusaka, Zambia: The University of Zambia, p. 74.

Keller, D.E., Fischer, A.M., Frei, C., Liniger, M.A., Appenzeller, C. and Knutti, R. (2015) Implementation and validation of a Wilks-type multi-site daily precipitation generator over a typical alpine river catchment. *Hydrology and Earth System Sciences*, 19(5), 2163–2177. https://doi.org/10.5194/hess-19-2163-2015.

Kilsby, C.G., Jones, P.D., Burton, A., Ford, A.C., Fowler, H.J., Harpham, C., James, P., Smith, A. and Wilby, R.L. (2007) A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, 22(12), 1705–1719. https://doi.org/10.1016/j.envsoft.2007.02.005.

Kim, J. and Ivanov, V.Y. (2015) A holistic, multi-scale dynamic downscaling framework for climate impact assessments and challenges of addressing finer-scale watershed dynamics. *Journal of Hydrology*, 522, 645–660. https://doi.org/10.1016/j.jhydrol.2015.01.025.

Kling, H., Stanelz, P. and Preisheuber, M. (2014) Impact modelling of water resources development and climate scenarios on Zambezi River discharge. *Journal of Hydrology: Regional Studies*, 1, 17–43. https://doi.org/10.1016/j.jhydrol.2014.05.002.

Kundzewicz, Z.W., Mata, L.J., Arnell, N.W., Doll, P., Jimenez, B., Miller, K., Oki, T., Sen, Z. and Shiklomanov, I. (2008) The implications of projected climate change for freshwater resources and their management. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 53(1), 3–10. https://doi.org/10.1623/hysj.53.1.3.

Lindvall, J. and Svensson, G. (2015) The diurnal temperature range in the CMIP5 models. *Climate Dynamics*, 44(1), 405–421. https://doi.org/10.1007/s00382-014-2144-2.

Locchbihler, K., Lenderink, G. and Siebesma, A.P. (2017) The spatial extent of rainfall events and its relation to precipitation scaling. *Geophysical Research Letters*, 44(16), 8629–8636. https://doi.org/10.1002/2017gl074857.

Luhungu, P., BotaI, J. and Kahimba, F. (2016) Evaluation of the performance of CORDEX regional climate models in simulating present climate conditions of Tanzania. *Journal of Southern Hemisphere Earth Systems Science*, 66(1), 32–54. https://doi.org/10.22499/3.6601.005.

Maraun, D. (2016) Bias correcting climate change simulations - a critical review. *Current Climate Change Reports*, 2(4), 211–220. https://doi.org/10.1007/s40641-016-0050-x.

Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brieden, S., Rust, H.W., Sauter, T., Thiemesl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M. and Thiele-Eich, I. (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48, RG3003. https://doi.org/10.1029/2009rg000314.

Mariotti, L., Coppola, E., Sylla, M.B., Giorgi, F. and Piani, C. (2011) Regional climate model simulation of projected 21st century climate change over an all-Africa domain: comparison analysis of nested and driving model results. *Journal of Geophysical Research-Atmospheres*, 116, D15111. https://doi.org/10.1029/2010JD015068.

Matoe, C.M.R., Yamazaki, D., Kim, H., Champathong, A., Vaze, J. and Oki, T. (2017) Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods. *Hydrology and Earth System Sciences*, 21(10), 5143–5163. https://doi.org/10.5194/hess-21-5143-2017.

McGregor, G.R. and Nieuwolt, S. (1998) *Tropical Climatology: An Introduction to the Climates of the Low Latitudes*. Chichester: John Wiley & Sons Ltd xi + 339 pp. pp.

Meinshausen, M., Smith, S.J., Calvin, K., Daniel, J.S., Kainuma, M.L.T., Lamarque, J.-F., Matsumoto, K., Montzka, S.A., Raper, S.C.B., Riahi, K., Thomson, A., Velders, G.J.M. and van Vuuren, D.P.P. (2011) The RCP greenhouse gas concentrations to 2300. *Climatic Change*, 109(1), 213–241. https://doi.org/10.1007/s10584-011-0156-z.

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E. and Houston, T.G. (2012) An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, 29(7), 897–910. https://doi.org/10.1175/jtech-d-11-00103.1.

Miao, H., Wang, X.C., Lui, Y.M. and Wu, G.X. (2019) An evaluation of cloud vertical structure in three reanalyses against CloudSat/cloud-aerosol lidar and infrared Pathfinder satellite observations. *Atmospheric Science Letters*, 20(7), e906. https://doi.org/10.1002/asl.906.

Millar, R.J., Fuglestvedt, J.S., Friedlingstein, P., Rogelj, J., Grubb, M.J., Matthews, H.D., Skeie, R.B., Forster, P.M., Frame, D.J. and Allen, A.R. (2017) Emission budgets and pathways consistent with limiting warming to 1.5 degrees C. *Nature Geoscience*, 10(10), 741. https://doi.org/10.1038/NGEO3031.

New, M., Hulme, M. and Jones, P. (1999) Representing twentieth-century space–time climate variability. Part I: development of a
1961–90 mean monthly terrestrial climatology. Journal of Climate, 12(3), 829–856. https://doi.org/10.1175/1520-0442(1999)012<0829:Rtscct>2.0.CO;2

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Buchner, M., Cerezo-Mota, R., Christensen, O.B., Deque, M., Fernandez, J., Hansler, A., van Meijgaard, E., Samuelsson, P., Sylla, M.B. and Sushama, L. (2012) Precipitation climate in an ensemble of CORDEX-Africa regional climate simulations. Journal of Climate, 25(18), 6057–6078. https://doi.org/10.1175/jcli-d-11-00375.1

Onyutha, C., Rutkowska, A., Nyeko-Ogiramoi, P. and Willems, P. (2019) What would it take to achieve the Paris temperature targets? Geophysical Research Letters, 12(10), 104011. https://doi.org/10.1002/2016gl069563.

Semenov, M.A. and Barrow, E.M. (1997) Use of a stochastic weather generator in the development of climate change scenarios. Climatic Change, 35(4), 397–414. https://doi.org/10.1023/A:1005342632279.

Shahin, M. (2006) Hydrology and Water Resources of Africa, Vol. 41. Dordrecht: Springer.

Singer, M.B. and Michaelides, K. (2017) Deciphering the expression of climate change within the lower Colorado River basin by stochastic simulation of convective rainfall. Environmental Research Letters, 12(10), 104011. https://doi.org/10.1088/1748-9326/aa8e50.

Sorup, H.J.D., Davidsen, S., Lowe, R., Thorndahl, S.L., Borup, M. and Arnbjerg-Nielsen, K. (2018) Evaluating catchment response to artificial rainfall from four weather generators for present and future climate. Water Science and Technology, 77(11), 2578–2588. https://doi.org/10.2166/wst.2018.217.

Stettz, S., Zaïtchik, B.F., Ademe, D., Musie, S. and Simane, B. (2019) Estimating variability in downwelling surface shortwave radiation in a tropical highland environment. PLoS One, 14(2), e0211220. https://doi.org/10.1371/journal.pone.0211220.

Suris, M., Suriyova, N. and Cebecauer, T. (2014) Solar Modelling Report. World Bank: Solar Resource Mapping in Zambia.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V. and Roo, A.D. (2012) Validation of satellite-based precipitation products over sparsely gauged African River basins. Journal of Hydrometeorology, 13(6), 1760–1783. https://doi.org/10.1175/jhm-d-12-032.1.

Tian, Y.D. and Peters-Lidard, C.D. (2010) A global map of uncertainties in satellite-based precipitation measurements. Geophysical Research Letters, 37, 37. https://doi.org/10.1029/2010GL046008.

Torralba, V., Doblas-Reyes, F.J. and Gonzalez-Reviriego, N. (2017) Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison. Environmental Research Letters, 12(11), 114019. https://doi.org/10.1088/1748-9326/aa8a58.

Van Vooren, S., Van Schaeybroeck, B., Nyssen, J., Van Ginderachter, M. and Termonia, P. (2019) Evaluation of CORDEX rainfall in Northwest Ethiopia: sensitivity to the model representation of the orography. International Journal of Climatology, 39(5), 2569–2586. https://doi.org/10.1002/joc.5971.

Warnatzsch, E.A. and Reay, D.S. (2019) Temperature and precipitation change in Malawi: evaluation of CORDEX-Africa climate simulations for climate change impact assessments and adaptation planning. Science of the Total Environment, 654, 378–392. https://doi.org/10.1016/j.scitotenv.2018.11.098.

Wasko, C., Sharma, A. and Westra, S. (2016) Reduced spatial extent of extreme storms at higher temperatures. Geophysical Research Letters, 43(8), 4026–4032. https://doi.org/10.1002/2016gl068509.

Wilby, R.L. and Yu, D. (2013) Rainfall and temperature estimation for a data sparse region. Hydrology and Earth System Sciences, 17(10), 3937–3955. https://doi.org/10.5194/hess-17-3937-2013.

Wilby, R.L., Dawson, C.W., Murphy, C., O’Connor, P. and Hawkins, E. (2014) The statistical DownScaling model - decision centric (SDSM-DC); conceptual basis and applications. Climate Research, 61(3), 259–276. https://doi.org/10.3354/cr01254.

Willems, P., Arnbjerg-Nielsen, K., Olsson, J. and Nguyen, V.T.V. (2012) Climate change impact assessment on urban rainfall extremes and urban drainage: methods and shortcomings.
Atmospheric Research, 103, 106–118. https://doi.org/10.1016/j.atmosres.2011.04.003.

World Bank, 2017. Global Solar Atlas. https://globalsolaratlas.info/. Last visited: June 2019.

Wu, P.L., Christidis, N. and Stott, P. (2013) Anthropogenic impact on Earth's hydrological cycle. Nature Climate Change, 3(9), 807–810. https://doi.org/10.1038/nclimate1932.

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F. and Lin, R. (2017) Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. Journal of Hydro-meteorology, 18(6), 1617–1641. https://doi.org/10.1175/jhm-d-16-0168.1.

Xu, C.Y. (1999a) Climate change and hydrologic models: a review of existing gaps and recent research developments. Water Resources Management, 13(5), 369–382. https://doi.org/10.1023/a:1008190900459.

Xu, C.Y. (1999b) From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. Progress in Physical Geography, 23(2), 229–249. https://doi.org/10.1177/030913339902300204.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Peleg N, Sinclair S, Fatichi S, Burlando P. Downscaling climate projections over large and data sparse regions: Methodological application in the Zambezi River Basin. Int J Climatol. 2020;40:6242–6264. https://doi.org/10.1002/joc.6578