



# Impacts of land surface model and land use data on WRF model simulations of rainfall and temperature over Lake Tana Basin, Ethiopia



Achenafi Teklay<sup>a,\*</sup>, Yihun T. Dile<sup>b</sup>, Dereje H. Asfaw<sup>c</sup>, Haimanote K. Bayabil<sup>d</sup>, Kibruyesfa Sisay<sup>e</sup>

<sup>a</sup> Ethiopian Institute of Water Resources, Department of Water Resources Engineering and Management, Addis Ababa University, Addis Ababa, Ethiopia

<sup>b</sup> Spatial Science Laboratory, Ecosystem Science and Management Department, Texas A & M University, College Station, TX, 77801, USA

<sup>c</sup> Addis Ababa Institute of Science and Technology, School of Environmental and Civil Engineering, Addis Ababa University, Addis Ababa, Ethiopia

<sup>d</sup> Agricultural and Biological Engineering, Tropical Research and Education Center, Institute of Food and Agricultural Sciences, University of Florida, Homestead, FL, 33031, USA

<sup>e</sup> Ethiopian Environment and Forest Research Institute, Addis Ababa, Ethiopia

## ARTICLE INFO

### Keywords:

Agriculture  
Environmental science  
Atmospheric science  
Agricultural engineering  
Agricultural water management  
Climatology  
Environmental impact assessment  
Environmental management  
Dynamical downscaling  
Updated land use  
USGS  
Simulations  
Experiment

## ABSTRACT

The Weather Research and Forecasting (WRF) model is one of the regional climate models for dynamically downscaling climate variables at finer spatial and temporal scales. The objective of this study was to evaluate the performance of WRF model for simulating temperature and rainfall over Lake Tana basin in Ethiopia. The WRF model was configured for six experimental setups using three land surface models (LSMs): Noah, RUC and TD; and two land use datasets: USGS and updated New Land Use (NLU). The performances of WRF configurations were assessed by comparing simulated and observed data from March to August 2015. The result showed that temperature and rainfall simulations were sensitive to LSM and land use data choice. The combination of NLU with RUC and TD produced very small cold bias (0.27 °C) and warm bias (0.20 °C) for 2m maximum temperature (Tmax) and 2m minimum temperature (Tmin), respectively. WRF model with RUC and NLU captured well the observed spatial and temporal variability of Tmax, while TD and NLU for Tmin. Moreover, rainfall simulation was better with NLU; especially NLU and Noah configuration produced the smallest mean bias (2.39 mm/day) and root mean square error (6.6 mm/day). All the WRF experiments overestimated light and heavy rainfall events. Overall, findings showed that the application of updated land use data substantially improved the WRF model performance in simulating temperature and rainfall. The study would provide valuable support for identifying suitable LSM and land use data that can accurately predict the climate variables in the Blue Nile basin.

## 1. Introduction

Rainfall and temperature affect various components of the hydrological cycle including, river flow, evapotranspiration, groundwater recharge and soil moisture. Reliable rainfall and temperature data at reasonable spatial and temporal resolutions are essential for quantifying the seasonal hydrological cycle (Dile et al., 2013; Tabari et al., 2016). However, attaining good quality climate data is challenging in developing countries like Ethiopia. In such cases, numerical models can be used to produce climate data at a fine resolution (Sisay et al., 2017). For example, such data can be used to predict future climate that can inform policy on climate change adaptation and mitigation (Gashaw et al., 2014).

In this regard, Regional Climate Models (RCMs) could be useful tools to downscale climate variables at a fine spatial and temporal resolutions

at a regional scale (Yhang et al., 2017). Moreover, RCMs realistically simulate climate information by incorporating land surface properties and detailed descriptions of physical processes (He et al., 2017). Recently, Weather Research and Forecasting (WRF) model has been applied as an RCM for research and operational purposes in many parts of the world, for example in Africa (Diaz et al., 2015), Europe (Banks and Baldasano, 2016), North America (Burakowski et al., 2016), and Asia (Cannon et al., 2017). Previous studies showed that the performance of WRF model is influenced by the selection of physical parameter schemes including, microphysics (Gao et al., 2017), radiation (Mooney et al., 2016), planetary boundary layer (Kim et al., 2015), cumulus (Mugume et al., 2017), and land surface model (Jain et al., 2017). Other model input such as initial boundary condition (Yang and Duan, 2016), land use data (Cheng et al., 2013), and domain size and resolution (Zeyaeyan et al., 2017) also affect the performance of WRF model.

\* Corresponding author.

E-mail address: [achenafi.teklay@gmail.com](mailto:achenafi.teklay@gmail.com) (A. Teklay).

Many studies in recent years have focused on the selection of suitable Land Surface Models (LSMs) for climate simulation at the mesoscale level. For example, [Zu-heng et al. \(2014\)](#) evaluated Noah, Pleim-Xiu, and Noah-MP LSMs for simulating extreme rainfall events over Southwest China. They found that WRF simulations with Noah-MP reproduced well the spatial and temporal variations of rainfall. [Lee et al. \(2016\)](#) reported that Noah LSM resulted in better simulations of observed climate variables compared to Thermal Diffusion (TD), Rapid Update Cycle (RUC) and Pleim-Xiu LSMs in a case study in South Korea. On the other hand, the Pleim-Xiu LSM simulated better temperature and wind speed over Delhi-Mumbai Industrial Corridor region in India ([Jain et al., 2017](#)), while RUC LSM was better for rainfall simulation in the Western Disturbances of India ([Thomas et al., 2014](#)). This discrepancy of LSM recommendation in the above studies indicates that the performance of LSM is highly dependent on the simulation region, studied climate variable, seasonality and considerations of other physical parameters. Therefore, site-specific studies are important to identify suitable LSM to simulate climate variables accurately.

Besides, the accuracy of the land surface datasets significantly affects the performance of WRF model ([Göndöcs et al., 2015](#)). The advancement of Remote Sensing (RS) and Geographic Information System (GIS) enable land surface datasets for WRF model at finer spatial and temporal resolutions ([Cannon et al., 2017](#); [Cao et al., 2017](#)), which help to improve WRF simulations. However, the WRF model takes most land surface properties from the default 1993 U.S. Geological Survey (USGS) land use data ([Yang and Duan, 2016](#)). This dataset has a certain degree of land use misrepresentations ([Cheng et al., 2013](#); [Puliafito et al., 2015](#)). To circumvent this problem, updated and fine resolution land use data has been employed in the WRF model ([Cao et al., 2015](#); [He et al., 2017](#)) which improve regional climate simulations. For example, the application of CORINE land use improved the WRF model performance compared to the USGS dataset ([De Meij and Vinuesa, 2014](#)). Also, [Lai et al. \(2016\)](#) showed that Global Land Cover 2009 data performed better than the default 1993 USGS dataset. The above reviews demonstrate that there have been numerous individual studies on the sensitivity of the WRF model to either LSM or land use data. However, few studies have focused on the influence of LSM and land use data combination on mesoscale climate simulations. This study contributes to filling this gap.

Although the climate of Lake Tana basin has a strong association with the topography and land use dynamics, most of the previous studies have applied statistical downscaling method ([Ayele et al., 2016](#); [Roth et al., 2018](#)); which overlooked the local landscape conditions mainly complex topography and accelerated land use change. Besides, RCM has rarely studied in this region, and suitable LSM and land use data configuration were not identified yet. Therefore, a focused study comparing climate variables on different LSMs along with and without updated land use information promises to provide valuable insight for WRF modelers and regional climate simulation.

In this study, six WRF experiments using three LSMs (Noah, RUC and TD) and two land use data (USGS and NLU) were considered to investigate the sensitivity of the WRF model to updated land use and identify the best performing LSM. This study is novel in the sense of that it evaluate the combined effect of LSMs and land use data on WRF model simulation performance. The WRF model was configured at fine grid spacing in order to include the complex topography and land use change dynamics. The fine grid spacing (<6 km) is adequate to depict mesoscale topography ([Goswami et al., 2012](#)), which resolve areas of high and low elevation. The result will provides useful insights into the advantages of using updated land use information in the WRF regional climate modeling applications. Moreover, the study will provide valuable information about the relative performance of land use data and LSMs for the regional climate simulation and inputs for WRF model improvement program. This study tested the hypotheses that updated land use improve the RCM at basin scale that strong spatial and temporal climate heterogeneity. Besides, the study addressed two major research questions: a) is there any difference among the land surface models for climate

simulation at finer spatial resolution? and b) does the choice of land use data result in substantial biases in rainfall and temperature? Model prediction performances were verified using observed rainfall and temperature data using different statistical and categorical verification methodologies. Thus, a thorough evaluation of the model performance and the best-identified parameters contributing to mesoscale climate simulation in this particular area results of general relevance.

## 2. Materials and methods

### 2.1. Study area

Lake Tana basin is located in Amhara region in Ethiopia and covers a catchment area of ~15,140 km<sup>2</sup>. The Lake Tana, which accounts for 20% of the basin area, is the largest freshwater body in Ethiopia and considered as the source of the Upper Blue Nile River. The basin is located at 10.95°N to 12.78°N and 36.89°E to 38.25°E ([Fig. 1](#)), and its elevation ranges from 1786 to 4109 meter above sea level.

The majority of the basin (75%) is located in temperate to cool semi-humid zone and the remaining 25% is found in cool to cold humid zone. In general, the climate in the basin is divided into rainy and dry seasons. The rainy season mainly spans from June to September ([Tigabu et al., 2018](#)), which account for 70–90% of the rainfall amount ([Enyew et al., 2014](#)). The dry season occurs from October to April ([Tigabu et al., 2018](#)) and characterized by no rainfall events. However, in the southern parts of the basin, less intense sporadic rainfall occurs in April and May.

There is high temporal and spatial rainfall variability in the Lake Tana basin because of extreme topographic variation and the presence of Lake water body ([Haile et al., 2009](#)). The mean annual rainfall within the basin varies from 964 mm to 2000 mm. The temperature is warmer around the Lake Tana and becomes cooler moving away from the Lake Tana. The mean annual maximum and minimum temperature of the basin is estimated at 27 °C and 15 °C, respectively.

### 2.2. WRF model configuration

In this study, Weather Research and Forecasting (WRF) model version 3.8 was used for climate simulation. The WRF model is developed by the National Center for Atmospheric Research (NCAR). This model is flexible and state-of-the-art atmospheric numerical simulation which is implementable parallel computing platforms ([Skamarock et al., 2008](#)) to predict weather on a mesoscale level for operational and research needs. The model was configured for a fine spatial resolution with three domains ([Fig. 1a](#)). The outer domain (D1) covers a total distance of 2775 km in the East and 1998 km in the North direction at 36 km grid spacing. This domain size was chosen to cover the Indian Ocean, red sea and East Africa region which influence rainfall patterns in Ethiopia ([Conway, 2000](#)). The intermediate domain (D2) covers a distance of 1225 km in the East and 1332 km in the North direction at 12 km resolution. The inner domain (D3) centered at Lake Tana and covers an area of 280 km in the East and 333 km in the North directions. This domain has 4 km resolution. This fine grid spacing was selected to resolve the mesoscale forcing associated with complex topography, lake and vegetation characteristics that influence local climate ([Goswami et al., 2012](#)). The three domains were run in one-way nesting mode with 35 vertical layers from near the surface to the model top at 50 hPa.

The WRF physical parameterizations were selected based on preliminary test and previous findings for the Africa region (e.g., [Pohl et al., 2011](#); [Abdallah et al., 2015](#); [Kerandi et al., 2017](#); [Mugume et al., 2017](#)). The parameterization schemes used include Thompson microphysics ([Thompson et al., 2008](#)), Rapid Radiative Transfer Model (RRTMG) for long-wave and shortwave radiation ([Hagos et al., 2014](#)), revised MM5 Monin-Obukhov surface layer ([Jiménez et al., 2012](#)), Yonsei University planetary boundary layer ([Diaz et al., 2015](#)), and Kain-Fritsch cumulus ([Berg et al., 2013](#)). However, the Kain-Fritsch cumulus scheme was not activated for the inner domain because cumulus parameterizations are

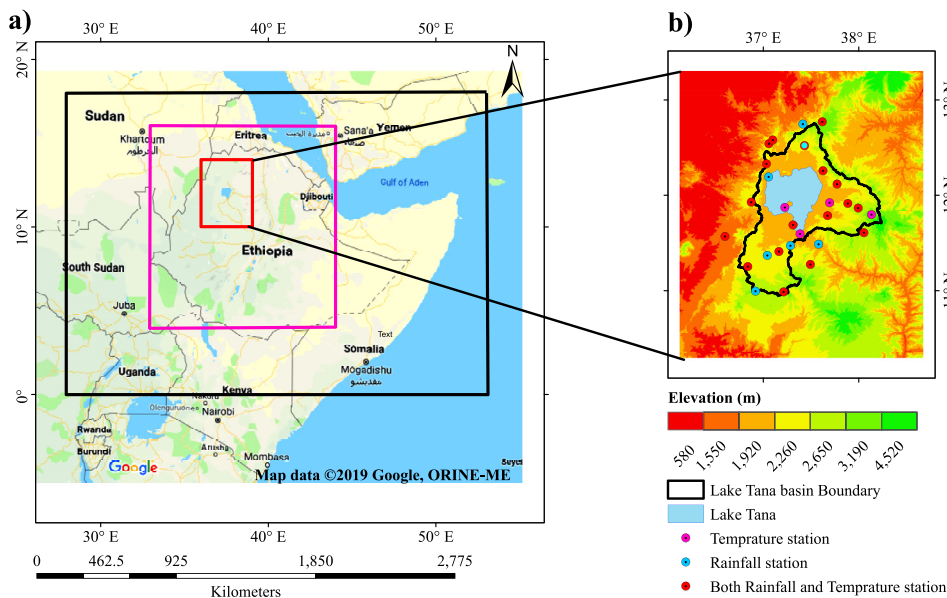


Fig. 1. Study area; a) three nested domains: the black rectangle represents the outer domain (D1) representing Eastern Africa, South Asia, and the Indian Ocean part, the purple rectangle shows the intermediate domain (D2) covering a large part of Ethiopia, and the red rectangle is the inner domain (D3) covering north western Amhara Region where the Lake Tana basin is located. b) Topographic features of the inner domain with the Lake Tana basin border (black dashed), Lake Tana, and meteorological stations that were used for model evaluation.

valid for coarse grid sizes, which is greater than 10 km (Skamarock et al., 2008).

The WRF model was initialized using boundary conditions obtained from the National Centers for Environmental Prediction (NCEP) global Final (FNL) analysis dataset. These datasets are available at a coarse resolution of  $1^\circ \times 1^\circ$  and 6 h intervals. The lateral boundary conditions in the WRF model were updated at 6 h intervals. The WRF model was forced to read-in the time-varying data and to update the sea surface temperature (SST) field. Due to the shortage of computational resource, the simulation period was six months (March–August 2015). These months

were selected purposively to represent the dry and wet season in the Lake Tana basin. The WRF simulation period was split into 6 days with 1 day of overlap between consecutive runs to allow for spin-up time. These short-run were considered to minimize the accumulation of errors as the forecast day's increases (Skamarock et al., 2008).

### 2.3. Experimental setup

Six WRF model experiments were configured using a factorial combination of two land use data and three LSMs. All the experimental

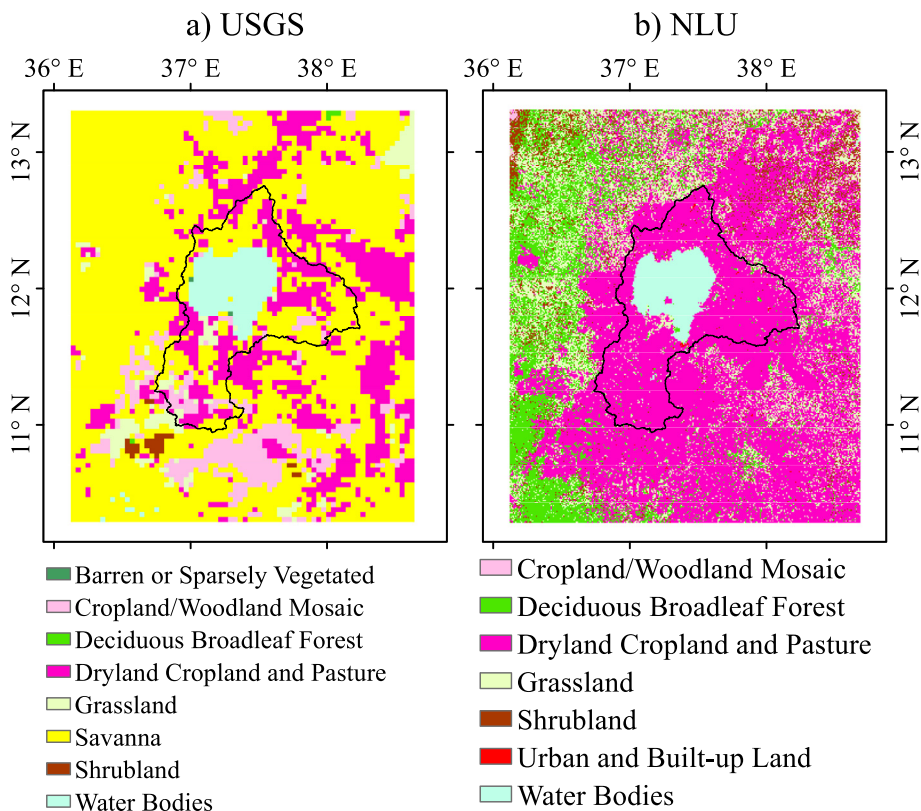


Fig. 2. Land use coverage for the inner domain; a) U.S. Geological Survey (USGS) land use, and b) new land use (NLU).

configurations used the same physical parameters except LSM and land use data. The effect of updated land use information on the performance of WRF model was studied using the U.S. Geological Survey (USGS) land use and new land use (hereafter NLU) data. The USGS land use data was generated in 1993 from Advanced Very High-Resolution Radiometer (AVHRR) data (Eidenshink and Faydeen, 1994). This data has 24 land use classes at ~900 m resolution. The inner domain USGS land use data is presented in Fig. 2a, where the larger part is covered by savanna and which failed to capture cropland expansion in 2015 (Teklay et al., 2018). Moreover, the USGS land use data did not show Alesmaga and Tara Gedam forest coverage in the Lake Tana basin, which covers ~19 km<sup>2</sup>. The NLU data was obtained from the Regional Centre for Mapping of Resources for Development (RCMRD) (<http://geoportal.rcmr.org>). This data was produced from 30 m Landsat thematic mapper using supervised classification method for 2016 Ethiopia land cover scheme II. The classification was satisfactory with an overall accuracy of 87.9% and a Kappa coefficient of 0.79 (Kindu et al., 2013). The NLU data represented well land use categories compared to USGS data. For example, the NLU captured cropland expansion and notable land use features such as Alesmaga and Tara Gedam forest coverage (Fig. 2b).

Major differences in land use distributions between USGS and NLU data were reflected in the forestland, cropland, grassland, and savanna (Table 1). The USGS data underestimated cropland coverage compared to previous research findings in the Ethiopian highlands (e.g., Teferi et al., 2010; Biru et al., 2015), while the NLU data was consistent with previous studies in capturing the urban, forest and cropland coverage.

Three land surface models (LSMs) namely; Noah, RUC and TD were also considered in the WRF experiments. The Noah LSM uses soil temperature and moisture at four layers with a total depth of 2 m. This scheme calculates the soil temperature, water-equivalent snow depth, soil ice and soil and canopy moisture (Skamarock et al., 2008). The RUC LSM estimates heat and moisture transfer for the nine layers from 0 to 3 m, and the model realistically represents soil moisture, soil temperature and snow (Smirnova et al., 2016). The RUC scheme accounts for the different phases of soil surface water, vegetation effects and canopy water. The RUC LSM has a thin layer feature that covers the ground surface including half of the first atmospheric layer and half of the topsoil layer to properly solve the energy budgets. The model accounts vegetation impact on evaporation capturing the role of canopy moisture and soil texture. The TD LSM calculates the energy budget from a 1-D equation assuming a linear temperature profile across 5 layers of soil. The TD LSM does not predict soil moisture (Skamarock et al., 2008). In TD model, initial soil moisture is fixed with the land use type and a season-dependent constant value. The bottom layer can be at a deep of 16 cm, where the average temperature is applied. The TD LSM does not explicitly account vegetation processes. Table 2 presents WRF experimental setup that combines LSMs and land use data.

2.4. Observation data

Observed climate data were used to validate the performance of WRF model. The observed daily temperature and rainfall data in the Lake Tana basin and nearby weather stations were obtained from the National

**Table 1**  
Summary of land use coverage for USGS and NLU data in the inner domain.

Land use type	USGS (%)	NLU (%)
Barren or Sparsely Vegetated	0.03	0
Cropland/Woodland Mosaic	7.99	0.03
Dryland Cropland and Pasture	20.97	65.65
Deciduous Broadleaf Forest	0.09	14.00
Grassland	3.35	14.71
Savannas	63.13	0
Shrubland	0.51	2.02
Urban and Built-up Land	0	0.02
Water Bodies	3.93	3.57

**Table 2**

The WRF experimental configurations based on land surface models (LSMs) and land use data.

Simulation number	Land surface model	Land use data	WRF experiment
1	Noah	USGS	Noah-U
2	Noah	NLU	Noah-N
3	RUC	USGS	RUC-U
4	RUC	NLU	RUC-N
5	TD	USGS	TD-U
6	TD	NLU	TD-N

Meteorological Agency of Ethiopia. Meteorological stations that have data more than 95% of the period from March to August 2015 were considered for model evaluation. Based on the availability and quality of the data, temperature and rainfall data were collected from 22 (red and purple circle, Fig. 1b) and 24 (red and blue circle, Fig. 1b) gauge stations, respectively.

2.5. Model evaluation

The WRF model performance was evaluated by examining the temporal and spatial agreement between simulated and observed values. The evaluation was done using the correlation coefficient (r), mean bias (MB) and root mean square error (RMSE). Correlation coefficient is widely used to quantify the linear relation between simulated and observed values (Eq. (1)). The r values near one indicate a strong positive correlation while values near zero indicate no correlation. The MB is a measure of the difference between simulated and observed mean (Eq. (2)); a positive value of MB indicates an overestimation and a negative value indicates an underestimation. The RMSE describes the amount of error in terms of the difference between simulated and observed values (Eq. (3)), the lower the RMSE values the smaller the errors. A Student's t-test was applied to assess the significance of the difference between simulated and observed values. The two-tail test was computed based on the two-sample assuming equal variance method at 95% confidence level.

$$r = \frac{\sum_{i=1}^n (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2 \sum_{i=1}^n (O_i - \bar{O})^2}} \tag{1}$$

$$MB = \frac{1}{n} \sum_{i=1}^n (S_i - O_i) \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2} \tag{3}$$

Where,  $S_i$  is the simulated climate variable,  $O_i$  is the observed climate variable,  $\bar{S}$  is the mean of simulated climate variable,  $\bar{O}$  is the mean of observed climate variable, and n is the total number of climate data.

Categorical statistics, which comprise Bias Score (BS), Probability of Detection (POD), False Alarm Ratio (FAR) and Critical Success Index (CSI), were also used to verify the WRF model's performance in simulating rain/no rain events at different thresholds. In this study, rainfall amount below 0.2 mm was considered as no rain (Schirmer and Jamieson, 2015). The simulated and observed pairs of a given threshold were classified into four categories and assigned in a 2 × 2 contingency table (Table 3).

The BS is a ratio between the frequency of simulated to the frequency of observed event (Eq. (4)). BS provides an evaluation of the model's tendency to overestimates (BS > 1) or underestimates (BS < 1) observed events. A BS value of one indicates that the number of simulated events is the same as the observed events.

**Table 3**  
Contingency table between observed and simulated rainfall events for a given threshold level.

		Observed		
		Yes	No	Total
Simulated	Yes	a	b	a + b
	No	c	d	c + d
	Total	a + c	b + d	n = a + b + c + d

Note a: hits, if both simulated and observed values are above threshold; b: false alarms, if simulated value is above threshold and observed value is below threshold; c: misses, if simulated value is below threshold and observed value is above threshold; d: correct non-events, if both simulated and observed values are below threshold, and n: total number of events.

$$BS = \frac{a + b}{a + c} \tag{4}$$

Where a is hits, b is false alarm and c is misses.

The POD estimates the proportion of rainfall events successfully simulated by the model (Eq. (5)). A POD value of one means a perfect rainfall event simulation.

$$POD = \frac{a}{a + c} \tag{5}$$

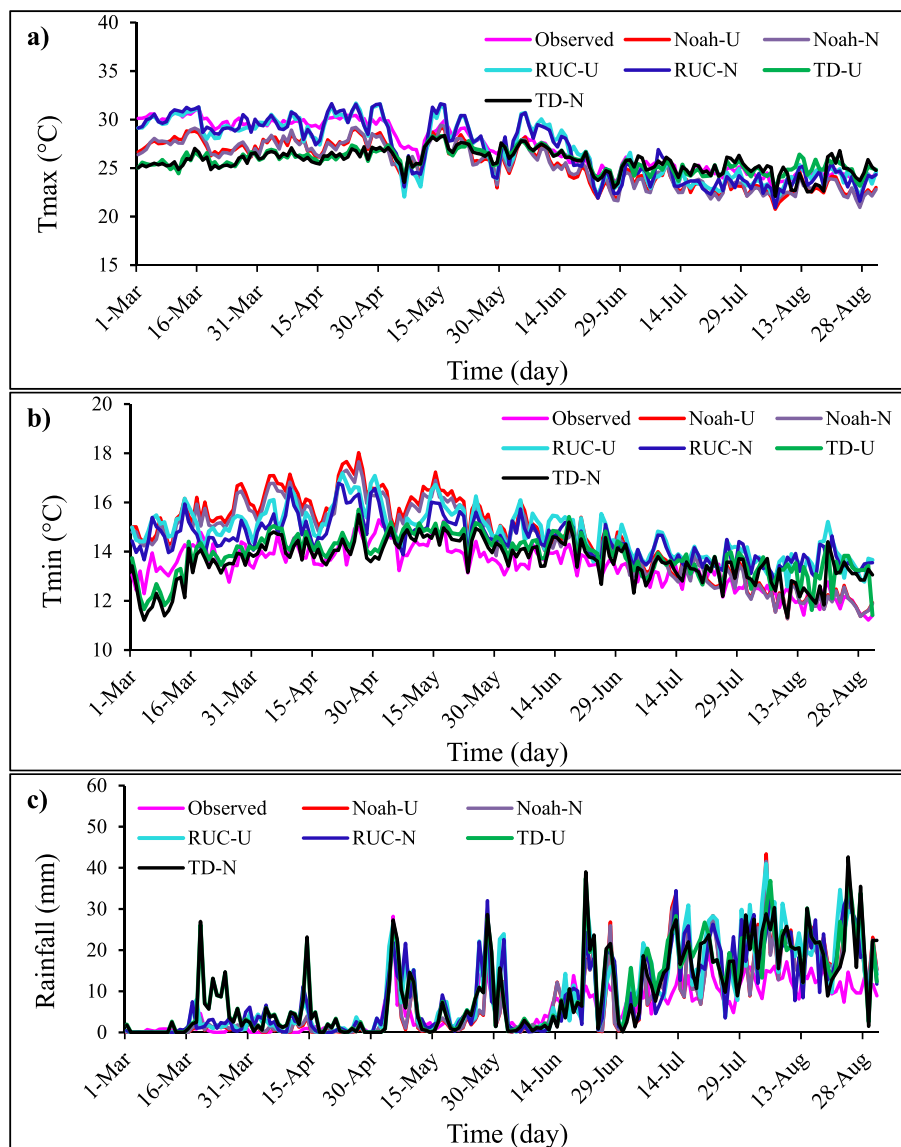
The FAR measures the number of simulated rainfall events when in fact there were no observed rainfall events (Eq. (6)). The best FAR score corresponds to zero.

$$FAR = \frac{b}{a + b} \tag{6}$$

The CSI is the fraction of observed rainfall that was correctly detected by the model (Eq. (7)).

$$CSI = \frac{a}{a + b + c} \tag{7}$$

Monthly and seasonal simulation performance of WRF model was evaluated by aggregating daily rainfall values. The model results were interpolated bi-linearly to the station location. The spatial patterns of WRF model simulation were also assessed through comparison of the observed and simulated map and Taylor diagram (Taylor, 2001). The station average temperature and rainfall values were interpolated to 4 km spatial resolution using the Kriging method. This method was selected



**Fig. 3.** The six WRF experimental simulation and observation; a) 2 m maximum temperature (Tmax), and b) 2 m minimum temperature (Tmin), and c) rainfall based stations average over the Lake Tana basin for the period March 2015 to August 2015.

because of its simplicity and reliability for spatial evaluation purposes (e.g., Pennelly et al., 2014; Zeyaayan et al., 2017).

### 3. Results

#### 3.1. WRF model evaluation

The WRF model was used to simulate temperature and rainfall at daily time steps spatially across the Lake Tana basin. The temporal and spatial simulation performance of the model was evaluated using the observed temperature and rainfall data.

##### 3.1.1. Daily temperature

Temporal variation of daily 2 m maximum temperature (Tmax) averaged for all meteorological observations and WRF simulations for the period March to August 2015 is shown in Fig. 3a. The comparison result indicated that Tmax simulation using RUC-U and RUC-N experiments showed generally good reproduction of the observed temporal variation in the Lake Tana basin. However, in some points, Tmax simulating value was not still high accurate (Fig. 3a). Both Noah and TD LSMs tended to underestimate the Tmax events from March to mid-May. During this period, Noah-N and Noah-U experiments slightly underestimated, whereas TD-U and TD-N experiment severely underestimated all the Tmax values. The comparison indicates that TD-N and TD-U simulated 2 m minimum temperature (Tmin) fairly well in general; the simulated and the observed values were coincident in most time profile. However, Tmin simulated by TD experiments slightly underestimated during the initial simulation periods and overestimated in sometimes profile in August. Both Noah and RUC LSMs experiments well captured the Tmin temporal fluctuation but had an extreme overestimation tendency in most time profile (Fig. 3b). Overall, the TD-N and TD-U experiments agreed well with the observed magnitude in most events in Tmin.

Comparisons between observed and simulated 2 m maximum and minimum temperature from six WRF experiments are presented in Table 4. The correlation coefficient values in Tmax were exceptionally high for all Noah and RUC experiments with values in the range from 0.91 to 0.94. The highest correlation was found at RUC-N and Noah-U experiments, and the lowest at RUC-U experiment. The LSM comparison showed that RUC experiments slightly underestimated the average observed Tmax, while Noah and TD simulations were considerably higher than the observed Tmax. The MB value for Tmax from RUC-N and RUC-U experiments were significantly different from other experiments ( $p < 0.05$ ). The RUC-N experiment resulted in the smallest cold bias (0.27 °C) for Tmax. On the other hand, considerable cold biases in Tmax were observed from the Noah and TD simulations with the worst performance from the Noah-U (MB = 1.78 °C) experiment. The root mean square error (RMSE) values for Tmax ranged from 1.21 to 2.48 °C, with the lowest value occurring from RUC-N and the highest value from TD-U experiment. This result showed that the observed Tmax was slightly closer to the simulated Tmax in RUC-N experiment. Although the difference between USGS and NLU data for Tmax simulations were insignificant ( $p > 0.05$ ), the NLU data along with Noah and RUC slightly improved the

**Table 4**  
Performance of WRF experiments for simulating daily 2 m maximum temperature (Tmax) and 2 m minimum temperature (Tmin) for the period March to August 2015 from 22 station average.

WRF experiment	Tmax			Tmin		
	r	MB (°C)	RMSE (°C)	r	MB (°C)	RMSE (°C)
Noah-U	0.94	-1.78	1.96	0.85	1.21	1.58
Noah-N	0.93	-1.76	1.94	0.85	1.02	1.39
RUC-U	0.91	-0.28	1.23	0.76	1.35	1.49
RUC-N	0.94	-0.27	1.21	0.75	1.07	1.22
TD-U	0.61	-1.42	2.42	0.63	0.46	0.84
TD-N	0.57	-1.47	2.48	0.65	0.20	0.77

Tmax simulation. In RUC LSM, replacing USGS through NLU data reduced Tmax RMSE value by 0.02 °C and slightly improved the WRF model performance. Overall, RUC-N experiment was the best in simulating Tmax compared to the other five experiments.

LSM selection had greater influence in simulating Tmin when compared to land use data selection (Table 4). The correlation coefficients for Tmin were good but slightly lower than Tmax except for TD experiments. The highest agreement was found from the Noah-N and Noah-U experiments, and the lowest from the TD-U experiment. In general, all the WRF experiments overestimated Tmin with a range from 0.20 to 1.35 °C. Although the warm biases were high in most experiments, TD-N simulation was very close to the observed Tmin (MB = 0.20 °C). Tmin RMSE values in NLU experiments were considerably lower than USGS. Similar to the Tmax simulation, updated land use information improved Tmin simulation and TD-N being the most consistent experiment to replicate the observed Tmin.

##### 3.1.2. Daily rainfall

Amongst all the experiments, relatively Noah-N and Noah-U experiments had the best performance in simulating rainfall (Fig. 3c). Overall, the Noah and RUC experiments reasonably captured the temporal fluctuations of low and moderate rainfall events. All the WRF experiments consistently overestimated the observed rainfall amounts in most time profiles, especially from June to August. However, TD-N and TD-U experiments severely overestimated the rainfall amount during the simulation period. From these, TD-N experiment simulated considerable rainfall amount as high as 42.4 mm and 42.2 mm on June 21 and August 24, respectively.

Comparisons between observed and simulated rainfall from the six WRF experiments are presented in Table 5. All the WRF experiments overestimated the average daily rainfall with a range of 2.39–4.53 mm/day. The lowest overestimation was found in Noah-N experiment and the highest in TD-U experiment. The impact of LSM and land use data on average daily rainfall was statistically significant. However, the LSM effect was much higher than the land use data. For example, Noah (Noah-U and Noah-N average value) LSM produced 1.91 mm/day lower rainfall than TD (TD-U and TD-N average value), whereas the NLU (Noah-N, RUC-N and TD-N average value) slightly reduced model bias by 0.31 mm/day compared to USGS (Noah-U, RUC-U and TD-U average value). The MB values in Noah-N and Noah-U experiments were significantly lower than other experiments. Also, Noah-N and Noah-U experiments produced significantly lower RMSE values than TD-U and TD-U. These results demonstrate that Noah rainfall simulations resulted in better RMSE values than and TD. The application of NLU data instead of USGS improved the WRF model performance by reducing the RMSE by 0.33, 0.59, 0.09 mm/day in Noah, RUC, and TD LSM, respectively. The correlation coefficients between simulated and observed rainfall were significant ( $p < 0.05$ ). Amongst all the LSM, Noah yielded very small mean bias and model error. Overall, the combination of Noah and NLU relatively captured well the observed rainfall amount compared to other experiments.

The high values of BS, POD, and CSI showed that the WRF experimental simulations could correctly detect and distinguish many more

**Table 5**  
Performance of WRF experiments for simulating daily rainfall for the period March to August 2015 from 24 station average.

WRF experiment	r	MB (mm/day)	RMSE (mm/day)	BS	POD	FAR	CSI
Noah-U	0.77	2.44	6.88	0.92	0.86	0.07	0.81
Noah-N	0.77	2.39	6.55	0.97	0.88	0.09	0.81
RUC-U	0.80	4.18	7.98	1.03	0.91	0.11	0.82
RUC-N	0.79	3.71	7.39	1.02	0.92	0.10	0.83
TD-U	0.76	4.53	8.24	1.01	0.91	0.10	0.83
TD-N	0.75	4.12	8.15	1.01	0.90	0.10	0.82

rainfall events (Table 5). Amongst all the experiments, Noah-N and Noah-U slightly underestimated the frequency of observed rainfall events, but the other experiments slightly overestimated. As can be seen, the WRF model detected most rainfall events as high as 0.92 in RUC-N experiment. Also, all the WRF experiments had very high CSI value ranging from 0.81 to 0.83, indicating more than 81% of the observed events are properly detected by the WRF model. However, the WRF experimental simulations missed some rain events or reported false events. From these; RUC-U experiment missed 11% of the rainfall events.

The WRF model performance for simulating different rainfall threshold values was assessed using four skill scores (Fig. 4). Amongst all the experiments, Noah slightly overestimated the frequency of light rainfall events (<5 mm/day), while RUC and TD slightly underestimated it (Fig. 4a). However, almost all the WRF experiments substantially underestimated and overestimated the frequency of moderate (5–15 mm/day) and high (>20 mm/day) rainfall events, respectively. For the rainfall threshold up to 15 mm/day, Noah-N experiment displayed the highest skill scores with BS close to one. The POD in the Noah-N experiment was appreciably high, especially for the low and moderate events (Fig. 4b). This experiment detected more than 22% of rain occurrences at a threshold of 10 mm/day, but it was incapable of detecting high threshold values. The POD values generally decreased in a wave pattern with as the rainfall threshold increased. This result demonstrated that detecting heavy rainfall events are more difficult than light rainfall events. Unlike to POD, FAR was increased with the increasing of threshold values. The WRF experiments had different FAR values in 10 and 15 mm/day threshold values (Fig. 4c). Both Noah-U and Noah-N experiments produced small FAR in moderate rainfall events. The CSI values were so high for low threshold and extremely low for medium and high threshold (Fig. 4d), which indicate that WRF experiments correctly estimated low threshold events than medium and high threshold events.

### 3.2. Monthly and seasonal WRF simulation

#### 3.2.1. Monthly and seasonal temperature

Fig. 5 shows a comparison between WRF experimental simulation and observation average monthly temperature and rainfall. The RUC (RUC-U and RUC-N) average monthly Tmax was very close to the observed Tmax in most of the months (Fig. 5a), which caused lower model biases. Contrarily Noah and TD simulation was considerably lower than the observed Tmax, especially in March and April. During the simulation

periods, most of the WRF experiments underestimated the observed Tmax with a range of 0.1 °C–4.4 °C. As can be seen in Fig. 5b, Noah-U and Noah-N experiment produced the highest (16.5 °C) and lowest (12 °C) average monthly Tmin in April and August, respectively. In general, Noah-N produced the lowest overestimation (0.2 °C) in July, while Noah-U produced the highest overestimation (2.1 °C) in April.

The rainy season in the study area occurs from June to September (Enyew et al., 2014; Tigabu et al., 2018). Thus, the months from June to August were considered as the wet season, and due to low rainfall amount, the months of March to May were considered as the dry season. Table 6 presents the comparison between simulated and observed Tmax on a monthly and seasonal basis. The correlation between simulated and observed average monthly Tmax was very strong for the RUC-N experiment. From the seasonal comparison, RUC-N and Noah-N experiment had very strong correlation during the dry and wet season, respectively. All the six experiments reasonably underestimated both dry and wet season Tmax except TD in the wet season. During the dry season, considerably high and low cold biases in Tmax occurred in the TD-N and RUC-N experiments, respectively. Overall, TD-N experiment well captured wet season Tmax. These results showed that dry season biases were significantly higher than the respective wet season biases except for RUC-N experiment. Based on RMSE values, RUC-N and TD-N experiment provided a better outcome during the dry season wet season, respectively. This result showed that the application of NLU data captured well the seasonal cycle in the study area. Overall, better monthly and dry season WRF simulation performance was obtained from the RUC-N experiment.

During monthly time steps, most of the experimental simulations had the best agreement with the observed Tmin (Table 7). However, the MB and RMSE in TD-N experiment were significantly different from others. This result showed that TD-N experiment resulted in the best WRF model configuration to simulate monthly Tmin. The seasonal comparison of Tmin showed that wet season RMSE values were lower than dry season RMSE. This result revealed that the WRF model had a better capability in simulating wet season Tmin compared to the dry season. Amongst all the experiments, TD-N produced significantly lower RMSE in both dry and wet season. The correlations between observation and WRF simulations were very strong in both seasons, but the dry season r values were slightly higher compared to wet season simulation (Table 7). Overall, TD-N experiment adequately captured the monthly and seasonal Tmin cycle.

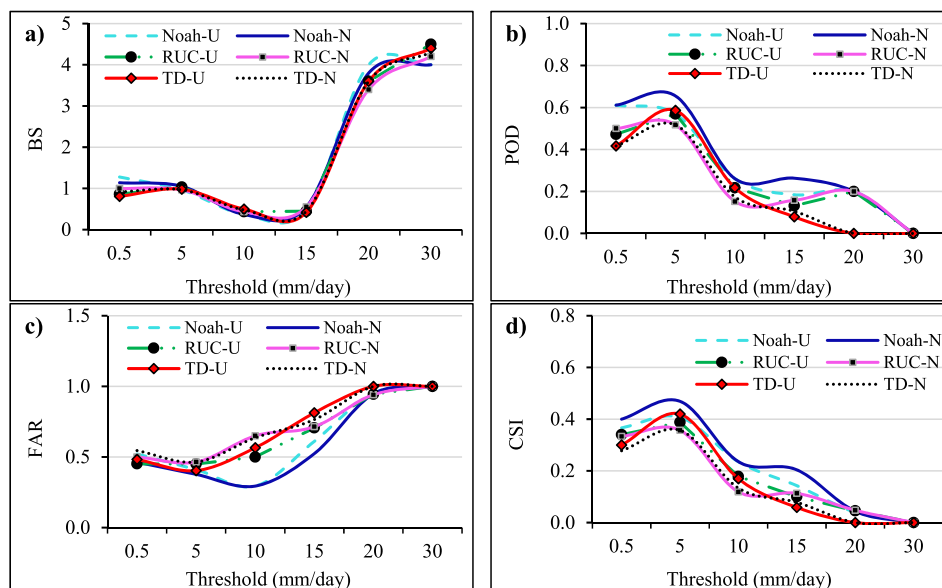


Fig. 4. Skill scores as a function of rainfall threshold for six WRF experiments for the periods March 2015 to August 2015; a) bias score (BS), b) probability of detection (POD), c) false alarm ratio (FAR), and d) critical success index (CSI).

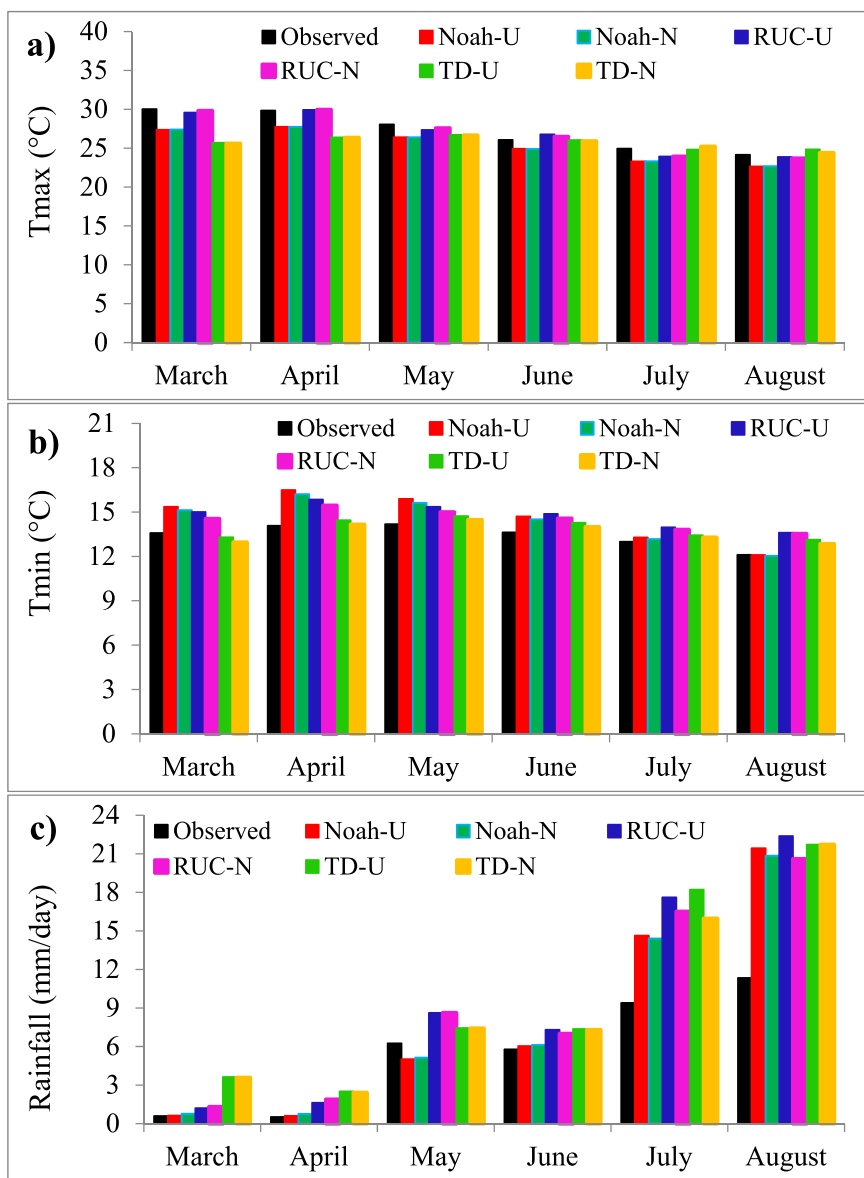


Fig. 5. Average monthly 2m maximum temperature (a), 2m minimum temperature (b), and rainfall (c) for the simulation period March–August 2015 over Lake Tana basin.

Table 6

Performance of WRF experiments for simulating monthly and seasonal 2 m maximum temperature (Tmax) for the period March to August 2015.

WRF experiment	Monthly			Dry season			Wet season		
	r	MB (°C)	RMSE (°C)	r	MB (°C)	RMSE (°C)	r	MB (°C)	RMSE (°C)
Noah-U	0.81	-1.78	2.76	0.73	-2.12	3.04	0.70	-1.42	2.45
Noah-N	0.81	-1.77	2.76	0.73	-2.11	3.03	0.71	-1.41	2.44
RUC-U	0.82	-0.28	2.14	0.74	-0.36	2.14	0.69	-0.20	2.14
RUC-N	0.83	-0.27	2.10	0.75	-0.19	2.09	0.69	-0.34	2.12
TD-U	0.64	-1.42	3.09	0.64	-3.01	3.85	0.68	0.17	2.07
TD-N	0.65	-1.47	3.11	0.63	-3.06	3.91	0.70	0.12	2.01

3.2.2. Monthly and seasonal rainfall

Observed and simulated average monthly rainfall for the period from March to August 2015 is shown in Fig. 5c. Amongst all the WRF experiments, RUC-U and Noah-U simulated the largest (22.4 mm/day) and the smallest (0.6 mm/day) average monthly rainfall in August and March, respectively. From March to July, the rainfall in the Noah (Noah-U and Noah-N) experiments was slightly lower than the other experiments. In August, RUC-N experiment produced the lowest rainfall compared to

others. The average monthly rainfall was remarkably overestimated, primarily in July and August, and, to a lesser extent in March and April. The lowest overestimation (0.1 mm/day) occurred in March and April in Noah-U experiment, while the highest overestimation (11 mm/day) in August in RUC-U.

The analysis at a monthly level showed that the correlation coefficient values were between 0.71 and 0.79 (Table 8). The monthly MB and RMSE value in the Noah-N experiment was significantly lower than



**Table 7**

Performance of WRF experiments for simulating monthly and seasonal 2 m minimum temperature (Tmin) for the period March to August 2015.

WRF experiment	Monthly			Dry season			Wet season		
	r	MB (°C)	RMSE (°C)	r	MB (°C)	RMSE (°C)	r	MB (°C)	RMSE (°C)
Noah-U	0.83	1.21	2.09	0.88	1.96	2.80	0.84	0.46	1.77
Noah-N	0.83	1.02	2.00	0.88	1.70	2.65	0.84	0.35	1.74
RUC-U	0.87	1.35	2.05	0.86	1.45	2.56	0.87	1.23	2.01
RUC-N	0.88	1.07	1.84	0.88	1.05	2.22	0.87	1.06	1.90
TD-U	0.88	0.46	1.73	0.89	0.21	2.04	0.85	0.70	1.83
TD-N	0.88	0.20	1.66	0.89	-0.08	1.99	0.86	0.46	1.73

others. This result indicated that the Noah-N experiment captured well the monthly rainfall cycle. Similar to temperature simulations, the use of NLU data produced better rainfall simulation compared to USGS data. Overall, the Noah-N experiment provided the best performance, while TD-U was the worst.

The wet season r values were slightly lower than the dry season values, indicating difficulties of the WRF model to reproduce the wet season rainfall. All the WRF experiments considerably overestimated the wet season rainfall with MB values between 153.8 to 215.1 mm/month. However, in the dry season, the Noah experiments slightly underestimated the observed rainfall. In the wet season, the MB values in the NLU experiment were significantly different from the MB values in the USGS. The NLU data produced so low MB values with the lowest value in the Noah-N experiment. In all the WRF experiments, the wet season RMSE values were considerably larger than the respective dry season values, which caused by the higher rainfall magnitude in the wet season. This result agreed with Jain et al. (2017) study in the Delhi-Mumbai Industrial Corridor. They found that the WRF model error was increased with rainfall amount increased. In the Noah simulations, the application of NLU data instead of USGS improved the WRF model performance by reducing the RMSE by 0.6 and 9.9 mm/month in the dry and wet season, respectively. However, during the dry season, the updated land use information with the TD LSM did not show model improvement. As can be seen, Noah and RUC LSM have better skill in responding to the updated land use information. Overall, the Noah-N experiment was the most accurate in replicating the monthly and seasonal temporal patterns (Table 8).

### 3.3. Spatial pattern

#### 3.3.1. Temperature

RUC-N experiment resulted in the highest spatial correlation pattern ( $r = 0.85$ ) for Tmax (Fig. 6a), while Noah-N experiment produced the highest r (0.96) for Tmin (Fig. 6b). Moreover, RUC-N and Noah-N provided much lower RMSE values in Tmax and Tmin simulation, respectively. From the six experiments, TD-N and TD-U standard deviations were slightly smaller than the observed Tmax standard deviation. In contrast, all the six WRF experiment standard deviations were considerably lower than the observed Tmin standard deviation. Overall, RUC-N experiment replicated well the observed Tmax spatial patterns with a higher spatial correlation and a lower RMSE value, whereas Noah-N experiment consistently simulated the observed Tmin spatial pattern. On the other hand, TD simulations poorly replicated both Tmax and Tmin

**Table 8**

Performance of WRF experiments for simulating monthly and seasonal rainfall (mm/month) for the period March to August 2015.

WRF experiment	Monthly			Dry season			Wet season		
	r	MB	RMSE	r	MB	RMSE	r	MB	RMSE
Noah-U	0.77	74.1	178.8	0.79	-11.6	61.1	0.58	161.0	229.3
Noah-N	0.77	72.7	171.5	0.79	-7.3	60.5	0.58	153.8	219.4
RUC-U	0.79	127.9	206.8	0.77	42.2	84.3	0.60	214.8	262.1
RUC-N	0.76	113.8	192.7	0.75	46.3	81.2	0.55	182.2	241.3
TD-U	0.75	138.6	217.3	0.61	63.3	101.6	0.59	215.1	269.8
TD-N	0.71	125.8	210.3	0.60	61.9	102.3	0.50	190.6	259.5

spatial patterns with lower correlation and higher RMSE value.

Tmax spatial pattern over Lake Tana basin is shown in Fig. 7, where Figs. 7a and b are averages for the simulation period from observation and WRF experimental simulations, respectively. Fig. 7c shows the difference between WRF experimental simulation (Fig. 7b) and observation (Fig. 7a). The observed Tmax magnitude and spatial pattern were represented well by RUC-N and RUC-U experiments compared to others. During the simulation period, the highest Tmax was found over the eastern part of the Lake Tana border with an average daily Tmax above 29 °C in the RUC-N and RUC-U experiments; while the lowest Tmax was observed in the northern and eastern mountainous part of the basin with an average daily Tmax below 22 °C in the Noah-N and Noah-U experiments. All the WRF experiments substantially underestimated Tmax over the entire basin with cold bias as high as 4 °C in Noah and TD experiments over the northern and southern mountainous part of the basin. From the six experiments, RUC-N produced considerable small cold biases ( $> -3$  °C) and warm biases ( $< 1$  °C) over the basin, which indicates that the combination of RUC with NLU data replicated well Tmax spatial patterns in the Lake Tana basin. In general, the comparison between NLU and USGS simulation showed that the consideration of real-time land use data slightly enhanced the spatial representation performance of the WRF model. This result agreed with the study conducted by Sertel et al. (2009) in the Marmara Region in Turkey. They found that the updated land cover data improved the WRF model performance in temperature simulations. In the central part (Bahir Dar) of the basin, the NLU data produced higher Tmax than USGS land use data. This variation may have been related to the urban and cropland representation of the area by NLU and USGS data, respectively. Higher temperature simulation was commonly observed in urban land use coverage (Li et al., 2018).

All the WRF experiments showed similar spatial patterns with the observed Tmin, especially the Noah-U experiment replicated well the higher Tmin and experiments with RUC and TD LSM captured well the lower Tmin (Fig. 8). Overall, the six WRF experiments produced reasonable cold biases of Tmin over the northern parts of the Lake Tana basin. The Noah and RUC experiments produced strong warm biases up to 3 °C in the southern part, while TD-N and TD-U experiments yielded strong cold biases as high as 4 °C in the northern part. In most parts of the basin, TD-N experiment produced slight warm biases. Overall, Tmin simulation with TD-N experiment was the best to replicate the spatial distribution of the observed data.

#### 3.3.2. Rainfall

The spatial rainfall pattern was extremely sensitive to LSM selection

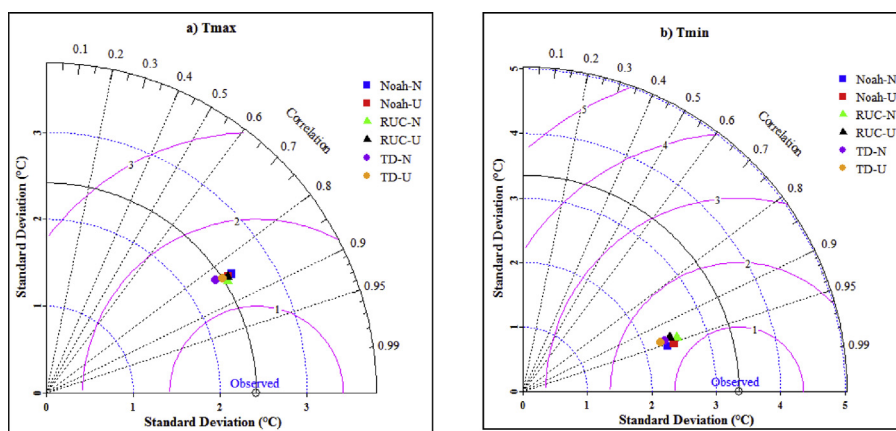


Fig. 6. Taylor diagram showing a statistical comparison of six WRF experimental simulations; a) 2 m maximum temperature (Tmax), and b) 2 m minimum temperature (Tmin) for the period from March 2015 to August 2015.

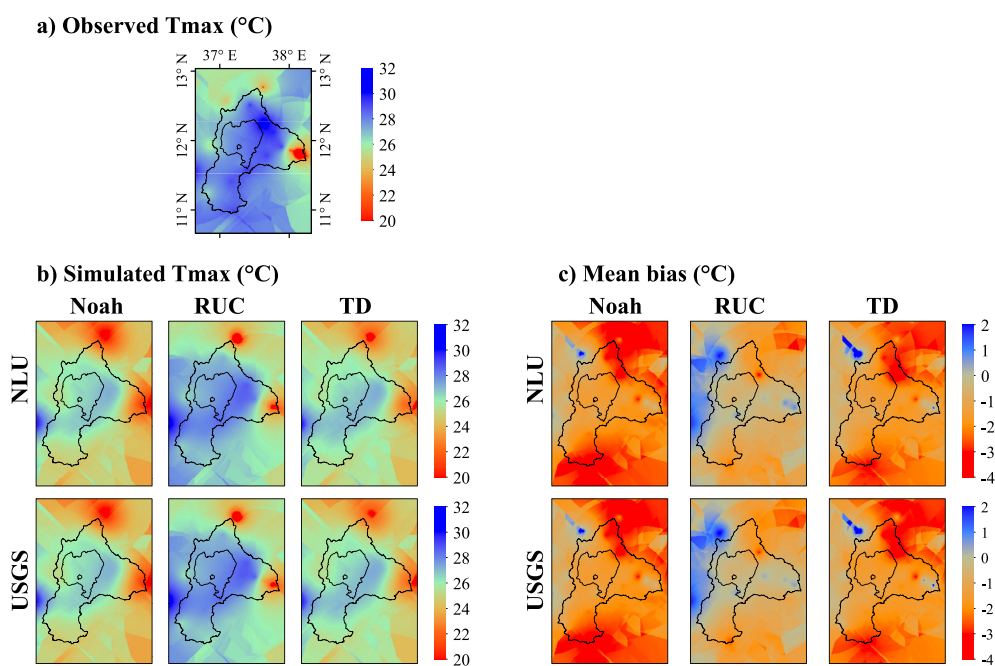


Fig. 7. Spatial distribution of average 2 m maximum temperature (Tmax) across the Lake Tana basin; a) Observed average, b) Simulated average from WRF experiment, and c) Mean bias, which is simulated minus observed.

compared to land use data (Fig. 9). The standard deviations in Noah and RUC experiments were slightly lower than the observed standard deviation (1.8 mm/day). The simulated pattern of average rainfall using NLU data agreed well with the observed rainfall compared to USGS data; especially the Noah-N experiment provided the best spatial agreement with observations. The worst performance of spatial rainfall distribution was found from TD-U experiment.

As can be seen in Fig. 10, the spatial pattern showed large discrepancies between observed and simulated rainfall in some parts of the basin. This may be due to the insufficient number of rainfall stations (Osuri et al., 2015). Simulated spatial rainfall in the Noah-N and Noah-U experiments represented the spatial patterns and magnitude well, perhaps more strongly in the mountainous parts of the Lake Tana basin. However, there were pockets of large overestimations by the Noah experiment over the central and southwestern part of the basin. This result demonstrated that Noah-N and Noah-U experiment was the best in capturing the spatial rainfall distribution with slight underestimation in the northern and eastern part and slight overestimation in the southern part of the basin. However, TD and RUC LSM considerably overestimated

the average rainfall in most parts of the basin. Overall, the Noah-N experiment was the most accurate configuration to reproduce the spatial pattern over the basin except around Lake Tana.

A summary of the spatial correlation coefficients between observation and simulation for the dry and wet season is given in Table 9. All spatial correlations were significant ( $p < 0.05$ ) except rainfall in the dry season. All the WRF experiments replicate both the dry and wet seasons Tmax spatial distribution very well with correlation as high as 0.74 to 0.87. Similarly, Tmin spatial correlation coefficients were extremely high in both dry and wet seasons. In the wet season, rainfall and Tmin spatial correlation from Noah-U and Noah-N were much higher than other experiments. On the other hand, RUC-N was consistent in representing the dry and wet season Tmax distribution. In general, coefficients from the dry season rainfall using TD were negative, indicating that TD LSM simulated high rainfall during the dry season. At the same season, Noah and RUC coefficients were generally smaller and positive. These results demonstrate that the WRF model had a better capability for simulating temperature compared to rainfall. Overall, six WRF experiments showed significant limitations in representing the dry season rainfall distribution

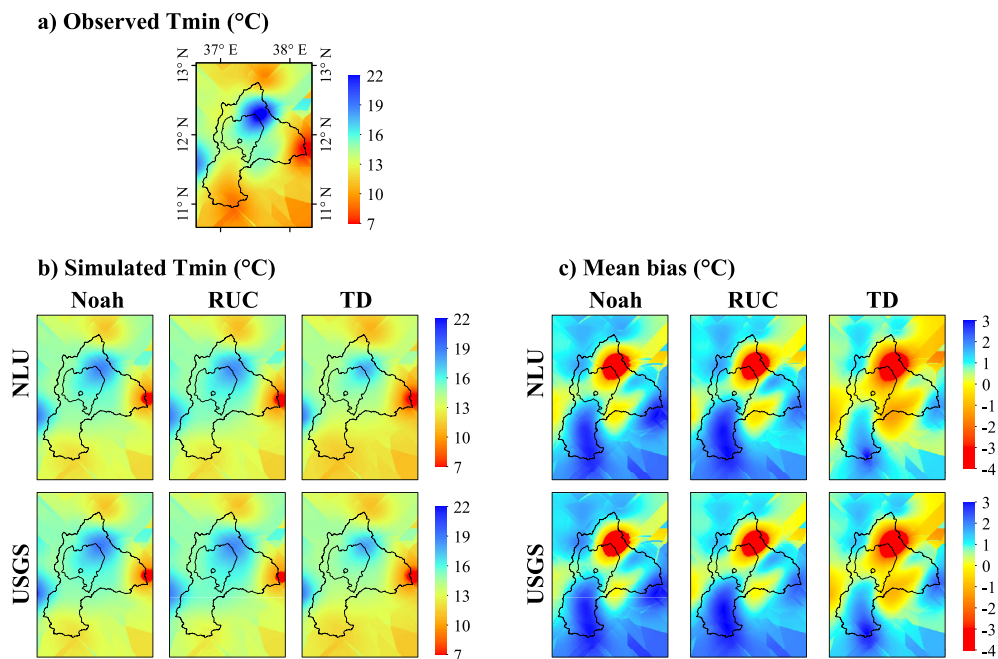


Fig. 8. Spatial distribution of average 2 m minimum temperature (Tmin) across the Lake Tana basin; a) Observed average, b) Simulated average from WRF experiment, and c) Mean bias, which is simulated minus observed.

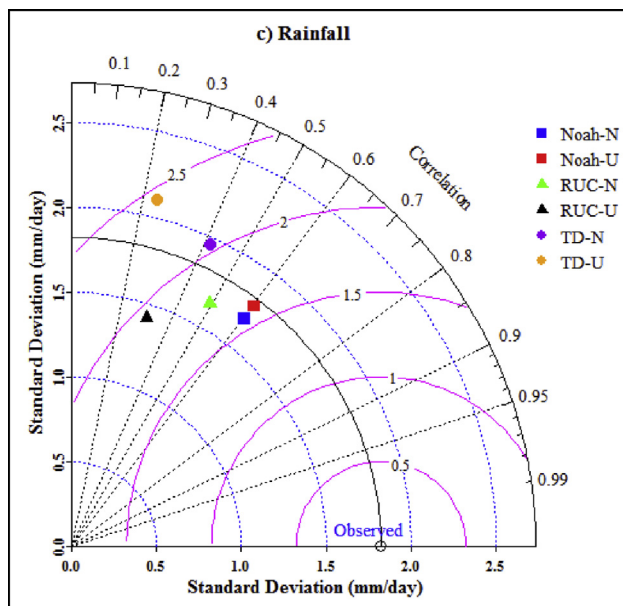


Fig. 9. Taylor diagram showing a statistical comparison of rainfall from six WRF experimental simulations for the period from March 2015 to August 2015.

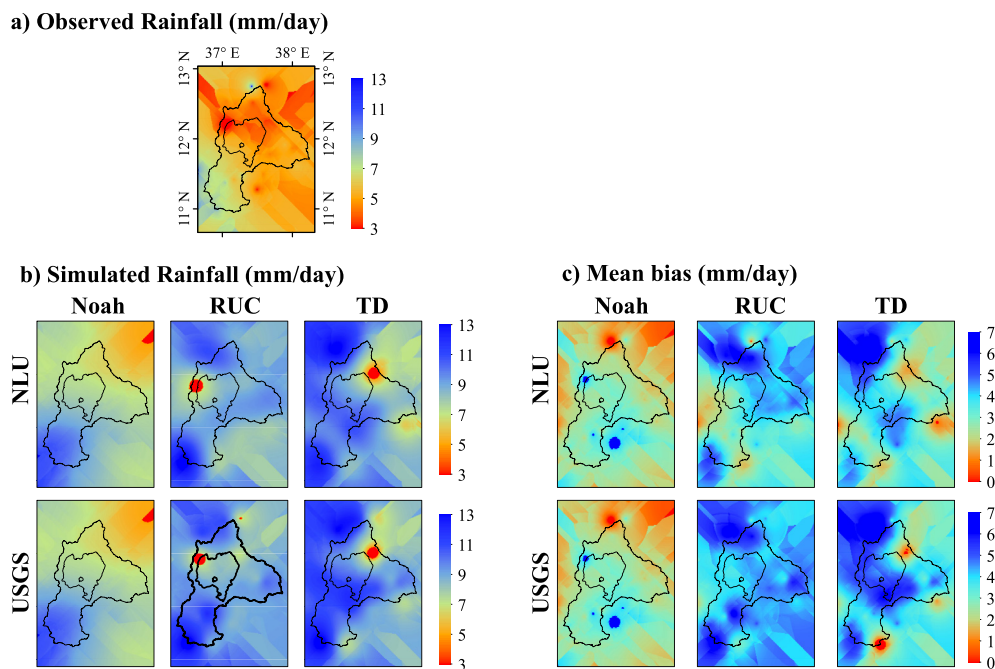
over the study area.

#### 4. Discussion

Although WRF is a widely used RCM, the choice of parameterization schemes is site-specific (Mooney et al., 2016; Jain et al., 2017; Mugume et al., 2017). The LSMs and land use data have been a substantial influence on local climate simulations, especially in areas where land surface characteristics dynamically changed (Deng et al., 2013). Because of this, the sensitivity of LSMs and land use data on WRF simulation over Lake Tana basin was demonstrated. In general, WRF rainfall and temperature simulation were sensitive to the choice of LSM and land use data. The

NLU with Noah and RUC simulated higher Tmax than USGS land use since the NLU includes more cropland area than the USGS data (Table 1 and Fig. 2) and therefore increase the temperature probably due to a lower vegetation greenness that reduces evapotranspiration. This is in agreement with the findings of Muller et al. (2014). The comparison results demonstrated that there was no single best WRF model configuration exists for simulations of dry and wet season climate in Lake Tana basin. The combination of updated NLU with RUC and TD well reproduce Tmax and Tmin, respectively. This performance variation of LSM may arise from the consideration of soil layers and land use information, which was consistent with the previous finding (Chen et al., 2014; Burakowski et al., 2016). The RUC LSM has a better capability in resolving the soil moisture and temperature at tiny layers (Smirnova et al., 2016), which result in lower model biases in Tmax simulation. As a result of fine spatial resolution and the real-time representation of land use categories, the NLU data provided reasonably better simulation than USGS. The WRF improvements resulted from updated land use data agree with Cheng et al. (2013), Cao et al. (2015) and He et al. (2017). However, during Tmax simulation, the combination of TD and NLU had slightly lower performance than TD and USGS combination. This may be related to the application of fixed soil moisture with land use type in TD LSM (Lee et al., 2016).

The rainfall simulated in the WRF experiments was much higher than the observed, which may have been caused by the excessive convection in the Indian Ocean (Diaz et al., 2015) and Kain-Fritsch cumulus scheme application (Ramarohetra et al., 2015; Chawla et al., 2018). Moreover, the presence of Lake Tana water body may influence the rainfall simulation in the WRF model (Haile et al., 2009). Although the WRF experiments poorly captured the average rainfall magnitude, the relative performance of Noah LSM and NLU data were significantly higher compared to others. The high performance of Noah LSM and NLU is consistent with Ramarohetra et al. (2015) and Gondöcs et al. (2015). In general, the deficiency of rainfall simulation using RCM is a common phenomenon (Chotamonsak et al., 2012; Chawla et al., 2018), which result from the incapability of the models to handle complex biogeochemical and biogeophysical processes (Pongratz et al., 2010), and the nonlinear interactivity of model parameters including microphysics, cumulus, and planetary boundary layer.



**Fig. 10.** Spatial distribution of average rainfall across the Lake Tana basin; a) Observed average, b) Simulated average from WRF experiment, and c) Mean bias, which is simulated minus observed rainfall.

**Table 9**

Spatial correlation coefficients between observation and WRF experimental simulations for the dry (March to May) and wet (June to August) season Tmax, Tmin, and rainfall over Lake Tana basin.

WRF experiment	Tmax		Tmin		Rainfall	
	Dry season	Wet season	Dry season	Wet season	Dry season	Wet season
Noah-U	0.86	0.75	0.92	0.93	0.15 <sup>ns</sup>	0.62
Noah-N	0.86	0.75	0.92	0.93	0.13 <sup>ns</sup>	0.62
RUC-U	0.87	0.74	0.90	0.92	0.09 <sup>ns</sup>	0.57
RUC-N	0.87	0.75	0.91	0.92	0.03 <sup>ns</sup>	0.38
TD-U	0.87	0.73	0.94	0.90	-0.21 <sup>ns</sup>	0.56
TD-N	0.87	0.74	0.93	0.90	-0.22 <sup>ns</sup>	0.38

Note: ns represent spatial correlation is non-significant at 95% confidence level.

Although the hypotheses were supported statistically, the WRF model had considerable limitation to reproduce the temporal and spatial variations of rainfall over Lake Tana basin. Future research should, therefore, include other physical parameterization such as microphysics, cumulus, and planetary boundary layers to evaluate in greater detail whether the rainfall magnitude and spatial coverage improved in the WRF model.

### 5. Conclusions

This study evaluated the Weather Research and Forecasting (WRF) model for regional climate simulations in the Lake Tana basin in Ethiopia. Six WRF experiments were configured to assess the performance of three LSM (Noah, RUC and TD) and two land uses (USGS and NLU) to simulate temperature and rainfall. The new land use data captured well the land use dynamics in the basin such as cropland and urban expansions, and reduction of Lake Tana water body. However, the USGS land use data extremely underestimated the cropland and forest coverage in the basin.

The temporal and spatial comparison showed that the impact of LSM choice was slightly higher than land use data. Differences in land use data provide a slight difference for both temperature and rainfall simulations

across space. On the other hand, the simulated average temperature and rainfall by the different LSMs were substantially different. The RUC-N and TD-N simulations yielded small cold bias for Tmax (0.27 °C) and small warm bias (0.20 °C) for Tmin, respectively. For rainfall simulations, the Noah-N and Noah-U experiment provided similar results; with slightly better spatial and temporal representation in favor of Noah-N. On average, TD produced 1.91 mm/day more rainfall when compared to Noah. The application of updated and accurate land use information improved the spatial and temporal representation of the WRF model in simulating temperature and rainfall. During rainfall simulation, updated land use improved model performance by reducing the RMSE by 0.34 mm/day when compared to USGS data. However, the effect of land use data on rainfall magnitude and distribution are complex due to the presence of Lake Tana water and the extreme topographic variability in the Lake Tana basin.

This study was able to show modeling approaches that can provide temperature and rainfall data in data-scarce environments and would provide valuable supports in identifying the suitable land use data and LSM on climate and hydrological process modeling. The findings suggested that the combination of new land use data with RUC, TD, and Noah LSM in the WRF model can be used to simulate Tmax, Tmin, and rainfall, respectively. Such data will be vital for climate change impact research for sustainable water resources management in the Lake Tana basin.

### Declarations

#### Author contribution statement

Achenafi Teklay: Conceived and designed the experiments; Performed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Yihun Dile: Conceived and designed the experiments; Performed the experiments. Dereje Asfaw Hailu: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Haimanote Bayabil, Kibruyesfa Sisay: Analyzed and interpreted the data.

### Funding statement

This work was supported by the International Foundation for Science [grant number W/5926-1].

### Competing interest statement

The authors declare no conflict of interest.

### Additional information

No additional information is available for this paper.

### Acknowledgements

National Meteorological Agency of Ethiopia is gratefully acknowledged for providing rainfall and temperature data. We also acknowledge the financial and software support provided by the University of Gondar and Addis Ababa University. The authors also thank the International Livestock Research Institute for facilitating internet service during the research period.

### References

- Abdallah, A.A., Eid, M.M., Wahab, M.M., 2015. Regional climate simulation of WRF model over north Africa: temperature and precipitation. *World Environ.* 5, 160–173.
- Ayele, H.S., Li, M.H., Tung, C.P., Liu, T.M., 2016. Assessing climate change impact on Gilgel Abbay and Gumara watershed hydrology, the upper blue Nile basin, Ethiopia. *Terr. Atmos. Ocean. Sci.* 27, 1005–1018.
- Banks, R.F., Baldasano, J.M., 2016. Impact of WRF model PBL schemes on air quality simulations over Catalonia, Spain. *Sci. Total Environ.* 572, 98–113.
- Berg, L.K., Gustafson, W.I., Kassianov, E.I., Deng, L., 2013. Evaluation of a modified scheme for shallow convection: implementation of CuP and case studies. *Mon. Weather Rev.* 141, 134–147.
- Biru, M.K., Minale, A.S., Debay, A.B., 2015. Multitemporal land use land cover change and dynamics of blue Nile basin by using GIS and remote sensing techniques, north-western Ethiopia. *Int. J. Environ. Sci.* 4, 81–88.
- Burakowski, E., Ollinger, S.V., Bonan, G.B., Wake, C.P., Dibb, J.E., Hollinger, D.Y., 2016. Evaluating the climate effects of reforestation in new England using a weather research and forecasting (WRF) model multiphysics ensemble. *J. Clim.* 29, 5141–5156.
- Cannon, F., Carvalho, L.M., Jones, C., Norris, J., Bookhagen, B., Kiladis, G.N., 2017. Effects of topographic smoothing on the simulation of winter precipitation in high mountain Asia. *J. Geophys. Res.* 122, 1456–1474.
- Cao, Q., Yu, D., Georgescu, M., Han, Z., Wu, J., 2015. Impacts of land use and land cover change on regional climate: case study in Northern China. *Environ. Res. Lett.* 10, 1–12.
- Cao, Y., Cervone, G., Barkley, Z., Lauvaux, T., Deng, A., Taylor, A., 2017. Analysis of errors introduced by geographic coordinate systems on weather numeric prediction modeling. *Geosci. Model Dev.* 10, 3425–3440.
- Chawla, I., Osuri, K.K., Mujumdar, P.P., Niyogi, D., 2018. Assessment of the Weather Research and Forecasting (WRF) model for simulation of extreme rainfall events in the upper Ganga Basin. *Hydro. Earth Syst. Sci.* 22, 1095–1117.
- Chen, F., Liu, C., Dudhia, J., Chen, M., 2014. A sensitivity study of high-resolution regional climate simulations to three land surface models over the western United States. *J. Geophys. Res.* 119, 7271–7291.
- Cheng, F.Y., Hsu, Y.C., Lin, P.L., Lin, T.H., 2013. Investigation of the effects of different land use and land cover patterns on mesoscale meteorological simulations in the Taiwan area. *J. Appl. Meteorol. Climatol.* 52, 570–587.
- Chotamonsak, C., Salathe, E.P., Kreasuwan, J., Chantara, S., 2012. Evaluation of precipitation simulations over Thailand using a WRF regional climate model. *Chiang Mai J. Sci.* 39, 623–628.
- Conway, D., 2000. The climate and hydrology of the upper blue Nile river. *Geogr. J.* 166, 49–62.
- De Meij, A., Vinuesa, J.F., 2014. Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. *Atmos. Res.* 143, 351–370.
- Deng, X., Zhao, C., Yan, H., 2013. Systematic modeling of impacts of land use and land cover changes on regional climate: a review. *Adv. Meteorol.* 2013, 1–11.
- Diaz, J.P., González, A., Expósito, F.J., Pérez, J.C., Fernández, J., García-Díez, M., Taima, D., 2015. WRF multi-physics simulation of clouds in the African region. *Q. J. R. Meteorol. Soc.* 141, 2737–2749.
- Dile, Y.T., Berndtsson, R., Setegn, S.G., 2013. Hydrological response to climate change for Gilgel Abay river, in the Lake Tana basin - upper blue Nile basin of Ethiopia. *PLoS One* 8, 1–13.
- Eidenshink, J.K., Faydeen, J.L., 1994. The 1 km AVHRR global land dataset: first stages in implementation. *Int. J. Remote Sens.* 15, 3443–3462.
- Enyew, B., Lanen, H., Van, L.A., 2014. Assessment of the impact of climate change on hydrological drought in Lake Tana catchment, blue Nile basin, Ethiopia. *Geol. Geosci.* 3, 1–18.
- Gao, J., Hou, W., Xue, Y., Wu, S., 2017. Validating the dynamic downscaling ability of WRF for East Asian summer climate. *Theor. Appl. Climatol.* 128, 241–253.
- Gashaw, T., Mebrat, W., Hagos, D., Nigussie, A., 2014. Climate change adaptation and mitigation mechanisms in Ethiopia. *J. Biol. Agric. Healthcare* 4, 148–152.
- Göndöcs, J., Breuer, H., Horváth Á., Ács F., Rajkai, K., 2015. Numerical study of the effect of soil texture and land use distribution on the convective precipitation. *Hungarian Geogr. Bull.* 64, 3–15.
- Goswami, P., Shivappa, H., Goud, S., 2012. Comparative analysis of the role of domain size, horizontal resolution and initial conditions in the simulation of tropical heavy rainfall events. *Meteorol. Appl.* 19, 170–178.
- Hagos, S., Leung, L.R., Xue, Y., Boone, A., de Sales, F., Neupane, N., Huang, M., Yoon, J.H., 2014. Assessment of uncertainties in the response of the African monsoon precipitation to land use change simulated by a regional model. *Clim. Dyn.* 43, 2765–2775.
- Haile, A.T., Rientjes, T., Gieske, A., Gebremichael, M., 2009. Rainfall variability over mountainous and adjacent lake areas: the case of Lake Tana basin at the source of the Blue Nile River. *J. Appl. Meteorol. Climatol.* 48, 1696–1717.
- He, J.J., Yu, Y., Yu, L.J., Liu, N., Zhao, S.P., 2017. Impacts of uncertainty in land surface information on simulated surface temperature and precipitation over China. *Int. J. Climatol.* 37, 829–847.
- Jain, S., Panda, J., Rath, S.S., Devara, P.C., 2017. Evaluating land surface models in WRF simulations over DMIC region. *Indian J. Sci. Technol.* 10, 1–24.
- Jiménez, P.A., Dudhia, J., González-Rouco, J.F., Navarro, J., Montávez, J.P., García-Bustamante, E., 2012. A revised scheme for the WRF surface layer formulation. *Mon. Weather Rev.* 140, 898–918.
- Kerandi, N.M., Laux, P., Arnault, J., Kunstmann, H., 2017. Performance of the WRF model to simulate the seasonal and interannual variability of hydrometeorological variables in East Africa: a case study for the Tana River basin in Kenya. *Theor. Appl. Climatol.* 130, 401–418.
- Kim, Y., Sarlet, K., Raut, J.C., Chazette, P., 2015. Influence of an urban canopy model and PBL schemes on vertical mixing for air quality modeling over Greater Paris. *Atmos. Environ.* 107, 289–306.
- Kindu, M., Schneider, T., Teketay, D., Knoke, T., 2013. Land use/land cover change analysis using object-based classification approach in Munessa-Shashemene landscape of the Ethiopian highlands. *Remote Sens.* 5, 2411–2435.
- Lai, A., Liu, Y., Chen, X., Chang, M., Fan, Q., Chan, P., Wang, X., Dai, J., 2016. Impact of land-use change on atmospheric environment using refined land surface properties in the Pearl River Delta, China. *Adv. Meteorol.* 2016, 1–15.
- Lee, C.B., Kim, J.C., Belorid, M., Zhao, P., 2016. Performance evaluation of four different land surface models in WRF. *Asian J. Atmos. Environ.* 10, 42–50.
- Li, X., Mitra, C., Dong, L., Yang, Q., 2018. Understanding land use change impacts on microclimate using Weather Research and Forecasting (WRF) model. *Phys. Chem. Earth* 103, 115–126.
- Mooney, P.A., Mulligan, F.J., Broderick, C., 2016. Diurnal cycle of precipitation over the British Isles in a 0.44° WRF multiphysics regional climate ensemble over the period 1990–1995. *Clim. Dyn.* 47, 3281–3300.
- Mugume, I., Waiswa, D., Mesquita, M., Reuder, J., Basalirwa, C., Bamutaze, Y., Twinomuhangi, R., Tumwine, F., Otim, J.S., Ngailo, T.J., Ayesiga, G., 2017. Assessing the performance of WRF model in simulating rainfall over western Uganda. *J. Climatol. Weather Forecast.* 5, 1–9.
- Muller, O.V., Berbery, E.H., Alcaraz-Segura, D., Ek, M.B., 2014. V regional model simulations of the 208 drought in southern South America using a consistent set of land surface properties. *J. Clim.* 27, 6754–6778.
- Osuri, K.K., Mohanty, U., Routray, A., Niyogi, D., 2015. Improved prediction of Bay of Bengal Tropical cyclones through assimilation of Doppler weather radar observations. *Mon. Weather Rev.* 143, 4533–4560.
- Pennelly, C., Reuter, G., Flesch, T., 2014. Verification of the WRF model for simulating heavy precipitation in Alberta. *Atmos. Res.* 135–136, 172–192.
- Pohl, B., Crétat, J., Camberlin, P., 2011. Testing WRF capability in simulating the atmospheric water cycle over Equatorial East Africa. *Clim. Dyn.* 37, 1357–1379.
- Pongratz, J., Reick, C.H., Raddatz, T., Claussen, M., 2010. Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change. *Geophys. Res. Lett.* 37, 1–5.
- Puliafito, S.E., Allende, D.G., Mulena, C.G., Cremades, P., Lakkis, S.G., 2015. Evaluation of the WRF model configuration for Zonda wind events in a complex terrain. *Atmos. Res.* 166, 24–32.
- Ramarohetra, J., Pohl, B., Sultan, B., 2015. Errors and uncertainties introduced by a regional climate model in climate impact assessments: example of crop yield simulations in West Africa. *Environ. Res. Lett.* 10, 1–16.
- Roth, V., Lemann, T., Zeleke, G., Subhatu, A.T., Nigussie, T.K., Hurni, H., 2018. Effects of climate change on water resources in the upper Blue Nile Basin of Ethiopia. *Heliyon* 4, e00771.
- Schirmer, M., Jamieson, B., 2015. Verification of analyzed and forecasted winter precipitation in complex terrain. *Cryosphere* 9, 587–601.
- Sertel, E., Robock, A., Ormeci, C., 2009. Impacts of land cover data quality on regional climate simulations. *Int. J. Climatol.* 123, 162–173.
- Sisay, K., Thurnher, C., Hasenauer, H., 2017. Daily climate data for the Amhara region in Northwestern Ethiopia. *Int. J. Climatol.* 37, 2797–2808.
- Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.Y., Wang, W., Power, J.G., 2008. A Description of the Advanced Research WRF Version 3, NCAR/TN-475+STR NCAR TECHNICAL NOTE.
- Smirnova, T.G., Brown, J.M., Benkamin, S.G., Kenyon, J.S., 2016. Modifications to the Rapid update cycle land surface model (RUC LSM) available in the weather research and forecasting (WRF) model. *Am. Meteorol. Soc.* 144, 1851–1865.
- Tabari, H., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Saeed, S., Brisson, E., Van Lipzig, N., Willems, P., 2016. Local impact analysis of climate change on precipitation

- extremes: are high-resolution climate models needed for realistic simulations? *Hydrol. Earth Syst. Sci.* 20, 3843–3857.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* 106, 7183–7192.
- Teferi, E., Uhlenbrook, S., Bewket, W., Wenninger, J., Simane, B., 2010. The use of remote sensing to quantify wetland loss in the Choke Mountain range, Upper Blue Nile basin, Ethiopia. *Hydrol. Earth Syst. Sci.* 14, 2415–2428.
- Teklay, A., Dile, Y.T., Setegn, S.G., Demissie, S.S., Asfaw, D.H., 2018. Evaluation of static and dynamic land use data for watershed hydrologic process simulation: a case study in Gummara watershed, Ethiopia. *Catena* 172, 65–75.
- Thomas, L., Dash, S.K., Mohanty, U.C., 2014. Influence of various land surface parameterization schemes on the simulation of western disturbances. *Meteorol. Appl.* 21, 635–643.
- Thompson, G., Paul, R.F., Roy, M.R., William, D.H., 2008. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization. *Mon. Weather Rev.* 136, 5095–5115.
- Tigabu, T.B., Hörmann, G., Wagner, P.D., Fohrer, N., 2018. Statistical analysis of rainfall and streamflow time series in the Lake Tana Basin, Ethiopia. *J. Water Clim. Change* 2018, 1–17.
- Yang, J., Duan, K., 2016. Effects of initial drivers and land use on WRF modeling for near-surface fields and atmospheric boundary layer over the Northeastern Tibetan Plateau. *Adv. Meteorol.* 2016, 1–16.
- Yhang, Y., Sohn, S.J., Jung, I.W., 2017. Application of dynamical and statistical downscaling to East Asian summer precipitation for finely resolved datasets. *Adv. Meteorol.* 2017, 1–9.
- Zeyaeyan, S., Fattahi, E., Ranjbar, A., Azadi, M., Vazifedoust, M., 2017. Evaluating the effect of physics schemes in WRF simulations of summer rainfall in north west Iran. *Climate* 5, 1–17.
- Zu-heng, H.U., Zhong-feng, X.U., Ning-Fang, Z., Zhu-Guo, M., Guo-Ping, L., 2014. Evaluation of the WRF model with different land surface schemes: a drought event simulation in southwest China during 2009-10. *Atmos. Oceanogr. Sci. Libr.* 7, 168–173.