Scientific Research Publishing

Assessing the Suitability of the WorldClim Dataset for Ecological Studies in Southern Kenya

Tim J. L. Wango¹*, Douglas Musiega¹, Charles N. Mundia²

¹Department of Geomatic Engineering & Geospatial Information Systems, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya
²Department of Geomatics and Geospatial Information Systems, Dedan Kimathi University, Nyeri, Kenya
Email: *timjlwango@gmail.com

Abstract

There have been numerous efforts to generate freely available climatic datasets for use in species distribution models, the most popular being the global climatic dataset known as WorldClim. The availability of such datasets is invaluable to scientists as many studies are performed in remote areas where no weather stations are found. However, many users do not critically assess the suitability of these datasets for their applications, and errors associated with global datasets are often assumed to be negligible. Understanding what a climate dataset can or cannot deliver requires the user to have a working knowledge of what the basic spatial climate-forcing factors are at the scale of his/her study, and to have a good understanding of the uncertainty in the dataset. In geographic studies, uncertainty is often described by the degree of error (uncertainty), or degree of accuracy (certainty) in data, and thematic uncertainty refers to the uncertainty in measures made for each variable, whereas temporal uncertainty refers to the uncertainty in time period represented by each variable. Here, we used climatic data from weather stations to investigate the climate-forcing factors in southern Kenya, and then used this weather station data to investigate the uncertainty in the WorldClim dataset. Results indicated that the nineteen core Worldclim variables, known as bioclimatic variables, accurately depicted the local variations in climate in the study area. However, whereas the monthly and seasonal temperature variables represented the same time period in different locations, the same was not true for the monthly and seasonal precipitation variables. The onset of rains is a key biological indicator, and scientists studying phenomena tied to the onset of rains need to keep in mind the temporal variations represented in the WorldClim dataset.

1. Introduction

The idea that variations in climate exert a strong influence on organisms' distribution is centuries old [1] [2] [3]. Climate has a direct influence on vegetation distribution as plants are unable to evade adverse climate by sheltering or migration, and are limited to areas with suitable climate [4] [5]. Animals respond to climate directly when they actively habituate specific climatic zones, or indirectly where an animal's distribution positively correlates to vegetation that is only found within specific climatic zones [4]. Whereas literature on the effects of climate on organism's distribution dates back to the 5th century, methods for generating climatic grid surfaces for use in species distribution models (SDMs) were first published in the mid-80s [3]. During this period advances in computer science enabled the creation of specialist tools [6] to study species distribution based on geocoded species distribution data and interpolated environmental variables [3]. Though not often described as such, these early specialist tools were Geographic Information Systems (GIS), and many are now found as stand-alone software (e.g. ANUSPLIN), as extensions for popular GIS software packages (e.g. QGIS), or as spatial analysis tools in popular statistical software (e.g. CRAN R). The BIOCLIM package, conceived by Henry Nix [3] [7] [8], generates organisms' distribution maps by delineating areas suitable for habituation based on sampled species distribution data and interpolated climatic variables.

There have been numerous efforts to generate freely available global and continental climatic datasets for use in SDMs [7] [8] [9]. Many of these datasets are based on the climatic variables conceived by Nix [3], the most popular being the global climatic dataset known as WorldClim [10] [11]. This dataset is free for download on the Internet (http://www.worldclim.org/), and the variables represented in the WorldClim dataset, known as bioclimatic variables, have been used in various ecological studies [12] [13] [14]. The availability of datasets such as WorldClim is invaluable to biological studies in areas where climatic records are sparse or non-existent. However, many users do not critically assess the suitability of these datasets for their applications [15], and errors associated with global datasets are often assumed to be negligible. Whereas in the past large datasets were produced by governments or large organizations with well laid out data standards [16], in recent years the Internet has enabled many non-traditional actors to publish and widely distribute spatial data, and the topic of uncertainty in these datasets has begun gaining prominence [17] [18]. Data quality is a very important factor in the processing of the spatial data, and the data needs of users vary from person to person, organization to organization, or

from application to application. Because of this it is the ultimate responsibility of the user to check the quality and suitability of data for their specific application. Spatial data which is suitable for one application need not be necessarily suitable for another [19]. Even in cases where data standards have been published, understanding what a climate dataset can or cannot deliver requires the user to have a working knowledge of what the basic spatial climate-forcing factors are at the scale of his/her study in order to correctly interpret trends and uncertainty in the data [15].

In GIS, uncertainty is often described by the degree of error (uncertainty), or degree of accuracy (certainty) in data [6] [17] [19]. Errors are inevitable in any in any spatial data [19], and should be recognized as an inherent part of any spatial dataset. Errors may creep in at any stage of data acquisition, transformation, and/or analysis [19]. Furthermore, GIS often deals with different layers of spatial data from numerous sources, collected using different sampling technics, geocoded using different map projections, and composed at different scales by the original authors. The combination of different data is one of the strengths of GIS [19], but for this very reason GIS models often have complex errors propagated from data collection through to analysis [19], and species distribution models are no exception. The correct conceptualization of uncertainties in individual data as well as the eventual propagation of uncertainties when the data is combined for used in SDMs is a growing consideration when presenting results of modeling [19].

Uncertainty has been the subject of much research in GIS and Remote Sensing, and is recognized as one of the priority areas in GIS research [17]. Various authors have looked at uncertainty as a result of processes used in generating spatial data [6] [17] [19], propagation of errors and uncertainty in processes used in spatial analysis [18] [19] [20], and uncertainty in the results of spatial models [21] [22]. Daly [15], in a review on the assessment of climatic datasets, discusses climatic datasets available to scientists and highlights the importance of familiarization with climate-forcing factors in one's study area, familiarization with methods of spatial interpolation used to generate climatic data for one's study area, and familiarization with errors inherent in the generated data.

Here, we first asses the relationship between climate-forcing factors and the climate of southern Kenya by assessing the relationship between the distance of a weather station from the ocean, its altitude, and the climate recorded. It has long been reported that a correlation exists between the altitude of a station, the rainfall and temperatures recorded [15] [23] [24]. Reports also indicate a correlation between a weather station's distance from large water bodied and the rainfall recorded [15] [25]. Second, we assess the thematic uncertainty, and finally the temporal uncertainty in the WorldClim data for southern Kenya. Thematic uncertainty refers to the uncertainty in measures made for each bioclimatic variable, whereas temporal uncertainty refers to the uncertainty in time period represented by each bioclimatic variable. Thematic uncertainty is normally

represented by one or more statistical measures of error or accuracy such as the Standard Error of Estimates (SEE) [6] [17]. Whereas thematic uncertainty has received a lot of attention in literature [17], uncertainty in temporal representation of data, a key aspect when considering climatic data, has not [16]. Time is an important ecological consideration, and the onset of different seasons is often tied to environmental conditions that affect availability of food. In semi-arid areas such as Amboseli National Park in Kenya stress levels of baboons have been are known to vary with seasons which occur in different times of the year [26] [27]. Migration of wildebeest also seasonal [28], and migration routes are associated with different times of the year. Time is often not dealt with explicitly in geospatial databases, and temporal information is often omitted except in databases designed explicitly for historical or timelapse/time-series studies [16].

2. Methodology

Kenya, situated approximately between latitudes 4°21'N and 4°28'S and between longitudes 34° and 42°E, shares her boarders with Tanzania, South Sudan, Somalia, Ethiopia and Uganda, with the Indian Ocean on the south-eastern edge (Figure 1). The country lies within the Inter Tropical Convergence Zone (ITCZ), which determines the main features of annual rainfall and its seasonal variations [29] [30] [31]. Because of the local influence from the different geographic features, Kenya has diverse bioclimatic zones that include tropical, temperate, arid, and desert. The Rift Valley, a low-lying area characterized by arid climate, is found between highlands characterized by tropical and temperate climate. Areas near large water bodies such as Lake Victoria and the Indian Ocean exhibit tropical monsoon climate. The country also has large areas that can be described as semi-arid, arid, and desert. The area adjacent to and south of the Equator, referred to here as southern Kenya, has numerous sites of interest to ecologists, and these include the Amboseli National Park, The Maasai Mara National Park, and Nairobi National Park among others. Many of these sites do not have weather stations within or nearby, and this means the best estimates for the climate of these areas are from interpolated data such as the WorldClim dataset.

To investigate the correlation between climate and the geography of southern Kenya, climatic data consisting of monthly precipitation, average monthly minimum temperature, and average monthly maximum temperature for the years 2001 to 2012 were collected from 11 weather stations (Mombasa, Malindi, Voi, Makindu, Machakos, Narok, Nakuru, Meru, Laikipia, Garissa and Kisumu), all monitored by the Kenya Meteorological Department (KMD) (Figure 1). Following the description by O'donnel and Ignizio [7], nineteen core bioclimatic variables were calculated for each weather station. A raster with grid values determined by the distance of each grid cell from the Kenyan coastline was generated using the Distance tool in ArcGIS 9.2 (ESRI, USA). Altitude data were downloaded from the United States Geological Survey (USGS) website
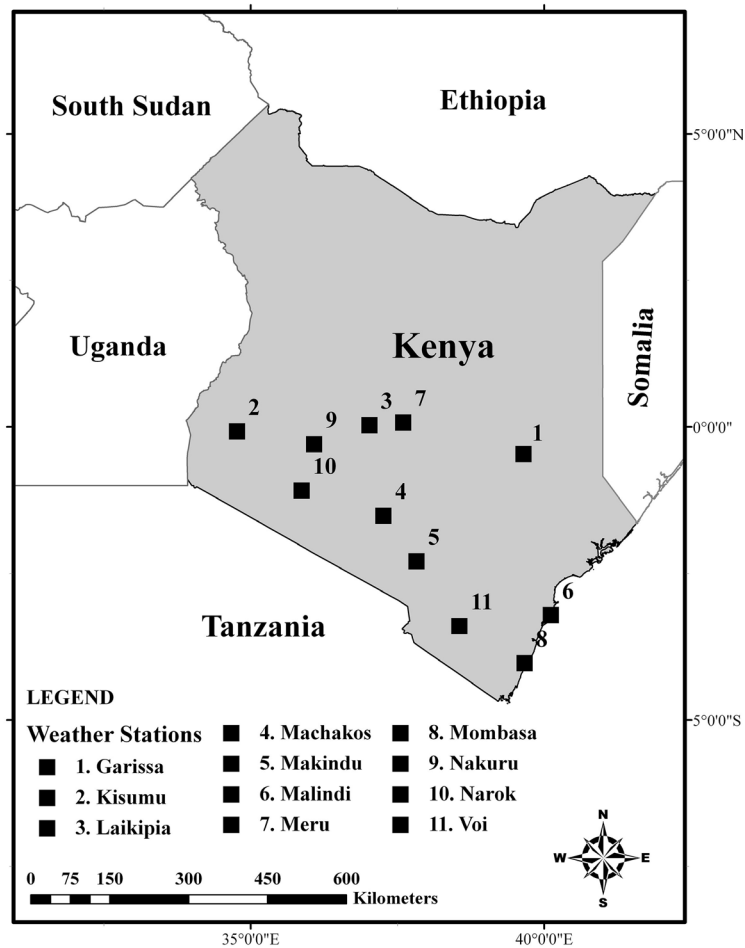
Figure 1. The map shows the distribution of Kenya Meteorological Weather Stations in southern part of the country.

(https://earthexplorer.usgs.gov/ accessed 25th July 2017) and imported into ArcGIS 9.2 as a raster dataset. By identifying the grid value where a weather station is located, the altitude of each weather station (ALT) and distance from the ocean of each weather station (DIST) were determined. Scatter analysis was then used to assess the relationship between the bioclimatic variables and the geography of each weather station (ALT and DIST).

The bioclimatic variables envisaged by Nix [3] highlight average annual climatic patterns, extreme climatic monthly patterns, as well as extreme climatic seasonal patterns. The maximum temperature of the hottest month (BIO5), minimum temperature of the coldest month (BIO6), precipitation of the wettest month (BIO13), and precipitation of the driest month (BIO14) represent monthly climatic extremes. Seasonal bioclimatic variables described by Hijmans *et al.* [10] are calculated on a quarterly basis (*i.e.* the coldest three consecutive months, the hottest three consecutive months, the wettest three consecutive months, and the driest three consecutive months). The mean temperature of the wettest quarter (BIO8), mean temperature of the driest quarter (BIO9), mean temperature of the warmest quarter (BIO10), mean temperature of the coldest

quarter (BIO11), precipitation of the wettest quarter (BIO16), precipitation of the driest quarter (BIO17), precipitation of the warmest quarter (BIO18) and precipitation of the coldest quarter (BIO19) represent seasonal climatic extremes.

To investigate the temporal uncertainty in the WorldClim dataset, we focused on generating time charts of the respective monthly and seasonal bioclimatic variables. Literature indicates that Kenya experiences two wet seasons, with precipitation experienced during the "long rains" (March-May) and also during the "short rains" (October-December), with two dry seasons in between [29], [32]. To highlight the distribution of climatic extremes experienced at each weather station and provide a comparison with patterns in literature, the time charts were based on a hydrological calendar which starts in October and ends in September.

To investigate thematic uncertainty in the WorldClim dataset, data from WorldClim version 2.0 [11] consisting of nineteen core bioclimatic variables were downloaded from the WorldClim repository (worldclim.org), and the score of each bioclimatic variable at each respective weather station was noted. As WorldClim 2.0 is generated from data collected between the years 1970 and 2000 [11], we used the bioclimatic variables calculated from each KMD weather station (years 2001 to 2012) as test data and calculated the Standard Error of Estimates, and the correlation (R^2) between the WorldClim dataset and the KMD data.

3. Analysis

3.1. Analysis of Climatic Patterns in Southern Kenya

Analysis indicated that temperatures in southern Kenya were mainly influenced by the altitude of an area, with the average annual temperature (BIO1), maximum temperature of the warmest month (BIO5), minimum temperature of the coldest month (BIO6), mean temperature of the wettest quarter (BIO8), mean temperature of the driest quarter (BIO9), mean temperature of the warmest quarter (BIO10), and mean temperature of the coldest quarter (BIO11) all showing strong correlation with altitude (all $R^2 > 0.9$) The maximum temperature of warmest month (BIO5) also showed strong correlation with the altitude ($R^2 = 0.7$) (Figure 2). Mean diurnal temperature range (BIO2), isothermality (BIO3), temperature seasonality (BIO4), and temperature annual range (BIO7) all showed strong correlation with the distance from the ocean (all $R^2 > 0.6$). Analysis indicated that the distance of a station from the ocean had an influence on the precipitation received; however the influence was not uniform and showed patterns interpreted to be the interplay between convectional and relief rainfall (Figure 3).

Annual total precipitation (BIO12) showed good correlation to the distance of a station to the ocean ($R^2 = 0.54$). The precipitation patterns indicated that the areas closer to the ocean received good precipitation and this dropped as one moved inland, and this can be interpreted as the influence of convectional rain-

fall. After about 300 km, the precipitation received picked up and slowly increased as one moves further inland, and this can be interpreted as the influence of relief rainfall. Precipitation of Driest Month (BIO14), Precipitation of Wettest Quarter (BIO16), Precipitation of Driest Quarter (BIO17), Precipitation of Coldest Quarter (BIO19) all showed similar patterns. Precipitation of Warmest Quarter (BIO18) exhibiting a linear pattern, with precipitation steadily rising as one moved away from the coastline.

The precipitation seasonality (BIO15) showed good correlation with the distance from the ocean ($R^2 = 0.7$), indicating that the further inland one moves, the smaller the variation between monthly precipitation recorded at weather stations. Notably, the precipitation of the wettest month (BIO13) showed poor correlation with altitude ($R^2 = 0.14$), and with the distance from the ocean ($R^2 = 0.3$).
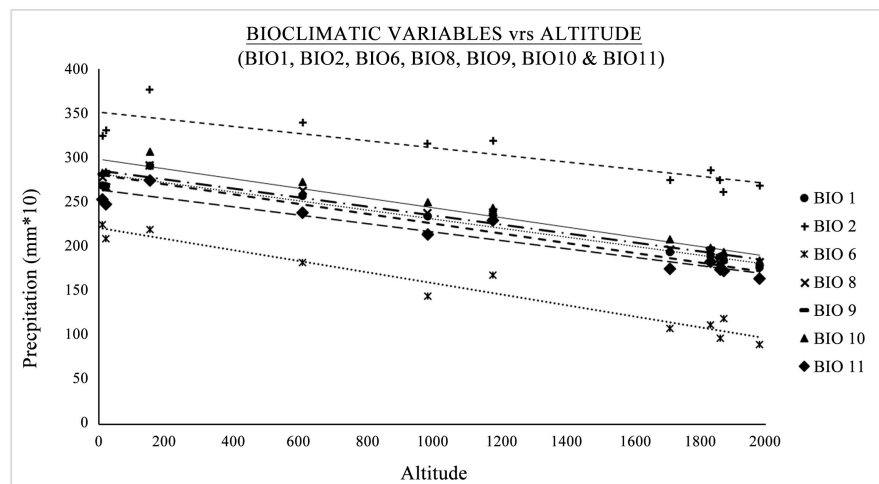

Figure 2. The scatter plots show the influence of the altitude on the temperatures experienced when calculation BIO1, BIO5, BIO6, BIO8, BIO19, BIO10 and BIO11.
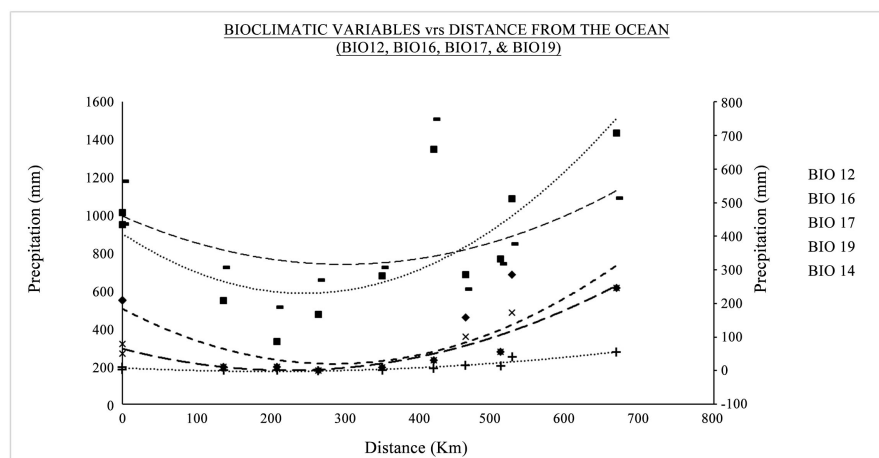

Figure 3. The scatter plot shows the influence of the distance from the ocean on precipitation. BIO12 is plotted against the primary vertical axis whereas BIO14, BIO16, BIO17 and BIO19 are plotted against the secondary vertical axis.

3.2. Analysis of Thematic Uncertainty in the WorldClim Dataset

The WorldClim dataset showed strong correlation to the KMD dataset with sixteen variables scoring correlation values ranging from 0.9 - 0.99. The remaining three variables scored correlation values ranging from 0.85 to 0.89. Error analysis indicated that the WorldClim dataset estimated climatic conditions well, with low SEE values for all nineteen variables (Table 1). Cross-correlation analysis between WorldClim variables indicated that there was strong correlation between variables within the study area, more so for the first eleven bioclimatic temperature variables (BIO1 to BIO11). There was good cross-correlation between the bioclimatic precipitation variables (BIO12 to BIO19) (Figure 4).

3.3. Analysis of Temporal Uncertainty

Analysis of monthly data indicated that five stations received the most precipitation in November (Voi, Garissa, Makindu, Machakos, and Meru), four received the most precipitation in April (Laikipia, Narok, Nakuru, and Kisumu), and two received the most precipitation in May (Mombasa and Malindi) (Figure 5).

Analysis indicated that six stations received the least precipitation in February (Malindi, Mombasa, Garissa, Laikipia, Nakuru, and Kisumu), one received the least precipitation in July (Meru), and the remaining four received the least precipitation in July (Voi, Makindu, Machakos, and Narok). Three stations experienced their coldest month in January (Maru, Laikipia, and Nakuru), six experienced their coldest month in July (Mombasa, Garissa, Makindu, Machakos, Narok, and Kisumu), with one experiencing it's coldest month in August (Voi) and one in September (Malindi). All weather stations experienced their hottest month between February and March.

Analysis of the seasonal data indicated that five weather stations experienced their wettest quarter during the "short rains" between October and January (Voi, Garissa, Makindu, Machakos,v and Meru), with the rest experiencing their wettest quarter during the "long rains" between March and June (Malindi, Mombasa, Laikipia, Narok, Nakuru, Kisumu). Mombasa, Malindi, Laikipia, Nakuru and Kisumu experienced their driest quarter between December and February. Voi, Garissa, Makindu, Machakos, Meru and Narok experienced their driest quarter between June and September. All stations experienced their hottest season between January and April, and all stations experienced their coolest season between May and September (Figure 6).

4. Discussion

4.1. Climate Forcing Factors in Southern Kenya

Correlation analysis between test KMD data and independent variables (DIST = Distance from the Ocean and ALT = Altitude) indicated that bioclimatic variables representing temperature (BIO1 to BIO11) showed consistently strong correlation with ALT (all $R^2 > 9$), and this was an indication that the altitude is a climate forcing factor to consider when assessing the climatic data. Generally,

Table 1. BioClimatic Variables.

Code	Description	Standard Error of Estimates	Correlation Coefficient
BIO1	Annual Mean Temperature	3.41	1.00
BIO2	Mean Diurnal Range	5.81	0.93
BIO3	Isothermality	2.39	0.90
BIO4	Temperature Seasonality	21.01	0.98
BIO5	Max Temperature of Warmest Month	2.00	0.99
BIO6	Min Temperature of Coldest Month	5.16	0.98
BIO7	Temperature Annual Range	4.06	0.94
BIO8	Mean Temperature of Wettest Quarter	3.58	0.99
BIO9	Mean Temperature of Driest Quarter	7.93	0.91
BIO10	Mean Temperature of Warmest Quarter	3.55	1.00
BIO11	Mean Temperature of Coldest Quarter	3.48	0.99
BIO12	Annual Precipitation	45.15	0.94
BIO13	Precipitation of Wettest Month	11.28	0.91
BIO14	Precipitation of Driest Month	3.21	0.86
BIO15	Precipitation Seasonality	2.56	0.96
BIO16	Precipitation of Wettest Quarter	23.19	0.96
BIO17	Precipitation of Driest Quarter	9.09	0.93
BIO18	Precipitation of Warmest Quarter	24.48	0.89
BIO19	Precipitation of Coldest Quarter	9.51	0.97

The table shows the code of each bioclimatic variable and gives a short description of the information captured by each variable. The table also shows the Standard Error of Estimates calculated for the WorldClim data using Kenya Meteorological Department data from the years 2001-2012 as test data, and the correlation (R^2) between the WorldClim data and the Kenya Meteorological Department data.
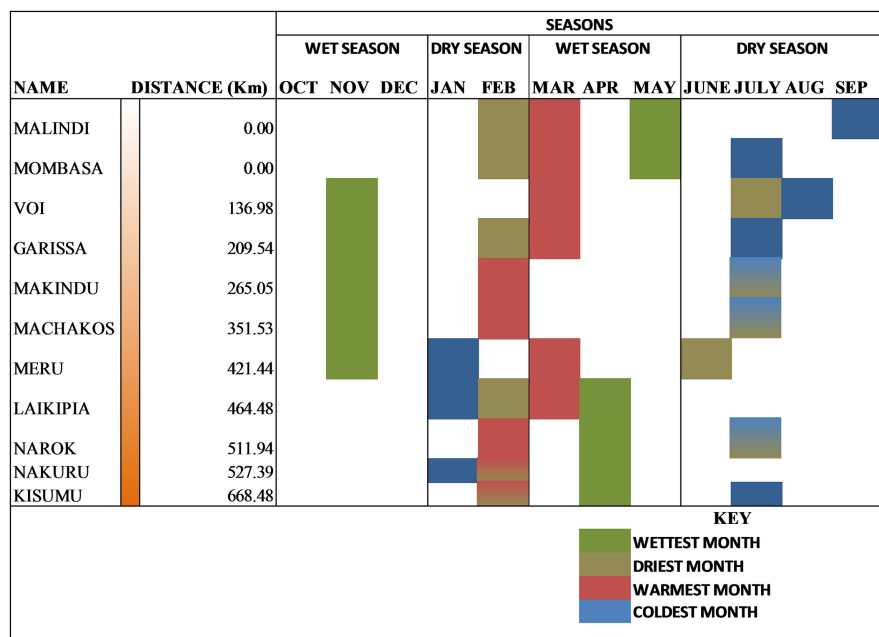

Figure 4. The chart indicates the spatial distribution of bioclimatic variables based on monthly climatic extremes.
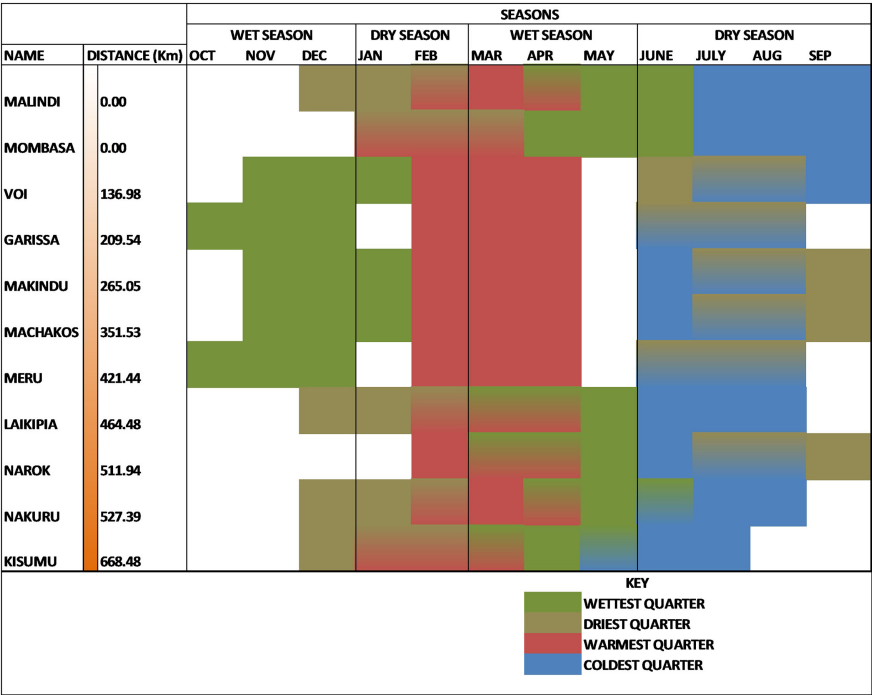
NAME	DISTANCE (Km)	SEASONS											
		WET SEASON			DRY SEASON		WET SEASON			DRY SEASON			
		OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUNE	JULY	AUG	SEP
MALINDI	0.00												
MOMBASA	0.00												
VOI	136.98												
GARISSA	209.54												
MAKINDU	265.05												
MACHAKOS	351.53												
MERU	421.44												
LAIKIPIA	464.48												
NAROK	511.94												
NAKURU	527.39												
KISUMU	668.48												

KEY
WETTEST QUARTER
DRIEST QUARTER
WARMEST QUARTER
COLDEST QUARTER

Figure 5. The chart indicates the spatial distribution of bioclimatic variables based on quarterly climatic extremes.
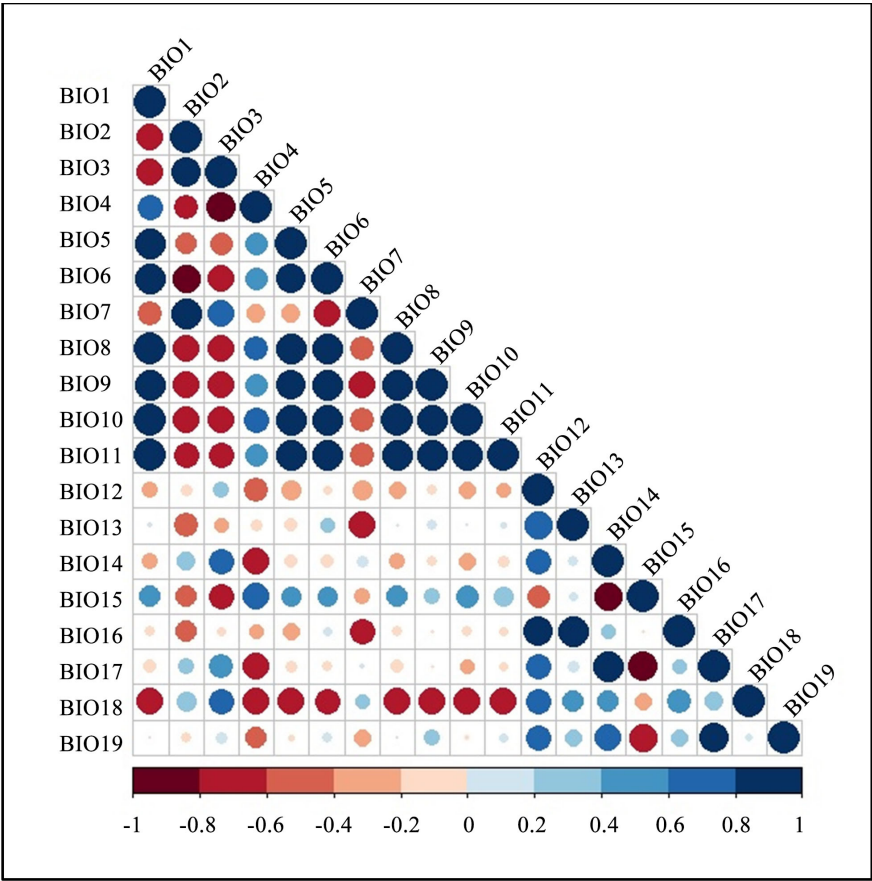

Figure 6. The chart shows strong cross-correlation between bioclimatic variables.

the higher the altitude the cooler an area will be, and this coincided with other scientific studies [15]. Seven bioclimatic variables representing precipitation (BIO12, BIO14, BIO15, BIO16, BIO17, BIO18 to BIO19) showed good correlation with DIST (all R^2 ~0.5 to 0.7), with one bioclimatic variable showing poor correlation with either ALT or DIST (BIO13). For BIO13, the low correlation value indicates that local variation in precipitation patterns measured by this variable are complex, and were not fully captured by the KMD weather stations sampled. Generally, areas closer to the ocean experienced high rainfall, and this steadily reduced as one moved inland. After 300 kilometers the precipitation steadily picked up and continued to steadily rise. These results indicated that the precipitation followed a global trend, and was not strongly influenced by local variations in altitude in the study area.

4.2. Thematic Uncertainty

It is often the case that where wildlife activity is vibrant there are no active weather stations, and this presents a challenge to ecologists studying in these areas. Southern Kenya, for example, has a sparse network of weather stations, most of which are located in irregularly spaced towns across the region. This presents a challenge for ecologists working far from town where the weather stations are located. For these scientists, interpolation is one option pursued in order to get estimation of climatic conditions in their study sites. However, many scientists lack the requisite skills to perform interpolation. Compounding this problem, when spatial data are sparse the assumptions made about the underlying variations that have been sampled and the choice of method to be used for interpolation and its parameters can be critical if one is to avoid misleading results (e.g. Burrow and McDonnell 1998). For this reason, grid climatic data found on the Internet offer an attractive alternative. Normally, the first concern is if the downloaded data gives close estimates to the conditions found on the ground, *i.e.* is the thematic uncertainty low. Results of analysis indicate that the WorldClim dataset gave close estimates of the average climatic conditions for the years 2001-2012, with low standard errors for all variables measured. This does indicate that the WorldClim dataset can be used to estimate climatic conditions in areas where there are no weather records within southern Kenya.

4.3. Temporal Uncertainty

Seasonal bioclimatic variables described by Hijmans *et al.* [10] are calculated on a quarterly basis, and this fits well with seasonal patterns experienced in temperate climates where four distinct seasons (winter, summer, spring and fall) are well defined. Kenya experiences two wet seasons with two dry seasons in between, and analysis of the seasonal quarters as described in the WorldClim data (hottest quarter, coldest quarter, wettest quarter, and driest quarter) showed great temporal variation in seasonal precipitation patterns. In contrast, analysis showed stable seasonality in temperature patterns. Temporal patterns observed

in the wettest and driest quarters implied three temporal zones. Areas close to the Indian Ocean (Mombasa, Malindi, and Voi) received the most precipitation (wettest quarter) during April, May and June, the period commonly referred to as the "long rains". Areas furthest from the Indian Ocean (Laikipia, Narok, Nakuru, and Kisumu) also received the most precipitation during March, April, May, and June. The areas in between (Garissa, Makindu, Machakos, and Meru) received the most precipitation between October and January, the period referred to as the "short rains". The temporal distribution of the driest quarter also showed a mixed pattern with Mombasa and Malindi, the stations closest to the Indian Ocean, Nakuru and Kisumu, stations furthest from the Indian Ocean, experiencing their driest quarter between December and March. Voi, Garissa, Makindu, Machakos, and Meru experienced their driest season between June and September. Laikipia and Narok seeming out of place when compared to the stations closest to them. This implies that any reference to the "wettest season" describes one thematic phenomena that occurs at different times in different regions of the country. Similarly, references to the "driest season" would have spatio-temporal implications. For example, in Nakuru the "driest season" would refer to the December-February season, whereas in Narok the "driest seasons" is experienced in July-September. Camberlin *et al.* [30] noted that the onset of the rainy season varied locally, and that where there was spatial coherence in precipitation events there was none in the intensity of the precipitation experienced, and analysis in this work support these findings.

The period described as the "hottest quarter" spanned four months throughout the country, and was experienced between January and April. The period described as the coldest quarter was experienced between June and September in all except one weather station (Kisumu). This indicates that the temporal pattern in temperatures observed at each weather station was distinct and uniform across the country. Temporal patterns exhibited in the monthly data (hottest month, coldest month, wettest month, and driest month) also showed complex climatic patterns, more so in precipitation experienced across the country. All stations experienced their hottest month between February and March. Three stations (Meru, Narok, and Nakuru) experienced their coldest month in January, and this fell out of the period described as the "coldest quarter" in this work as well as in literature [32]. The rest of the stations experiencing their coldest month between July and September, and this coincided with the seasonal analysis in this work and in literature. Voi, Garissa, Machakos, Makindu, and Meru experienced their wettest month in November, with the rest experiencing their wettest month between April and May. Malindi, Mombasa, Garissa, Laikipia, Nakuru, and Kisumu experienced their driest month in February, with the rest experiencing their driest month between June and July. As with the seasonal analysis, the monthly analysis showed that the precipitation patterns were complex, and that different areas experienced heavy precipitation during different months as observed by Chamberlin *et al.* [30]. Analysis of the coldest month in-

dicated that three stations (Meru, Laikipia and Nakuru) experienced their coldest month outside the coldest quarter (or cold season), and it can be assumed that the local geography had a strong influence on the temperatures experienced at these three stations in January.

5. Conclusion

In summary, we can conclude that the variations in climate observed in southern Kenya were influenced by the altitude the distance from the Indian Ocean, and that the influence of these climate forcing factors was well captured in the WorldClim dataset. Whereas the patterns in temperature, mainly influenced by altitude, showed strong local variation, patterns in precipitation showed global trends and were mainly influenced by the distance of an area from the ocean. The study concluded that the WorldClim dataset closely estimated average climatic conditions found in southern Kenya, and that thematic uncertainty was not a major concern when using the WorldClim dataset. However, the temporal uncertainty in the dataset, more so for bioclimatic variables that measure different aspects of precipitation, would be a concern for some. Scientists hoping to use the WorldClim dataset for species distribution models should carefully consider both the temporal characteristics of their study species and the temporal uncertainty of the WorldClim dataset before using the WorldClim dataset in their respective studies. Activities that are closely coupled with seasonality in organisms include migration and dispersal, and these activities often have an influence on the distribution of different species. Migration, describes as a seasonal long distance movement of individual or groups, is often tied to seasonal triggers. Dispersal in animals is described as the movement from one area to a breeding site, and is often tied to seasonal triggers. In plants seed dispersal and germination are closely related to seasonality, and only occur when climatic conditions are favorable. When studying a species that is found in a wide geographic area, the different patterns observed in precipitation patterns indicate that key behavior that is triggered by seasonal changes may occur at different times in different areas, and this may not be clear when using the WorldClim dataset.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Yan, H., Nix, H.A., Hutchinson, M.F. and Booth, T.H. (2005) Spatial Interpolation of Monthly Mean Climate Data for China. *International Journal of Climatology*, **25**, 1369-1379. https://doi.org/10.1002/joc.1187

[2] Broitman, A.B.R., Szathmary, P.L., Mislan, K.A.S. and Blanchette, C.A. (2018) Nordic Society Oikos Predator-Prey Interactions under Climate Change: The Importance of Habitat vs. Body Temperature. http://www.jstor.org/stable/40235630

[3] Booth, T.H., Nix, H.A., Busby, J.R. and Hutchinson, M.F. (2014) Bioclim: The First Species Distribution Modelling Package, Its Early Applications and Relevance to Most Current MaxEnt Studies. *Diversity and Distributions*, **20**, 1-9. https://doi.org/10.1111/ddi.12144

[4] Hirzel, A.H. and Le Lay, G. (2008) Habitat Suitability Modelling and Niche Theory. *Journal of Applied Ecology*, **45**, 1372-1381. https://doi.org/10.1111/j.1365-2664.2008.01524.x

[5] Guisan, A., Guisan, A., Zimmermann, N.E. and Zimmermann, N.E. (2000) Predictive Habitat Distribution Models in Ecology. *Ecological Modelling*, **135**, 147-186. https://doi.org/10.1016/S0304-3800(00)00354-9

[6] Longley, P.A., Goodchi Ld., M.F., Maguire, D.J. and Rhind, D.W. (2011) Geographical Information Systems and Science. John Wiley & Sons, Hoboken.

[7] O'Donnell, M.S. and Ignizio, D.A. (2012) Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States. *United States Geological Survey Data Series*, **691**, 10.

[8] Wagner, M., Trutschnig, W., Bathke, A.C. and Ruprecht, U. (2018) A First Approach to Calculate BIOCLIM Variables and Climate Zones for Antarctica. *Theoretical and Applied Climatology*, **131**, 1397-1415. https://doi.org/10.1007/s00704-017-2053-5

[9] Perry, M. and Hollis, D. (2005) The Generation of Monthly Gridded Datasets for a Range of Climatic Variables over the UK. *International Journal of Climatology*, **25**, 1041-1054. https://doi.org/10.1002/joc.1161

[10] Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. (2005) Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology*, **25**, 1965-1978. https://doi.org/10.1002/joc.1276

[11] 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**, 4302-4315. https://doi.org/10.1002/joc.5086

[12] Stone, O.M.L., Laffan, S.W., Curnoe, D. and Herries, A.I.R. (2013) The Spatial Distribution of Chacma Baboon (*Papio ursinus*) Habitat Based on an Environmental Envelope Model. *International Journal of Primatology*, **34**, 407-422. https://doi.org/10.1007/s10764-013-9669-9

[13] Winder, I.C. (2014) The Biogeography of the *Papio baboons*: A GIS-Based Analysis of Range Characteristics and Variability. *Folia Primatologica*, **85**, 292-318. https://doi.org/10.1159/000362545

[14] Makori, D., *et al.* (2017) Predicting Spatial Distribution of Key Honeybee Pests in Kenya Using Remotely Sensed and Bioclimatic Variables: Key Honeybee Pests Distribution Models. *ISPRS International Journal of Geo-Information*, **6**, 66. https://doi.org/10.3390/ijgi6030066

[15] Daly, C. (2006) Guidelines for Assessing the Suitability of Spatial Climate Data Sets. *International Journal of Climatology*, **26**, 707-721. https://doi.org/10.1002/joc.1322

[16] Veregin, H. (1999) Data Quality Parameters. *Geographical Information Systems*, **1**, 177-189.

[17] Foody, G.M. and Atkinson, P.M. (2002) Uncertainty in Remote Sensing and GIS. John Wiley & Sons, Hoboken. https://doi.org/10.1002/0470035269

[18] Heuvelink, G.B.M. (1999) Propagation of Error in Spatial Modeling with GIS. *Geogr. Inf. Syst. Vol. 1 Princ. Tech. Issues*, **2**, 207-217.

[19] Devendran, A. and Lakshmanan, G. (2014) A Review on Accuracy and Uncertainty of Spatial Data and Analyses with Special Reference to Urban and Hydrological Modelling. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, **2**, 171-178. https://doi.org/10.5194/isprsannals-II-8-171-2014

[20] Arbia, G., Griffith, D. and Haining, R. (1998) Error Propagation Modelling in Raster Gis: Overlay Operations. *International Journal of Geographical Information Science*, **12**, 145-167. https://doi.org/10.1080/136588198241932

[21] Suriyaprasit, M. and Shrestha, D.P. (2007) Deriving Land Use and Canopy Cover Factor From Remote Sensing and Field Data in Inaccessible Mountainous Terrain for Use in Soil Erosion Modelling. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, **37**, 1747-1750.

[22] Fielding, A.H. and Bell, J.F. (1997) A Review of Methods for the Assessment of Prediction Errors in Conservation Presence/Absence Models. *Environmental Conservation*, **24**, 38-49. https://doi.org/10.1017/S0376892997000088

[23] Morton, D., De Fries, R., Shimabukuro, Y. and Hansen, M. (2005) Rapid Assessment of Annual Deforestation in the Brazilian Amazon using MODIS Data. *Earth Interactions*, **9**, 1-22. https://doi.org/10.1175/EI139.1

[24] Boone, R.B. (1997) Modeling the Climate of Maine for Use in Broad-Scale Ecological Analyses. *Northeastern Naturalist*, **4**, 213-230. https://doi.org/10.2307/3858607

[25] Hayward, D. and Clarke, R.T. (1996) Relationship between Rainfall, Altitude and Distance from the Sea in the Freetown Peninsula, Sierra Leone. *Hydrological Sciences Journal*, **41**, 377-384. https://doi.org/10.1080/02626669609491509

[26] Alberts, S.C., *et al.* (1993) Seasonality and Long-Term Change in a Savanna Environment. *Cambridge Studied in Biological and Evolutionary Anthropology*, **44**, 157.

[27] Gesquiere, L.R., *et al.* (2008) Coping with a Challenging Environment: Effects of Seasonal Variability and Reproductive Status on Glucocorticoid Concentrations of Female Baboons (*Papio cynocephalus*). *Hormones and Behavior*, **54**, 410-416. https://doi.org/10.1016/j.yhbeh.2008.04.007

[28] Musiega, D. and Kazadi, S.-N. (2004) Simulating the East African Wildebeest Migration Patterns using GIS and Remote Sensing. *African Journal of Ecology*, **42**, 355-362. https://doi.org/10.1111/j.1365-2028.2004.00538.x

[29] Camberlin, P. and Okoola, R.E. (2003) The Onset and Cessation of the "Long Rains" in Eastern Africa and Their Interannual Variability. *Theoretical and Applied Climatology*, **54**, 43-54.

[30] Camberlin, P., Moron, V., Okoola, R., Philippon, N. and Gitau, W. (2009) Components of Rainy Seasons' Variability in Equatorial East Africa: Onset, Cessation, Rainfall Frequency and Intensity. *Theoretical and Applied Climatology*, **98**, 237-249. https://doi.org/10.1007/s00704-009-0113-1

[31] Camberlin, P. and Olivier, P. (2018) Coastal Precipitation Regimes in Kenya. *Geografiska Annaler. Series A, Physical Geography*, **79**, 109-119. http://www.jstor.org/stable/521423

https://doi.org/10.1111/j.0435-3676.1997.00010.x

[32] Altmann, J.A., Alberts, S.C. and Altmann, S.A. (2002) Dramatic Chance in Local Climate Patterns in the Amboseli Basin, Kenya. *African Journal of Ecology*, **40**, 248-251. https://doi.org/10.1046/j.1365-2028.2002.00366.x