Eco-meteorological characteristics of the southern slopes of Kilimanjaro, Tanzania

Tim Appelhans, a, * Ephraim Mwangomo, a, b Insa Otte, a Florian Detsch, a Thomas Nauss, a
and Andreas Hemp b

a Environmental Informatics, Faculty of Geography, Philipps-University Marburg, Germany
b Tanzania National Parks, Arusha, Tanzania

ABSTRACT: This study introduces the set-up of a new meteorological station network on the southern slopes of Kilimanjaro, Tanzania, since 2010 and presents the recorded characteristics of air temperature, air humidity and precipitation in both a plot-based and area-wide perspectives. The station set-up follows a hierarchical approach covering an elevational as well as a land-use disturbance gradient. It consists of 52 basic stations measuring ambient air temperature and above-ground air humidity and 11 precipitation measurement sites, with recording intervals of 5 min. With respect to precipitation observations, the network extends the long-term recordings of A. Hemp who has installed and maintained up to 117 multi-month accumulating rainfall buckets in the region since 1997. The meteorological characteristics of the study region based on the derived data since 2010 are mostly in line with previous studies, although we see increased precipitation amounts at higher elevations during these years when compared with long-term means. We furthermore identify a mean annual condensation level at about 2300 m a.s.l. which has not been reported before. Finally, this is the first study to provide high resolution maps of mean monthly and mean annual temperature, humidity and precipitation for Kilimanjaro, which are of great value for geographically oriented meteorological or ecological investigations. Detailed performance statistics of the geo-statistical and machine learning techniques used for the gap filling of the recorded meteorological time series and their regionalization to the Kilimanjaro region indicate that the presented data sets provide reliable measurements of the meteorological reality at Kilimanjaro.

KEY WORDS meteorology; climatology; Tanzania; Kilimanjaro; kriging; machine learning

Received 29 March 2015; Revised 25 September 2015; Accepted 4 October 2015

1. Introduction

Biodiversity and its underlying ecosystem processes and services are closely linked to local climatic conditions. In this context, changes in local climate characteristics and dynamics are mainly determined by (1) global climate and (2) local land-cover change. With regard to ecological importance, changes in the hydrological cycle are likely to be most detrimental or beneficial on a regional scale, as the hydrological cycle is closely linked to the carbon cycle through the water use efficiency (e.g. Schuur, 2003; Mu et al., 2007; Lu and Zhuang, 2010; Mu et al., 2011). However, many tropical regions lack consistent observations of even the most basic meteorological parameters such as temperature, humidity or precipitation. This holds particularly true for areas of complex terrain, such as the region of Kilimanjaro, Tanzania. On the other hand, it is exactly these regions that are of high ecological interest and value with regard to biodiversity and related ecosystem services, often termed ‘global biodiversity hot-spots’. Being a solitary mountain, surrounded by savanna ecosystems, the slopes of Kilimanjaro exhibit a wide range of highly adapted floral and faunal communities in response to a steep climatic gradient. For a detailed description of the most important vegetation habitats at Kilimanjaro, see Hemp (2006).

Based on the extensive works of A. Hemp (e.g. Hemp, 2005a, 2005b; Hemp, 2006) and utilizing the aforementioned climatic gradient in some sort of space-for-time approach, a research group named Kilimanjaro ecosystems under global change: Linking biodiversity, biotic interactions and biogeochemical ecosystem processes was established in 2010 by the German Research Foundation (DFG) that investigates responses of local biodiversity patterns, ecosystem functioning and ecosystem and bio-geo-chemical processes to changes in global and local climate, as well as local land-cover. To provide detailed climatological, biological, paedological and geo-chemical characteristics of the most prominent regional habitat types as well as their interconnecting processes and feedback links, 60 study sites covering the 12 most prominent habitat types were selected along climatic (elevation) and (anthropogenic) disturbance gradients on the southern slopes of Kilimanjaro.

© 2015 The Authors. International Journal of Climatology published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
With regard to previous work identifying climatic characteristics of Kilimanjaro, several studies have been conducted over the past two decades. Yet, the overall data basis is fragmentary at best with consistent observations being almost non-existent. Furthermore, much of the work has been focused on the glaciated summit region of the mountain. Using cumulative rain gauges along seven south-side transects which have been installed since 1997 between 800 and 4000 m.a.s.l., Hemp (2001) identified a height level of about 2200 m.a.s.l. for maximum precipitation. Although no objective quality correction could be applied, the use of accumulated rainfall should generally minimize errors and data inhomogeneity. The data show a rather linear increase in precipitation from about 900 mm at 800 m.a.s.l. to 2700 mm at 2200 m.a.s.l., followed by a slightly exponential drop to about 750 mm at the freezing point level of about 3750 m.a.s.l. (Hemp, 2005a, see also Hemp, 2006) – and continuing to decline above this elevation. These findings have been backed by Røhr and Kellingtveit (2003), who used ten stations of the regular meteorological network in the southern vicinity of the mountain (ending in 1996–2000) and 12 temporarily installed stations along the Mweka climbing route up to the Kibo plateau at 4571 m.a.s.l. for the period of October 1999 to September 2001. After a necessary and quite rigid pre-processing of the data to ensure homogeneity and reliability, they identified a height level of about 2200 m.a.s.l. for maximum precipitation.

Vertical variability of precipitation during the course of the year is analysed by Mölg (2009). They used data from weather stations located at the summit of Kibo, Tropical Rainfall Measurement Mission (TRMM) satellite precipitation information and National Centers for Environmental Prediction / National Center for Atmospheric Research (NCEP/NCAR) re-analysis between 2005 and 2006 for an analysis of the seasonal variation. Using idealized model runs from the regional atmospheric modelling system (RAMS), they show a vertical shift in the maximum precipitation level during extreme daily precipitation events (caused by high moisture in the boundary layer) from 2000 to 4500 m.a.s.l. However, no precipitation gradient measurements are presented to validate the model results, although similar conditions are reported from Mt. Kenya. Intra-seasonal snowfall variability has been analysed by Chan et al. (2008) using snow height data from 2000 to 2005 retrieved from an automatic weather station at the Kibo summit. They link snowfall dynamics to the large-scale atmospheric circulation with west to east and east to west moisture transport during the long and short rainy season, respectively.

Regarding additional moisture advection to the upper region of Kilimanjaro, Pepin et al. (2010) analysed the impact of the montane circulation by comparing temperature and humidity measurements along a height gradient between September 2004 and July 2008 with NCEP/NCAR re-analysis data as free-air reference. Although five of the ten utilized stations show quite large data gaps between >1 and 2.5 years, the data basis was still sufficient to identify a large-scale contrast between Kilimanjaro and the surrounding atmosphere related to day/night heating of (primarily) non-forest areas and an upslope transport of moisture towards the summit. However, no conclusion can be made regarding differences in vertical moisture dynamics according to different vegetation types or regarding horizontal gradients due to different land-cover types/disturbances. The same logger data set has also been used by Duane et al. (2008) to detect a non-linear gradient in mean air temperature due to the dampening effect of the forest belt. They additionally show a non-linearity in the transition zone between the upper rock and glacier area and emphasize that summit climate cannot be retrieved from sub-summit measurements.

Tackling the impact of climate change on local land-cover, Hemp (2005b) shows a replacement of Erica forest by Erica bush or Helichrysum vegetation which was mainly attributed to increasing forest fire frequencies due to decreasing rainfall trends. However, these trends were based on data from two stations located at the lower slope (1430 m.a.s.l.) and one in Moshi (830 m.a.s.l.), which might not be fully representative for the upper slope conditions. Further studies investigating the influence of local land-cover change on local climate conditions include the numerical modelling studies of Fairman et al. (2011) and Mölg et al. (2012).

In summary, two patterns emerge regarding regional meteorological studies at Kilimanjaro, (1) data collection is spatially and temporally diffuse, which means that (2) process-oriented numerical modelling studies are restricted in their scope due to a lack of evaluation data for certain times and places. The data presented in this study are intended to address these issues by being spatially and temporally consistent, which is why the study at hand focuses on the collection and statistical modelling of the data to provide base-line meteorological information for all study sites of the research project. We provide in-depth specifications of the installed sensor network as well as detailed descriptions of the approaches taken to model missing data along with respective model performance evaluations. We also provide preliminary meteorological/climatological descriptions of the study area based on the presented data along with tentative generalizations of the derived observations.

2. KiLi surface meteorology network

The implementation design of the multi-disciplinary Kilimanjaro research unit introduced above combines joint observations from seven sub-project groups on 60 research plots of about 2500 m² size. The plots are structured along both an elevational gradient from 867 to 4550 m a.s.l. and a disturbance gradient for each elevational zone ranging from (near-natural) to (anthropogenically) disturbed land-cover/land-use types (see Table 1).

Following this study design, we installed 100+ meteorological stations across the 60 research plots and some additional locations within the study region between 2010 and 2013. The general set-up follows a hierarchical approach...
providing base-line data of ambient air temperature and relative humidity (TRH) for all study sites of the research unit (basic stations). An extended station set-up, additionally measuring precipitation (PRCP), is used once for each of the 12 habitats (extended stations), and detailed atmospheric observations are collected at four locations outside the forest belt (advanced stations). The latter are automatic weather stations which also measure wind direction, wind speed, atmospheric pressure, four component radiation balance and soil temperature. Here, we only provide information on data that is or can be made available for all research plots of KiLi, namely TRH and PRCP. Figure 1 gives an overview of the spatial distribution for the respective station types. Regarding the elevational distribution of the stations, the inset in Figure 8 shows that the locations can be considered to be representative of the general elevational distribution of the study area. To facilitate semi-automated data handling for all observations, we have developed the software suite julendat (https://code.google.com/p/julendat/) which provides consistent data handling, quality control, temporal aggregation and basic visualization of the recorded parameters.

2.1. Basic stations

In order to get basic information on ecologically important atmospheric parameters, we installed combined TRH sensors on all 60 study plots. The sensors are installed at about 2 m above ground, depending on suitable locations (such as tree branches). According to the manufacturer Driesen + Kern GmbH, the temperature sensor has a measurement range between −30 and +70 °C with a resolution of 0.01 K and a nominal accuracy of ±0.5 °C, while the relative humidity sensor covers the range of 0–100% with a resolution of 0.01% and an accuracy of ±2% (DK3XX, 2014). For both sensors, measurement accuracy decreases towards the extreme ends of the measurement range. These ‘DK320 Humilog rugged’ sensors with integrated loggers (hereafter referred to as loggers) are radiation shielded with coated standard household funnels (see Figure 2). Security reasons and the huge number of installations forced us to consider low-budget options for radiation shields that are easy to obtain locally and small enough not to draw too much attention, as petty theft is common in the area. The loggers are placed inside the funnels so that the sensor is pointing downward ensuring both shielding

Table 1. Overview of the general characteristics of the research plots of the KiLi research unit.

| Plots   | Habitat                      | Disturbance | Elevation TRH | PRCP |
|---------|------------------------------|-------------|---------------|------|
| mai1-5  | Maize field                  | Yes         | 869–1008      | 5    | 1   |
| sav1-5  | Savanna                      | No          | 883–1139      | 4    | 1   |
| cof1-5  | Coffee plantation            | Yes         | 1134–1669     | 5    | 1   |
| hom1-5  | Chagga homegarden            | Yes         | 1171–1824     | 5    | 1   |
| gra1-5  | Grassland                    | Yes         | 1266–1754     | 5    | 1   |
| flm1-4,6| Lower montane forest         | No          | 1780–2054     | 5    | 1   |
| fod1-5  | Ocotea forest disturbed      | Yes         | 2185–2568     | 5    | 1   |
| foc1-5  | Ocotea forest                | No          | 2150–2741     | 2    | 1   |
| fp1-5   | Podocarpus forest            | No          | 2752–3009     | 3    | 1   |
| fpd1-5  | Podocarpus forest disturbed  | Yes         | 2756–2996     | 5    | 1   |
| fer0-4  | Erica forest                 | No          | 3420–3956     | 3    | 1   |
| hel1-5  | Helichrysum                  | No          | 3849–4548     | 5    | 0   |

The last two columns denote how many basic (TRH) and extended (PRCP) stations are installed in each habitat.

© 2015 The Authors. International Journal of Climatology published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.
from direct sunlight, as well as sufficient ventilation. The placement of the loggers inside the funnel allows heated air to escape through the opening above the logger unit. However, it needs to be mentioned that this design does not prevent reflected short-wave radiation from vegetated or snow-covered surfaces below.

The recording interval is 5 min, and data are collected manually by a local field assistant roughly every 4 weeks for each plot. The data are then transferred to a server at Philipps Universität Marburg, Germany, where it is processed using the *julendat* software framework mentioned above.

The first logger was installed in December 2010. Since then, every study site of the research unit has been equipped with one of these basic stations. Due to theft, however, not all sites have continuous recordings since the first installation. In fact, for some sites, no observations are available at all, as loggers were repeatedly stolen before the first data collection (if theft of a logger occurred twice, we decided to not equip the respective site anymore). As of April 2014, 52 basic stations are currently operational (see Figure 1).

2.2. Extended stations

As mentioned earlier, there are 12 different habitat types that are being investigated in the framework of the KiLi research unit. For each of these, one designated ‘focal-plot’ was established, where more elaborate data are gathered by each group of the research unit. With regard to meteorological measurements on these plots, we additionally collect precipitation and partly short-wave downward radiation (which we will not elaborate here). For precipitation measurements, we use the ‘Aerodynamic Raingauge ARG100’ of Environmental Measurements LTD, a standard tipping-bucket precipitation gauge with a funnel diameter of 254 mm. One tip corresponds to approximately 0.2 mm precipitation, however, each gauge is calibrated by the manufacturer and individual calibration factors are supplied. On plots outside the forest belt, these gauges are installed at least 50 cm above the canopy to prevent back-splash. Inside the montane forest zone of Kilimanjaro, we placed the gauges in clearings in close vicinity to the KiLi research sites. Finding suitable locations inside the forest is challenging, as usually a compromise between clearing size and distance to the research plot has to be made. If the clearing is too small, the measurements can hardly be considered representative of field precipitation, yet if the clearing is too far away, the measurement cannot be considered representative for the research site. Installation of the rain gauges started in April 2011. As some research plots are quite remote, and identification of suitable locations was time intensive, the final installation of the network was established in October 2012 (see Figure 1 for the currently operational rain gauge network). As with the basic stations, data are recorded at 5 min intervals, collected manually approximately every 4 weeks and processed with *julendat*.

To complete this set-up, 6 of the 12 plots have also been extended by throughfall measurement set-ups which consist of 30 cumulative gauges each that are read manually about once a week. While 29 of these gauges are placed below the canopy, one is also installed on a clearing vis-a-vis the automated gauges introduced above. The plots are located along the Machame climbing route between 1800 and 3880 m a.s.l. Fog is collected on 6 of the 12 plots with both cumulative gauges attached to collection nets following the design of Schemenauer and Cereceda (1994) and recording buckets attached to Juvik style collectors (Juvik and Ekern, 1978). The corresponding data presentation and analysis will be subject of another publication.

3. Data and methods

Due to the integrative and hierarchical nature of the study design of the KiLi research group, it is imperative that data are provided consistently for all designated sites. For the meteorological base-line data encompassing mean monthly and annual observations of temperature, humidity and precipitation, these need to be made available for all 60 research sites. As mentioned earlier, it was not possible to equip all 60 sites with TRH loggers, and precipitation is only collected at 12 locations in the study area. In order to provide data for the complete suite of study sites, we used spatial and machine learning modelling techniques to fill the gaps. Temperature and humidity were modelled using universal kriging with external drift variables, whereas annual precipitation was modelled using a gradient boosting model with smooth splines as the base learner from the R add-on package ‘bst’ (Zhu, 2011). Monthly precipitation was modelled using the ‘Cubist’ machine learning algorithm available through the R add-on package ‘Cubist’ (Kuhn et al., 2014). Detailed descriptions of these procedures are given below.

Prior to the data aggregation and spatial modelling, a range and a step check are applied to the data set. The thresholds for the valid range of the individual parameters are based on the 0.001 and 0.999 percentile of the entire time series without taking values outside the sensor range into account (e.g. the latter can occur if batteries
are running out). For the step check, the minimum and maximum changes between two subsequent recordings are based on expert knowledge supported by summary statistics of the overall data set.

3.1. Air temperature and air humidity

As shown in Figure 1, 52 of the 60 study sites are currently equipped with TRH loggers. This is, however, a snap-shot that is valid only for April 2014. For the entire research period, this number varies from month to month (due to manifold reasons including logger malfunctioning and petty theft, to name the most important). In order to provide consistent data for all months and all plots, we used a multi-step data imputation approach that can be summarized as follows (for temperature):

1. Aggregate all available 5-min measurements between 2011 and 2014 to hourly observations.
2. Fill any existing gaps of up to 1 year through multivariate regression using the five nearest stations (any remaining gaps will not be filled at this stage).
3. Next, aggregate gap-filled hourly data to daily values for all days that show at least 22 h of valid records.
4. Subsequently, aggregate these to monthly values for months with at least 20 valid daily records.
5. Interpolate monthly records in space using universal kriging with elevation, aspect, slope, sky-view factor and mean monthly normalized difference vegetation index (NDVI) as external drift variables.
6. Aggregate maps from step 5 to multi-year (i.e. 2011–2014) mean monthly and annual maps.
7. Extract multi-year mean monthly and annual data for stations where no observations are available.

Kriging performance is evaluated in a repeated random sub-sampling approach using 15 iterations for each month, and the data have been split so that interpolation is based on 80% of the available stations and the remaining 20% are used for evaluation. The sampling was stratified so that at least one station is taken for each of the 12 habitats for the interpolation. The evaluation statistics of this procedure gave a root mean square error (RMSE) of 1.18 K and a coefficient of determination of 0.96. For humidity, the procedure is similar, with the exception that in step 5 we substituted elevation information with the previously generated temperature maps for each month. With an RMSE of 7.71% and a coefficient of determination of 0.59, performance of the humidity interpolation is worse than for temperature. The spatial resolution of the interpolated maps is 30 m × 30 m, which is the nominal resolution of the utilized digital elevation model (DEM, digitized from topographical maps by J. A. Onginjo, C. Lambrecht and A. Hemp). The NDVI data are mean monthly averages and were resampled from 250 m × 250 m to 30 m × 30 m using bilinear interpolation. For further description of the utilized NDVI data, see Appelhans et al. (2015a).

3.2. Precipitation

Because precipitation changes much less inertially in space and time than temperature or humidity, the approach described above is not feasible to provide spatially and temporally consistent precipitation information for the study sites mainly because of the limited number of observation sites for the temporally high resolution precipitation records (12 stations) and the limited time span of just about 2 years. In order to derive the respective monthly and multi-year information about the recent precipitation distribution in the study area, we therefore followed a two step approach:

1. A climatic-scale annual precipitation map is computed based on the long-term observations of the multi-month accumulating rain gauge network of A. Hemp using universal kriging with elevation as external drift.
2. The climatic-scale values are adjusted to the recently recorded observations using machine learning resulting in a multi-year (i.e. 2012–2014) mean monthly and annual map set which in turn is used for the extraction of the respective precipitation values at each study site.

As mentioned in the Section 1, Hemp (2005a) provides long-term records of precipitation for the study area encompassing up to 117 cumulative rain gauges along seven transects which have been installed since 1997 between 800 and 4000 m a.s.l. Based on this, a first interpolated mean annual precipitation map has been created by Hemp (2006) with a resolution of 30 m × 30 m in order to provide an overview of the climatic background in a vegetation dynamics-oriented publication. Although this map already exists, we decided to compute the climatic-scale and area-wide precipitation distribution again because of two reasons: (1) the underlying observation points used for the 2005 map are rather non-homogeneous regarding their spatial distribution and (2) the computational approach from 2005 does not allow for cross-validation statistics. In order to address both issues, we thinned the available points using a grid of 7500 m × 7500 m for which we randomly selected one observation point out of the full set within each grid cell. This results in a subset of 28 observation points and ensures spatial homogeneity, one of the basic requirements for kriging. Given that the point selection is random, we computed ten different such thinned-point data sets for which we calculated cross-validation statistics using a leave-one-out approach. In summary, this approach provided 10 times 28 validation pairs of predicted and observed mean annual precipitation and revealed a RMSE of 367 mm and a coefficient of determination of 0.61. The final climatic-scale precipitation map shown in Figure 3 is then defined as the maximum value composite of all ten random point interpolations (this time utilizing all 28 observations for each map). For the kriging approach, a bi-linearly smoothed DEM (see description above) has been used as external drift information. We decided to smooth the DEM to avoid unrealistic precipitation gradients across valleys.

Utilizing this climatic-scale map, we used a spline-based gradient boosting algorithm to adjust these long-term records to our recent measurements using the 12 automated
stations as well as the 6 accumulating gauges installed on the clearings of the throughfall plots introduced above. To create mean annual precipitation records for these 6 manual gauges, the readings are evenly distributed over all hours between the actual reading and the one before (i.e. if the precipitation has been accumulated from 0900 12 February to 0800 19 February, then 1/168 of the accumulated sum is mapped to every hour in between). Subsequently, the hourly values are aggregated to daily values. Then, a 365-day moving sum is calculated (i.e. if the time series consists of 370 days, five 365-day moving sums are calculated). Mean annual precipitation is then simply taken as the mean of all the sums present in the moving time series.

The 12 tipping buckets are aggregated to average monthly precipitation records following the aggregation scheme outlined above (steps 1, 3 and 4) and subsequently only summated if records for 12 sequential months are available.

Finally, for all plots equipped with both a tipping-bucket and a manual rain gauge, we averaged those to obtain one value for each plot. This results in 13 locations for which we have mean annual precipitation records. These are then used as the response to train a boosted gradient descent model using smooth splines as the base learner. Explanatory variables for this model are elevation and the respective climatic-scale averages extracted for the plot locations from the aforementioned interpolated rainfall map (Figure 3).

Figure 4 gives an overview of the predicted precipitation (FIT: black) along with the interpolated precipitation from the map (INT: red) and the precipitation observation from the 13 training stations (OBS: blue) as a function of elevation. Given the low number of response values for the model, we rely on internal bootstrapping statistics to assess predictive performance. Using 25 bootstrapping repetitions, an RMSE of 429 mm and coefficient of determination of 0.77 have been derived. These statistics, however, need to be interpreted with care, as they only show how well the model is able to reproduce the 13 stations using the climatic-scale mean precipitation. A more detailed discussion on the meaningfulness of these data is given in later sections.

Monthly precipitation information is modelled with Cubist by predicting monthly contribution weights of the annual sums as a function of mean annual precipitation, month of the year as well as x, y and z coordinates of the site location. Performance of the model is assessed through leave-one-out cross-validation and yields an RMSE of 93 mm and a coefficient of determination of 0.63. Figure 5 provides a graphical overview of the prediction performance for each leave-one-out case. The algorithm is obviously able to predict the seasonal cycle rather well in terms of timing and amount of minima and maxima. Regarding the annual sums of the predictions (printed in the top right of each panel), these correspond very well with the actual observations. Therefore, we are confident that this approach produces reliable estimates of mean
Figure 4. Modelled annual precipitation (FIT: black) in comparison to field observations (OBS: blue) and interpolated values from Hemp (2006) (INT: red) as a function of elevation. Lines are loess fits with a span of 0.75, shaded areas show 95% confidence bands of the smooth fit.

Figure 5. Mean monthly precipitation values for selected plots of the research unit. Predicted values are shown in blue (solid), observed values in black (dashed). The annual sums of the predictions are given at the top of each panel along with the plot ID.

monthly precipitation for all study sites of the research unit for the considered time period between 2012 and 2014.

4. Recent meteorological characteristics of the Kilimanjaro region

This section provides multi-year mean monthly and mean annual meteorological characteristics of the southern slopes of Kilimanjaro as derived from the data introduced above. Both area-wide and plot-based information are presented. Point observations allow for investigation and formulation of processes that relate biotic and abiotic ecosystem components of the system, while area-wide data sets are able to provide a geographic perspective enabling inference of the locally identified processes to a larger, regional extent. In the following sections, the latter
are summarized as aggregated values over the considered habitats, while individual information for each research plot can be found in Figure S1–S3.

4.1. Air temperature

Ambient air temperature is a major driver of ecosystem functioning, as it is tightly linked to the level of available energy within a certain system. One obvious aspect of this is its potentially limiting influence on metabolic rates and subsequently activity, abundance and diversity of ectothermic organisms (Classen et al., 2015). In this sense, spatial patterns of ambient air temperature play a major role in controlling species distribution of certain organisms which makes area-wide data indispensable for spatial ecological analyses such as species distribution modelling.

Figure 6 shows the recent multi-year mean annual air temperature for the southern slopes of Kilimanjaro between 2011 and 2014. It ranges from about 25°C in the surrounding savannah to about −8°C at the summit and generally depicts a linear decrease of temperature with elevation. Figure 7 provides an overview of monthly and annual temperatures for the 12 major habitats revealing a seasonal amplitude of about 2.5–5.5 K, depending on elevation, with a general trend of more damped amplitudes towards higher elevations (although diurnal amplitudes increase with elevation to a certain degree, not shown). This is mainly due to the decreasing amount of heatable surface area towards the summit, thus approximating temperatures of the surrounding free atmosphere. Amplitude damping in mid-elevations is largely a result of the forest belt surrounding the mountain. It is also in the forest belt that we find a break point in mean annual lapse rate (see Figure 8). At a height level of about 2320 m a.s.l., the mean annual lapse rate changes from more or less dry-adiabatic to moist-adiabatic denoting the mean annual cloud base/condensation level for the southern slopes of Kilimanjaro. To investigate seasonal variations, we calculated this breakpoint along with the respective lapse rates for the two dry and wet seasons according to Yang et al. (2014). Its variation in altitude throughout the year is rather low (approximately 65 m, which is within the estimated standard error of the annual breakpoint of 70 m), being most elevated in the so-called ‘short-rains’ during October to December (OND). It is lowest during the colder dry period from July to September (JJAS). The remaining two seasons, January to February (JF) and the so-called ‘long rains’ between March and May (MAM) exhibit similar breakpoint elevations close to the annual mean. The lapse rates follow the general regional thermal seasonality being steeper during the warmer periods of the year (JF and OND). These findings are in line with Pepin et al. (2010).

Breakpoint estimation was carried out using the R package ‘segmented’ which uses an iterative approach to find optimal breakpoints (Muggeo, 2003).

4.2. Air humidity

Mean annual relative humidity between 2011 and 2014 is shown in Figure 9. The moist forest belt is clearly
Figure 7. Mean seasonal temperature profiles for all habitats of the KiLi research unit. Mean annual values are at the top of each panel. Habitats are ordered according to their mean elevation (shown at the bottom of each panel) from left to right. Time period is 2011–2014.

Figure 8. Mean annual (black) and seasonal (coloured) temperature lapse rates derived from all plot observations. Filled circles are observed, empty circles are modelled values. The mean annual (black) and seasonal (coloured) condensation levels are indicated by the horizontal lines. The top right inset shows the elevational distribution of the observation sites in comparison to the theoretical elevations distribution derived from a DEM. The bottom left inset gives a tabular overview of the identified seasonal values for the lower and upper lapse rates (LLR, ULR) along with the identified breakpoints. Row colouring provides the colour key for the rest of the figure. Abbreviations are LLR, lower lapse rate (unit Kelvin); ULR, upper lapse rate (unit Kelvin); SE, standard error of the estimated breakpoint (unit: metres); JA, January/February; MAM, March/April/May; JJAS, June/July/August/September; OND, October/November/December; CDF (unitless), cumulative distribution function; DEM (unit: metres), digital elevation model.

visible exhibiting sharp edges towards the areas above and below. These areas are significantly drier with the driest areas found in the savannah and agricultural lowlands surrounding Kilimanjaro. Due to lack of information and a result of the inherently linear nature of the kriging prediction, humidity of the Kibo summit around the caldera is very likely and highly overestimated and should not be considered realistic. Below the forest, we see a wide belt of intermediate humidity values of around 80% which roughly corresponds to the area of the traditional local agro-forestry system called ‘Chagga homegardens’ (Hemp, 2006). Further below, we see a more or less...
Figure 9. Relative humidity map predicted through universal kriging with external drift between 2011 and 2014. Black line denotes the current national park border. Grey lines are elevation contour lines. For details on the kriging method, see respective sections in the text.

Figure 10. Relative humidity profiles for all habitats of the KiLi research unit. Mean annual values are at the top of each panel. Habitats are ordered according to their mean elevation (shown at the bottom of each panel) from left to right. Time period is 2011–2014.

gradual change towards the driest areas in the region which are dominated by savannah or agricultural crop lands such as maize and sunflower fields. Above the forest, we see a gradual decrease in humidity towards the alpine zone. Figure 10 provides an overview of moisture seasonality for the main habitats. It becomes evident that seasonality is much more pronounced outside the forest belt, especially at lower elevations. Highest humidity levels can be found at around 2400 m a.s.l. which is in line with the lapse rate change point identified earlier. Another notable feature is that below this height level, the most humid month is May; whereas above this point, the peak in atmospheric humidity is usually reached a month earlier and the difference between April and May is reduced. This indicates that higher elevations are influenced earlier and more consistently by the enhanced convection during the so-called ‘long rains’ in boreal spring. Lower areas receive their peak moisture during a narrower window usually during May, just after the April precipitation peak. The August peak which is especially obvious at higher elevations is caused by unusually high rainfall in the region in early August 2013 (see also Figures 12 and S3). Therefore, it remains debatable whether this is a climatological feature.
4.3. Precipitation

As mentioned earlier, the hydrological component plays a major role in ecosystems through its close link to the carbon cycle via the water use efficiency. With regard to atmospheric water availability, precipitation is generally the most important input parameter. The climate-scale observations of Hemp (2005a) reveal wetter southern and drier eastern slopes (Figure 3). The former are further divisible into two distinct regions with enhanced precipitation amounts separated by an area of drier conditions. This pattern to some degree may be a result of the interpolation, yet, in light of the dominant prevailing air flow from the south-east, these patterns do make physical sense. There are two distinct spur features in the topography in the south-eastern region of the study area which may well be responsible for the observed rain shadow. The same holds true for the highest precipitation area in the west, which is located just east of a similar, even bigger morphological feature (see Figure 1). With regard to elevation, precipitation is distributed in an inverse parabolic-like fashion along elevation with a peak around 2200 m.a.s.l. just below the mean annual condensation level (Figure 4).

In recent years (i.e. between 2012 and 2014 – Figure 11), the general pattern of precipitation distribution along the southern slopes is quite similar to the climatological pattern based on the data from Hemp (2005a). However, the region of highest precipitation is shifted upward slightly and confined to a narrower band. Furthermore, higher elevations are markedly wetter in comparison to the long-term average. Driest areas are found in the surrounding savannah lowland regions with annual precipitation values of less than 1000 mm. Precipitation at the summit plateau is, similar to relative humidity levels, expected to be unrealistically overestimated by the kriging prediction due to lack of observations during the last 2–3 years.

The observed precipitation increase at higher elevations in particular has strong implications with regard to potential climate change scenarios for the region. Although (Otte et al., 2015; personal communication) have not found any signs of decreasing precipitation over the last three decades in the region, recent years indicate drier conditions in the lowlands. In light of the findings here, we can extend the picture by adding the fact that this drying is somewhat compensated by wetter conditions in the higher areas of Kilimanjaro. Socio-economically this is still disadvantageous, as agricultural use of the lowlands provides the livelihood for a vast portion of the local population. Further resolving details of such elevational fluctuations in precipitation should be a priority for future research. It remains to be seen how usual such elevational fluctuations in precipitation really are.

Seasonal precipitation distributions for the 12 major habitats are shown in Figure 12. The general pattern of the bimodal seasonality is clearly apparent with the major rainy season in March/April/May, the so-called ‘short-rains’ centred around November and the driest period around September. This is completely in line with...
the widely recognized regional pattern, and the long-term climatology reported from Kilimanjaro airport (see Otte et al., 2015; personal communication). Regarding the elevational distribution, it becomes apparent that the mid-elevation precipitation peak is quite dependent on the season. It is most pronounced during the rainy seasons, whereas it is much less apparent during the drier times of the year. For the timing of the wettest month, we see an elevational signal with lower regions generally showing peak precipitation around April, while the higher parts (above the lower forest border – flm) tend to receive more precipitation during May. A look at the annual precipitation distributions for all plots additionally indicates that this shift is more pronounced in the west (see Figure S3), which is in line with frequent oral reports from locals who state that the rainy seasons tend to start ‘a few weeks’ earlier in the eastern parts of the mountain. Driest conditions are found during September, regardless of elevation.

5. Discussion

In this article, we have outlined the set-up, distribution and characteristics of a meteorological station network to provide base-line atmospheric data for ecological studies at Kilimanjaro. Methods to produce spatially and temporally consistent data sets from this observation network are presented and evaluated. The methods are shown to produce reliable estimates where no data are collected, even for highly variable data such as precipitation. The recent meteorological characteristics of Kilimanjaro’s southern slopes are in line with previously reported findings which indicate reliable performance of the utilized methods and the validity of the thus created data sets. This is supported by the presented model validation results which are all within acceptable error margins.

Elevational temperature gradients found here agree very well with those reported by Duane et al. (2008), both monthly and annually. Lapse rates, however, are somewhat different in both slope and location/elevation of the mean annual condensation level. Duane et al. (2008) reported an average lapse rate of $-5.1$ K km$^{-1}$ with a shallower rate of $-1.4$ K km$^{-1}$ within the rain forest and postulated a change to a steeper rate throughout the alpine zone at the upper tree line at about 3500 m a.s.l. Our much higher resolution observations indicate a rather constant lapse rate of $-8.4$ K km$^{-1}$ until the Ocotea zone where it changes to $-4.3$ K km$^{-1}$ until the upper regions of the alpine zone. One aspect of the differing results might be found in the far less extensive station network of Duane et al. (2008), with ten stations along just one transect above 1800 m a.s.l.

Regarding mean annual humidity, the elevational distribution found here is in line with the values reported by Duane et al. (2008) at comparable elevation levels. This is, however, not true for the summit region (i.e. the region above the highest observation point), as humidity values modelled for this area are much higher than those measured by Duane et al. (2008). There is no physical reason for humidity to increase towards the summit. On the contrary, atmospheric conditions at higher elevations should increasingly approximate those of the free atmosphere, thus a continuous decrease of humidity towards the summit should be expected. This is indeed the case, based on more than a decade of summit observations (D. Hardy, 2015; pers. comm.) and therefore these values are without a doubt merely the result of the interpolation technique. This provides evidence that, even though the utilized classical geo-statistical approach is able to provide acceptable estimates for the region covered by our measurement locations, the technique is not able to provide realistic predictions beyond the region covered by observation sites. As mentioned earlier, kriging is an inherently linear method while the elevational distribution of above-ground air humidity is not a linear function of elevation in tropical regions. Therefore, non-linear approaches, e.g. spline-based models or non-parametric approaches such as generalized additive models or decision tree-based models might be more suitable to capture the underlying spatial process adequately. Such methods have been evaluated for monthly temperatures in Appelhans et al. (2015b) where it is shown that tree-based methods perform better than kriging. Nonetheless, at least for the region covered by observation points, the estimates should be acceptable.
Apart from the good agreement with regard to annual mean values compared with Duane et al. (2008), their reported increase in variability towards the summit is also indicated in our data (see Figure 10). This is mostly a function of the land-cover, with the forest significantly decreasing variability in above-ground air humidity by modulating the fluctuations through more or less constant plant transpiration throughout the year.

As mentioned in Section 4.3, our data indicate wetter/drier conditions at high/low elevations in recent years compared with climate-scale measurements. This is very interesting, as it highlights that in order to understand and interpret changes in the local precipitation climatology, a further (elevational) dimension needs to be considered. Even if future predictions may indicate wetter conditions in the wider East African region (IPCC, 2013), our findings indicate that this may not affect all elevations of Kilimanjaro equally. A first step towards a better understanding of these modulations is to investigate how common such precipitation variations between low and high elevations really are. Another crucial investigation in terms of precipitation patterns at Kilimanjaro is to get better insights into the spatial distribution of precipitation amounts at and around the mountain. This could ideally be achieved through satellite retrievals and/or numerical modelling studies in order to validate the observed patterns of orographic lifting and corresponding precipitation shadows on a process level. For such analyses, the presented station network would provide an ideal training and validation basis.

6. Conclusion and outlook

Despite the semi-artificial nature of the presented data sets, we are very confident that these provide accurate representations of the recent meteorological reality at Kilimanjaro. Regarding their climatic-scale validity, we will need to see how these data will change over time, as we gather more ground observations over the coming years to refine the models. A modification of the absolute values is certainly expected, yet, especially for temperature observations, it can be expected that this change will not be drastic and may well be within the error margins of the sensors. For precipitation, clearly different patterns for recent years when compared with the long-term mean are observable, but it cannot be concluded if these differences are part of a longer trend or just short-term fluctuations. With respect to the utilized methods, there is certainly scope for improvement, especially regarding the spatial predictions of relative humidity. As mentioned in Section 5, non-linear and/or non-parametric models should likely provide better approximations of the observed values, especially along the elevational gradient. Preliminary results from a comparison study of different machine learning techniques for the spatial prediction of temperature in the study region indicate superior model performances for a number of algorithms when compared with kriging, which strongly suggests that a similar approach should also be able to enhance humidity predictions (Appelhans et al., 2015b).

Regarding precipitation measurements, in the first half of 2014, we have started deploying more tipping-bucket gauges in the study region at locations that will greatly increase homogeneity and density of the network and thus should enable more robust precipitation modelling for the southern slopes of Kilimanjaro.

Acknowledgements

This work was carried out in the frame work of DFG-Research Unit 1246 KiLi – ‘Kilimanjaro ecosystems under global change: Linking biodiversity, biotic interactions and bio-geochemical ecosystem processes’, funded by the DFG (funding id Ap 243/1-2, Na 783/5-1, Na 783/5-2). We thank Jimmy Ndyamkama and the rest of our field assistant team who are constantly looking after the installations and facilitate data collection in the field. Finally, we express our gratitude towards the reviewer whose comments greatly helped improve this manuscript.

Supporting Information

The following supporting information is available as part of the online article:

Figure S1. Temperature profiles for all plots of the KiLi research unit. Mean annual values are at the top of each panel. Plots are ordered according to their elevation from bottom (low) to top (high). Time period is 2012–2014.

Figure S2. Relative humidity profiles for all plots of the KiLi research unit. Mean annual values are at the top of each panel. Plots are ordered according to their elevation from bottom (low) to top (high). Time period is 2012–2014.

Figure S3. Precipitation profiles for all plots of the KiLi research unit. Mean annual values are at the top of each panel. Plots are ordered according to their elevation from bottom (low) to top (high). Time period is 2012–2013.

References

Appelhans T, Detsch F, Nauss T. 2015a. Remote: empirical orthogonal teleconnections in R. J. Stat. Softw. 65(10): 1–9.

Appelhans T, Mwangomo E, Hardy DR, Hemp A, Nauss T. 2015b. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. Spatial Stat. (in press), doi: 10.1016/j.spasta.2015.05.008.

Chan RY, Vuille M, Hardy DR, Bradley RS. 2008. Intraseasonal precipitation variability on Kilimanjaro and the East African region and its relationship to the large-scale circulation. Theor. Appl. Climatol. 93: 149–165.

Classen A, Peters MK, Kindeleta WJ, Appelhans T, Eardley CD, Gikangwa MW, Hemp A, Steffan-Dewenter I. 2015. Temperature versus resource constraints: which factors determine bee diversity on Mt. Kilimanjaro, Tanzania? Glob. Ecol. Biogeogr. 24: 642–652, doi: 10.1111/geb.12286.

DK3XX. 2014. Technical description of Driesen + Kern DK 3xx logger series. http://www.driesen.kern.de/downloads/dk3xloggerserie.pdf (accessed 20 October 2014).

Duane WJ, Pepin NC, Losleben ML, Hardy DR. 2008. General characteristics of temperature and humidity variability on Kilimanjaro, Tanzania. Arct. Antarct. Alp. Res. 40(2): 323–334, doi: 10.1657/1523-0430(06-127).

Fairman JG, Nair US, Christopher SA, Mölg T. 2011. Land use change impacts on regional climate over Kilimanjaro. J. Geophys. Res. 116(D3): 1–24, doi: 10.1029/2010JD014712.
Hemp A. 2001. Ecology of the pteridophytes on the southern slopes of Mt. Kilimanjaro. Part II: habitat selection. *Plant Biol.* **3**: 493–523.

Hemp A. 2005a. Continuum or zonation? Altitudinal gradients in the forest vegetation of Mt. Kilimanjaro. *Plant Ecol.* **184**(1): 27–42.

Hemp A. 2005b. Climate change-driven forest fires marginalize the impact of ice cap wasting on Kilimanjaro. *Glob. Change Biol.* **11**: 1013–1023.

Hemp A. 2006. Vegetation of Kilimanjaro: hidden endemics and missing bamboo. *Afr. J. Ecol.* **44**: 305–328, doi: 10.1111/j.1365-2028.2006.00679.x.

IPCC. 2013. *Climate Change 2013: The Physical Science Basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds). Cambridge University Press: Cambridge, UK and New York, NY, 1535 pp.

Juvik JO, Ekern PC. 1978. A climatology of mountain fog on Mauna Loa, Hawai’i Island. Honolulu (HI). WRRC Technical Report No. 118, Water Resources Research Center, University of Hawaii at Manoa.

Kuhn M, Weston S, Keefer C, Coulter N. 2014. *C* code for Cubist by Ross Quinlan. Cubist: rule- and instance-based regression modeling. *R* package version 0.0.18.

Lu X, Zhuang Q. 2010. Evaluating evapotranspiration and water-use efficiency of terrestrial ecosystems in the conterminous United States using MODIS and AmeriFlux data. *Remote Sens. Environ.* **114**(9): 1924–1939, doi: 10.1016/j.rse.2010.04.001.

Mølg T, Chiang JHC, Gohm A, Cullen NJ. 2009. Temporal precipitation variability versus altitude on a tropical high mountain: observations and mesoscale atmospheric modeling. *Q. J. R. Meteorol. Soc.* **135**: 1439–1455.

Mølg T, Grothhauser M, Hemp A, Hofer M, Marzeion B. 2012. Limited forcing of glacier loss through land-cover change on Kilimanjaro. *Nat. Clim. Change* **2**(2): 1–5, doi: 10.1038/nclimate1390.

Mu Q, Heinsch FA, Zhao M, Running SW. 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens. Environ.* **111**: 519–536.

Mu Q, Zhao M, Running SW. 2011. Evolution of hydrological and carbon cycles under a changing climate. *Hydrol. Processes* **25**: 4093–4102.

Muggo VMR. 2003. Estimating regression models with unknown break-points. *Stat. Med.* **22**: 3055–3071.

Pepin NC, Duane WJ, Hardy DR. 2010. The montane circulation on Kilimanjaro, Tanzania and its relevance for the summit ice fields: comparison of surface mountain climate with equivalent reanalysis parameters. *Glob. Planet Change* **74**(2): 61–75.

Rohr PC, Kilingtveit A. 2003. Rainfall distribution on the slopes of Mt Kilimanjaro. *Hydrol. Sci. J.* **48**(1): 65–77.

Schemenauer RS, Cereceda P. 1994. A proposed standard fog collector for use in high-elevation regions. *J. Appl. Meteorol.* **33**(11): 1313–1322.

Schuur EAG. 2003. Productivity and global climate revisited: the sensitivity of tropical forest growth to precipitation. *Ecology* **84**: 1165–1170.

Zhu W. 2011. HingeBoost: ROC-based boost for classification and variable selection. *Int. J. Biostat.* **7**(1): 1–30.

© 2015 The Authors. *International Journal of Climatology* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.