Advances in
Science & Research
Open Access Proceedings

Evaluation of gridding procedures for air temperature over Southern Africa

Kai-Uwe Eiselt[1,2], **Frank Kaspar**[1], **Thomas Mölg**[2], **Stefan Krähenmann**[1], **Rafael Posada**[1], **and Jens O. Riede**[1]

[1]Deutscher Wetterdienst, Frankfurter Str. 135, 63067 Offenbach, Germany
[2]Friedrich-Alexander Universität Erlangen-Nürnberg, Institute for Geography, Erlangen, Germany

Correspondence to: Frank Kaspar (frank.kaspar@dwd.de) and Jens O. Riede (jens-olaf.riede@dwd.de)

Abstract. Africa is considered to be highly vulnerable to climate change, yet the availability of observational data and derived products is limited. As one element of the SASSCAL initiative (Southern African Science Service Centre for Climate Change and Adaptive Land Management), a cooperation of Angola, Botswana, Namibia, Zambia, South Africa and Germany, networks of automatic weather stations have been installed or improved (http://www.sasscalweathernet.org). The increased availability of meteorological observations improves the quality of gridded products for the region. Here we compare interpolation methods for monthly minimum and maximum temperatures which were calculated from hourly measurements. Due to a lack of longterm records we focused on data ranging from September 2014 to August 2016. The best interpolation results have been achieved combining multiple linear regression (elevation, a continentality index and latitude as predictors) with three dimensional inverse distance weighted interpolation.

1 Introduction

Precise monitoring of climate variability and climate change are challenges for many regions in the world. For Africa it was noted that the lack of adequate data and observation systems seriously hinders the ability of scientists to assess the past and current state of climate (ACC, 2013). This applies to several developing regions in the world. Beside others, Southern Africa lacks historical and current ground-based climate records, which, in turn, hinders the capacity of understanding climate variability in the region (Niang et al., 2014).

As part of the SASSCAL initiative (Southern African Science Service Center for Climate Change and Adaptive Land Management), 148 automatic weather stations (AWS) have been installed since 2013 to improve this situation (Kaspar et al., 2015; Posada et al., 2016). Most of them have been installed in Angola, Botswana, Namibia and Zambia and few of them in South Africa, and cover an area of approximately 3.4 million km^2. SASSCAL is a multidisciplinary initiative which aims to improve knowledge in several different scientific areas such as agriculture, biology, and hydrology. In these disciplines there is a strong need for climate information at high spatial resolution. A gridded climatological dataset of the SASSCAL AWS Network could support climate monitoring and climate impact studies in the region. Such datasets are an established tool for climate monitoring in other regions (e.g. Kaspar et al., 2013).

In an earlier study, Krähenmann et al. (2013) published an observational reference dataset for daily minimum and maximum temperature at 0.22° resolution for the African continent except for some regions over Southern Africa, Somalia and Libya due to a lack of observational data. Here, we have used observational data from the SASSCAL AWS Network (Kaspar et al., 2015) and evaluated the quality of several temperature interpolation methods for Southern Africa to fill some of the identified data gaps. For evaluation purposes a dataset of monthly mean minimum and maximum temperatures with a resolution of 0.22° was produced covering the period between September 2014 and August 2016.

The selection of an adequate interpolation method to produce a gridded dataset can be challenging since various factors, such as station density, data coverage and orographic features have to be considered (Krähenmann and Ahrens,

2013). Several interpolation methods have been proposed in the past (e.g. Meyers, 1994; Bolstad et al., 1998; Daly, 2006; Stahl et al., 2006), but, as Tveito et al. (2006) suggested, the interpolation of monthly mean temperature is simple – at least in comparison to precipitation data – considering its high correlation with elevation or land surface characteristics.

Here, three often used different algorithms have been compared using cross-validation (Isaaks and Srivastava, 1989; Wackernagel, 2003): (1) regression kriging (RK), (2) regression combined with two-dimensional inverse distance weighting (2D-IDW) (R2D, distance measure based on geographical coordinates) and (3) regression with three-dimensional inverse distance weighting (3D-IDW) (R3D, distance measure based on geographical coordinates and elevation). Kriging and 2D-IDW are common interpolation methods in geostatistics and earth sciences in general (Bartier and Keller, 1996; Oliver and Webster, 2015), whereas 3D-IDW is a newly developed approach (Krähenmann et al., 2013)

2 Methods

2.1 Study region

The study region (11 and 34° E, 4 and 34° S) is located in Southern Africa (Fig. 1) and includes the four countries Angola, Botswana, Namibia and Zambia. It covers 3.4 million km^2. The northern region is dominated by tropical climates while the southern part is dominated by subtropical semi-arid/arid climate. Large parts of the region are elevated between 1000 to 2000 m, while the western parts close to the coast in Namibia and Angola are lower and separated from the plateau by the Great Escarpment (Stock, 2013). The climate of the coastal region in Namibia and southern Angola is influenced by the cold Benguela Current. The region encounters a dry season (April to September) and a rainy season (September to April). Most precipitation falls between November and March.

2.2 Dataset

Our dataset includes 148 stations (Fig. 1) from the SASS-CAL Weathernet (http://www.sasscalweathernet.org/) providing data at a temporal resolution of either 15 m or 1 h. We used only data for the time period from September 2014 until August 2016. Some station data are available for a longer period, however they cover mainly parts of Namibia. Note that there are only a few stations in eastern Angola. The daily and monthly means have been calculated for each station. Stations with more than one quarter of missing values for a specific month have been discarded. Therefore, the number of stations used in the interpolation process for each month was reduced, to a minimum of 69 stations for September and October 2014, and a maximum of 113 in February 2016 (for sta-
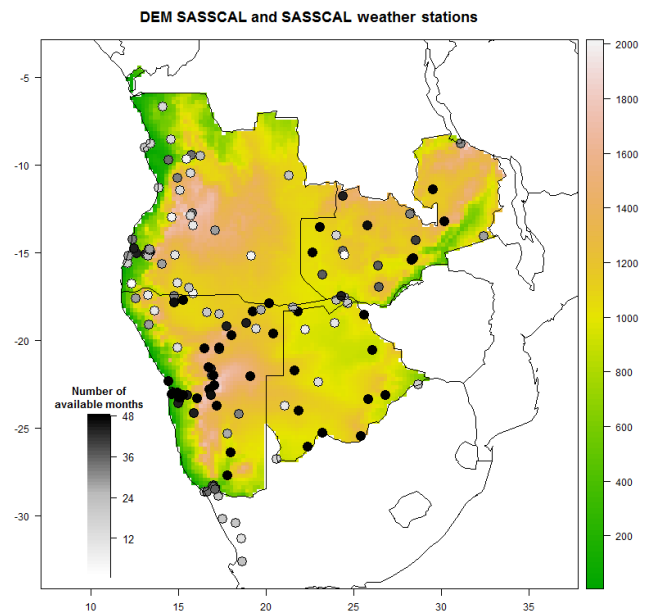

Figure 1. Digital elevation model (color coded) for the study area. Circles show the location of the SASSCAL Weathernet stations. The circle filling illustrates the data availability from black for data completely available over the whole period to empty circle for no data available for the period.

tion data availability see Fig. 1). Other freely available data for region have not been available for the time period.

2.3 Interpolation methods

We compared three different hybrid interpolation methods that combine multiple linear regression with residual interpolation. The general procedure of the three hybrid algorithms is a three-step approach: (a) linear regression using specific predictors to explain the variation of the monthly variable, (b) interpolation of the monthly regression residuals to account for the unexplained variation and (c) summation to yield the final result. The following methods have been applied to perform residual interpolation: two-dimensional inverse distance weighting (2D-IDW), three-dimensional IDW (3D-IDW), and simple kriging (SK). For all statistics, methods and plots we used R version 3.1.2 (R Development Core Team, 2016, http://www.r-project.org).

The regression coefficients necessary for the hybrid methods can be estimated from point data (observations and predictors at station locations) if the regression function applied is linear (Heuvelink and Pebesma, 1999). Multiple linear regression amounts to the following equation:

$$Y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_i x_i + \ldots + b_N x_N \tag{1}$$

where Y is the dependent variable, x_i are the independent variables, b_0 is the intercept coefficient, b_i are the regression coefficients and N the number of independent variables. The

regression coefficients are derived by minimising the squared differences between the observed variable and its estimate.

2.3.1 Inverse distance weighting (2D-IDW and 3D-IDW)

IDW and kriging interpolation are based on the assumption that points closer together are more alike than points further apart. For IDW, the values are estimated from the weighted linear combination of nearby observational stations (Naoum and Tsanis, 2004):

$$z_{x,y} = \frac{\sum_{i=1}^{n} z_i w_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

where $z_{x,y}$ are the interpolated values at coordinates x and y, the z_i are the observed values, and the w_i are the weights determining the influence of the respective observed value z_i. The w_i can be defined as:

$$w_i = d_{x,y,i}^{-\beta} \tag{3}$$

where $d_{x,y,i}$ is the distance between $z_{x,y}$ and z_i, and β is an exponent (Bartier and Keller, 1996), which was determined using cross-validation to minimize the interpolation error.

Cross-validation revealed 2 as the overall optimal beta value (which is the distance weighting power). Therefore, cross-validation was run over the whole period and for several weighting powers ranging from 1.5 to 3. To prevent over-fitting and due to the low number of available observing stations a constant distance weighting power (β) of 2 was chosen for the whole interpolation period.

3D-IDW is based on an Euclidean distance measure, which is expanded by accounting for the elevation z:

$$d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i \cdot F - z_j \cdot F)^2} \tag{4}$$

with $d_{i,j}$ being the distance between points i and j, x_i and y_i the coordinates of point i, and x_j and y_j the coordinates of point j, and z_i and z_j the height above sea level of points i and j, respectively. F is a factor to penalise the elevation difference.

As with the 2-D case, cross validation revealed a distance weighting power of 2 as most suitable. F is used to modify the distance between a grid node and neighboring stations as a function of elevation difference. Depending on the weather condition at hand F may vary considerably to find the minimum interpolation error. In fact F is considerably larger than one. For the interpolation procedure F values ranging from 50 to 2000 (the elevation distance is transferred to degrees to match the unit of the geographical coordinates) have been tested and cross validation revealed 1000 as the most suitable F value.

2.3.2 Simple Kriging

Simple Kriging involves solving a set of linear equations to minimize the mean squared error of the residuals from the interpolating surface. To ensure spatial homogeneity of the residuals, which is a requirement to solve this least squares problem, a normal score transformation was applied to the data before interpolation (Deutsch and Journel, 1998). For the Simple Kriging process, we applied a spherical variogram model including a nugget variance. The accurate estimation of a variogram model requires a large sample size. However, the station coverage over Southern Africa is poor. Therefore, we have chosen to derive a global variogram from all available stations. Additionally, the previous regression is used to remove major geographic features and to yield a more homogeneous distribution of the residuals.

The variogram parameters were determined using an approach suggested by Ahrens and Beck (2008). We used a data pool containing all monthly values of the period September 2014 to October 2015, and estimated a range of approximately 12 °C and a nugget variance of 0.3 °C. Subsequent to the residual interpolation using Simple Kriging, the gridded residuals were back transformed to the original distribution and combined with the regression-based estimates.

2.4 Validation statistics

For the validation of the three interpolation methods we employed a leave-one-out cross-validation approach, which involves leaving out each station in turn and estimating its value from the remaining observations using an interpolation method. The estimated value is then compared with the actually observed value, which provides an estimate of the model error at this point (Oliver and Webster, 2015).

The Root Mean Square Error (RMSE) is an often used skill score in model evaluation. Small RMSE values indicate high model accuracy with theoretical optimum at a value of 0 denoting a perfect model prediction. The RMSE is calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_i - Z_i)^2} \tag{5}$$

where N is the number of observations, P_i the model value at the coordinates of station i, and Z_i the observed value at the coordinates of station i.

As a second performance criterion we used a measurement for the variance preserved in the interpolation data compared to the variance in the observations "VARI" (Krähenmann et al., 2013). VARI is the ratio of the variance of the interpolated values (using cross-validation) and the observed values. It ranges from 0 to 1, with high values indicating that a large portion of variance has been retained in the interpolated data. A VARI greater than one depicts an enhancement of the spatial temperature variation, whereas a value between 0

and 1 implies the observed temperature variability is reduced. The explained variance, however, explains how well the used method/parameters explain the spatial temperature variability.

The VARI equates to:

$$\text{VARI} = \frac{\sigma^2(P)}{\sigma^2(Z)}$$

where P are the predictions and Z the observed values.

3 Results and discussion

3.1 Validation results

The suitability of the selected predictors for multiple linear regression for minimum temperature ($T_{\min}$) and maximum temperature ($T_{\max}$) has been validated using the RMSE. Thereby, the RMSE was calculated from the regression residuals. For single predictors, the RMSE errors were highest for longitude (x) and elevation (z), lower for zonal mean temperature (b) while continentality index (K) and latitude (y) performed best (Table 1). Successive combination of predictors reduces the RMSE values, and the combination of all available predictors performs best.

The VARI was used to assess how much variance of the data was retained by the model, with low values indicating little and high values high retention of variance. As a single predictor, K clearly performs best, for y and b a lower fraction of the spatial variability could be retained, and z and x performed worst. Predictor combinations show improved VARI values with the combination of all available predictors performing best (Table 1).

In conclusion, the combination of all predictors yields the best result in both, RMSE and VARI. However, the more predictors are used within the linear regression, the greater the risk of over-fitting becomes. Previous studies reported high correlation between air temperature and elevation (Benavides et al., 2007; Kurtzman and Kadmon, 1999; Hudson and Wackernagel, 1994) and air temperature and latitude (Wackernagel, 1994), respectively. One reason for the poor elevation performance in the study region could be the effect of other variables (e.g. continentality) and latitudinal effects masking the elevation dependence of air temperature. This can be explained by the vast size of the study area. This issue may be reduced by splitting of the target area into several more homogeneous climate regions which is, however, not feasible due to the low station number. Furthermore, application of a non-linear temperature profile may significantly improve the gridding result. Yet, besides the low station density also the underrepresentation of higher elevations prevents its application.

Furthermore, we found an autocorrelation between longitude (x) and continentality index (K), since K outperformed x as single predictor, K has been preferred over x.

Table 1. RMSEs and VARI averaged over first year of the interpolation period (September 2014–August 2015). Predictors: elevation (z), continentality index (K), zonal mean temperature (b), longitude (x), latitude (y).

Predictor	RMSE $T_{\min}$ in °C	RMSE $T_{\max}$ in °C	VARI $T_{\min}$	VARI $T_{\max}$
$\sim z$	2.9	3.4	0.08	0.08
$\sim K$	2.7	3.0	0.21	0.25
$\sim b$	2.8	3.1	0.10	0.18
$\sim x$	2.9	3.4	0.09	0.04
$\sim y$	2.7	3.0	0.14	0.16
$\sim z + K$	2.5	3.0	0.34	0.28
$\sim K + b$	2.5	2.7	0.30	0.40
$\sim y + K$	2.4	2.7	0.35	0.40
$\sim z + K + b$	2.3	2.6	0.44	0.44
$\sim y + z + K$	2.2	2.6	0.50	0.44
$\sim x + z + K + b$	2.3	2.5	0.47	0.53
$\sim y + z + K + b$	2.1	2.6	0.53	0.48
$\sim x + y + z + K$	2.2	2.5	0.52	0.53
$\sim x + y + z + K + b$	2.1	2.4	0.55	0.57

Hence, in this study the combination ($y + z + K$) was chosen.

3.2 Model evaluation

While the RMSEs (Fig. 2) for all methods exhibit a small decline over the whole interpolation period for $T_{\max}$ and little variation in general this is very different for $T_{\min}$. Here the RMSEs show a distinct annual cycle. They are lowest between December and March (i.e. summer), with minimum values around 1.0 °C. However, they are considerably higher between May and August, with values between 2.5 and 3.5 °C. For $T_{\min}$, the hybrid methods perform similarly well and generally better than the mere linear regression (Fig. 2). For $T_{\max}$, R3D performs generally better than RI and RK, while all hybrid methods clearly outperform linear regression. Overall, in terms of the RMSE, R3D performs best, particularly for $T_{\max}$.

In general, the evaluated interpolation methods show similar performance in terms of VARI (Fig. 3) as previously observed regarding the RMSE. For $T_{\min}$ high values (i.e. a large proportion of observed variance preserved) occur from October 2014 until April 2015 and from October 2015 until March 2016. VARI was low from May to September 2015 and April to August 2016. For $T_{\max}$ there is less variation over the year with a remarkable performance drop of linear regression and RK in April and May 2015, which is much less profound for RI and especially for R3D. As with the RMSE, all three hybrid methods outperform linear regression, with R3D again performing best. Yet, unlike with the RMSE, RI also outperforms RK, which is evident from the annually averaged values (Fig. 3).
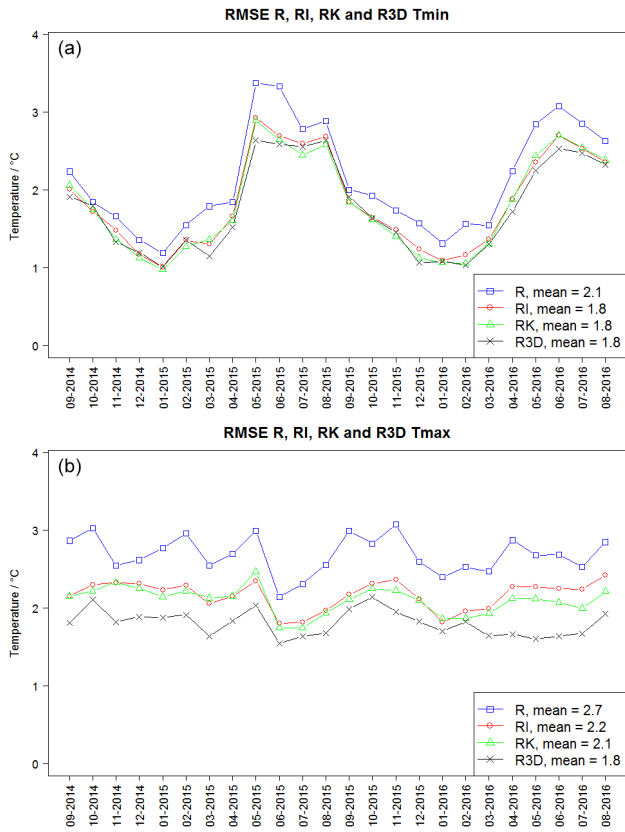
Figure 2. Results of the RMSE calculation for all interpolation models, where R = Regression, RI = Regression + two dimensional IDW, RK = Regression + Kriging, R3D = Regression + three dimensional IDW, for minimum temperature **(a)** and maximum temperature **(b)**.

Figure 3. Results of the VARI calculation for all interpolation models, where R = Regression, RI = Regression + two dimensional IDW, RK = Regression + Kriging, R3D = Regression + three dimensional IDW, for minimum temperature **(a)** and maximum temperature **(b)**.

Overall, there is a strong seasonal cycle in the RMSE for minimum temperature, whereas there is no pattern in maximum temperature.

The southern winter season (June, July, and August) is also the dry period, when lowest minimum temperatures occur, particularly in lower elevations. This is due to the lower relative humidity in this season. As a result, temperatures drop particularly in valleys and depressions, whereas air temperatures do not drop as much in elevated areas.

This cascades by a temperature inversion with non-linear temperature profiles to an increase in regression residuals, whose interpolation is compound with increased interpolation errors.

In such cases, modelling temperatures profiles using nonlinear regression or varying regression predictors (Hofer et al., 2012) would be of benefit. However, the low station density (Fig. 1) and the under representation of summits and elevated areas does not allow for a robust estimate of non-linear temperature profiles.

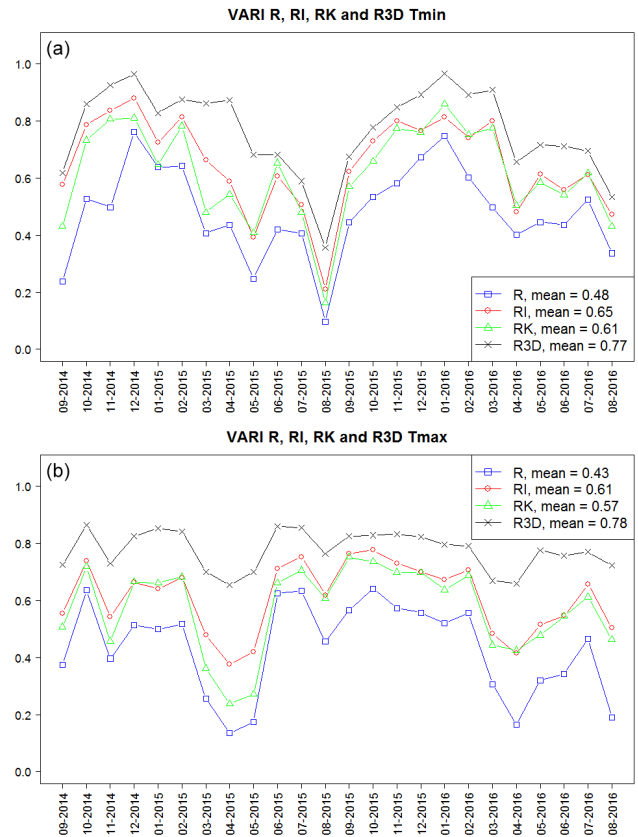Maximum temperature is less affected by changes in relative humidity. During southern winter months, near sur-face areas get high insolation rates and air temperature rises markedly. The air temperature rises therefore more strongly in low areas than in elevated areas, where temperature remains more constant. This removes temperature inversion leading to a more linear temperature gradient during day time. Thus linear regression and residual interpolation yield better interpolation results and lower interpolation errors for T_{max}.

As recommended by Tveito et al. (2006) the interpolation results were also visually examined. Figure 4 gives an example of the interpolated residuals using SK, 2D- and 3D-IDW for July 2015, where all the characteristics of the respective methods can be observed. In the case of the 2D-IDW interpolation (Fig. 4, top panels), small circular high- or low-value structures can be observed, with the measuring stations at the centres, while in the area between stations the values quickly tend to zero. In comparison, the SK (Fig. 4, middle panels) yields structures of larger extent and more gradually changing values. Also the value range of SK is smaller than that of 2D-IDW, which can be attributed to the nugget effect (i.e. non-explained variance in the short distances) and thus
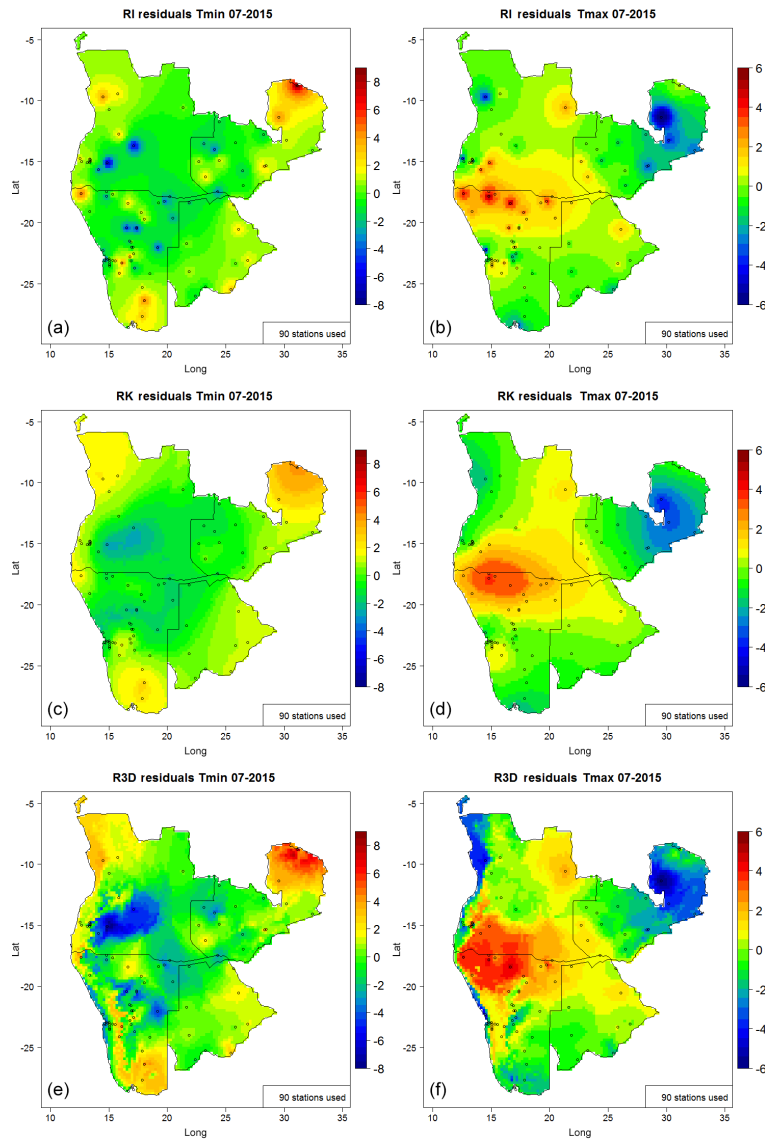
Figure 4. Maps of the interpolated residuals of RI **(a, b)**, RK **(c, d)** und R3D **(e, f)** in July 2015, for minimum temperature **(a, c, e)** and maximum temperature **(b, d, f)**.
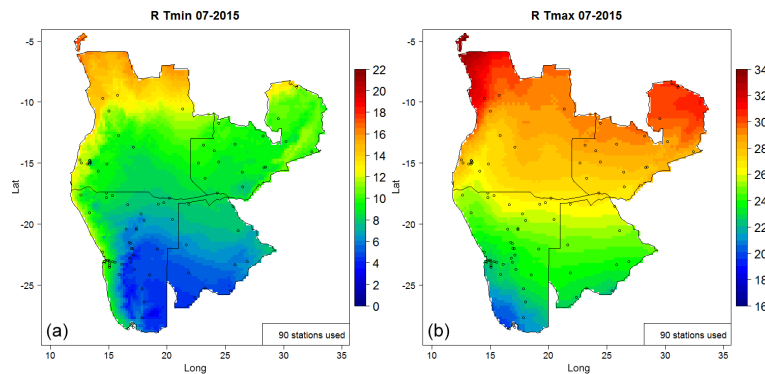

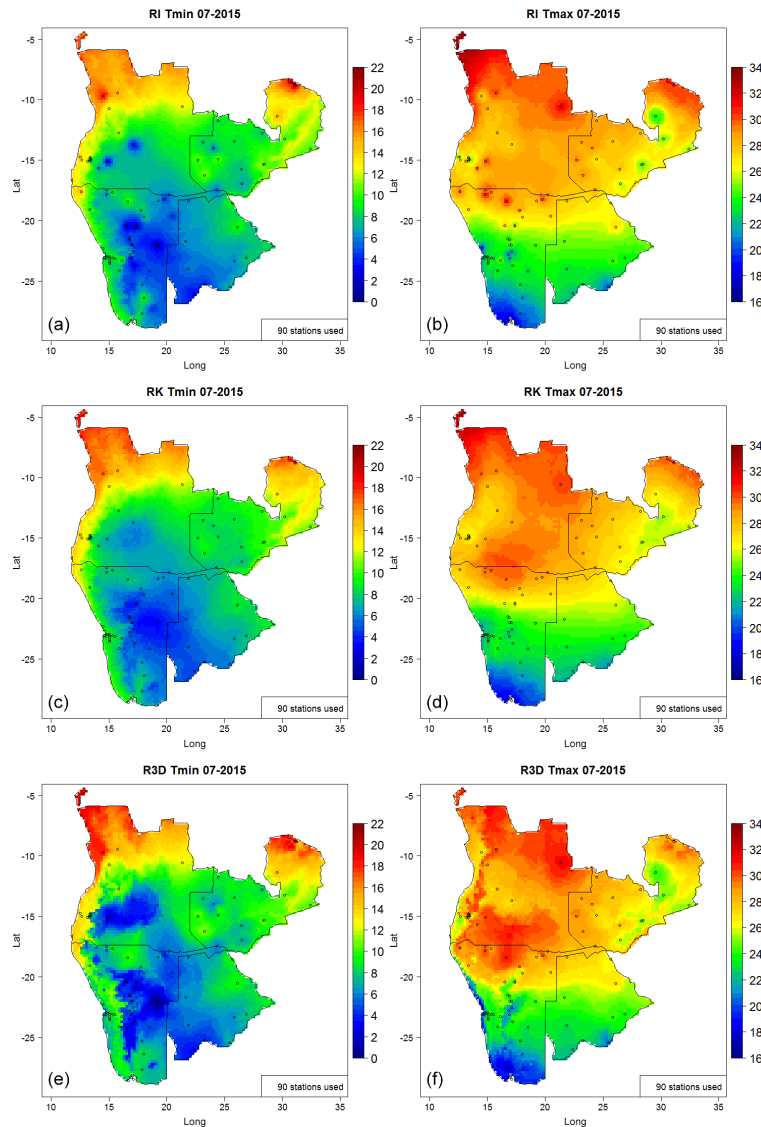**Figure 5.** Temperature map of July 2015 on basis of the linear regression. **(a)** T_{min}, **(b)** T_{max}. Unit: °C.

Figure 6. Temperature map of July 2015 on basis of RI **(a, b)**, RK **(c, d)** and R3D **(e, f)**.

leads to stronger smoothing. The 3D-IDW (Fig. 4, bottom panels) yields a larger value range and also smaller spatial structures than SK, which tend to follow terrain structures such as mountains and the coast line.

Examining the final interpolation results, linear regression (Fig. 5) and RK (Fig. 6, middle panels) exhibit the least small-scale variation, where temperature patterns follow the continentality-index and the DEM. Some of these terrain following patterns may also be found in the RI-grids (Fig. 6, top panels), however, the RI-based residuals maps contain also some circular small-scale structures surrounding observing stations. R3D-based grids (Fig. 6, bottom panels) exhibit some small-scale structures, however, they are mostly noncircular and the station distribution is less evident. Overall, the R3D method performed best in terms of RMSE, VARI

and visual inspection, and was thus chosen for the gridding of the $T_{\min}/T_{\max}$ maps over Southern Africa.

3.3 Application results

In this study, we generated a 0.22° gridded dataset of monthly averaged daily maximum and minimum temperature covering the period September 2014–August 2016.

In this period, temperature values range from −3 to 41 °C ($T_{\min}$ between −3 and 26 °C, $T_{\max}$ between 16 and 41 °C).

Analysing the interpolation maps for $T_{\min}$ (Fig. 7), the following structure is evident in most of the months: In the southwest part of the study region (Namib dessert) the minimum temperatures are relatively low, while further north in the coastal region of Angola the values are generally higher. East of the Namib, there is a north–south oriented strip of rel-
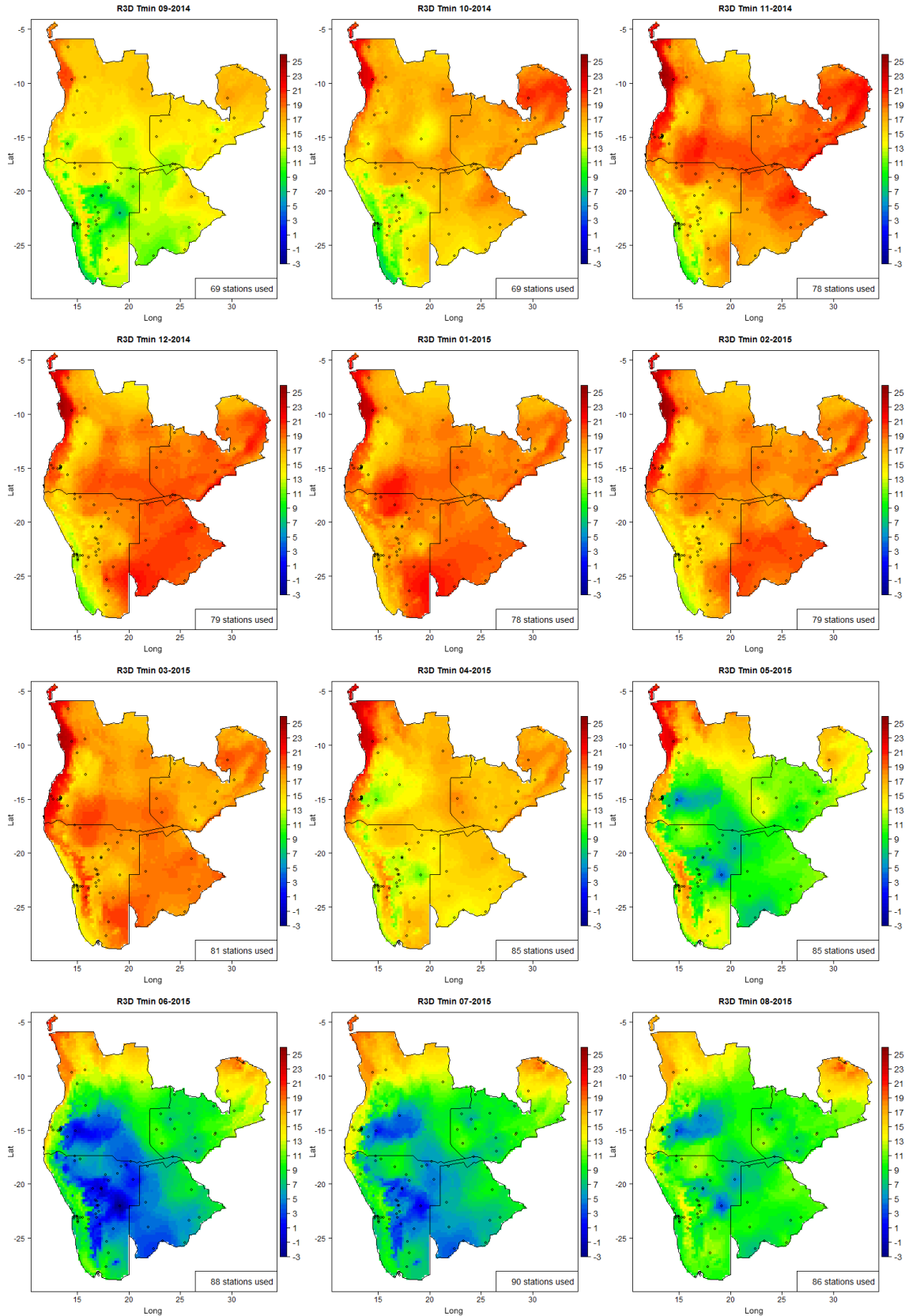
Figure 7. Minimum temperature maps on basis of three dimensional inverse distance weighting method from September 2014 until August 2015. Unit: °C.
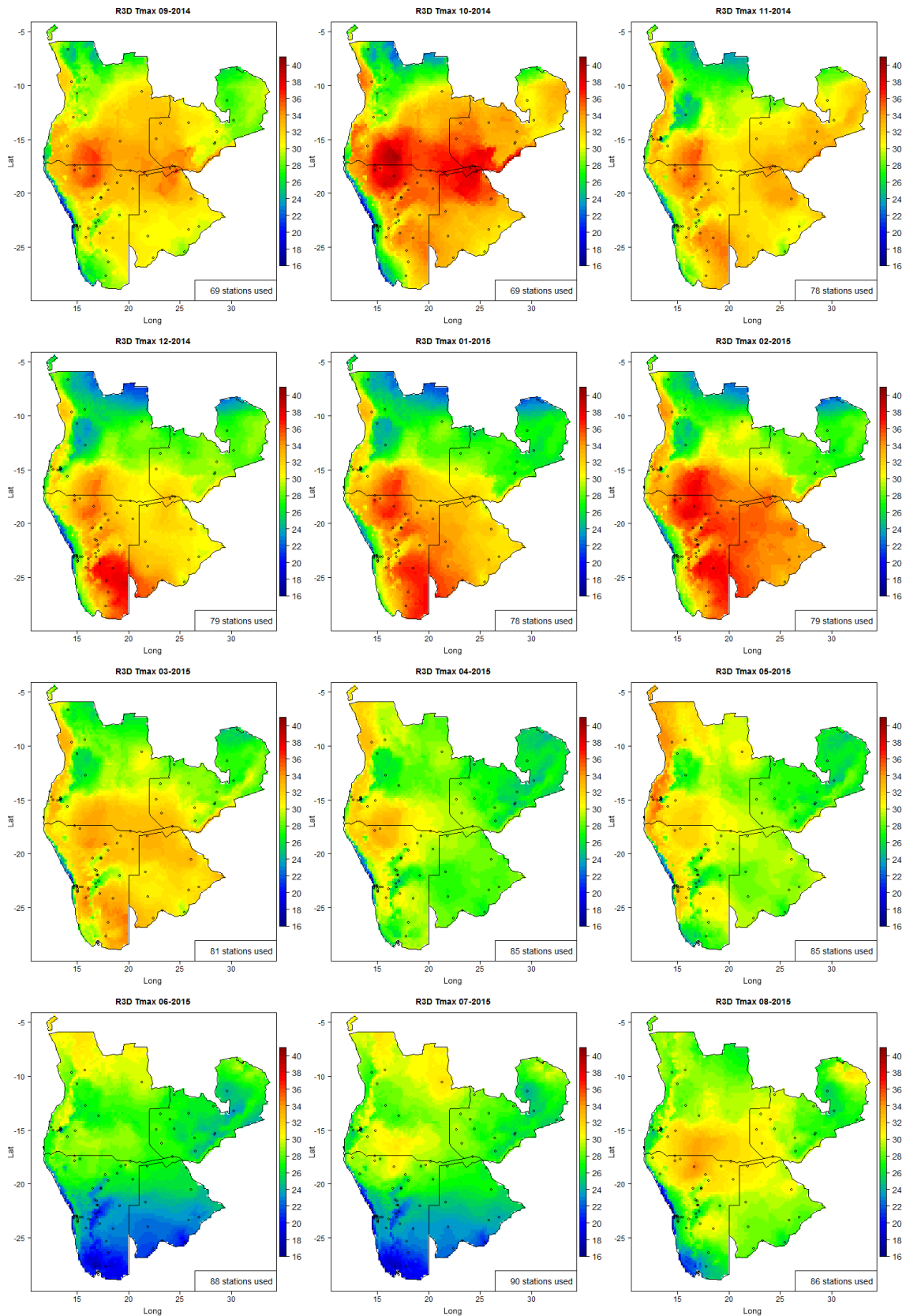
Figure 8. Maximum temperature maps on basis of three dimensional inverse distance weighting method from September 2014 until August 2015. Unit: °C.

atively high temperature, while in Angola the temperatures are high along the coast, yet, just off the coast, temperatures are considerably lower. Over the eastern part of the study region, temperatures are predominantly high.

In the maps of T_{max} (Fig. 8) the cold Namib is also apparent as well as the higher temperatures on the Angola coast. In the north of Angola and Zambia the temperatures are quite homogenous while further south is an evident seasonal pattern with hot temperatures during the rainy season and cold temperature during the dry season. In all maps it is evident that some regions are still lacking ground based observation data. Moreover, from August 2015 onwards no data were available for east and north Angola. However, there was a concurrent increase in the number of stations over Zambia, which partly compensates the data decrease over Angola.

A general problem is the station data availability; to get as reliable results as possible, we used all available data for each month. Since several stations did not deliver data within the whole period, we had to calculate the REM for each single map. However, since the region is sparse in data, this method was a compromise between data density and homogeneity over time. More stations, especially in Angola, and a data transmission would improve the model outcome. Furthermore, longer time series over a period of ten years are required to evaluate the model results more robustly. Using a multi model approach could improve the model result further with regard to the seasonality of dry/wet seasons in the region (e.g. Hofer et al., 2012). Our data set, therefore, encounters typical problems of gridded data (as widely documented in the literature) and, hence, information on the number of stations entering the gridding procedure is vital for the user to assess the limitations for a specific place and time (e.g. Mitchell and Jones, 2005). All data used in this publication are available at (http://www.sasscalweathernet.org).

4 Conclusions

A gridded dataset of monthly mean daily maximum (T_{max}) and minimum (T_{min}) temperature was developed for Southern Africa, covering the period September 2014 to August 2016. Several interpolation methods and predictor sets have been evaluated to determine R3D as the most suitable interpolation method for the study area. This dataset has a spatial resolution of 0.22° and provides opportunity for different kinds of applications, such as drought index, heat stress maps and many more. Note that the length of the dataset in this study is too short to provide climatological statements. Overall the model performs well for T_{max}, while for T_{min} there are some patterns in the seasonal cycle where the model performs poorer (May–August).

The splitting of the interpolation process using regression combined with IDW increases robustness of the interpolation, since the regression and IDW distance parameters can be separately modified. Multi-dimensional IDW distance al-

lows explicitly separating residuals in elevated areas from those in valleys and depressions. Although several caveats remain, it was decided to use one model for the whole time period to reduce the computational time (and therefore to make later applications of the tool more feasible).

The topography of the regions is characterized mainly by a flat area. Only at the Atlantic coast, the north of Angola and the valley of the Sambesi River there are more complex topological patterns. We assume that the results would not improve significantly by using a more complex model.

In general the aim of this work was to implement simple and easy to implement methods for temperature interpolations, which require only station observations, geographical coordinates, elevation and continentality – to provide an interpolation tool for local meteorological services.

Further studies could show the information increase with more complex models such as different model approaches for seasons or non-Euclidean distance measures.

Also the splitting of the target area and the application of a non-linear temperature profile would be of great benefit. However, this is not feasible due to the low station density and the under representation of elevated areas.

References

ACC: Final Declaration, Africa Climate Conference 2013, 15–18 October 2013, Arusha, Tansania, available at: http://www.wmo.int/amcomet/sites/default/files/field/doc/events/acc2013_final_declaration.pdf (last access: 13 June 2017), 2013.

Ahrens, B. and Beck, A.: On upscaling of rain-gauge data for evaluating numerical weather forecasts, Meteorol. Atmos. Phys., 99, 155–167, 2008.

Bartier, P. M. and Keller, C. P.: Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW), Comput. Geosci., 22, 795–799, 1996.

Benavides, R., Montes, F., Rubio, A., and Osoro, K.: Geostatistical modelling of air temperature in mountainous region of Northern Spain, Agr. Forest Meteorol., 146, 173–188, 2007.

Bolstad, P. V., Swift, L., Collins, F., Regniere, J.: Measured and predicted air temperatures at basin to regional scales in the southern Appalachian Mountains, Agr. Forest Meteorol., 9, 167–176, 1998.

Daly, C.: Guidelines for assessing the suitability of spatial climate data sets, Int. J. Climatol., 26, 707–721, 2006.

Deutsch, C. V. and Journel, A. G.: GSLIB: geostatistical software library and users's guide, 2nd Edn., Oxford University Press, New York, 1998.

Heuvelink, G. B. M. and Pebesma, E. J.: Spatial aggregation and soil process modelling, Geoderma, 89, 47–65, 1999.

Hofer, M., Marzeion, B., and Mölg, T.: Comparing the skill of different reanalyses and their ensembles as predictors for daily air temperature on a glaciated mountain (Peru), Clim. Dynam., 39, 1969–1980, https://doi.org/10.1007/s00382-012-1501-2, 2012.

Hudson, G. and Wackernagel, H.: Mapping temperature using kriging with external drift: theory and example from Scotland, Int. J. Climatol., 14, 77–91, 1994.

Isaaks, E. H. and Srivastava, R. M.: An introduction to applied geostatistics, Oxford University Press, New York, 1989.

Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., and Deutschländer, T.: Monitoring of climate change in Germany – data, products and services of Germany's National Climate Data Centre, Adv. Sci. Res., 10, 99–106, https://doi.org/10.5194/asr-10-99-2013, 2013.

Kaspar, F., Helmschrot, J., Mhanda, A., Butale, M., de Clercq, W., Kanyanga, J. K., Neto, F. O. S., Kruger, S., Castro Matsheka, M., Muche, G., Hillmann, T., Josenhans, K., Posada, R., Riede, J., Seely, M., Ribeiro, C., Kenabatho, P., Vogt, R., and Jürgens, N.: The SASSCAL contribution to climate observation, climate data management and data rescue in Southern Africa, Adv. Sci. Res., 12, 171–177, https://doi.org/10.5194/asr-12-171-2015, 2015.

Krähenmann, S. and Ahrens, B.: Spatial gridding of daily maximum and minimum 2 m temperatures supported by satellite observations, Meteorol. Atmos. Phys., 120, 87–105, 2013.

Krähenmann, S., Kothe, S., Panitz, H.-J., und Ahrens, B.: Evaluation of daily maximum and minimum 2-m temperatures as simulated with the Regional Climate Model COSMA-CLM over Africa, Meteorol. Z., 22, 297–316, 2013.

Kurtzman, D. and Kadmon, R.: Mapping of temperature variables in Israel: a comparison of different interpolation methods, Clim. Res., 13, 33–43, 1999.

Meyers, D. E.: Spatial interpolation: an overview, Geoderma, 62, 17–28, 1994.

Mitchell, T. and Jones, P.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, 2005.

Naoum, S. and Tsanis, I. K.: Ranking spatial interpolation techniques using a GIS-based DSS, Global Nest: Int. J., 1, 1–20, 2004.

Niang, I., Ruppel, O. C., Abdrabo, M. A., Essel, A., Lennard, C., Padgham, J., and Urquhart, P.: Africa, in: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects, Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Barros, V. R., Field, C. B., Dokken, D. J., Mastrandrea, M. D., Mach, K. J., Bilir, T. E., Chatterjee, M., Ebi, K. L., Estrada, Y. O., Genova, R. C., Girma, B., Kissel, E. S., Levy, A. N., MacCracken, S., Mastrandrea, P. R., and White, L. L., Cambridge University Press, Cambridge, UK and New York, NY, USA, 1199–1265, 2014.

Oliver, M. A. and Webster, R.,: Basic Steps in Geostatistics: The Variogram and Kriging, Springer, Heidelberg, 2015.

Posada, R., Nascimento, D., Neto, F. O. S., Riede, J., and Kaspar, F.: Improving the climate data management in the meteorological service of Angola: experience from SASSCAL, Adv. Sci. Res., 13, 97–105, https://doi.org/10.5194/asr-13-97-2016, 2016.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/ (last access: 13 June 2017), 2016.

Stahl, K., Moore, R. D., Floyer, J. A., Asplin, M. G., and McKendry, I. G.: Comparison of approaches for spatial interpolation of daily temperature in a large region with complex topography and highly variable station density, Agr. Forest Meteorol., 139, 224–236, https://doi.org/10.1016/j.agrformet.2006.07.004, 2006.

Stock, R. F.: Africa south of the Sahara. A geographical interpretation, 3rd Edn., The Guilford Press, New York, 2013.

Tveito, O. E., Wegehenkel, M., van der Wel, F., and Dobesch, H.: COST Action 719: The Use of Geographic Information Systems in Climatology and Meteorology: Final Report, EUR-OP, Office for Official Publications of the European Communities, Luxembourg, 254 pp., 2008.

Wackernagel, H.: Cokriging versus kriging in regionalized multivariate data analysis, Geoderma, 62, 83–92, 1994.

Wackernagel, H.: Multivariate geostatistics, 3rd edn., Springer, Berlin, 387 pp., 2003.