Hydrology and Earth System Sciences

Köppen versus the computer: comparing Köppen-Geiger and multivariate regression tree climate classifications in terms of climate homogeneity

A. J. Cannon

Meteorological Service of Canada, Environment Canada, 201–401 Burrard Street, Vancouver, V6C 3S5, BC, Canada

Correspondence to: A. J. Cannon (alex.cannon@ec.gc.ca)

Abstract. A global climate classification is defined using a multivariate regression tree (MRT). The MRT algorithm is automated, hierarchical, and rule-based, thus allowing a system of climate classes to be quickly defined and easily interpreted. Climate variables used in the MRT are restricted to those from the Köppen-Geiger classification system. The result is a set of classes that can be directly compared against those from the traditional system. The two climate classifications are compared at their 5, 13, and 30 class hierarchical levels in terms of climate homogeneity. Results indicate that both perform well in terms of identifying regions of homogeneous temperature variability, although the MRT still generally outperforms the Köppen-Geiger system. In terms of precipitation discrimination, the Köppen-Geiger classification performs poorly relative to the MRT. The data and algorithm implementation used in this study are freely available. Thus, the MRT climate classification offers instructors and students in the geosciences a simple instrument for exploring modern, computer-based climatological methods.

1 Introduction

The goal of a global climate classification is to divide the world's land surface into areas with homogeneous climate conditions. Intra-annual variability in temperature, precipitation, and other climate variables should be as similar as possible within a given class, irrespective of where on the globe a station belonging to said class is located. With this context, this paper pits the Köppen-Geiger climate classification system, which was originally published by Wladimir Köppen in 1918 and subsequently updated by Köppen, Rudolf Geiger, and others (see Wilcock, 1968), versus a modern computer-based clustering algorithm. At face value, this may seem a somewhat unfair match – after all, the Köppen-Geiger system was originally developed more than 90 years ago as a means of differentiating climate classes based on vegetation groups and not as a general climate classification. The fact remains that the Köppen-Geiger system has become one of the standard all-purpose regionalizations of the world's climates, and, in this context, it has stood the test of time, despite rapid improvements in computer processing power and the attendant expansion of global climate observing networks. For example, it has, among other applications, been used as a diagnostic tool to evaluate the performance of general circulation models (Lohmann et al., 1993), to detect and characterize climate change (Beck et al., 2005), and to identify and explain continental differences in runoff (Peel et al., 2001). Reasons for the longevity of the Köppen-Geiger classification have been explored by a number of authors (see, for example, Peel et al., 2007 and references therein). In short, these stem from the system's "historical inertia", which is primarily due to its use in teaching first courses in physical geography, meteorology, and climatology; its definition in terms of simple climate variables; and its rule-based nature.

The Köppen-Geiger climate classification is defined in terms of rules applied to variables derived from long-term monthly mean temperatures and precipitation amounts (see Table 1). Other manual classifications, for example the "rational classification" of Thornthwaite (1948), follow a similar rule-based approach, although the climate variables are not as conceptually simple or as widely monitored as those

Table 1. Köppen-Geiger climate variables and classification rules (after Peel et al., 2007 and Kottek et al., 2006). Rules for climate class E must be applied first, then those for class B, and finally those for classes A, C, and D. For locations that satisfy both Cs and Cw (or Ds and Dw) rules, the w classification is applied if summer precipitation exceeds winter precipitation. In the following, temperature variables are measured in °C, precipitation variables are measured in mm, and summer/winter is defined as the warmer/cooler of the six month period of October-March/April-September. MAP = mean annual precipitation; MAT = mean annual temperature; T_{hot} = temperature of the hottest month; T_{cold} = temperature of the coldest month, T_{mon10} = number of months above 10 °C; P_{dry} = precipitation of the driest month; P_{sdry} = precipitation of the driest month in summer; P_{wdry} = precipitation of the driest month in winter; P_{swet} = precipitation of the wettest month in summer; P_{wwet} = precipitation of the wettest month in winter; $P_{threshold}$ = precipitation threshold, which varies according to the following rules: if 70 % of MAP occurs in winter, then $P_{threshold} = 2 \times$ MAT, if 70 % of MAP occurs in summer, then $P_{threshold} = 2 \times$ MAT + 28, otherwise $P_{threshold} = \times$ MAT + 14.

1st	2nd	3rd	Description	Rule
A			Tropical	$T_{cold} \geq 18$
	f		– Rainforest	$P_{dry} \geq 60$
	m		– Monsoon	Not(Af) & $P_{dry} \geq 100 -$ MAP/25
	w		– Savannah	Not(Af) & $P_{dry} < 100 -$ MAP/25
B			Arid	MAP $< 10 \times P_{threshold}$
	W		– Desert	MAP $< 5 \times P_{threshold}$
	S		– Steppe	MAP $\geq 5 \times P_{threshold}$
		h	– Hot	MAT ≥ 18
		k	– Cold	MAT < 18
C			Temperate	$T_{hot} \geq 10$ & $0 < T_{cold} < 18$
	s		– Dry Summer	$P_{sdry} < 40$ & $P_{sdry} < P_{wwet}/3$
	w		– Dry Winter	$P_{wdry} < P_{swet}/10$
	f		– Without Dry Season	Not(Cs, Cw)
		a	– Hot Summer	$T_{hot} \geq 22$
		b	– Warm Summer	Not(a) & $T_{mon10} \geq 4$
		c	– Cold Summer	Not(a, b)
D			Cold	$T_{hot} \geq 10$ & $T_{cold} \leq 0$
	s		– Dry Summer	$P_{sdry} < 40$ & $P_{sdry} < P_{wwet}/3$
	w		– Dry Winter	$P_{wdry} < P_{swet}/10$
	f		– Without Dry Season	Not(Ds, Dw)
		a	– Hot Summer	$T_{hot} \geq 22$
		b	– Warm Summer	Not(a) & $T_{mon10} \geq 4$
		c	– Cold Summer	Not(a, b, d)
		d	– Very Cold Winter	Not(a, b) & $T_{cold} < -38$
E			Polar	$T_{hot} < 10$
	T		– Tundra	$T_{hot} > 0$
	F		– Frost	$T_{hot} \leq 0$

in the Köppen-Geiger system. Thornthwaite, for example, includes potential evapotranspiration as a key classification variable. With the advent of the computer, however, climate classifications have typically been performed using automated clustering algorithms (see, for example, Fovell and Fovell, 1993). While a cluster analysis may define a set of well-separated, homogeneous climate regions, the simplicity and interpretability of traditional rule-based methods is, to an extent, sacrificed. Determining rules that define climate classes obtained using cluster analysis techniques is often not possible, and, as a result, clear interpretation of the results can be somewhat challenging. This is the case because most

automated clustering methods, including the majority of hierarchical and partitioning cluster analysis techniques are polythetic: i.e., cases are assigned to clusters based on information from all variables (Kaufman and Rousseeuw, 1990). In contrast, the use of simple rules in traditional climate classifications makes their interpretation relatively straightforward.

While less common than polythetic algorithms, monothetic algorithms have also been used for cluster analysis. Monothetic algorithms separate cases via a hierarchy of decision rules. Each new rule is defined using a single variable and adds a cluster to the analysis; the combined set of rules can be used to assign each case to exactly one cluster.

This style of cluster analysis is much closer in spirit to traditional climate classifications than polythetic methods. The most common monothetic algorithm is, however, restricted to clustering binary data; Williams and Lambert (1959) designed association analysis, a divisive hierarchical algorithm, to cluster ecological data based on binary variables indicating the presence or absence of species at sites. On the other hand, classification and regression trees (Breiman et al., 1984), more generally referred to as recursive partitioning trees, work in a similar manner but are more flexible and are not restricted to binary data. Despite being conceptually similar to association analysis, recursive partitioning models have generally been restricted to supervised classification or regression tasks. In such examples, variability in a predictand variable is explained by splitting cases according to rules defined using a separate set of predictor variables. Multivariate regression trees (MRTs), which are recursive partitioning models with multiple predictands, have also been proposed and applied to a diverse set of prediction problems (Segal, 1992; Chavent, 1998; Yu and Lambert, 1999; De'ath, 2002), including those in synoptic climatology (Cannon et al., 2002; Cannon, 2008, 2012).

The MRT is, in theory, very well suited to climate classification. It is automated, which speeds development and testing; it is hierarchical, which allows a series of nested classes to be defined; and it is rule-based, which allows climate classes to be unambiguously defined and easily interpreted. The goals of this study are to develop a rule- and computer-based global climate classification using the MRT algorithm and to compare its performance against the Köppen-Geiger system in terms of climate homogeneity. Given that the original intent of the Köppen-Geiger system involved discrimination of vegetation types, this could be restated as: "what level of performance in terms of climate discrimination could be gained if a rule-based classification were developed exclusively with climate variables?"

2 Köppen-Geiger climate classification and data

The rules that define the Köppen-Geiger climate classification are applied to variables derived from long-term climate normals of monthly mean temperatures and precipitation amounts. Version 1.4 (release 3) of the WorldClim 1950–2000 climate normals. Station observations and methods used to develop the dataset are described by Hijmans et al. (2005). Although WorldClim consists of globally gridded climate normals at a resolution of 30 arc s, the 10 arc min version of the dataset ($>500\,000$ grid points) is used here to lighten the computational burden of running the classifications.

The Köppen-Geiger climate classification used in this paper follows the rules given by Peel et al. (2007) with the order of application recommended by Kottek et al. (2006). Table 1 lists the classification rules and definitions of the 10 climate

variables (MAP, MAT, T_{hot}, T_{cold}, T_{mon10}, P_{dry}, P_{sdry}, P_{wdry}, P_{swet}, and P_{wwet}) used in the classification system. The system is hierarchical. The first level has 5 classes, the second has 13 classes, and the third has 30 classes.

3 Multivariate regression tree

A MRT is structured as a binary tree with nodes defined by simple decision rules applied to predictors X_{il}, where i and l are indexes over cases and predictor variables respectively. All cases start out assigned to a single node. Cases in this top-level node are divided into two groups by a decision rule defined by one of the predictors (e.g., MAT $< 12\,°\text{C}$). Depending on the outcome of the decision, cases follow one of the two branches from the node. New decision rules are created at nodes by a splitting algorithm until one or more stopping criteria are met. Cases are assigned to classes based on the terminal (unsplit) nodes they reach in the tree.

A splitting algorithm selects the decision variables and their thresholds. At each step it is responsible for determining the decision rule that will best partition the remaining cases into classes that are as homogeneous as possible. Homogeneity is measured with respect to a set of predictands Y_{ij}, where j is an index over the predictand variables. This requires a quantitative measure of error (or heterogeneity) for nodes in the tree to be defined. Each node is characterized by the multivariate predictand mean, i.e., the centroid, of its assigned cases

$$Y_{\bullet jk} = \frac{1}{N_k} \sum_{i=1}^{N_k} Y_{ijk}. \tag{1}$$

where Y_{ijk} is the value of the j-th of J predictand variables for the i-th of N_k cases assigned to the k-th node.

The error measure for the k-th node is the within-cluster sums of squared deviations from the centroid

$$\text{WSS}_k = \sum_{j=1}^{J} \sum_{i=1}^{N_k} \left(Y_{ijk} - Y_{\bullet jk} \right)^2 \tag{2}$$

and the overall error for a tree is calculated by summing the WSS_k values over the K terminal nodes

$$\text{WSS} = \sum_{k=1}^{K} \text{WSS}_k. \tag{3}$$

To build the tree, splits are chosen to maximize the decrease in WSS_k between the existing parent node WSS_A and the new child nodes WSS_B and WSS_C

$$\Delta\text{WSS} = \text{WSS}_\text{A} - (\text{WSS}_\text{B} + \text{WSS}_\text{C}). \tag{4}$$

The search for the best split is exhaustive, considering each value of the predictor variables as a potential threshold. Data are split until each terminal node contains a preset minimum number of cases, until no further splits can be made because
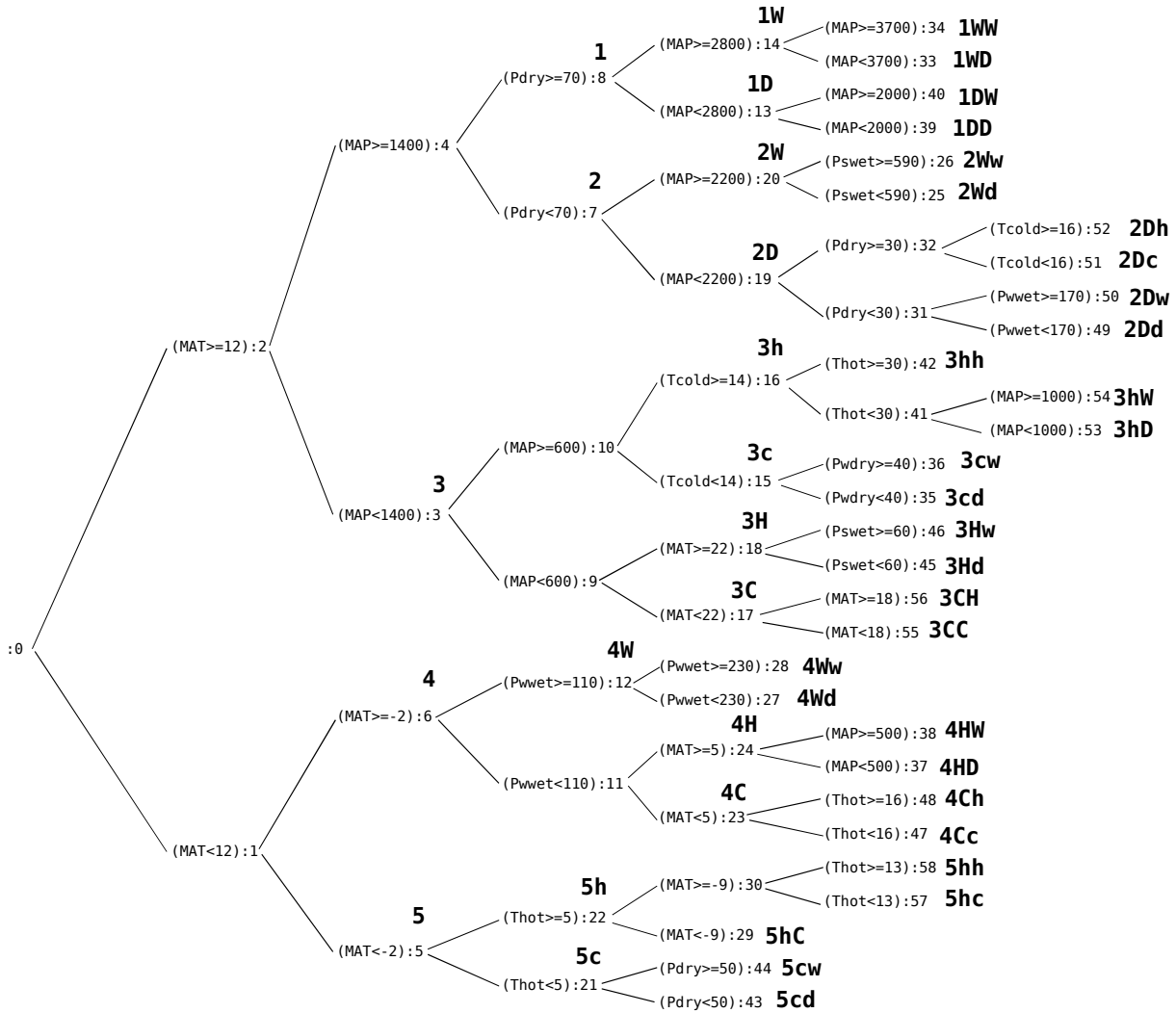
Fig. 1. Binary tree representation of the MRT climate classification rules. Climate variable names are given in Table 1. At each node, the decision rule is listed preceding the colon, while the order in which binary split was made is denoted by the integer value following the colon. Class names are given in bold. The 5 class level is assigned a number from 1–5. The 13 class and 30 class levels are denoted by adding letters to the numbers. If the rule immediately preceding the terminal node involves a precipitation variable, aW, w, D, or d is added, where the W or w indicates wetter conditions and the D or d indicates drier conditions. If the preceding rule involves a temperature variable, aH, h, C, or c is added, where the H or h indicates hotter conditions and the C or c indicates cooler conditions. Upper case letters are reserved for mean annual variables, whereas lower case letters are used for annual extreme or seasonal extreme variables. A map of class locations is given in Fig. 2

.

the error measure has been minimized, or until a desired level of complexity has been reached. A hierarchy of nested subtrees is created through this procedure. Once a tree has been built, the proportion of explained predictand variance EV can be calculated as

$$EV = 1 - \frac{WSS}{WSS_T} \tag{5}$$

where WSS_T is the total within-cluster sums of squared deviations for the top-level node.

Splits in a MRT can be used to gain insight into the importance of predictor variables. Direct analysis of the splits present in a tree may not, however, be informative, as variables that happen to be highly correlated with the splitting variables may not appear in the tree at all. Instead, Breiman et al. (1984) suggest another measure of variable importance calculated using the best splits and a set of alternate splits defined at each node. Once the best split has been found for a given node, surrogate splits using each of the remaining predictor variables are also investigated. Surrogate splits are
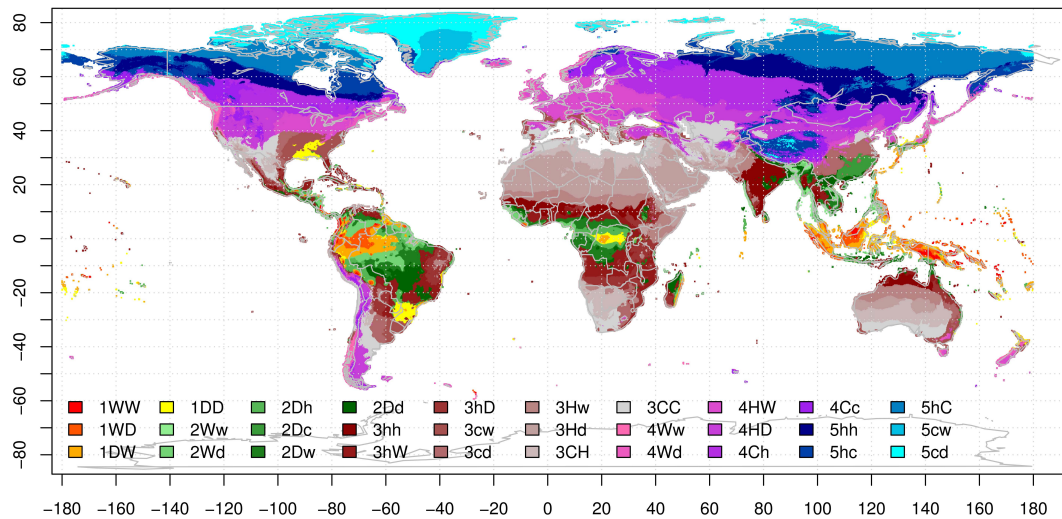
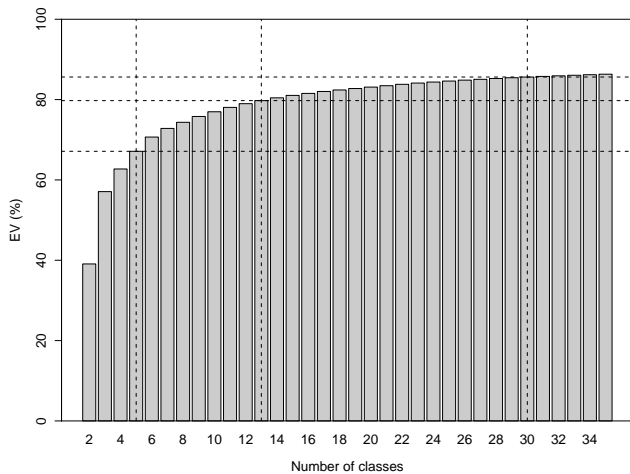**Fig. 2.** Global map of MRT climate classification. The naming convention is given in Fig. 1.

Fig. 3. Combined monthly temperature and precipitation EV for the MRT climate classification (2–35 classes). The 5, 13, and 30 class levels are indicated by dashed lines.
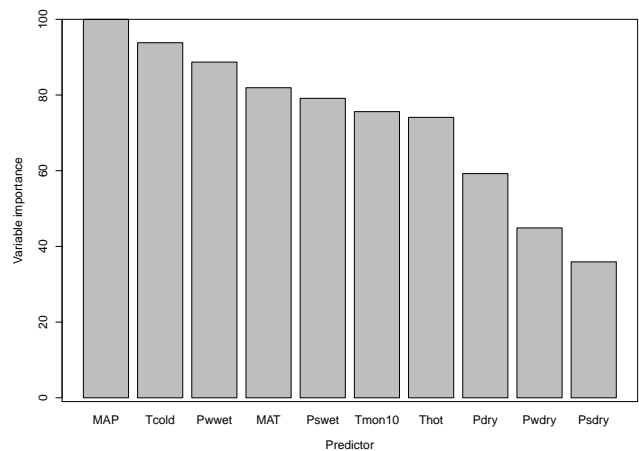

Fig. 4. MRT climate classification predictor importance calculated from ΔWSS values for primary and surrogate splits (see Sect. 3). The measure of predictor importance has been scaled to range from 0 (the predictor does not appear in any primary or surrogate splits) to 100 (the predictor is responsible for the greatest decrease in WSS).

chosen to mimic the behavior of the best split (i.e., send cases to the same nodes as the best split). Candidates are compared against a naive rule that sends data to the node that received the majority of cases from the best split. Splits that are better predictors than the naive rule are ranked and the best is selected as the surrogate split for a given variable. The ΔWSS value associated with the surrogate split is then used as the measure of variable importance at that node. Variable importance for the tree is taken as the sum of the ΔWSS values for surrogate splits and best splits at all nodes. Surrogate splits also allow cases to be assigned to clusters when variables involved in the primary splits are missing. The best surrogate rule for the non-missing variables is substituted for the primary rule. This allows cases with missing information to progress down the tree and be assigned to classes.

Once the tree has been created, new cases can be assigned to classes using the splits defined at the decision nodes. As mentioned above, each node is characterized by the centroid of cases assigned to it during fitting. These values serve as predictions for new cases, and estimates of predictive error, explained variance, etc. can be estimated for the model. This predictive aspect of the MRT is shared by some partitioning techniques, for example k-means clustering. Unlike these methods, however, the MRT only requires values to be measured for predictor variables involved in the decision rules. This ability to identify a subset of important predictor variables is another advantage of the MRT approach to clustering.
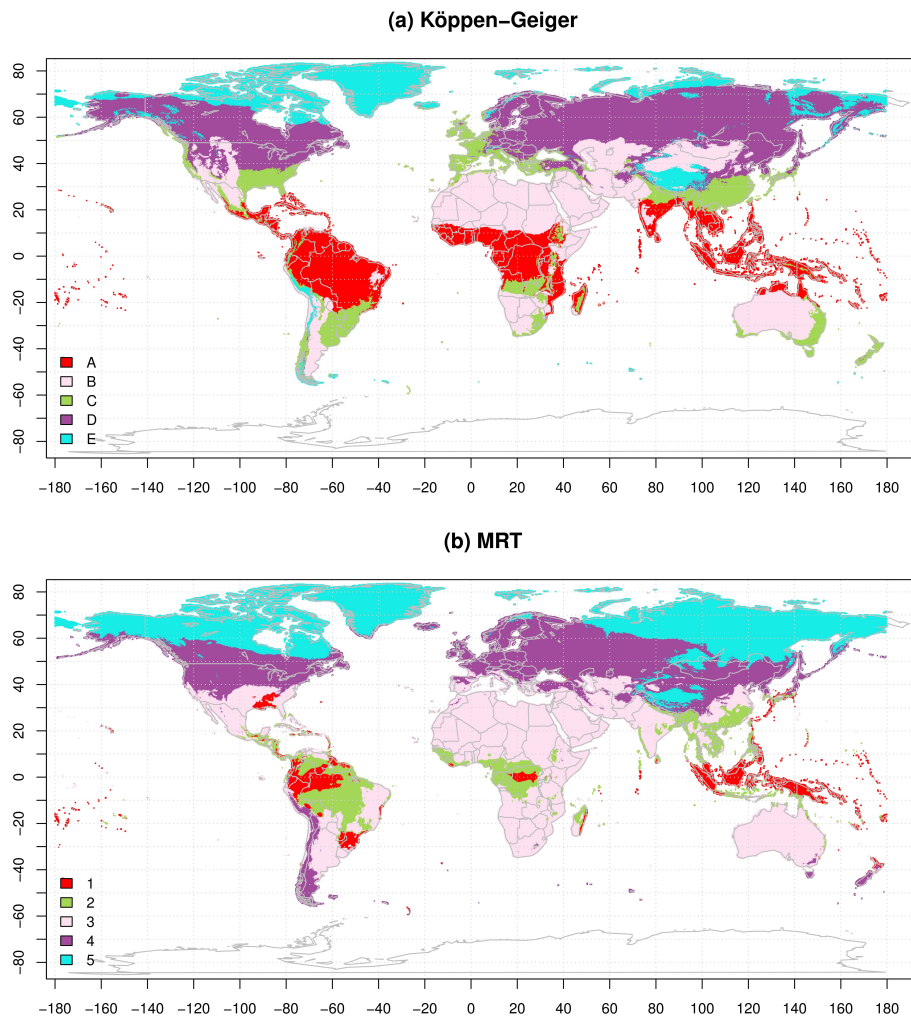
(a) Köppen–Geiger

(b) MRT

Fig. 5. Global maps of **(a)** Köppen-Geiger and **(b)** MRT climate classifications at the 5 class level.

4   MRT climate classification

The MRT algorithm is used to define a hierarchical, rule-based climate classification modelled after the Köppen-Geiger climate classification. To this end, the 10 climate variables from the Köppen-Geiger climate classification are used as predictor variables in the MRT model. For consistency with the Köppen-Geiger rules, which are based on rounded climate variables, MAT, T_{hot}, and T_{cold} are rounded to the nearest °C; MAP is rounded to the nearest 100 mm; and P_{dry}, P_{sdry}, P_{wdry}, P_{swet}, and P_{wwet} are rounded to the nearest 10 mm. As the ultimate goal is to separate the globe into climate classes with distinct intra-annual temperature and precipitation patterns, predictand variables are the 12 monthly mean temperature normals and the 12 monthly precipitation normals. For locations in the northern (southern) hemisphere, predictands are specified starting from January (July) and proceeding to December (June). To account for the poleward decrease in geographical area of grid points

in the WorldClim dataset, the Mollweide equal area projection is applied prior to classification. Without reprojection, a residual error at a gridpoint near the poles would be given more weight than the same error at a more equatorial latitude. Finally, predictands are rescaled to zero mean and unit standard deviation so that temperature and precipitation are given approximately equal weight in the calculation of WSS. (Note that while MRT models are fitted using computer code written by the author, results are verified using the mvpart package, Therneau et al., 2010 for the R programming language, R Development Core Team, 2009, which is free and open source software.)

The MRT algorithm is terminated after making 29 splits. While the MRT provides a complete hierarchy of solutions, in this case from 2 to 30 classes, three hierarchical levels (5, 13, and 30 classes) are defined for consistency with the Köppen-Geiger climate classification. The resulting MRT climate classification is summarized as a binary tree diagram in Fig. 1. Locations of climate classes are shown in Fig. 2.
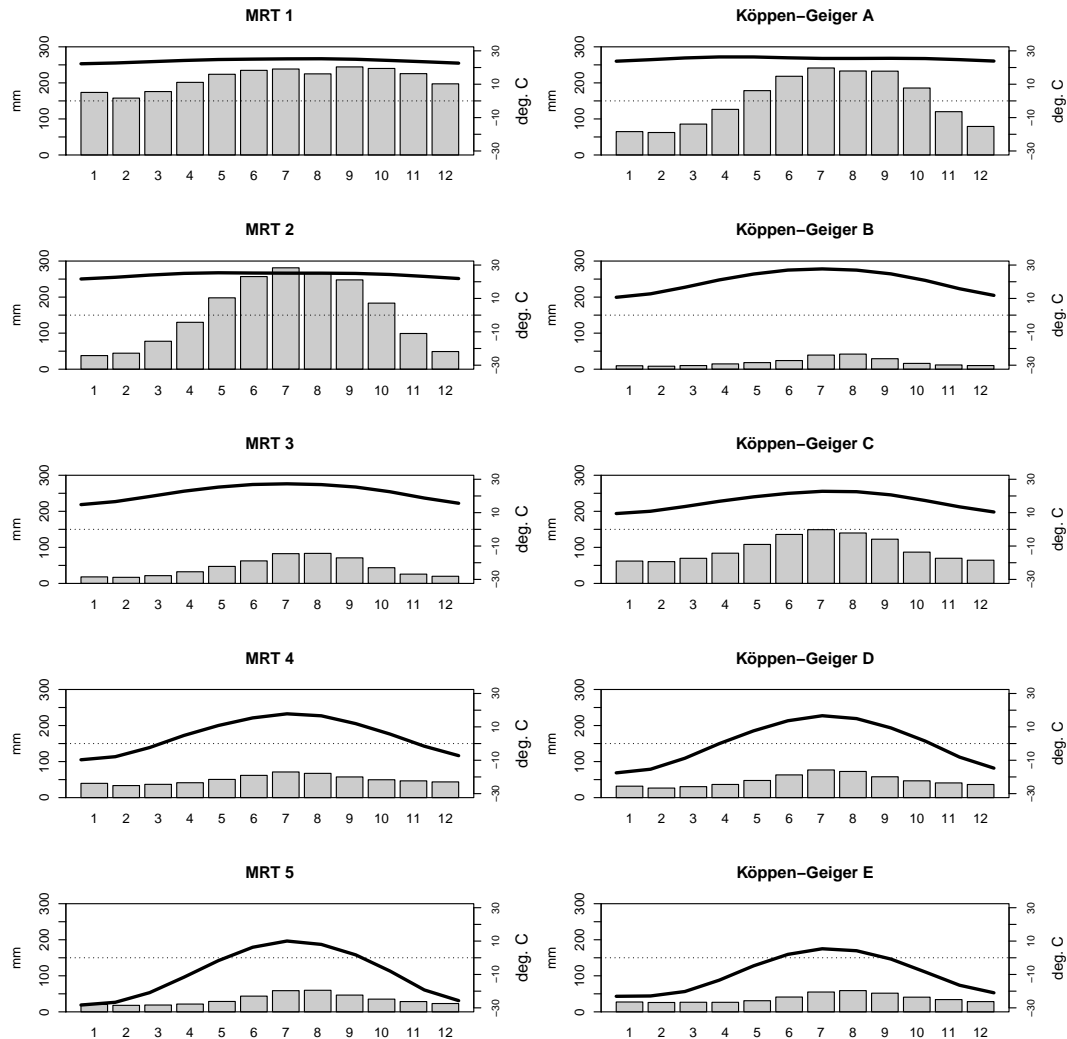
Fig. 6. Mean monthly temperatures (black line) and precipitation amounts (gray bars) for the Köppen-Geiger climate classification (right column) and MRT climate classification (left column) at the 5 class level.

Values of EV for each number of classes are shown in Fig. 3. The 5, 13, and 30 class levels explain, respectively, 67.1 %, 79.7 %, and 85.6 % of predictand variance. Little improvement would accrue with more classes (e.g., 35 classes results in 86.3 % EV) suggesting that the 30 class level is a reasonable level of complexity for the classification system.

All variables except for T_{mon10} are featured in primary splits of the 30 node MRT. When accounting for both primary and surrogate splits, all variables are assigned non-zero values of predictor importance, with MAP, T_{cold}, and P_{wwet} ranked as the three most important predictors in the MRT (Fig. 4).

5 Comparison of climate classifications

Before attempting to objectively compare the classification performance of the Köppen-Geiger and MRT climate classifications, it is instructive to consider global maps (Fig. 5), monthly distributions of temperature and precipitation (Fig. 6), and a cross-tabulation of classes (Table 2) at the simplest 5 class level. The two classifications bear some subjective similarities. When looking across classes, the largest cross-tabulation counts match for Köppen-Geiger class A (Tropical) and MRT class 2; classes B (Arid) and 3; and classes D (Cold) and 4. Monthly mean temperatures and precipitation amounts for these cases are, as expected, quite similar. On the other hand, while Köppen-Geiger class C (Temperate) matches best with MRT class 3, it also shows modest association with MRT classes 1, 2 and 4. Köppen-Geiger class E (Polar) is strongly associated with MRT class 5, but
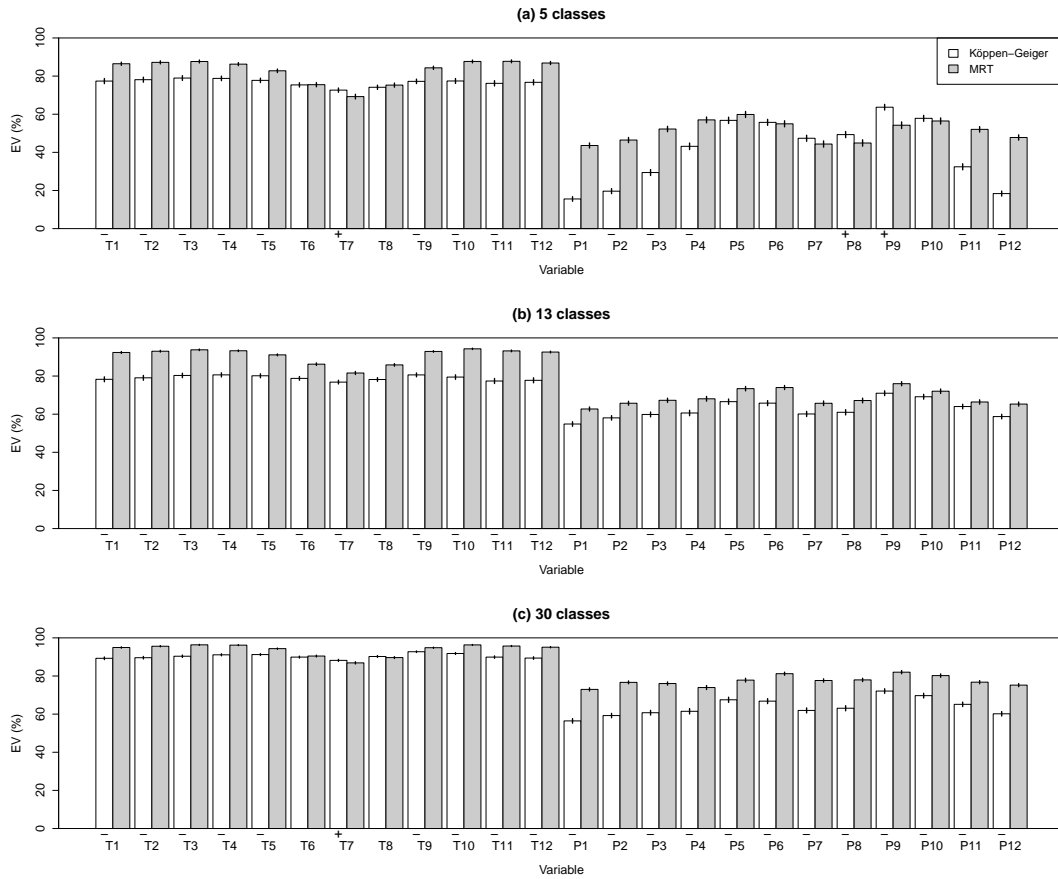
Fig. 7. EV (monthly temperature, T1 to T12, and precipitation, P1 to P12) of the Köppen-Geiger and MRT climate classifications at the **(a)** 5, **(b)** 13, and **(c)** 30 class levels. Error bars indicate 95 % confidence interval (CIs); ± signs along the horizontal axis indicate variables for which the 95 % CI for the Köppen-Geiger classification lies outside (above/below) the CI for the MRT classification.

Table 2. Cross-tabulation of grid cells assigned to the Köppen-Geiger (A to E) and MRT climate classifications (1 to 5) at the 5 class level. Italicized (bold) entries indicate the largest value in a given row (column).

Class	1	2	3	4	5	Total
A	**26 590**	**_48 930_**	36 058	0	0	111 578
B	0	0	**_128 058_**	23 795	442	152 295
C	6049	11 582	_41 749_	9724	0	69 104
D	7	25	3053	**_84 022_**	45 614	132 721
E	0	0	0	7023	_34 272_	41 295
Total	32 646	60 537	208 918	124 564	80 328	506 993

class 5 shows slightly better correspondence with Köppen-Geiger class D.

What is driving these differences? Of the 5 primary Köppen-Geiger classes, only class B is decided by a rule involving precipitation. The remaining 4 classes are discriminated in terms of temperature. Contrast this with the MRT, in

which rules leading to 3 of the 5 classes (classes 1, 2, and 3) feature precipitation variables and where the measure of predictor importance indicates that MAP is the most important discriminatory variable (Fig. 4). Are there then systematic differences in performance of the classification systems with respect to temperature and precipitation? If so, do these differences persist from the 5 class solution to the 13 and 30 class solutions?

To answer these questions, values of EV and their 95 % confidence intervals are calculated for the monthly temperature and precipitation variables at the 5, 13, and 30 class levels. Confidence intervals are based on an approximation of the standard error from Olkin and Finn (1995)

$$ \text{SE} = \sqrt{\frac{4}{n_{\text{eff}}}\, \text{EV}\,(1 - \text{EV})^2} \qquad (6) $$

in which the number of samples has been replaced by the effective number of samples n_{eff}, which takes into account spatial autocorrelation of nearest neighbours (Bretherton et al., 1999). Results are shown in Fig. 7. Averaged over the year, the EV for monthly temperatures is 76.7 %
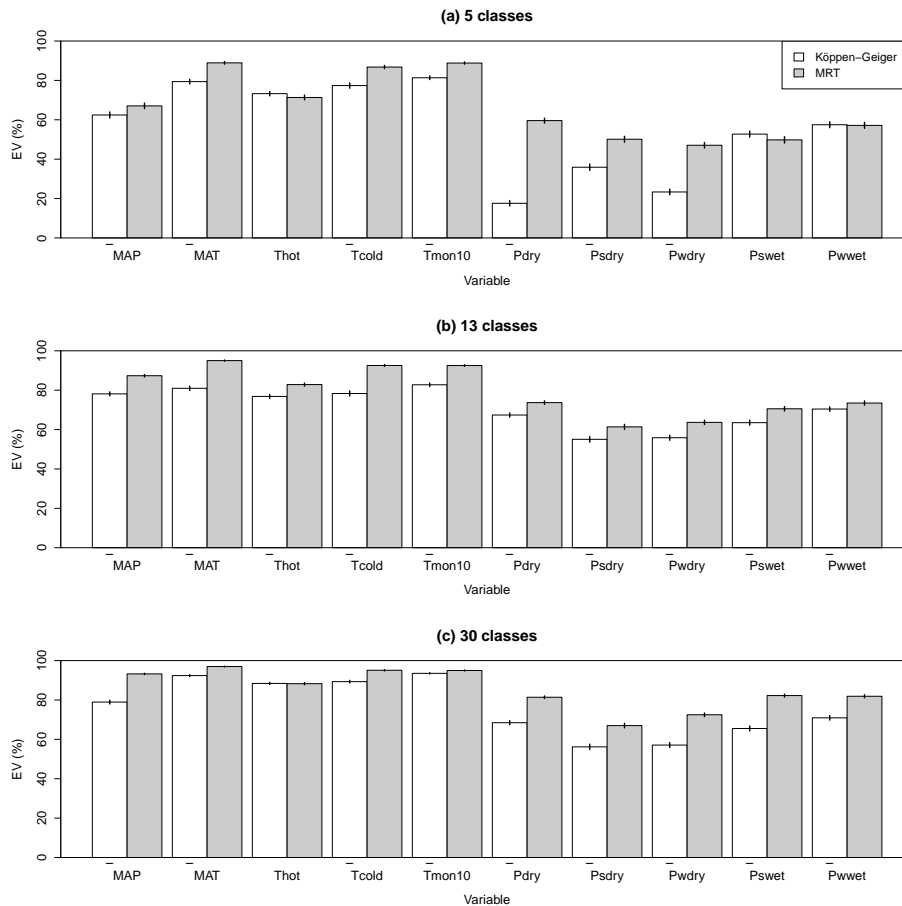
Fig. 8. As in Fig. 7, but for climate variables defined in Table 1.

(5 classes), 79.0 % (13 classes), and 90.3 % (30 classes) for the Köppen-Geiger classification, versus 83.1 % (+6.4 % difference), 90.8 % (+11.8 % difference), and 93.9 % (+3.6 % difference) for the corresponding MRT solutions. Values of EV for monthly precipitation are 40.8 %, 62.5 %, and 63.7 % for the Köppen-Geiger classification, versus 51.2 % (+10.4 %), 68.6 % (+6.1 %), and 77.4 % (+13.7 %) for the MRT. On a month by month basis, results for the MRT are generally significantly better than those for the Köppen-Geiger classification. In 60 of the 64 cases in which confidence intervals do not overlap, results indicate better performance by the MRT. Notable exceptions include better discrimination of northern hemisphere July (southern hemisphere December) temperatures (5 and 30 classes) and northern hemisphere August and September (southern hemisphere January and February) precipitation amounts (5 classes) for the Köppen-Geiger classification.

One potential criticism of the analysis above is that the values of EV are calculated over the monthly variables used as predictands in the MRT. As a result, the MRT may hold an unfair advantage over the Köppen-Geiger classification. While this is somewhat difficult to avoid, one alternative

is to calculate values of EV for the suite of 10 variables used to define the classification rules for both the Köppen-Geiger and MRT classifications. Since these variables are derived from the monthly variables, the problem, to an extent, remains. Nevertheless, results are shown in Fig. 8. The Köppen-Geiger classification performs worse than the MRT in all 26 instances in which significant differences are found. To illustrate the consequences of differences in EV between the two classification systems, Fig. 9 shows scatterplots of observed MAT and MAP versus predicted class centroids at the 30 class hierarchical level. As expected, spread about the one-to-one line is similar between the classification systems for MAT, but the MRT does a much better job at discriminating wet from dry climates than does the Köppen-Geiger classification. For example, the 99.9th percentile of observed MAP exceeds 4800 mm (maximum of 8000 mm), while MAP for the wettest Köppen-Geiger class centroid is less than 2700 mm. The corresponding value for the wettest MRT class exceeds 4300 mm.

Finally, to provide a more level playing field, restricted MRT climate classifications – ones in which a continent is removed during model fitting – are also developed. Following
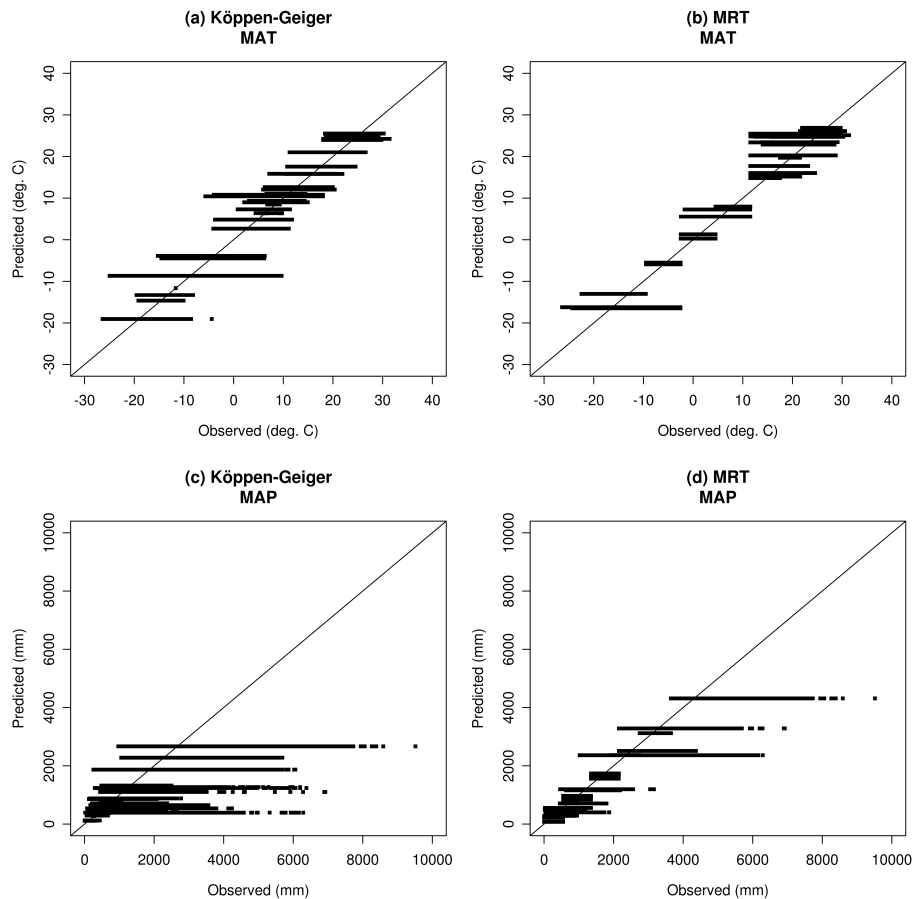
Fig. 9. Scatterplots of observed MAT and MAP versus predicted (a, c) Köppen-Geiger and (b, d) MRT class centroids.

model fitting, withheld points are classified according to the MRT rules and predictions are made for monthly temperatures and precipitation amounts. Average values of EV over all months are then calculated and compared against those for the classifications based on all grid points. As rules and centroids for both the Köppen-Geiger and original MRT classes are calculated based, in part, on data from the withheld continents, the restricted MRT models are at a substantial disadvantage in the "leave-one-continent-out" validation. Mean results for monthly temperature and precipitation at the 30 class level are shown in Fig. 10 for North America, South America, Africa, and Australia. While, as expected, the restricted MRT models perform worse than the original MRT, they typically explain more variance than the Köppen-Geiger classification. For temperature, differences in EV between the restricted MRT and the Köppen-Geiger classification are +2.7 %, −2.8 %, +13.1 %, and +0.9 % for North America, South America, Africa, and Australia respectively; for precipitation, the respective differences are +17.3 %, +6.9 %, +14.8 %, and +1.9 %.

6   Summary and conclusions

A global climate classification is defined using a MRT model – an automated, hierarchical, and rule-based clustering algorithm – and results are compared against the Köppen-Geiger system in terms of its ability to delineate homogeneous temperature and precipitation regions. Given that the Köppen-Geiger system was not originally designed as a general climate classification, but rather to differentiate climate classes based on vegetation groups, its performance is found to be extremely good, especially for temperature. The overall analysis of classification performance does, however, show that the MRT performs significantly better than the Köppen-Geiger classification, in particular for measures of precipitation. The Köppen-Geiger classification may thus be suboptimal for applications that are sensitive to spatial variations in precipitation, for example explaining differences in hydrologic variability (e.g., Peel et al., 2001). Likewise, assessments of climate change using the Köppen-Geiger system may be overly sensitive to changes in temperature rather than precipitation.

While this paper focused exclusively on an evaluation of climate classification systems in terms of monthly
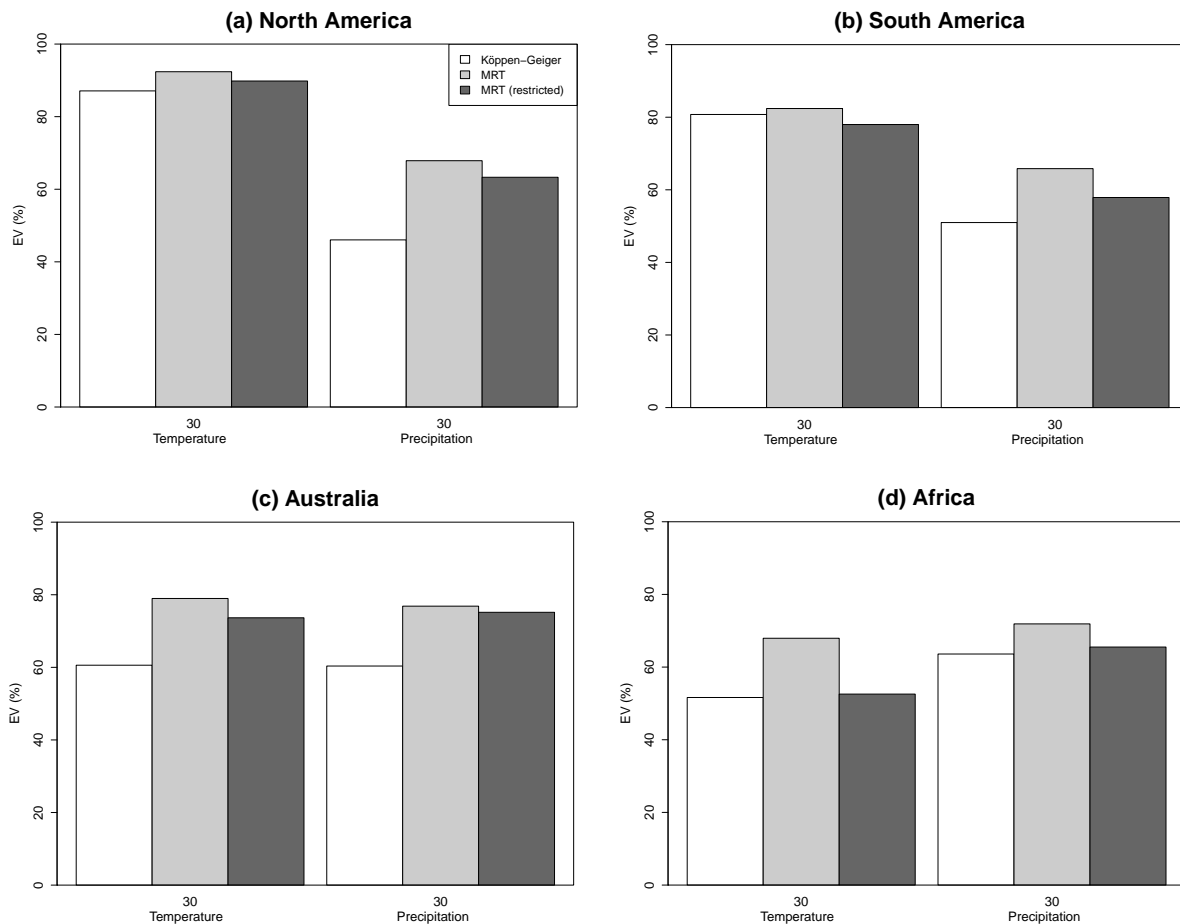
Fig. 10. Average monthly EV values for Köppen-Geiger, MRT, and restricted MRT climate classifications for temperature and precipitation variables at grid **(d)** Africa.

temperature and precipitation, the ability of the MRT technique to easily incorporate other sources of relevant information means that its classifications can be customized for a particular study or goal. For example, an MRT analogue of the Thornthwaite rational classification could be constructed simply by using measures of potential evapotranspiration as target variables.

For simplicity, the MRT climate classifications in this paper only consider binary splits based on a single boolean operator. Linear combinations of predictor variables can be used in MRT splits, although this would tend to lead to a less easily interpreted tree structure. The trade-off between interpretability and performance has not been fully explored. Similarly, the "greedy" optimization of splitting rules in a MRT can lead to a sub-optimal classification. As each split in the tree is chosen only to minimize the squared error at a particular node, the final tree is not guaranteed to reach the global minimum value of WSS. One way of improving the final solution is to revise previous splits as new splits are added. Chavent (1998) introduced a single-level revision step that adjusts the previously added split so that the WSS

of the affected classes is minimized. While this improves the final solution, there is still room for improvement as the search only considers the previous split, not all splits in the tree. Alternatively, one can either adjust all splits after a tree is built or build a tree from scratch using some form of global search algorithm. Such options are left as avenues for future research.

Note also that the final partitioning of cases into classes by a MRT is insensitive to monotonic transformations of the predictor variables. For instance, cases will still be sent down the same branches of the tree regardless of the units in which variables are measured. Scaling of predictand variables (and, similarly, redundancy between predictands), however, will affect the final classification. Fovell and Fovell (1993) and Mimmack et al. (2001) discuss biases related to scaling and redundancy in climatological cluster analyses. Redefinition of predictands using methods such as truncated principal component analysis may be warranted, depending on the problem. The MRT is somewhat unique because variables can be scaled or transformed differently depending on whether or not they are acting as predictors or predictands.

For example, variables expressed in their original units could be used as predictors and a set of standardized principal components of the same variables could be used as predictands in the classification. In this study, monthly temperature and precipitation predictand variables are standardized to zero mean and unit standard deviation, but no further attempt is made to reduce redundancy or differentially weight temperature and precipitation.

The goal and format of this paper prevents a full exploration of the MRT climate classification, especially at the 13 and 30 lass levels. As a means of fostering further work, Walter-Lieth climate diagrams for each class (Heinrich and Helmut, 1967; Guijarro, 2011) and a gridded global dataset of the system are provided as supplementary material. It is hoped that the results will be explored further using these resources.

Finally, data and software packages mentioned in this study are freely available, which means that they can also be employed by instructors and students as a means of interactively exploring modern, computer-based climatological methods. Irrespective of whether or not the MRT climate classification outlined here gains any measure of "historical inertia", it is hoped that the underlying tools and datasets will provide a rich environment for exploration, both by researchers and students.

Supplementary material related to this article is available online at: http://www.hydrol-earth-syst-sci.net/16/217/2012/ hess-16-217-2012-supplement.zip.

References

Beck, C., Grieser, J., Kottek, M., Rubel, F., and Rudolf, B.: Characterizing global climate change by means of Köppen climate classification, Climate Status Report 2005, Deutscher Wetterdienst, 139–149, 2005.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: Classification and Regression Trees, The Wadsworth statistics/probability series, Wadsworth International Group, Belmont, CA., 1984.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., and Blade, I.: The effective number of spatial degrees of freedom of a time-varying field, J. Climate, 12, 1990–2009, 1999.

Cannon, A. J.: Probabilistic multisite precipitation downscaling by an expanded Bernoulli-gamma density network, J. Hydrometeorol., 9, 1284–1300, doi:10.1175/2008jhm960.1, 2008.

Cannon, A. J.: Semi-supervised multivariate regression trees: putting the "circulation" back into a "circulation-to-environment" synoptic classifier, Int. J. Climatol., doi:10.1002/joc.2417, in press, 2012.

Cannon, A. J., Whitfield, P. H., and Lord, E. R.: Automated, supervised synoptic map-pattern classification using recursive partitioning trees, in: Preprints, The 16th Conference on Probability and Statistics in the Atmospheric Sciences, American Meteorological Society, Boston, J210–J216, 2002.

Chavent, M.: A monothetic clustering method, Pattern Recog. Lett., 19, 989–996, 1998.

De'ath, G.: Multivariate regression trees: a new technique for modeling species-environment relationships, Ecology, 83, 1105–1117, 2002.

Fovell, R. G. and Fovell, M. Y. C.: Climate zones of the conterminous United States defined using cluster-analysis, J. Climate, 6, 2103–2135, 1993.

Guijarro, J. A.: climatol: some tools for climatology: series homogenization, plus windrose and Walter & Lieth diagrams, http://CRAN.R-project.org/package=climatol, last access: 28 April 2011, R package version 2.0, 2011.

Heinrich, W. and Helmut, L.: Klimadiagramm-Weltatlas, VEB Gustav Fischer Verlag, Jena, 1967.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol., 25, 1965–1978, doi:10.1002/Joc.1276, 2005.

Kaufman, L. and Rousseeuw, P. J.: Finding Groups in Data: An Introduction to Cluster Analysis, Wiley-Interscience, New York, 1990.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World map of the Köppen-Geiger climate classification updated, Meteorol. Z., 15, 259–263, 2006.

Lohmann, U., Sausen, R., Bengtsson, L., Cubasch, U., Perlwitz, E., and Roeckner, E.: The Köppen climate classification as a diagnostic tool for general circulation models, Clim. Res., 3, 177–193, 1993.

Mimmack, G. M., Mason, S. J., and Galpin, J. S.: Choice of distance matrices in cluster analysis: Defining regions, J. Climate, 14, 2790–2797, 2001.

Olkin, I. and Finn, J. D.: Correlations redux, Psychol. Bull., 118, 155–164, 1995.

Peel, M. C., McMahon, T. A., Finlayson, B. L., and Watson, F. G. R.: Identification and explanation of continental differences in the variability of annual runoff, J. Hydrol., 250, 224–240, 2001.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci., 11, 1633–1644, doi:10.5194/hess-11-1633-2007, 2007

R Development Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org, last access: 28 April 2011, ISBN 3-900051-07-0, 2009.

Segal, M. R.: Tree-structured methods for longitudinal data, J. Am. Stat. Assoc., 87, 407–418, 1992.

Therneau, T. M., Atkinson, B., Ripley, B., Oksanen, J., and De'ath, G.: mvpart: multivariate partitioning, http://CRAN.R-project. org/package=mvpart, last access: 28 April 2011, R package version 1.3-1, 2010.

Thornthwaite, C. W.: An approach toward a rational classification of climate, Geogr. Rev., 38, 55–94, 1948.

Wilcock, A. A.: Köppen after fifty years, Ann. Assoc. Am. Geogr., 58, 12–28, 1968.

Williams, W. T. and Lambert, J. M.: Multivariates methods in plant ecology I: association-analysis in plant communities, J. Ecol., 47, 83–101, 1959.

Yu, Y. and Lambert, D.: Fitting trees to functional data, with an application to time-of-day patterns, J. Comput. Graph. Stat., 8, 749–762, 1999.