Coherence of Global Hydroclimate Classification Systems

Kathryn L. McCurley Pisarello[1], James W. Jawitz[1]

[1]Soil and Water Sciences Department, University of Florida, Gainesville, Florida 32611, USA

Correspondence to: James W. Jawitz (jawitz@ufl.edu)

Abstract. Climate classification systems are useful for investigating future climate scenarios, water availability, and even socioeconomic indicators as they relate to climate dynamics. These classification systems typically utilize various forms of water and energy indicators to create zone boundaries. However, there has yet to be a classification framework that includes evapotranspiration (ET) rates as a governing principle, nor has there been an effort to simultaneously compare the structure and function of multiple existing classification schemes. Here, we developed three new classification systems based on ET rates and one new system based on precipitation and potential evapotranspiration, and we compared these four new systems against four previously established climate classification systems. The within-zone similarity, or coherence, of long-term water budget components was evaluated for each system based on the premise that the application of a climate classification framework should correspond to those variables that are most coherent. Additionally, the complexity of zone boundaries in each system was assessed. The most frequently used system, Köppen-Geiger, had high hydroclimate coherence but also high spatial complexity. This study produced classification systems of improved coherence for individual water budget components, lower spatial complexity, and fewer parameters needed for their construction. The Water-Energy Clustering classification system is the primary framework proposed here for future investigations in which regions of interest include zones of differing hydrologic dynamics.

1 Introduction

A variety of classification schemes have been introduced to categorize specific biophysical characteristics of Earth systems, including those based on climatic behavior (Beck et al., 2018; Berghuijs and Woods, 2016; Holdridge, 1967), biodiversity (Olson et al., 2001), plant-climate interactions (Papagiannopoulou et al., 2018), or plant hardiness (Magarey et al., 2008; McKenney et al., 2007). These frameworks classify elements of a system based on common atmospheric or terrestrial characteristics to maximize their within-zone similarity, or coherence, which allows for a transfer of understanding across zones of similar attributes (Lanfredi et al., 2019). This study focuses specifically on climate classification schemes, which have provided a climatic context for a variety of applications, including socioeconomic assessments of human health conditions (Boland et al., 2017; Jagai et al., 2007; Lloyd et al., 2007), economic development (Mellinger et al., 2000; Richards et al., 2019), and evaluating anticipated biophysical and climatic changes (Chen and Chen, 2013; Tapiador et al., 2019).

Hydrology and
Earth System
Sciences
Discussions

Different climate classification systems have emerged based on framework-specific suites of hydroclimatic variables

30   used to define the climate zone boundaries. Therefore, users should consider how potential classification system application corresponds to the variables used to create it (Knoben et al., 2018; Meybeck et al., 2013). Climate classification systems are usually based in part on annual and seasonal water-energy budgets (Beck et al., 2018; Berghuijs and Woods, 2016; Holdridge, 1967; Knoben et al., 2018; Meybeck et al., 2013). The Köppen-Geiger classification system, the most widely used climate framework (Peel et al., 2007), was developed to regionalize climatic variables (specifically accounting for seasonal

35   precipitation and temperature) and is often employed to compare the output of global climate models (Peel et al., 2007; Tapiador et al., 2019). Another common system is the Holdridge Life Zones scheme, which was created to classify land area with respect to vegetation and soil (Holdridge, 1967). This system subdivides zones based on thresholds of annual precipitation (P), potential evapotranspiration (PET), biotemperature (growing season length and temperature), and latitude and altitude.

Recent work has extended climate classification frameworks to specifically encompass hydrological attributes, since

40   water resources-based analyses should take place within relevant hydrologic boundaries (Knoben et al., 2018; Meybeck et al., 2013). For example, Meybeck et al (2013) proposed a global zoning system that was primarily based on the mean temperatures and average locally generated runoff (Q) of river basins. They compared the resulting boundaries against the Köppen-Geiger and Holdridge frameworks to assess zone overlaps. Those authors also evaluated the within-zone coherence of mean annual temperature, P, and Q, concluding that the latter two were most coherent in dry zones and least coherent in equatorial zones,

45   while temperature was most coherent in equatorial zones. However, Meybeck et al. (2013) did not compare their zone coherence to that of previously established systems. Similarly, Knoben et al. (2018) formed zone boundaries based on climate indices (average aridity, seasonality of aridity, and P as snow) with the objective of minimizing within-zone Q variability (i.e., maximizing Q coherence). Those authors compared their results to the Köppen-Geiger framework and found theirs to be more coherent with respect to Q, but they did not evaluate other water budget components or other climate classification systems.

50   Although the P and Q components of the long-term water budget have been extensively considered in climate classification schemes (Beck et al., 2018; Berghuijs and Woods, 2016; Holdridge, 1967; Knoben et al., 2018; Meybeck et al., 2013), notably absent is a system that is directly based on actual evapotranspiration (ET) rates. This gap is likely because ET traditionally has been the least empirically identified element of regional to global water budgets (Zhang et al., 2016). Moreover, there has been no comparison of within-zone hydroclimate coherence across climate classification systems, with

55   evaluation particularly lacking in considering ET rates. Furthermore, the spatial complexity of climate classification systems has not been systematically examined. Assessing the spatial structure of a biophysical system is a concept that most notably originates from landscape ecology (O'Neill et al., 1988), which provides a suite of shape metrics that can be cross-disciplinarily applied. Quantifying shape pattern and spatial configuration of climate classification systems is important for understanding the interactions between governing hydroclimatic characteristics.

60   This work seeks to provide empirical support for application-dependent selection among possible climate classification systems. We suggest that a classification system should have high within-zone coherence for variables that are related to the system's intended use, as well as relatively low shape complexity across zones, which is useful for ease of

interpretation within management and policy contexts. As such, an overarching hypothesis was postulated that for a given climate classification system, within-zone hydrologic coherence and inter-zone shape complexity will be closely related to the

65   organizing principle of that system. For example, the Köppen-Geiger and Meybeck et al. (2013) systems are based in large part on P and Q, respectively, and therefore these systems should show high coherence for these variables. Similarly, zone shape complexity will be lower in classification systems that include spatial contiguity in the organizing criteria (e.g., Meybeck et al., 2013). Given the major gap regarding the inclusion of ET in climate classification systems, we also propose a series of ET-based global classifications that should yield comparatively higher ET coherence than other systems.

70   We tested our hypothesis by evaluating within-zone coherence of long-term water budget components (mean annual ET, P, and Q) and synchronous P and PET seasonality, as well as zone shape complexity for our four new global classification systems, and further compared these against four previously established systems (Beck et al., 2018; Holdridge, 1967; Knoben et al., 2018; Meybeck et al., 2013). The primary zone shape complexity metrics were zone area and zone fragmentation (i.e., number of patches comprising each zone). This work presents novel approaches to determine appropriate applications and

75   boundary complexities of classification frameworks. Understanding the relevance of a climate classification system is important since such frameworks are used in multi-disciplinary contexts to examine hydrological, ecological, and societal phenomena.

2 Methods

2.1 Database construction

80   We evaluated global gridded monthly P and PET and mean annual ET and Q between 1980 and 2018 at a 0.5° x 0.5° spatial resolution. The Climate Research Unit TimeSeries V4.04 supplied monthly P and PET (Harris et al., 2020), while mean annual ET and Q were constructed from aggregated TerraClimate daily data (Abatzoglou et al., 2018). Long-term mean values were used to mute interannual variability. Annual ET and Q were resampled from their original 1/24° x 1/24° resolution to the 0.5° x 0.5° resolution of P and PET. Spatial analysis R packages *raster* (Hijmans, 2017)*, sp* (Bivand et al., 2013) and *ncdf4*

85   (Pierce, 2017) were used to build the database of long-term monthly and annual averages. The spatial extent of this study comprised all global land areas, excluding Antarctica, which resulted in a total of 61,701 pixels.

2.2 Sinusoidal functions as a descriptor of seasonality

      The seasonal dynamics of monthly P and PET were also considered in this analysis, as they are also included in the KPG framework, which considers temperature as a general proxy for PET (Beck et al., 2018). Sine functions, and their

90   corresponding parameters, can be used to describe intra-annual climate behavior. Sine functions were fitted to the long-term monthly distribution (following Berghuijs and Woods, 2016)

$$y_m = \bar{y}\left[1 + r_y \sin\left(\frac{2\pi(m-\theta_y)}{12}\right)\right] \tag{1}$$

95   where y is P or PET (mm month^{-1}) for each month, m, with monthly mean denoted by the overbar, and r is dimensionless amplitude. Phase angle, θ, is the offset (months) from the reference time, January ($m = 1$), with absolute value of phase difference $|\Delta\theta| \leq 6$). Phase difference, $\Delta\theta$, describes the synchronization of P and PET throughout the year as

$$\Delta\theta = \begin{cases} \theta_{PET} - \theta_P, & -6 \leq \theta_{PET} - \theta_P \leq 6 \\ \theta_{PET} - \theta_P - 12, & \theta_{PET} - \theta_P > 6 \\ \theta_{PET} - \theta_P + 12, & \theta_{PET} - \theta_P < 6 \end{cases} \tag{2}$$

100

Equation 1 showed overall good fits to the long-term mean monthly distributions of P and PET, with R^2=0.67±0.28 and 0.85±0.17, respectively (mean±standard deviation across all pixels). To constrain $-6 \leq \Delta\theta \leq 6$, 12 was either added to or subtracted from $\Delta\theta$ values outside of these bounds (e.g., $\Delta\theta = 8$ months is translated to -4 months). Because a constant reference time of January does not describe the water year for each zone, the $\Delta\theta$ distributions were normalized by centering around the mode and correcting to contain only positive values (0 to 12 months).

105

2.3 Established climate classification systems

    Four previously established climate classification schemes were assessed in this analysis. We included two veteran schemes, Köppen-Geiger (KPG) and Holdridge Life (HDL) zoning systems, and two recently proposed frameworks, here referred to as Meybeck Hydroregion (MHR, Meybeck et al., 2013) and Knoben Hydroclimate (KHC, Knoben et al., 2018) 110   systems. Note that the original KHC zones created by Knoben et al. (2018) were not delineated by discrete boundaries but were instead represented as a probability continuum of belonging to a zone. In the present study, KHC pixel values were rounded to create 30 separate zones, equal to the number of KPG zones, and similar to HDL (n=38) and MHR (n=27). This KHC zone distinction was made so that each zone could be independently evaluated. Note that the very small KPG zones "Csc" and "Cwc" did not appear in the 0.5° x 0.5° resolution KPG output created by Beck et al. (2018) that was used in this 115   study, resulting in 28 KPG presently analyzed zones. As in other climate classification studies (Knoben et al., 2018; Meybeck et al., 2013), KPG was considered here to be the standard to which other systems are primarily compared.

2.4 Proposed univariate ET climate classification systems

    This study establishes and verifies ET-relevant climate classification frameworks by creating zones primarily based on ET rates and comparing ET coherence between systems. Three of the four systems proposed in this study were univariate 120   (formatted from global mean annual ET rates) and uni-conditional (incorporating an additional system-dependent single condition). The additional conditions were included to emphasize a specific optimization goal.

The first two proposed univariate classification systems were based on the global ET empirical cumulative distribution function (CDF). The first classification system, ET Area-optimizing (ETA), was created with the objective of having an equal number of pixels in each ET-based zone. This was motivated based on the relatively high spatial non-

125 uniformity of the KPG system, resulting in highly variable relevance for regional analyses. Additionally, it is useful to have a simple baseline framework upon which to compare other systems. This type of spatial condition is similar to the prioritizations of the MHR framework that state zones should ideally be "delineated in one piece" (Meybeck et al., 2013). The cumulative probability interval [0,1] was divided into 15 equal parts, and the corresponding upper and lower bounds of ET thresholds for each zone were determined from the CDF of mean annual ET for all global land pixels (Figure S1-A). The number of ETA

130 zones was chosen based on the number of zones in previously established systems, the relative improvement of ET coherence with the addition of more zones (Figure S1-B), and the objective of having equal or fewer zones than the standard KPG framework.

The second proposed classification system, ET Variability-optimizing (ETV), was based on the principle of maximizing within-zone ET coherence to the greatest feasible extent, considering the tradeoffs with increasing complexity by

135 adding zones. By fitting the empirical CDF with a continuous distribution, zone boundaries can be determined analytically for the minimum desired CV_{min}. For simplicity, and also supported by empirical evidence (Figure S1), we fitted a uniform distribution, which is characterized by lower and upper bounds a and b, with $CV = (b-a)/[\sqrt{3}(b+a)]$. The ET limits defining each zone, i, were then determined directly from this relation as

140
$$a_i = \frac{b_i(1-\sqrt{3}CV_{min})}{1+\sqrt{3}CV_{min}}$$
(3)

where the upper and lower limits of sequential zones are shared (i.e., $b_{i-1} = a_i$). The largest value of $b = 1,454$ mm yr^{-1} was based on the maximum ET for all pixels, and $CV_{min} = 0.075$ was chosen based on marginal decrease with increasing number of zones (Figure S2), which resulted in 29 zones. This method produces nearly equal CV in all zones.

145 The third univariate scheme proposed here is the ET Clustering (ETC) classification system. Previous analyses have used clustering techniques for climate classification purposes (Knoben et al., 2018; Tapiador et al., 2019), and here k-means clustering based on the Hartigan and Wong (1979) algorithm was employed using the kmeans function in the R package *stats* (R Core Team, 2018). Zones were built by forming clustering centers iteratively until the within-zone sum of squares of mean annual ET, based on Euclidean distances, was reduced. The final number of clustering centers (i.e., zones) was 20, which was

150 chosen because it is the fewest number of zones with ET mean CV ≤ 0.1 (Figure S3).

2.5 Proposed multivariate climate classification systems

The final proposed system in this study is a multivariate climate clustering framework, which was created from the same k-means clustering method as the ETC framework. This new climate classification system included multiple

hydroclimate variables and was designed for comparison with previously established systems that were similarly formed with
155 multiple variables, as well as to assess against the univariate ET classification frameworks. A suite of potential multivariate
systems was generated using mean annual P, PET, Δθ, ET, and Q. The best performing multivariate classification scheme was
chosen from the batch of potential systems by evaluating criteria corresponding to KPG system characteristics. The primary
criteria considered were water budget coherence (mean CV of ET, P, and Q) and zone complexity. Some flexibility was
allowed for the final multivariate system, with water budget coherence constrained to be within 50% of KPG coherence values.
160 The specific values for these elimination thresholds are listed in Supporting Information, and the resulting eligibility of the
multivariate climate classification systems is shown in Tables S1 and S2. Ultimately, the climate classification system formed
from clustering mean annual P and mean annual PET was chosen as the representative multivariate framework. This
classification scheme was named the Water-Energy Clustering (WEC) climate classification system.
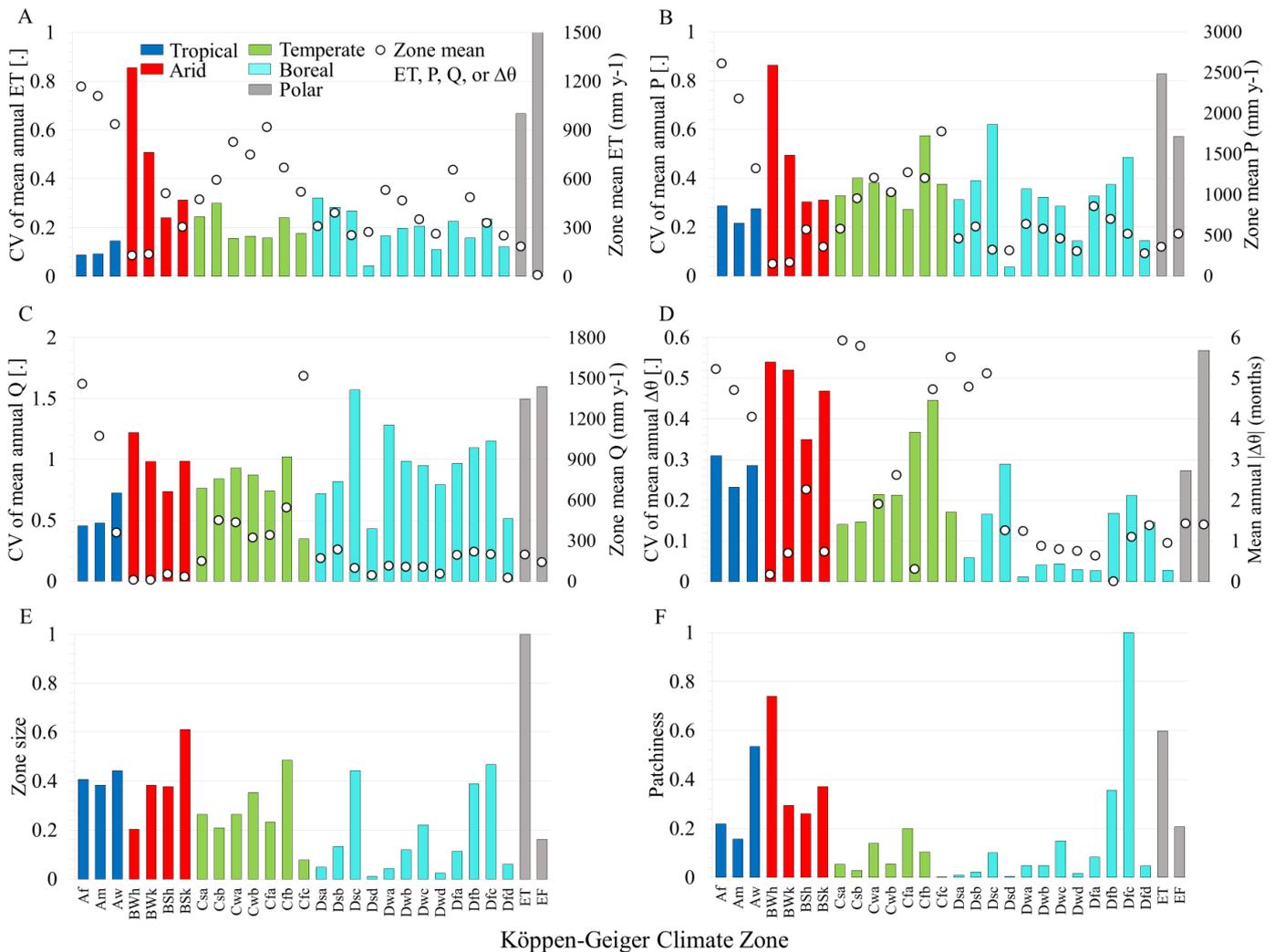
2.6 Coherence and complexity metrics

165 Variable coherence is defined by within-zone variability, represented by the intra-zone coefficient of variation (CV)
of the variable of interest. Lower CV corresponds to higher coherence. Complexity metrics were defined with the objective to
use the smallest number of metrics to characterize shape structure (Cushman et al., 2008). Classification system complexity
metrics were based on three principles: 1) Classification systems should consist of a relatively even distribution of pixels across
zones, avoiding disproportionately large or small zones, 2) Zones should be as hydrologically continuous as possible (Meybeck
170 et al., 2013), minimizing patchiness or fragmentation, and 3) Classification systems should comprise less than or equal to the
number of zones in the KPG framework. Therefore, complexity was assessed based on the inter-zone distribution of the number
of pixels (zone area evenness, CV_z) and the number of patches in each zone (zone fragmentation). Note that the subscript z is
added to differentiate between-zone complexity from the above metrics which emphasize intra-zone coherence. The number
of patches is the only primary coherence or complexity metric in which CV is not used, since the objective here was to minimize
175 the degree of fragmentation, and not the similarity of fragmentation across zones. The number of patches was determined using
the R function ClassStat in package *SDMTools* (VanDerWal et al., 2019). For each hydroclimate and complexity variable,
differences between the KPG framework and the other classification systems were determined based on the Kolmogorov-
Smirnov (K-S) test.

3 Results

180 This study compared four previously established climate classification systems (KPG, HDL, MHR, KHC) and four
potential new climate classification systems (ETA, ETV, ETC, and WEC) to assess for hydroclimate coherence as well as zone
boundary complexities. The coherence and complexity metrics for the KPG system, the standard used in this study, are shown
in Figure 1 for all zones. The KPG system had relatively high hydroclimate coherence, as all zones had CV < 1 for all variables
except Q (Figure 1). However, the KPG spatial complexity was also relatively high. Zones in the KPG system had high

185   variability in the number of pixels, with several Boreal zones less than 5% the size of the largest zone, polar tundra zone "ET"

(Figure 1E).

Figure 1: Water budget coherence, quantified as intra-zone CV (A-D), and complexity metrics of zone size and patchiness (E-F) for
190   **each zone for the Köppen-Geiger climate classification system. Mean hydroclimate values for each zone are also shown (circles, A-D). Zone size (number of pixels) and patchiness (number of patches) are normalized to their maximum values.**

Coherence and complexity results for the six best performing of the eight climate classification systems are shown in

Table 1. The performance compared to KPG, in order of worst-to-best, was generally HDL, ETV, KHC, MHR, ETC, ETA,

195   and WEC. The HDL and ETV systems performed either worse than or not statistically different from KPG in all metrics,

except higher ET coherence for ETV, and these results are therefore not shown in Table 1 (see Table S3). The established

7

KHC and MHR systems were also poor-performing, with overall worse or similar hydroclimate coherence compared to KPG. However, KHC and MHR had improved pixel coherence and fewer patches, respectively (Table 1). The proposed ETA and ETC systems had much better ET coherence than KPG and similar P coherence to KPG. The ETA system also had the fewest

200 number of zones and very high coherence for zone size (by far the lowest CV_z), but both ETA and ETC had lower coherence than KPG for the remaining hydroclimate variables. Finally, the proposed WEC system had similar or better performance than the KPG system in all coherence and complexity metrics, except for $\Delta\theta$ coherence, and can be concluded as a strong contender for the overall best system.

205 **Table 1: Hydroclimate coherence (intra-zone CV for mean annual ET, P, Q, $\Delta\theta$, and PET) and complexity (inter-zone CV_z for pixels, and numbers of patches and zones) in established and proposed climate classification systems. Mean(standard deviation), with significantly higher ($\uparrow$) or lower ($\downarrow$) values than KPG determined based on K-S tests. Bold indicates the best overall system for each metric (more than one system had statistically similar results for some metrics).**

Metric	Established systems			Proposed systems		
	KPG	MHR	KHC	ETA	ETC	WEC
CV(ET)	0.27(0.21)	0.44(0.26)$^\uparrow$	0.59(0.18)$^\uparrow$	**0.12(0.17)**$^\downarrow$	**0.08(0.15)**$^\downarrow$	0.31(0.23)
CV(P)	0.38(0.20)	0.55(0.24)$^\uparrow$	0.51(0.10)$^\uparrow$	0.56(0.42)	0.47(0.35)	**0.25(0.23)**$^\downarrow$
CV(Q)	**0.88(0.30)**	1.30(0.57)$^\uparrow$	**0.87(0.30)**	1.45(0.50)$^\uparrow$	1.29(0.53)$^\uparrow$	**0.75(0.38)**
CV($\Delta\theta$)	**0.24(0.18)**	0.31(0.14)$^\uparrow$	0.39(0.11)$^\uparrow$	0.38(0.08)$^\uparrow$	0.37(0.09)$^\uparrow$	**0.33(0.14)**
CV(PET)	0.20(0.13)	0.20(0.09)	0.48(0.13)$^\uparrow$	0.53(0.24)$^\uparrow$	0.42(0.25)$^\uparrow$	**0.14(0.06)**$^\downarrow$
CV_z(zone areas)	1.21	0.78	0.37$^\downarrow$	**0.02**$^\downarrow$	0.48$^\downarrow$	0.55$^\downarrow$
patches	46(36)	**10(9)**$^\downarrow$	275(148)$^\uparrow$	153(70)$^\uparrow$	150(52)$^\uparrow$	59(36)
zones	28	27	30	**15**	20	22

210 The coherence and complexity metrics for the WEC system are shown in Figure 2 for all zones. Like KPG, coherence was high (CV < 1) in all zones for each hydroclimate variable apart from Q, but WEC had even higher hydroclimate coherence than KPG overall. Zone area was more equally distributed in WEC, but with similar patchiness to KPG (Table 1). The KPG system groups 30 zones into 5 categories, and here the WEC zones were similarly divided into 5 groups (Figure 3) based on aridity index (ϕ = P/PET) and minimizing the variability of zone areas. Groups were organized to comprise near-equal pixel

215 distributions (11,367 to 13,895 pixels in each group) across zones of decreasing aridity index, with G1 to G5 $\bar{\phi}$ = {2.4, 1.1, 0.83, 0.46, 0.12}. In WEC zones, there were positive linear relationships (p<0.05) between aridity index and CV of each water budget variable: ET (R^2=0.25), P (R^2=0.53), and Q (R^2=0.69), indicating higher spatial variability (lower coherence) with

Hydrology and
Earth System
Sciences
Open Access
EGU
Discussions

increasing aridity. While WEC groups represented similar total areas, they comprised different numbers of zones, from G3
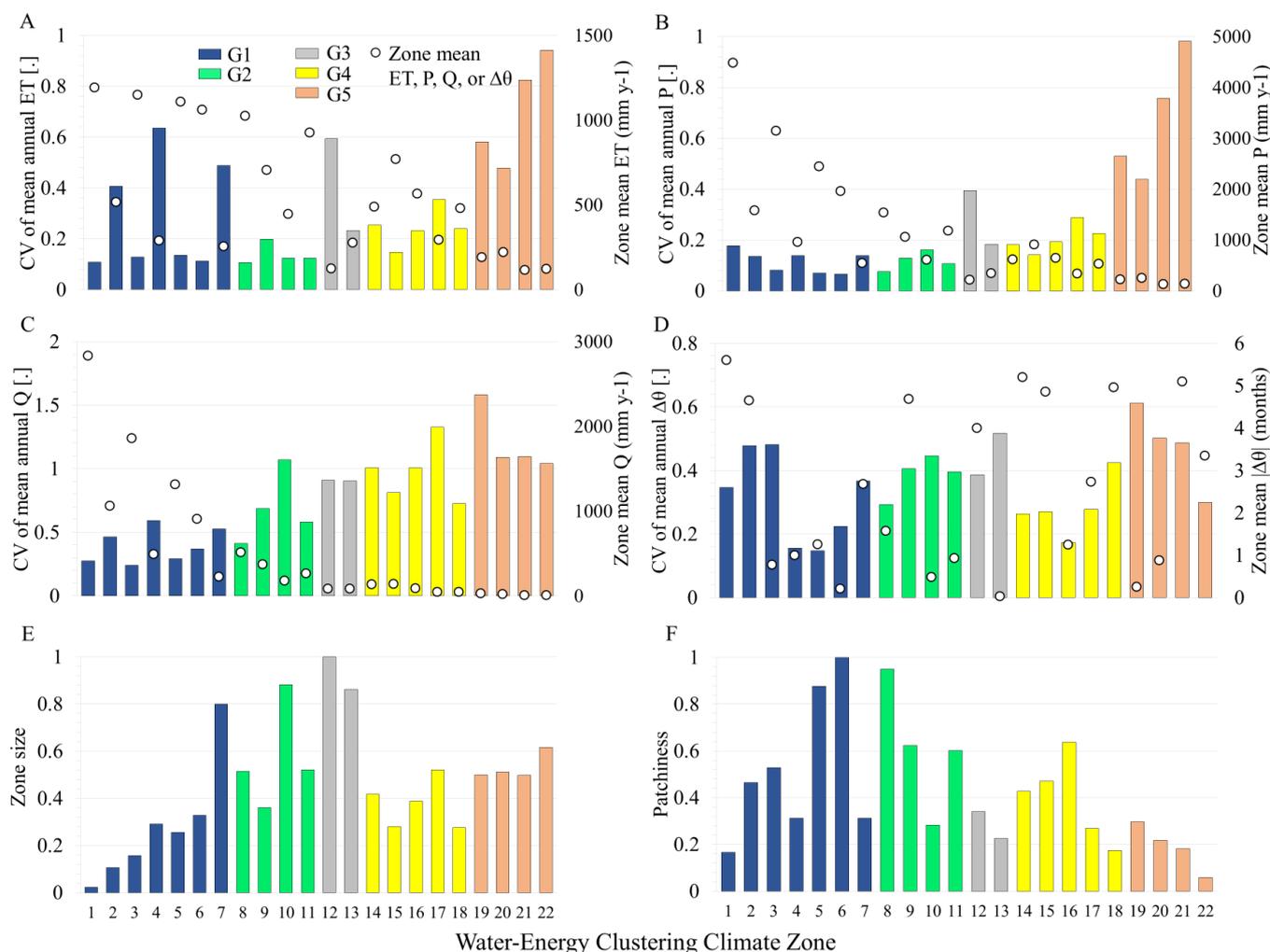with only two large zones, to G1 with 7 zones.

220

Figure 2. Water budget coherence (intra-zone CV, A-D) and complexity metrics (zone size and patchiness, E-F) for each zone for the Water-Energy Clustering (WEC) climate classification system. Mean hydroclimate values for each zone are also shown (circles, A-D). Zone size (number of pixels) and patchiness (number of patches) are normalized to their maximum values.

225

Maps of the boundaries for the proposed WEC system and the standard KPG framework are compared in Figure 3.
While there were some similarities (e.g., see the Iberian Peninsula in Figure 3), most regions are divided differently. For
example, parts of northern Europe are divided into three KPG zones but five WEC zones. Similarly, the southeastern United
States, excluding south Florida, is mostly one KPG Temperate zone, but is separated in the WEC system into two distinct G2

230 zones. Clustering centers for the WEC climate classification system are listed in Table S4.
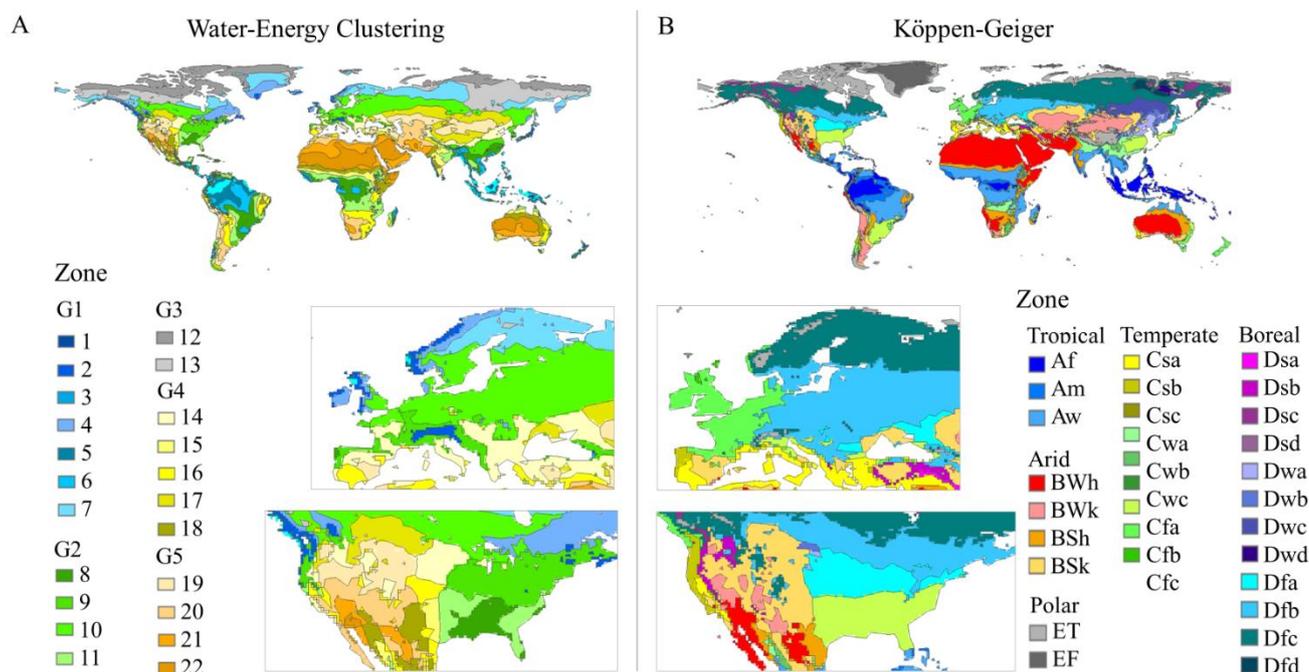
Figure 3. Spatial distribution of WEC (A) and KPG (B) classification systems. Europe and North America are magnified.

4 Discussion

We hypothesized that variable coherence and zone shape complexity would be related to the governing principles of the classification systems, which was mostly supported by the results of this study. Of the four previously established systems, KPG was the most hydroclimatically coherent, but had high variability in pixel distribution across zones (Table 1). The KPG and WEC frameworks had the overall highest $\Delta\theta$ coherence of the eight total compared systems, which is reasonable since KPG was the only system that accounted for monthly variability of water (P) and energy (temperature), resulting in 24 parameters (Beck et al., 2018). The WEC system was also based on water (P) and energy (PET), but from a mean annual perspective, thus requiring only two parameters. It is important to highlight that the number of required parameters, a notable aspect of system complexity, was lower for all systems proposed here (between 1 and 3 parameters) compared to the KPG system (24 parameters).

When comparing all eight systems, WEC had the highest P and PET coherence and similar ET, $\Delta\theta$, and Q coherence to KPG. It is not surprising that the WEC classification system yielded highest P and PET coherence, given these were the variables used to draw its zone boundaries, but WEC also had much more uniform pixel distribution, similar zone fragmentation, fewer zones, and required substantially fewer parameters than KPG. The MHR system used mean Q as a

governing principle (Meybeck et al., 2013), and the KHC framework considered Q as independent validation for their zones
250   (Knoben et al., 2018), but neither system was comparatively high in Q coherence. However, the principle of contiguity in the
MHR system led to the lowest patchiness of all systems evaluated, so this system could be useful when continuous boundaries
are important for ease of implementation or interpretation purposes. Lastly, the three univariate ET-based classification
systems had the highest ET coherence, while ETA (which additionally optimized equal zone area) also had the most uniform
pixel distribution across zones.

255   Evaluating hydroclimate coherence is important for understanding water availability distribution within a group of
related zones and within individual zones to make informed management decisions. For KPG, water budget dynamics are most
uniform in Tropical zones and least coherent in Polar and Arid zones (Table S5). Arid zones had the most uniform pixel
distribution, while the Boreal group was least fragmented with the lowest mean number of patches, suggesting these zones are
interrupted neither by other zones nor by water bodies. For WEC, group G1 had the highest mean P and Q coherence, but also
260   the lowest mean PET coherence. The latter is likely because of the temperature variation across G1 zones, which encompass
both equatorial and subarctic regions (Figure 3). Group G5, comprising the most arid zones, had the lowest water budget
coherence and highest PET coherence, indicating relative uniformity in (low) rainfall and (high) temperature. Pixel distribution
was most uniform in G3 and G5, while G5 was least fragmented. It is valuable to note the structural attributes of zone
boundaries because these boundaries are expected to change over time (Beck et al., 2018; Knoben et al., 2018).

265   Of the water budget components, Q was the least coherent while ET was the most coherent across all systems except
KHC and WEC (Table 1). This overall high coherence suggests that the variability of the drivers of ET (water and energy
budget components) are mostly captured, even if ET itself is not a governing principle in the framework. However, there was
still room for improvement within the established classification systems with respect to optimizing ET variability. Based on
the climate classification system comparison presented here, it can be concluded that water and energy ET drivers are important
270   considerations for broad hydroclimate analyses, since water budget coherence is mostly achieved when P and PET are included
as governing principles. However, when specifically evaluating ET dynamics, using an ET framework is most appropriate.
Depending on which spatial complexity metric is favored, MHR (mean zone patches = 10) and ETA ($CV_z = 0.02$) were the
least complex systems. Based on both ET coherence and spatial complexity, the ETA system established here is suggested for
ET-focused questions such as large-scale assessments of crop productivity (Howell et al., 2015).

275   This study is limited by a few factors. First, distinct climate zone boundaries, although useful in practice, do not exist
in the physical system (Knoben et al., 2018). Second, this study compared averaged metrics that were applied across zones
within each classification system, although some individual zones were better or worse than others with respect to coherence
and complexity. Third, the focus on long-term mean annual hydroclimate attributes for zone formation does not account for
interdecadal climate dynamics. It is also important to recall that Q here is based on locally generated runoff (P-ET), rather than
280   the accumulated runoff from upstream contributing areas, which would be representative of gaged streamflow.

The KPG system is the most widely used climate classification system, and this analysis revealed that it indeed has
high hydroclimatic coherence. However, WEC was either better than or not statistically different from the KPG framework in

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

all coherence and complexity metrics. Ultimately, the WEC climate classification system is suggested for large spatial scale hydroclimate analyses, as it has lower zone area variability than KPG and requires fewer parameters and zones to create a

285 hydroclimatically coherent classification system. Applying the most relevant framework in which hydroclimate similarities are captured by appropriate boundaries is critical to address questions of large spatial-scale climate variability and water availability (Meybeck et al., 2013). The WEC and ETA boundary systems proposed here can be useful in determining environmental management decisions in which land area encompasses multiple climate zones.

290 **Author Contribution**

KLMP performed the analyses and led the manuscript preparation. JWJ conceived and directed the study.

Competing Interests

The authors declare that they have no conflict of interest.

295

Acknowledgements

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high resolution global dataset
300   of monthly climate and climatic water balance from 1958-2015. *Scientific Data*, 5, 170191, 2018.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen
        Geiger climate classification maps at 1 km resolution. *Scientific Data*, 5,180214, 2018.

305 Berghuijs, W. R., and Woods, R. A.: A simple framework to quantitatively describe monthly precipitation and
        temperature climatology. *International Journal of Climatology*, 36(9), 3161 3174, 2016.

Bivand, R., Pebesma, E., and Gomez-Rubio, V.: *Applied Spatial Data Analysis with R*, Second edition. Springer, NY, 2013.

310 Boland, M. R., Parhi, P., Gentine, P., and Tatonetti, N. P.: Climate classification is an important factor in assessing quality-of-
        care across hospitals. *Scientific Reports*, 7(1), 1-6, 2017.

Chen, D., and Chen, H. W.: Using the Köppen classification to quantify climate variation and change: An example for
        1901–2010. *Environmental Development*, 6, 69-79, 2013.
315
Cushman, S. A., McGarigal, K., and Neel, M. C.: Parsimony in landscape metrics: strength, universality, and
        consistency. *Ecological indicators*, *8*(5), 691-703, 2008.

FAO (Food and Agriculture Organization of the United Nations – with UNESCO and WMO). *World Map of
320   Desertification*. Food and Agricultural Organization (FAO), Rome, 1977.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high resolution gridded multivariate

climate dataset. *Scientific Data*, 7(1), 1-18, 2020.

325   Hartigan, J. A., and Wong, M. A.: Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108, 1979.

Hijmans, R.: Raster: Geographic Data Analysis and Modeling. R package version 2.6-7, 2017.

330   Holdridge, L. R.: "Life zone ecology", *Life Zone Ecology*. Tropical Science Center, San Jose, Costa Rica, 1967.

Howell, T. A., Evett, S. R., Tolk, J. A., Copeland, K. S., and Marek, T. H.: Evapotranspiration, water productivity and crop coefficients for irrigated sunflower in the US Southern High Plains. *Agricultural Water Management*, 162, 33 46, 2015.

335

Jagai, J. S., Castronovo, D. A., and Naumova, E. N.: The use of Köppen climate classification system for public health research. *Epidemiology*, 18(5), S30, 2007.

Knoben, W. J., Woods, R. A., and Freer, J. E.: A Quantitative Hydrological Climate Classification Evaluated With
340   Independent Streamflow Data. *Water Resources Research*, 54(7), 5088-5109, 2018.

Lanfredi, M., Coluzzi, R., Imbrenda, V., Macchiato, M., and Simoniello, T.: Analyzing Space–Time Coherence in Precipitation Seasonality across Different European Climates. *Remote Sensing*, 12(1), 171, 2020.

345   Lloyd, S. J., Kovats, R. S., and Armstrong, B. G.: Global diarrhea morbidity, weather and climate. *Climate Research*, 34(2), 119-127, 2007.

Magarey, R. D., Borchert, D. M., and Schlegel, J. W.: Global plant hardiness zones for phytosanitary risk analysis. *Scientia Agricola*, 65(SPE), 54-59, 2008.

350

McKenney, D. W., Pedlar, J. H., Lawrence, K., Campbell, K., and Hutchinson, M. F.: Beyond traditional hardiness zones: using climate envelopes to map plant range limits. *BioScience*, 57(11), 929-937, 2007.

Mellinger, A. D., Sachs, J. D., and Gallup, J. L.: Climate, coastal proximity, and development. The Oxford handbook of
355   economic geography, 169, 194, 2000.

Meybeck, M., Kummu, M., and Dürr, H. H.: Global hydrobelts and hydroregions: improved reporting scale for water related issues? *Hydrology and Earth System Sciences*, 17(3), 1093 1111, 2013.

360   O'Neill, R. V., Krummel, J. R., Gardner, R. E. A., Sugihara, G., Jackson, B., and DeAngelis, D. L.: Indices of landscape pattern. *Landscape Ecology*, *1*(3), 153-162, 1988.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., and Underwood: Terrestrial Ecoregions of the World: A New Map of Life on Earth, A new global map of terrestrial ecoregions provides an
365   innovative tool for conserving biodiversity. *BioScience*, 51(11), 933-938, 2001.

Papagiannopoulou, C., Gonzalez Miralles, D., Demuzere, M., Verhoest, N., and Waegeman, W.: Global hydro-climatic biomes identified via multitask learning. *Geoscientific Model Development*, 11(10), 4139-4153, 2018.

370   Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen Geiger climate classification. *Hydrology and Earth System Sciences Discussions*, 4(2), 439-473, 2007.

Pierce, D.: ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. R package version 1.16, 2017.

375 R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/, 2018.

Richards, D., Masoudi, M., Oh, R. R., Yando, E. S., Zhang, J., and Friess, D. A.: Global Variation in Climate, Human Development, and Population Density Has Implications for Urban Ecosystem Services. *Sustainability*, 11(22), 6200,
380 2019.

Van der Ent, R. J., Savenije, H. H., Schaefli, B., and Steele-Dunne, S. C.: Origin and fate of atmospheric moisture over continents. *Water Resources Research*, *46*(9), 2010.

385 VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., and Storlie, C.: SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises. R package version 1.1 221.1. https://CRAN.Rproject.org/package=SDMTools, 2019.

Wang-Erlandsson, L., Fetzer, I., Keys, *P*. W., Van Der Ent, R. J., Savenije, H. H., and Gordon, L. J.: Remote land use
390 impacts on river flows through atmospheric teleconnections. *Hydrology and Earth System Sciences*, 22(8), 4311 4328, 2018.

Willmott, C. J., and Feddema, J. J.: A more rational climatic moisture index. *The Professional Geographer*, 44(1), 84 88, 1992.
395
Zhang, K., Kimball, J. S., and Running, S. W.: A review of remote sensing based actual evapotranspiration estimation. *Wiley Interdisciplinary Reviews: Water*, 3(6), 834-853, 2016.