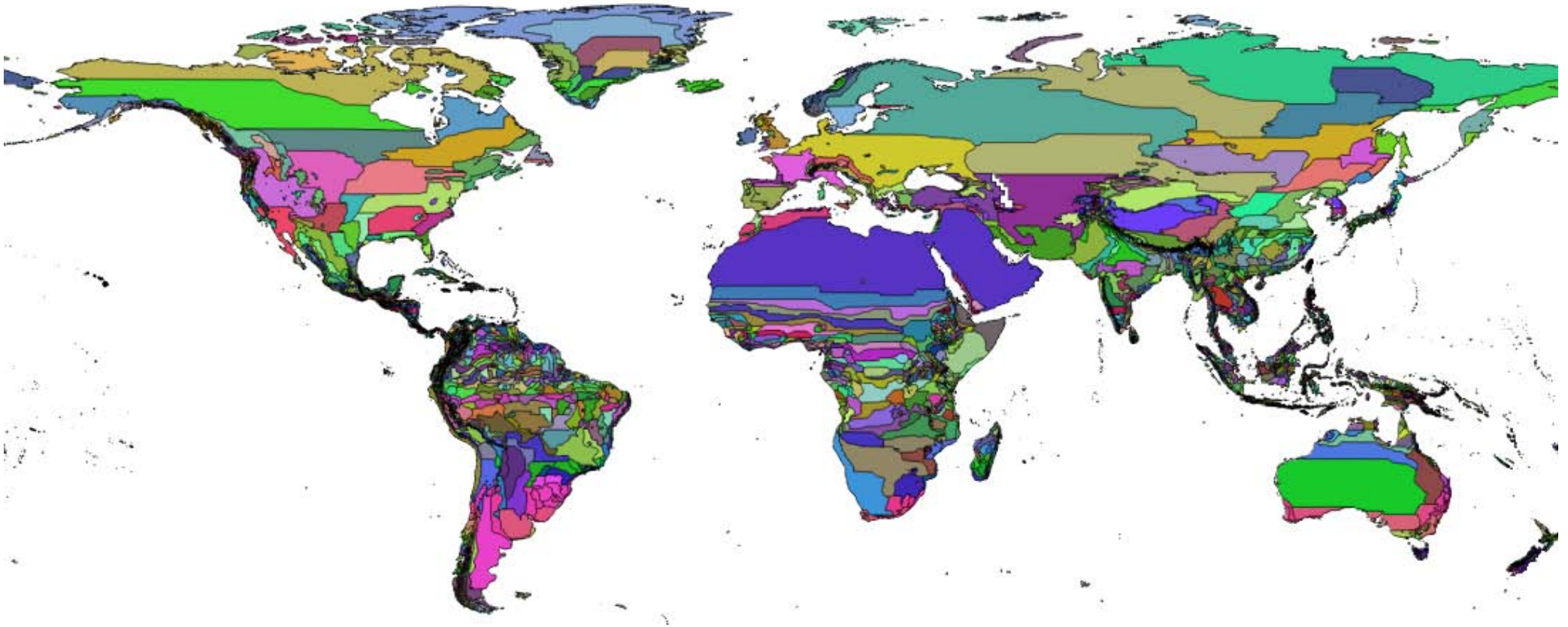
Segmentation of WorldClim dataset reveals new insight into spatial variability of global climate

Pawel Netzel and Tomasz Stepinski

Space Informatics Lab, University of Cincinnati
sil.uc.edu



GeoComputation 2017, Leeds, UK

What is global climate regionalization?

Global climate regionalization aims at delineation of mutually exclusive land units in a way that maximizes the intra-unit climate homogeneity and inter-unit climate disparity.

heuristic regionalization

Köppen-Geiger

data-driven regionalization

(1) clustering

(2) segmentation (new)

Why is it important?

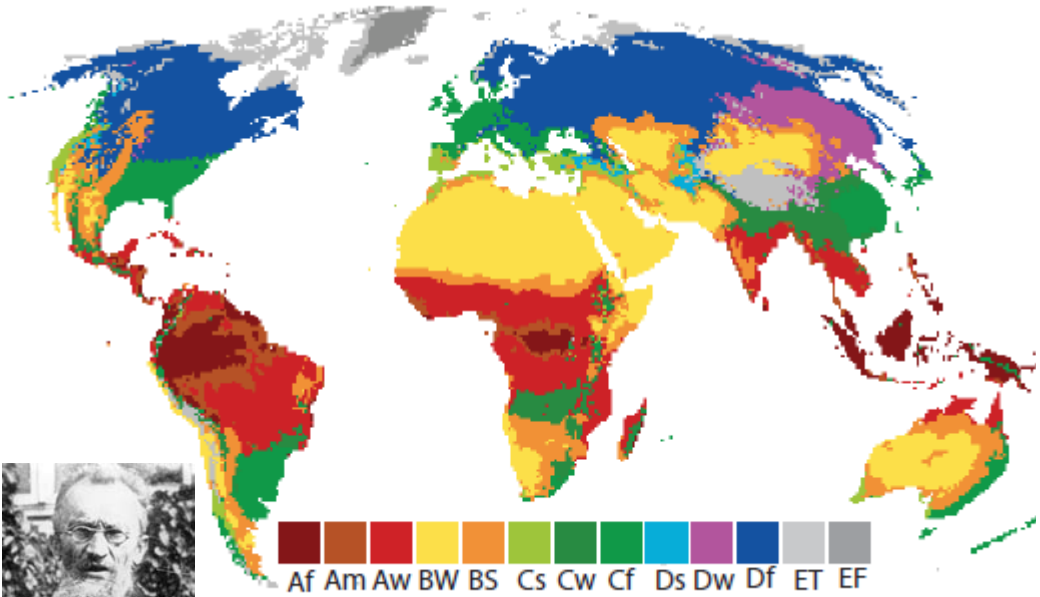
simplifies spatial representation of climates into a form that is more meaningful and easier to analyze.

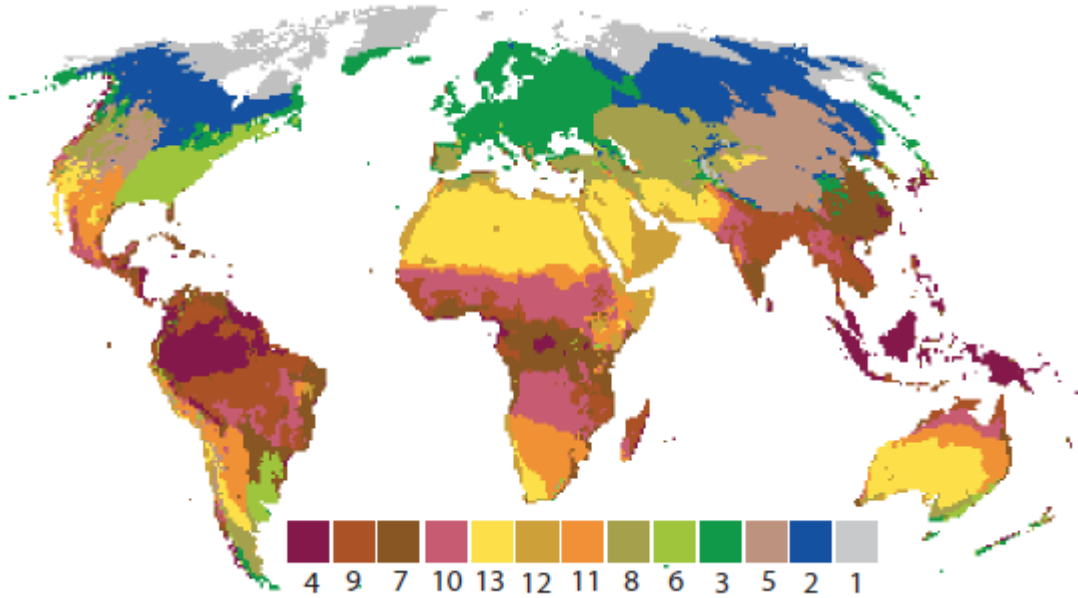
visualize spatial distributions of past and future climates

provides insight into the relationships between climate and Earth's physical and biological systems

Illustrate climate change in terms of shifting geographical boundaries of major climate types.

Köppen-Geiger vs. clustering





Netzel & Stepinski (20160, J. Climate, 29(9), pp. 3387–3401

Köppen-Geiger

>100 years old scheme

based on vegetation types as proxies for climate

boundaries between units expressed in climatic variables

used by everybody

clustering climate data

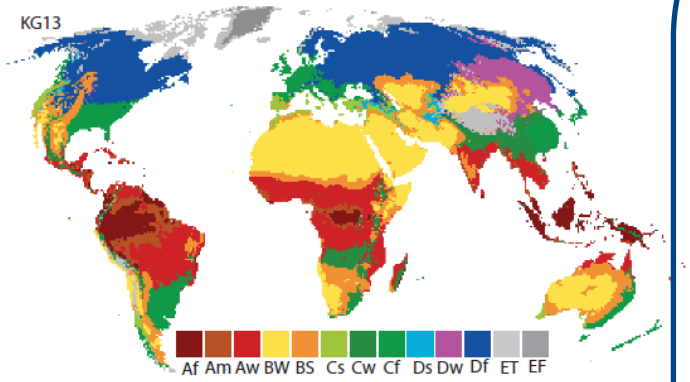
uses global gridded climate data, local climate = time series of local values of temperature and precipitation.

utilizes dissimilarity measure between time series

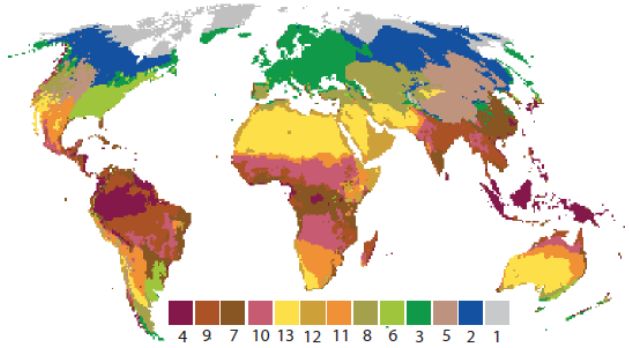
Uses clustering algorithm to obtain land units

Why Köppen-Geiger and clustering give similar results and why clustering is not the best data-driven method to regionalize global climates?

Köppen-Geiger



clustering climate data



Köppen-Geiger and clustering methods result in similar regionalizations of global climates because researchers **assume** the number of clusters to be **equal** or similar to the number of classes in the KG classification.

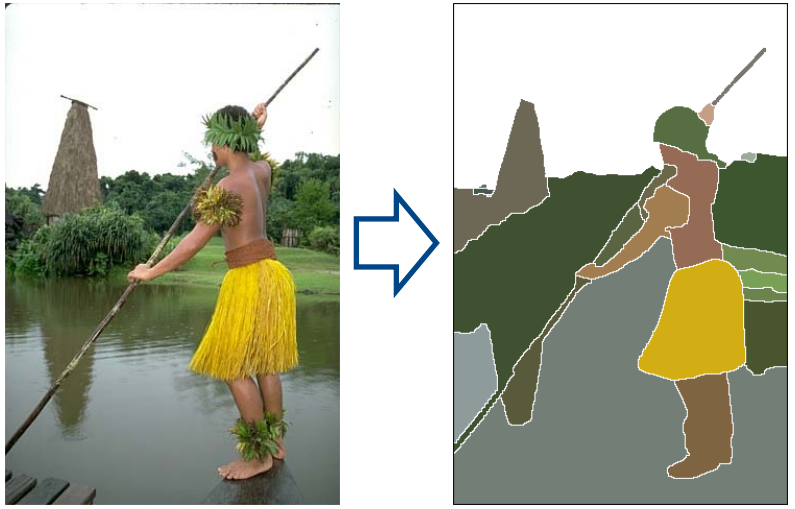
Clustering **does not ensure** that different clusters have the same level of homogeneity of local climates – **some clusters maybe more climatically homogeneous than others.**

There is no way to determine the **“correct”** number of clusters.

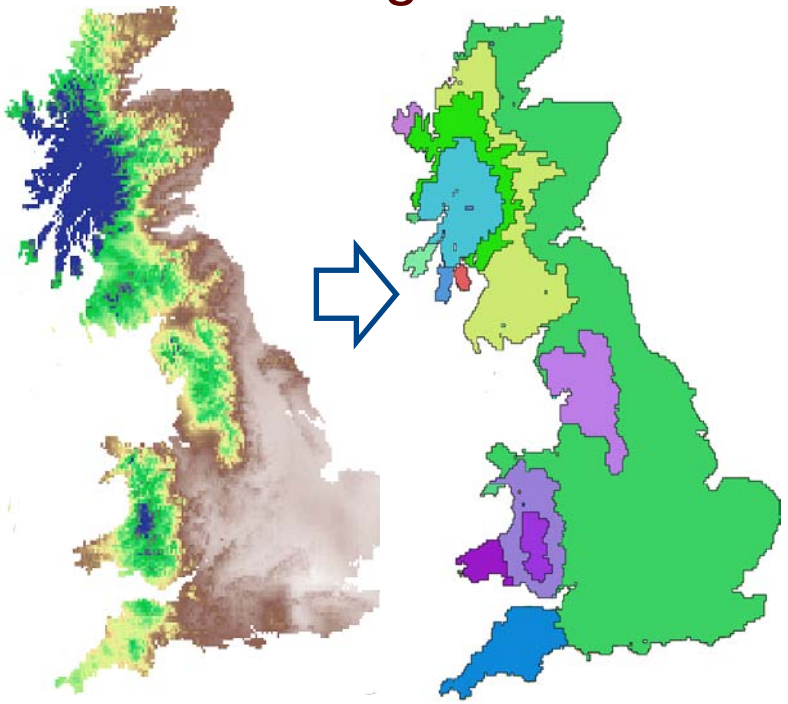
We didn’t gain much **new knowledge** from clustering than we have already known from the100 years old heuristic regionalization.

Climate segmentation

image segmentation



climate segmentation



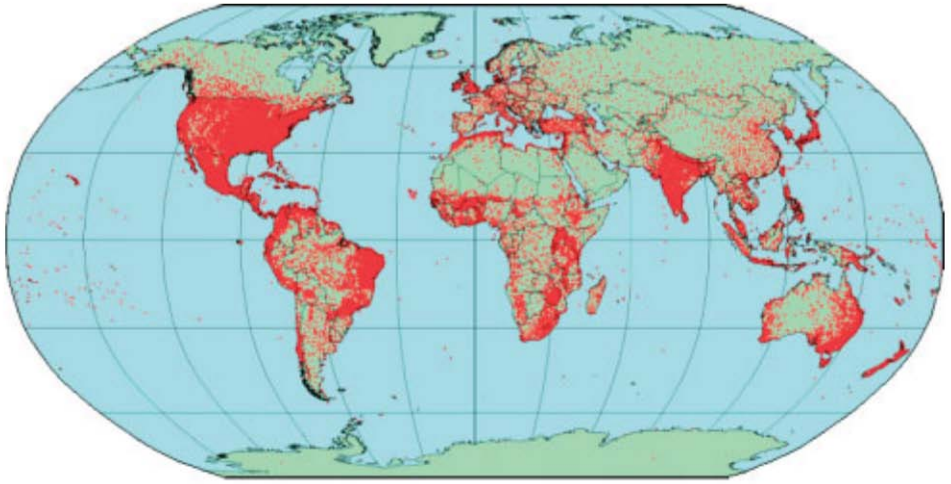
The result of grid segmentation is a set of segments that collectively cover the entire grid. Cells in a segment are similar. Adjacent segments are different.

Climate segmentation vs. clustering

issues	clustering	segmentation
numerical representations of climate	as bivariate time series	as bivariate time series
dissimilarity measure	dynamic time warping (required)	time-shift invariant Euclid. distance
homogeneity	different clusters are characterized by different climate homogeneities	all segments are characterized by the same level of homogeneity
map units	non-contiguous	contiguous
algorithm	Partitioning Around Medoids (PAM) clustering (slow)	Fast Scan segmentation (fast)

Global climate data

weather stations collect data



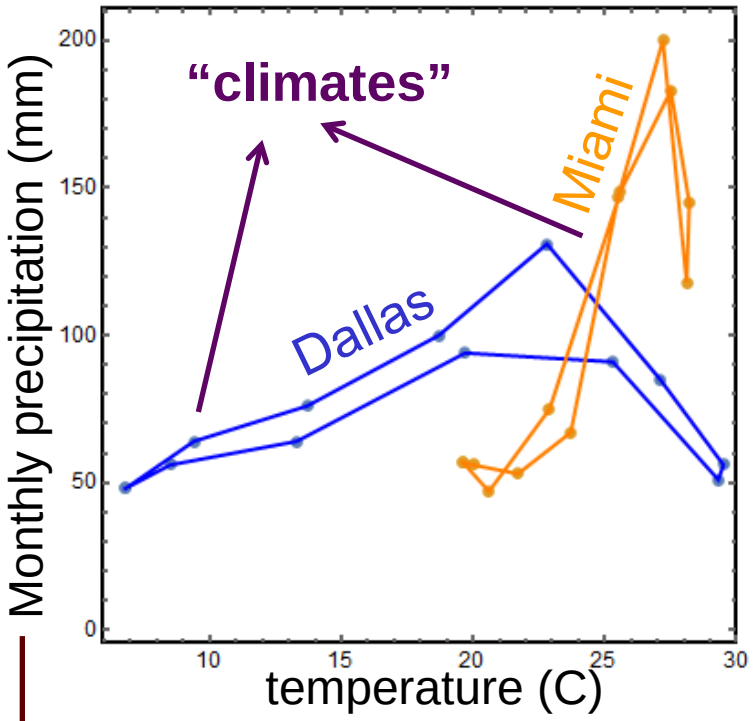
Long term data averages of monthly variables:

temperature
precipitation
maximum temperature
minimum temperature

Global gridded climate data
We use WorldClim monthly sum of precipitation (P) and average temperature (T) at 2.5 arc second (~5 km at the equator)
www.worldclim.org/

A climate for a given grid cell is mathematically represented by a bivariate cyclic time series C = $\{(T^1 , P^1) , ..., (T^{12} , P^{12})\}$, where the time series progresses through 12 months.

Climate variables vs. climate

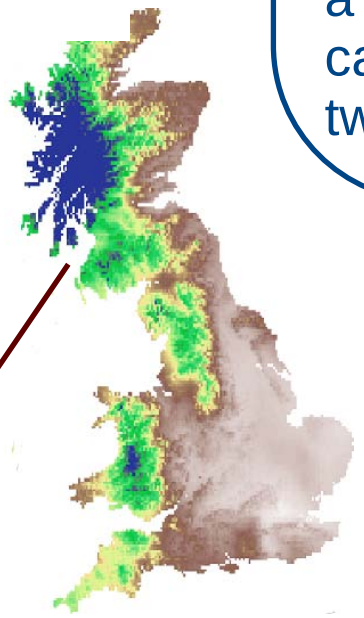


Climate: the statistics of weather over long periods of time. This is **not an actionable** definition.

To compare climates in two places patterns of **climatic variables** (for example, temperature, precipitation) are compared **separately**, but it would be better to compare climates **as single entities**.

We use a mathematical model (a bivariate time series) that **joins** temperature and precipitation into a single entity called a **"climate"**. This allows us to calculate **a single-valued dissimilarity** between two climates.

Climatic variables

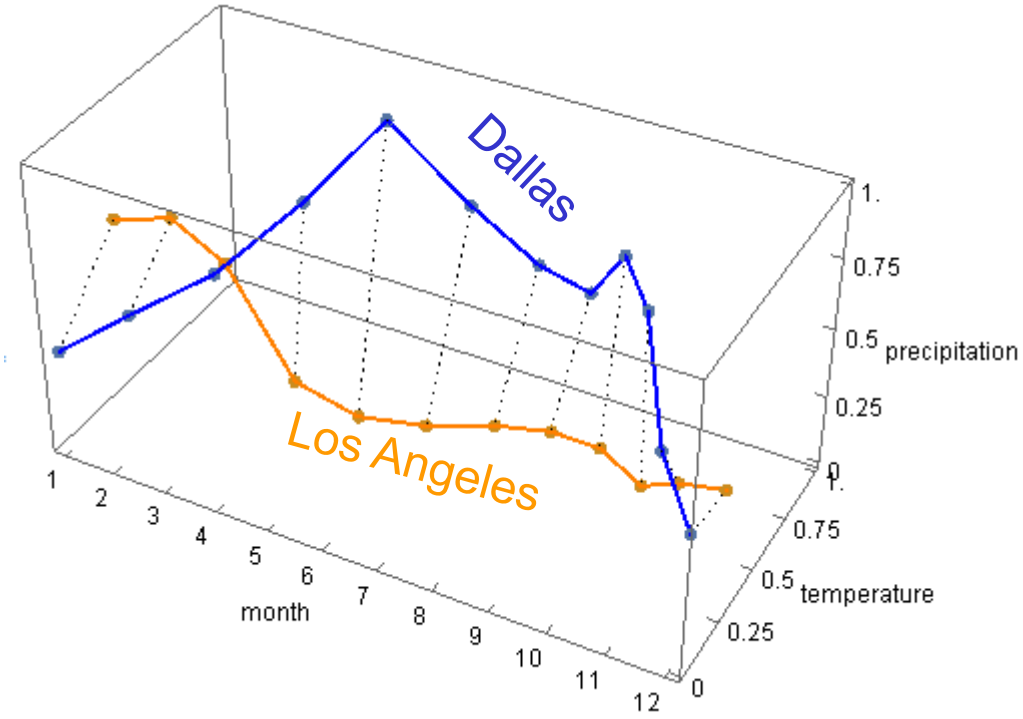


Map of climates – color indicates a dissimilarity of climate from the climate in London

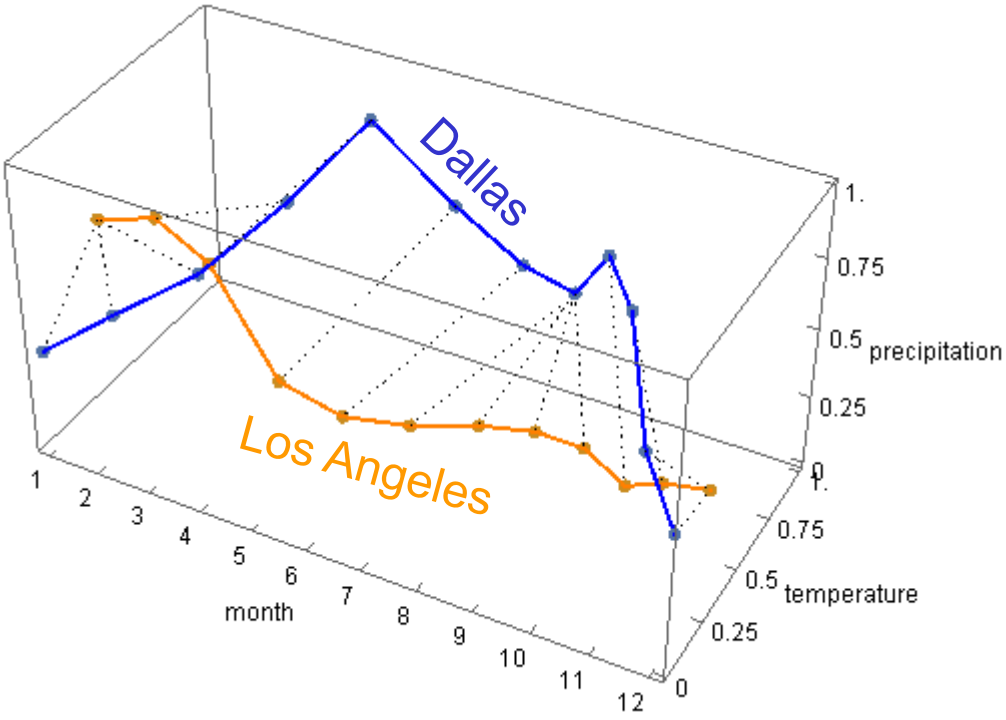
An analogy: In relativistic physics the space-time is a mathematical model that **joins** space and time into a single idea called a **continuum**. This allows calculating a single-valued distance between points in space-time

Climate dissimilarity measures

Euclidean distance



Dynamic time warping (DTW)



Euclidean distance is sufficient in segmentation because only climates in geographically restricted locations are compared.

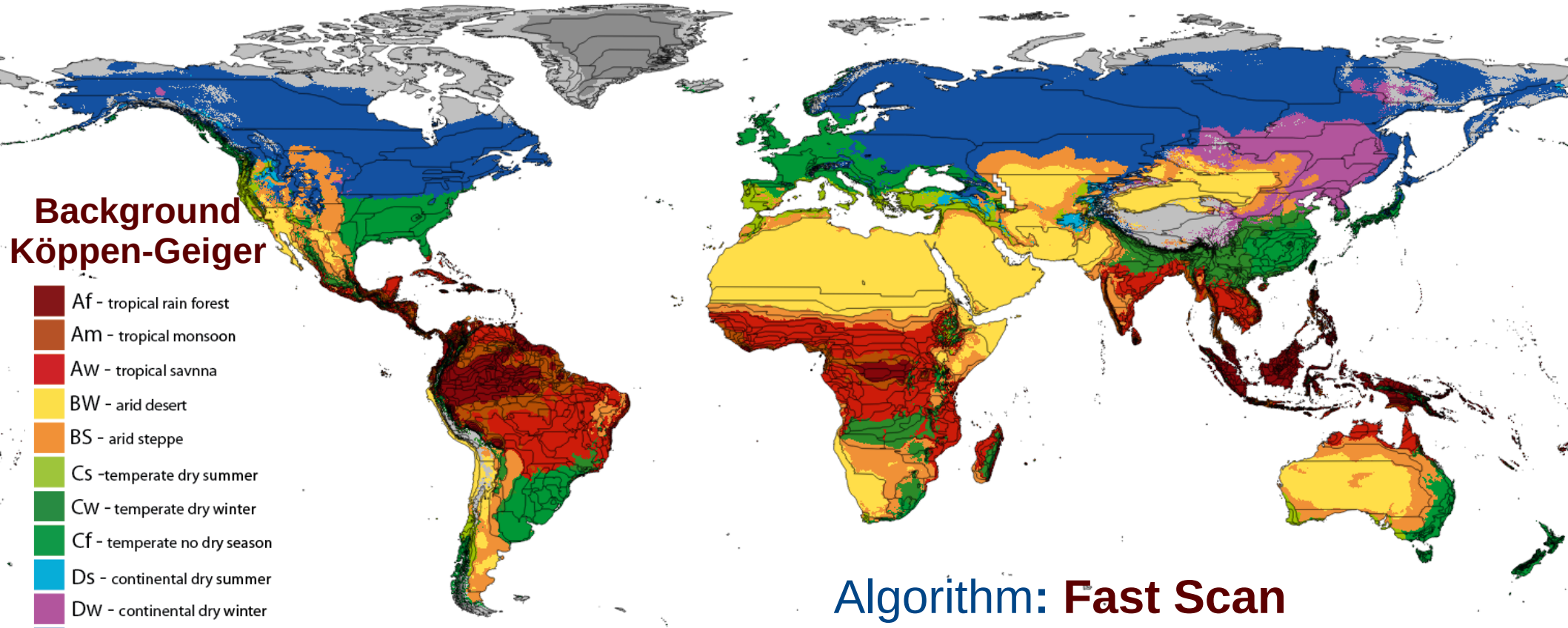
We use an improved “time-shift invariant” version of Euclidean distance

Netzel & Stepinski (20160, J. Climate, 29(9), pp. 3387–3401 .

DTW stretches or compresses two time series which are similar but locally out of phase to calculate **an optimal** match between them.

DTW is necessary in **clustering** because climates from **far away** locations (for example, at different hemispheres) need to be compared.

Segmentation of global climate



Background Köppen-Geiger

- Af - tropical rain forest
- Am - tropical monsoon
- Aw - tropical savnna
- BW - arid desert
- BS - arid steppe
- Cs -temperate dry summer
- Cw - temperate dry winter
- Cf - temperate no dry season
- Ds - continental dry summer
- Dw - continental dry winter
- Df - continental no dry season
- ET - polar tundra
- EF - polar frost

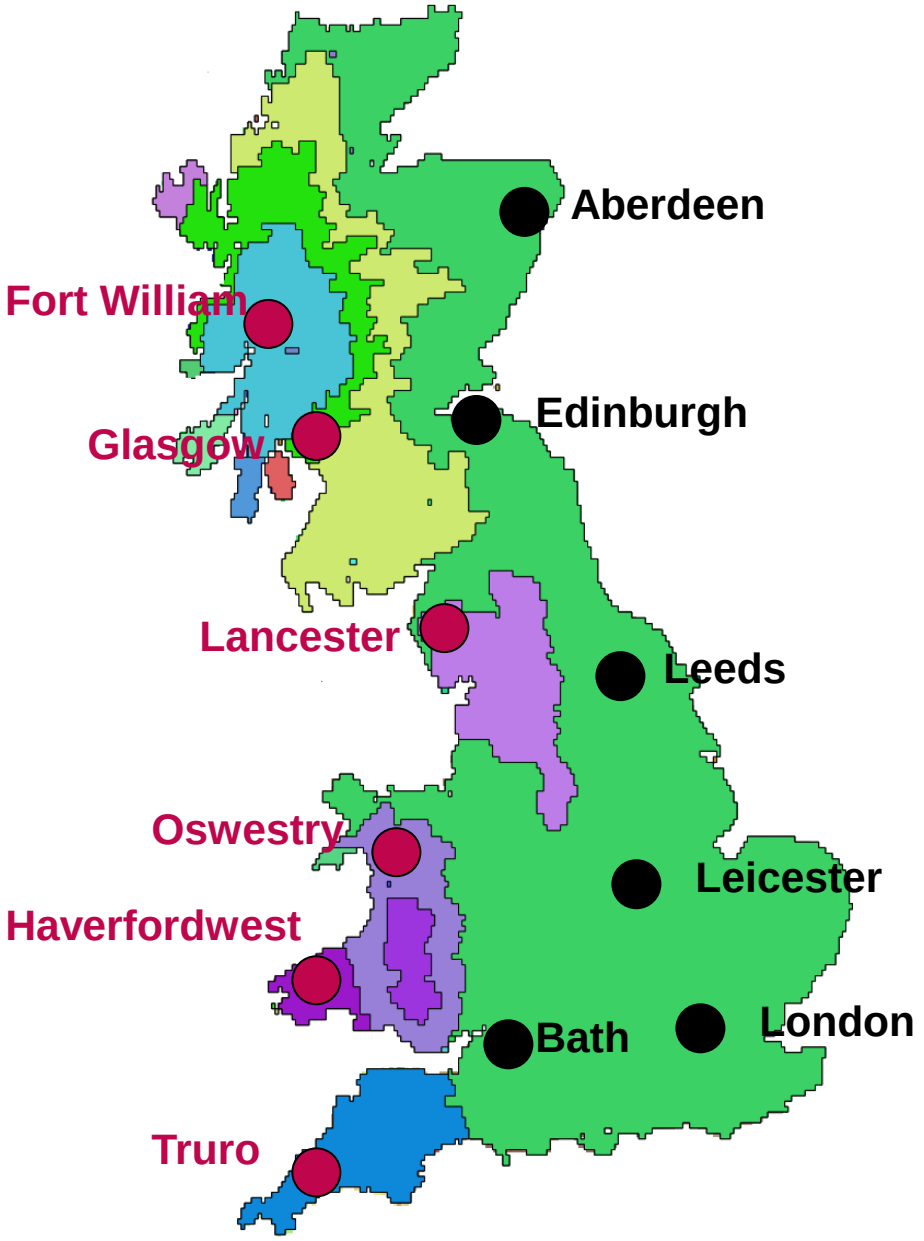
First observations

Climate changes on **smaller** spatial scale than indicated by Köppen-Geiger (especially in tropics)

In some regions segments lines are predominantly horizontal **(artifact of the Fast Scan algorithm)**

Algorithm**: Fast Scan**
Internal dissimilarity: **0.11**
External dissimilarity: **0.37**
Segments: **20,488**
Average segment area: **6700 km^2**

Climate dissimilarity in context



Dissimilarity of climate in London to climates in cities located in the same segment

D(London,Bath)=0.29
D(London,Leicester)=0.11
D(London, Leeds)=0.32
D(London,Edinburgh)=0.33
D(London,Aberdeen)=0.34

Dissimilarity of climate in London to climates in cities located in different segments

D(London,Truro)=0.70
D(London,Haverfordwest)=0.45
D(London, Oswetry)=0.40
D(London,Lancester)=0.70
D(London,Glasgow)=0.73
D(London, Ford William)=1.00

http://sil.uc.edu/webapps/climateex

Interpreting results of segmentation

Differences in spatial variability of climate shown by segmentation and Köppen-Geiger may be **surprising**.

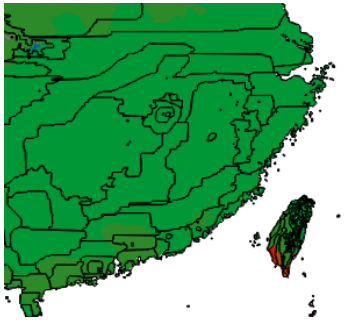
They stem from **differing definitions** of climate.

Köppen-Geiger – determined by **vegetation zones**

Segmentation – determined by how similar the **climates-as-time-series** are.

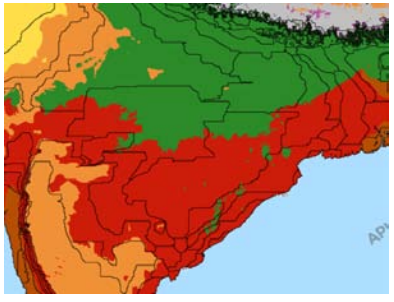
Largest differences between segmentation and Köppen-Geiger are when:

Climates in segments differ by precipitation range



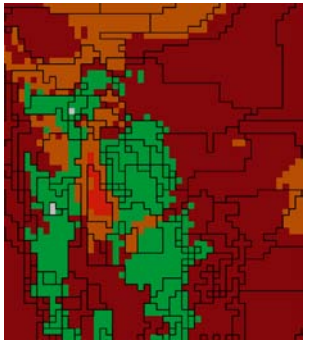
Southeast China (green = Köppen-Geiger climate Cf – temerate without dry season)

Climates in segments differ by precipitation and temperature ranges



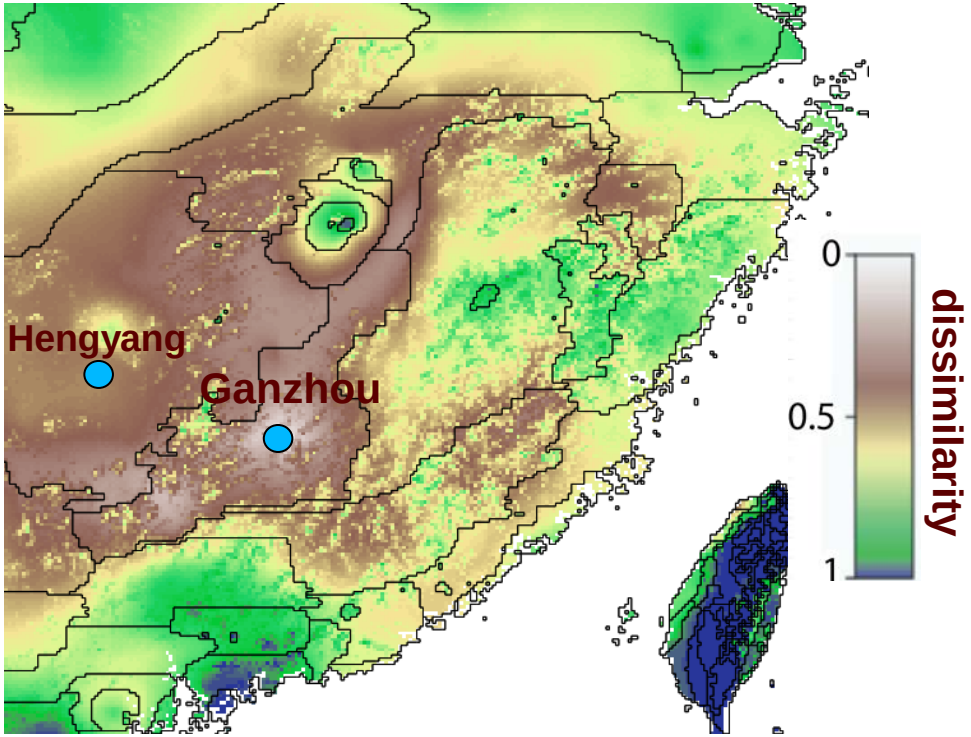
Southeast India (red = Köppen-Geiger climate Aw – tropical, savanna, wet)

In mountainous areas

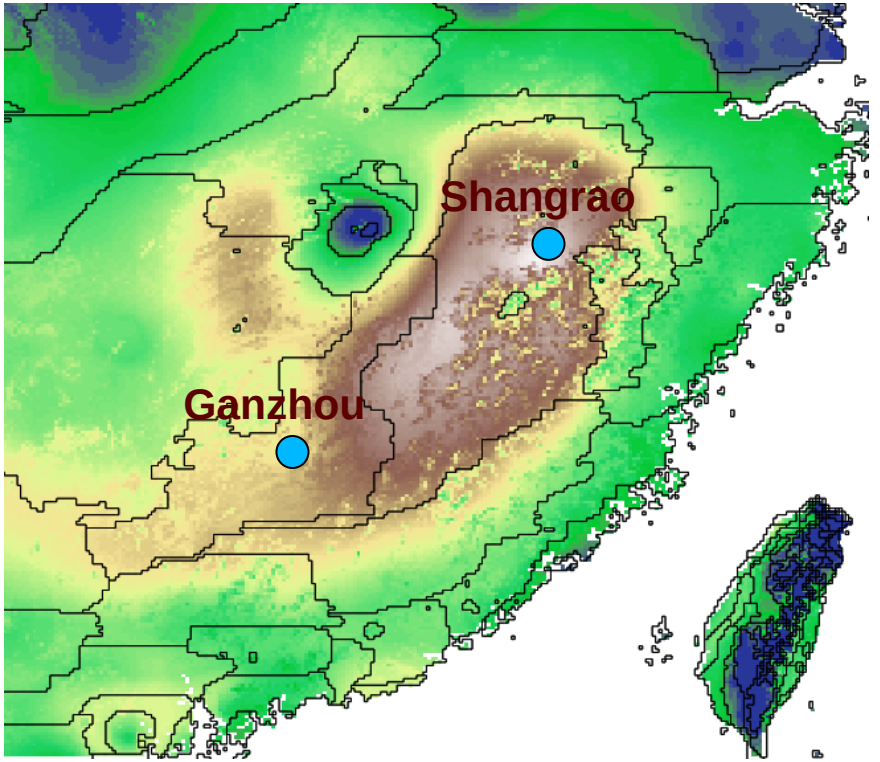


Columbian Andies (green = Köppen-Geiger climate Cf – temperate without dry season)

Climates differ by precipitation

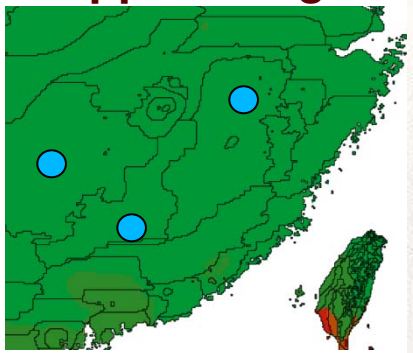


Map of climates relative to Ganzhou

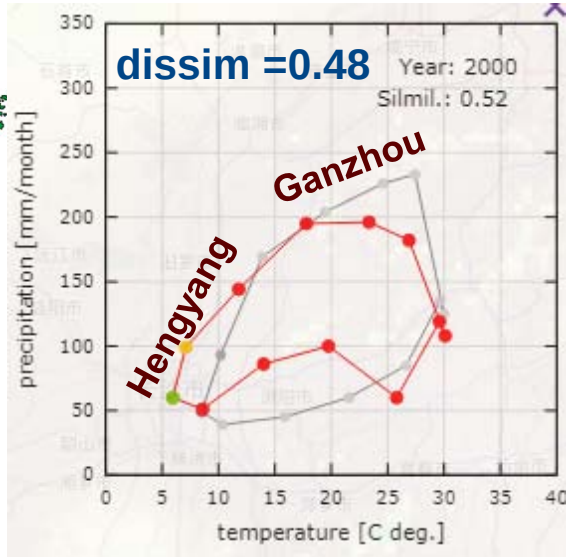


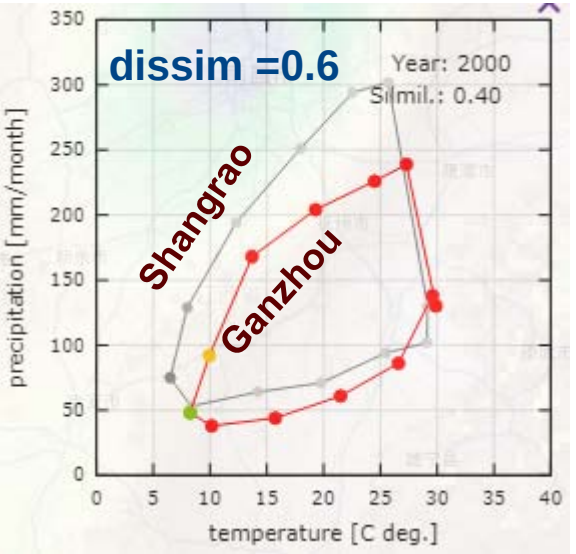
Map of climates relative to Shangrao

Köppen-Geiger

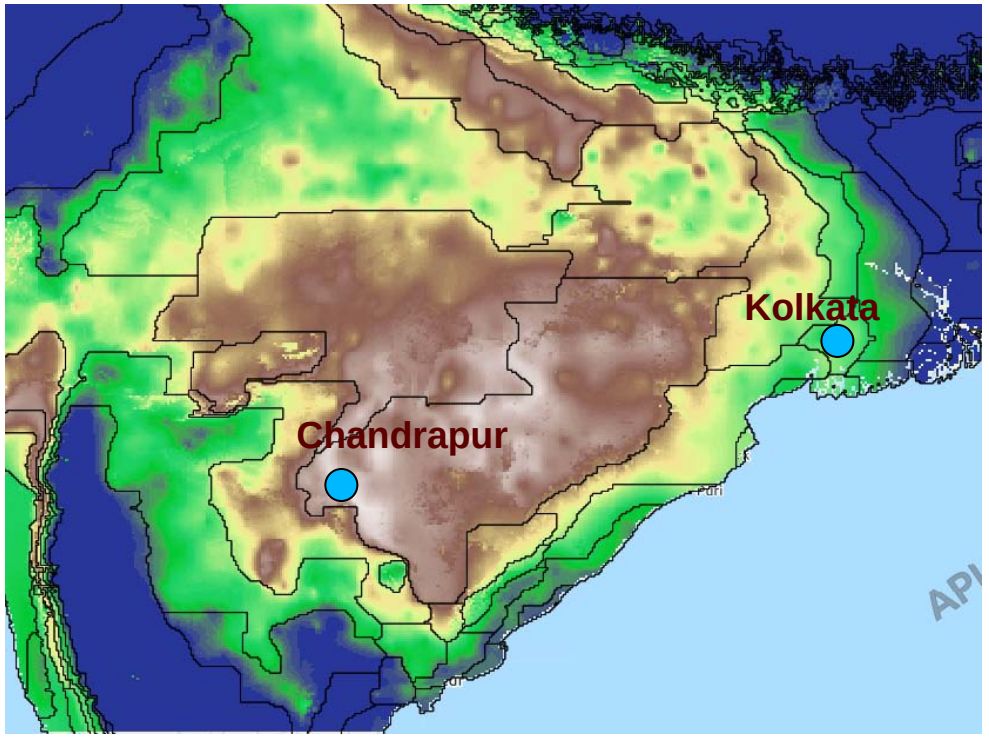


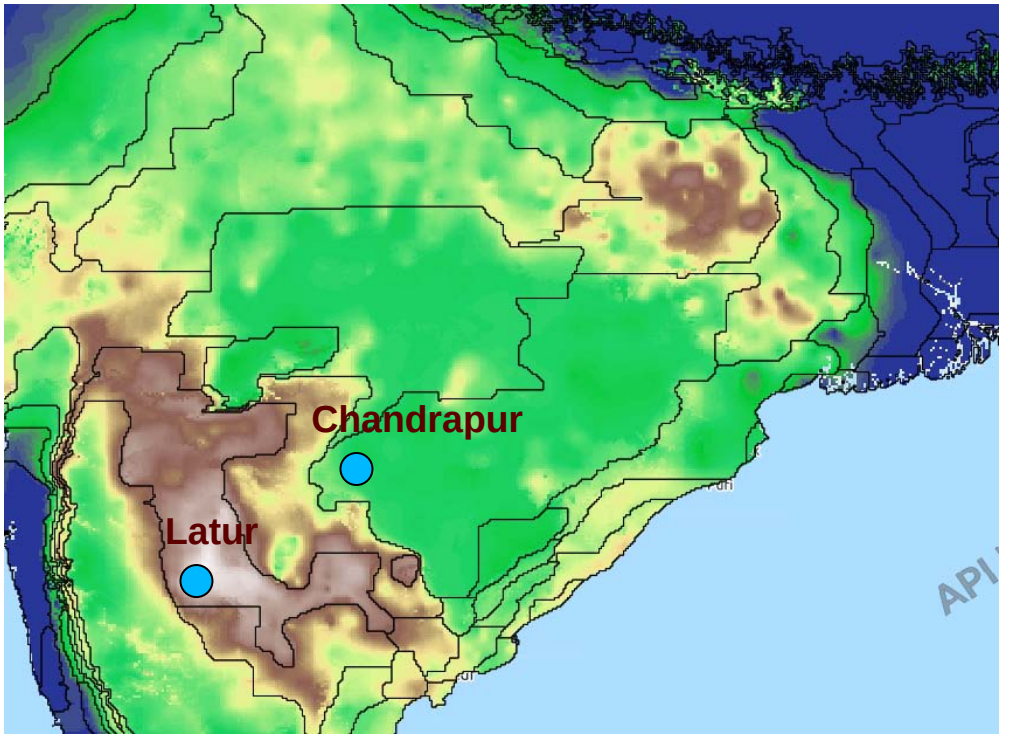
Green – Cf = temperate without dry season





Climates differ by precipitation and temperature

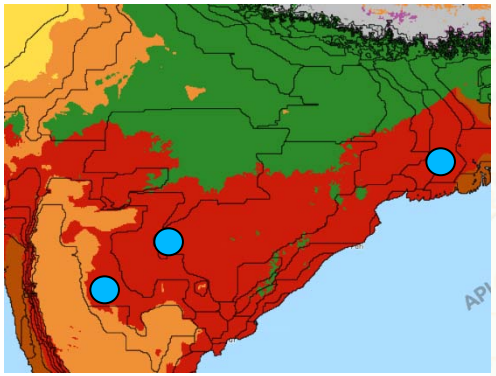




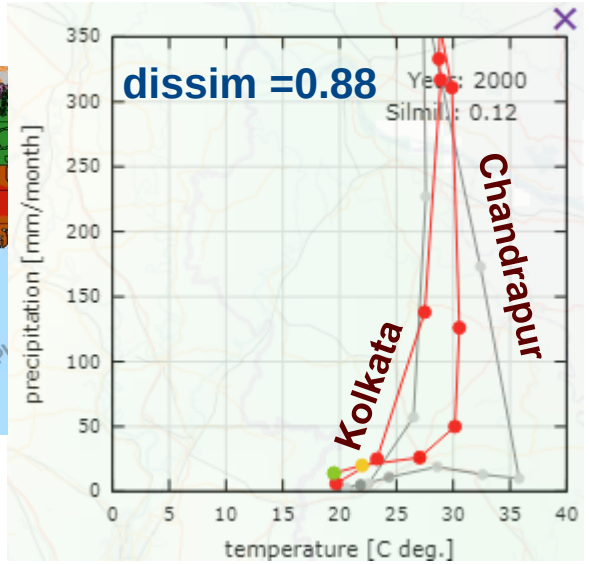
Map of climates relative to Chandrapur

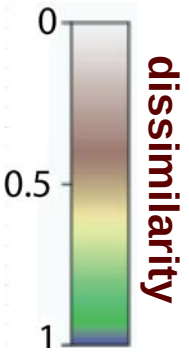
Map of climates relative to Latur

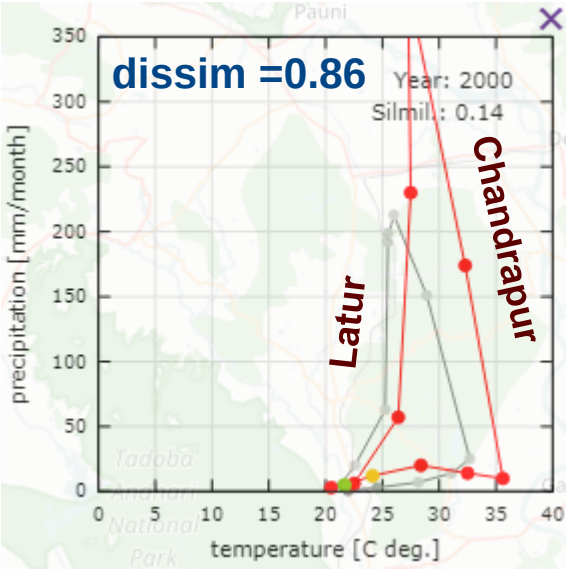
Köppen-Geiger



Red – Aw = tropical, savanna, dry

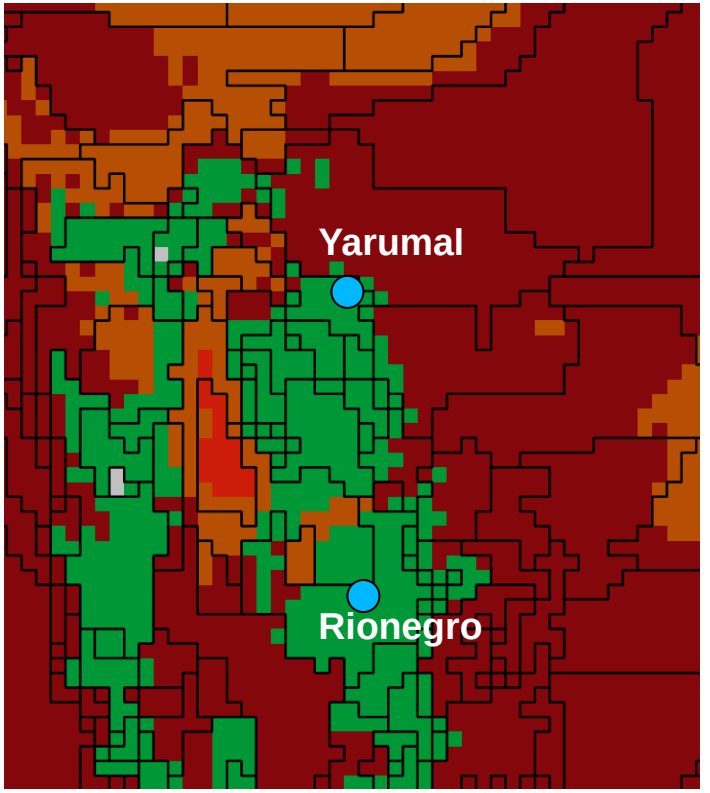




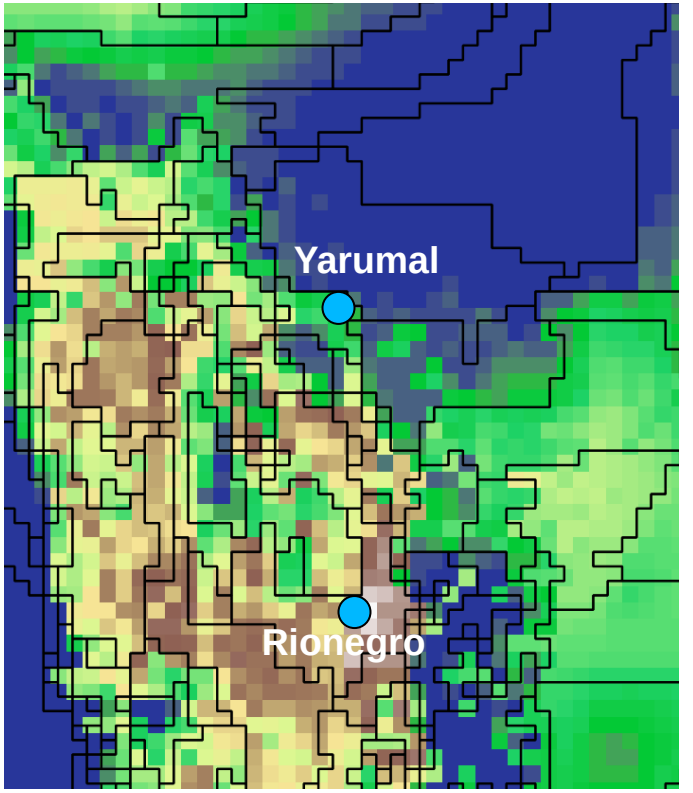


Climates differ in mountainous areas

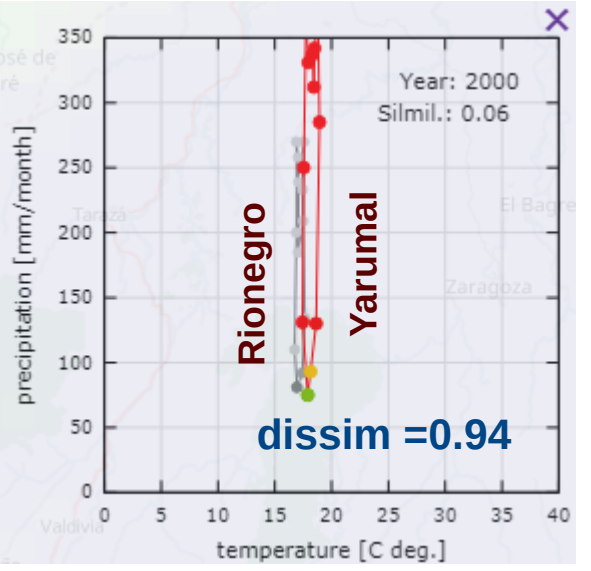
Map of Medallin region in Columbian Andies



Map of Köppen-Geiger climates
Green – Cf = temperate without dry season



Map of climates relative to Rionegro



In mountainous areas climate changes on small spatial scales. Köppen-Geiger classification puts two towns Rionegro and Yaramul in the same climate class (Cf) but the two climates differs a lot in precipitation during the wet season. Also note difference between and climates form climates in southeastern China also labeled as Cf by Köppen-Geiger

Regionalization of climates: KG vs. segmentation

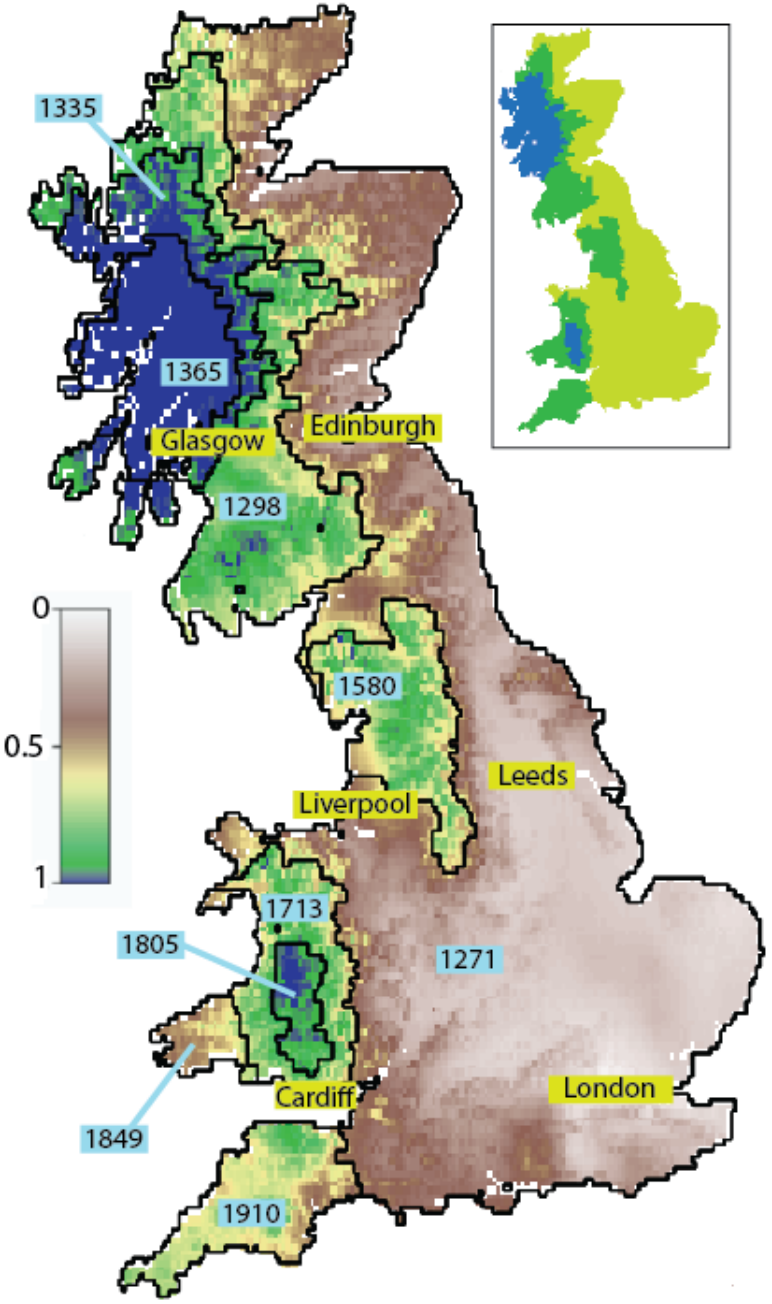
issues	Köppen-Geiger	segmentation
Climate definition	Broad, empirical and descriptive, does not take progression of seasons into account	Specific, takes into account progression of seasons and values of temperature and precipitation
Climate regions	Generally large, give qualitative description of global climate distribution	Sometimes large when there is small and gradual climate gradient. Generally small, give quantitative description of local climates
Notion of quantitative and single valued climate similarity	Does not exist. Climate is divided into 13 classes which are grouped into 5 larger classes	Numerical similarity value can be calculated between any two climates. Web exploration tool available at http://sil.uc.edu/webapps/climateex
Are climates in a single region highly similar	No – most of the time	Yes - always

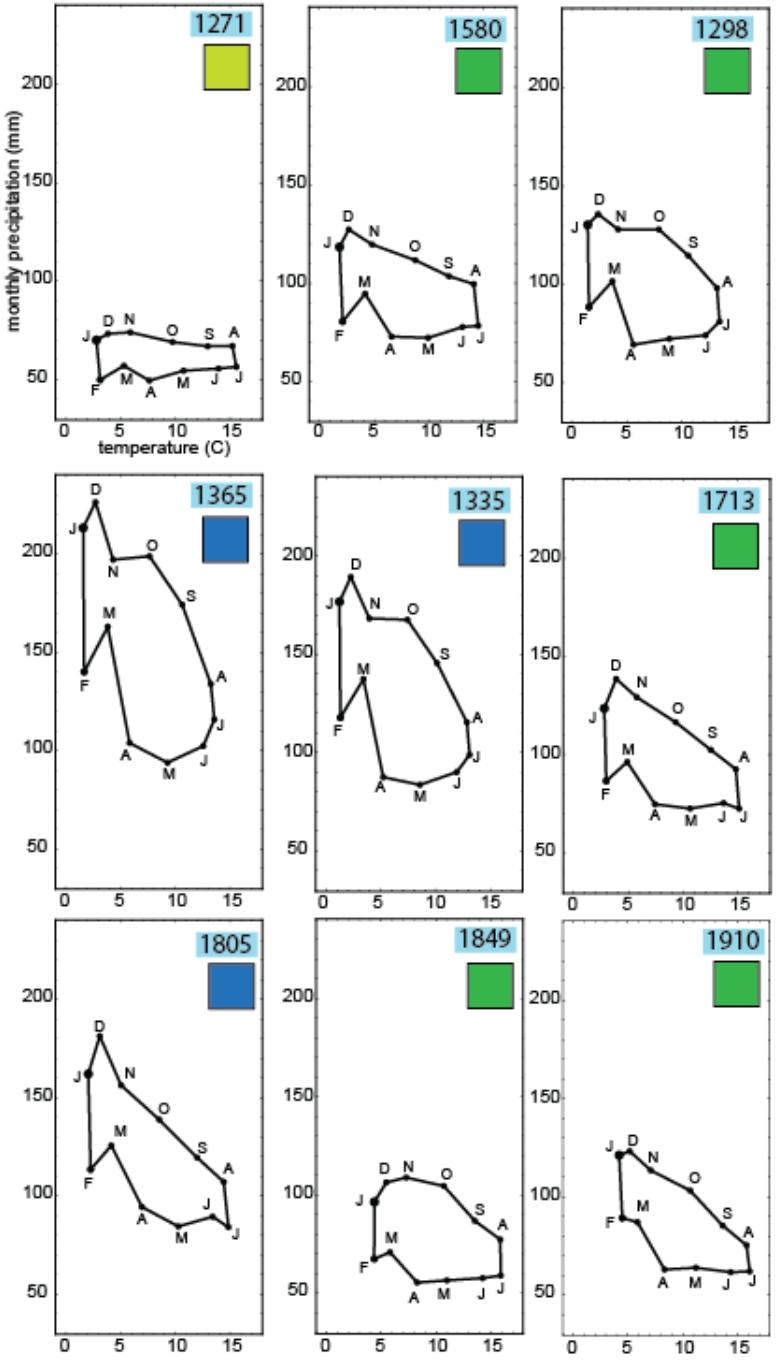
Application: Map of climates in UK

According to Köppen-Geiger the entire UK has **a single climate** – Cf.

Using global climate segmentation we can derive **more specific** map of climates in UK.

There are **nine** segments in UK, which we can optionally group into **three climate classes**





Conclusions

- Segmentation of global climate yields **a detailed map** of regional **climatic land units**.
- A combination of **global extent** and **high spatial precision** is what sets segmentation technique apart from the Köppen-Geiger classification and climate clustering.
- At present segmentation yields **the best regionalization** of terrestrial landmass into **land units of homogeneous climate**.
- Whereas Köppen-Geiger provides **a first order qualitative** information about climatic zones, segmentation provides **quantitative delineation** of climatic land units.
- Using segmentation-generated regionalization **detailed maps** of climatic land units **in smaller regions** (see the UK example) may be produced via **optional classification** of segments.