Week 4: Exercises on principal component analysis

Exercise 1:

You are given daily precipitation data for Alberta and British Columbia from 2006 – 2015. In total the dataset contains 3862 days of precipitation. Start by running the script **Lab4_part1.m** (**Lab4_part1.ipynb**) in order to load the dataset and plot some initial figures.

Specific questions:

1) Find the eigenvectors and PCs. What do the PCs of the first mode mean?

2) How many modes are enough to effectively reconstruct the data? [*Hint: you need to decide yourself what 'effectively' means, i.e. what you consider to be a successful reconstruction. The way to decide is to look into how much variance in total is explained by the first several modes*.]

Exercise 2:

High Mountains Asia (HMA) is Asia's "water tower" and hosts the largest glacier concentration in the world (outside of the ice sheets). These glaciers are important contributors to many prominent Asian rivers in one of the most populated areas of the world. Many studies have shown that there is no uniform response of HMA glaciers to climate change as the region experiences heterogeneity in climate setting, climate change signal and sensitivity of glaciers to climate change. A majority of glaciers in HMA experience mass loss that is most pronounced for glaciers sitting in a 'maritime' climate regime of Himalaya and Nyainqêntanglha. On the other hand, glaciers in a more 'continental' climate regime of Karakoram/Pamir region experience mass gain. To better understand the spatial and temporal differences in glacier mass balance, you are given the following data-set: reconstructed time series of annual mass balance for 1952-2014 for each of 45 glaciers in the region. The reconstruction is done by empirical modelling with all available mass balance and meteorological observations in the region (see Figure 1).

The data is given in **mass_balance_series.xlsx**. Open **Lab4_part2.m (Lab4_part2.ipynb)** for more instructions on how to load the data and do some basic plotting.

Specific questions:

1) What are the most characteristic temporal patterns of glacier mass balance for this region? How are these temporal patterns expressed spatially, i.e. for each of these 45 glaciers? [*Hint: eigenvectors here are temporal patterns (m=63 years), while PCs are glaciers (n=45)*].

2) What are the most characteristic spatial patterns of glacier mass balance? How do these patterns evolve in time throughout the period 1952-2014? [*Hint: eigenvectors here are spatial patterns, i.e. patterns across the glaciers; Hint: m=45, n=63*]

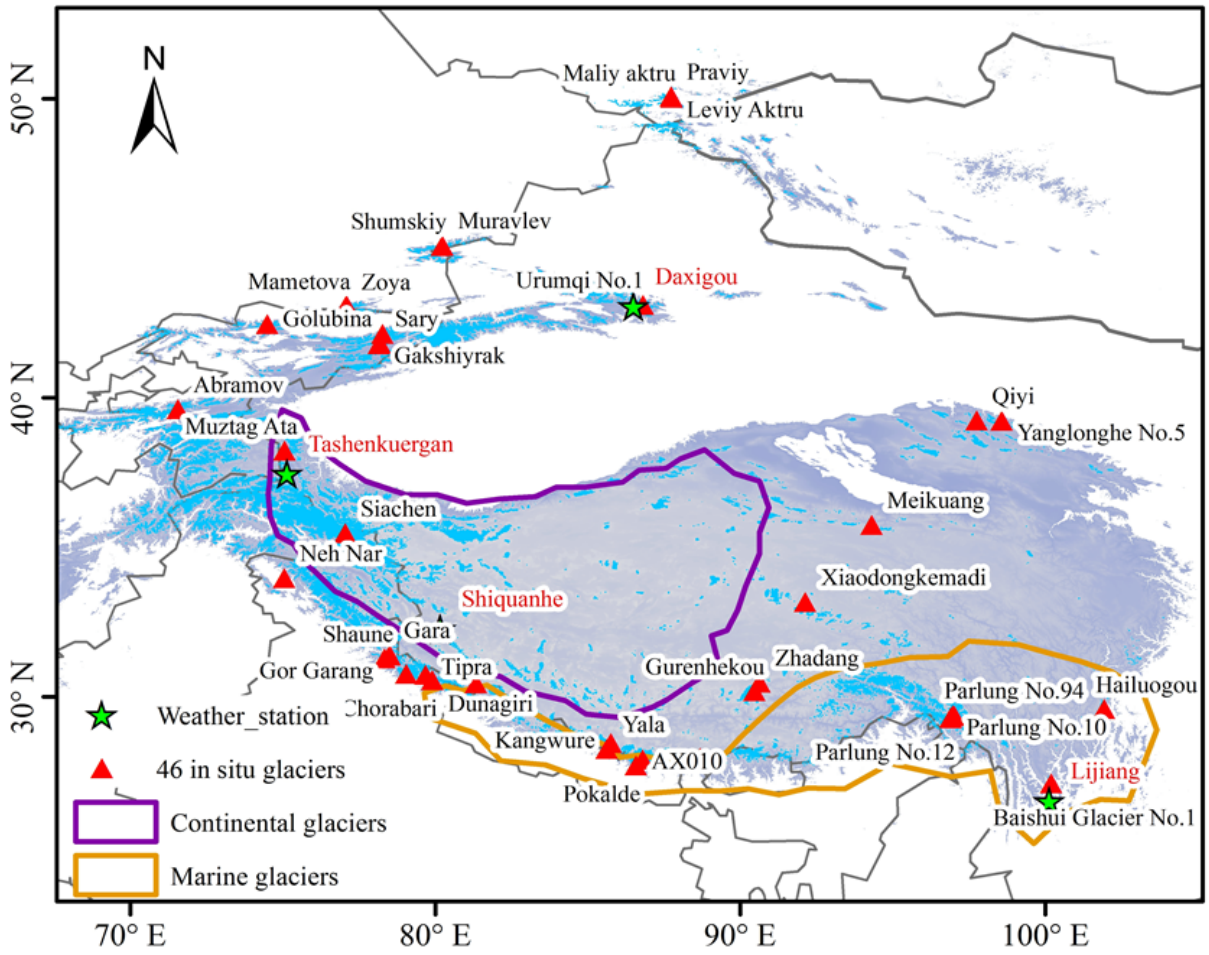

Figure 1 (copyright @Rongjun Wang): Spatial distribution of 45 glaciers in HMA (red triangles), 4 national weather stations (green star) and different types of glacier boundary (maritime glaciers with deep purple, continental glaciers with orange). Grey shading depicts the areas with elevation above 2500 m.