Part 1:
Winter temperature in the Swiss Alps


Figure 1: Montana. (left) and Zermatt (right). Source: `https://www.ilikeswitzerland.com` and Source: `https://www.zermatt.ch`.

Given the Alpine regions' susceptibility to extreme weather events, there is a clear rationale for studying such climatic occurrences, as they can have substantial consequences for the local population and infrastructure, often resulting in significant financial burdens. Instances of warm winter spells, characterized by prolonged periods of unusually high temperatures during winter, can have profound impacts on natural ecosystems, agriculture, and water supply. We are interested to model the relationship between extreme winter temperatures (specifically in December, January, and February) at two locations in the Swiss Alps: Montana, situated at an elevation of 1427m, and Zermatt, located at an elevation of 1638m. These sites are approximately 37km apart. We are also interested to analyse the extremal dependence between one of the resort station and the NAO. The NAO index represents the normalized pressure difference between Iceland and the Azores and is recognized to have a significant direct impact on temperatures in the alpine region, particularly during the winter season.

(a) Read in the data and select only the winter temperature data. You can use the code lines of the file `GetData.R` to get the data and select only the winter values.

(b) Represent the data and especially scatterplots of Zermatt values against Montana's values. Same with Montana values against NAO values. Comment.

(c) Test the correlation between these two pairs of series.

(d) Now concentrate on the extreme level and tail dependence. Represent χ and $\bar{\chi}$ for two pairs of series (Montana vs Zermatt and Montana vs NAO). You may want to use the function `chiplot` of the package `evd`. What do you observe?

Part 2:
Copula in finance

In a famous article from 2009 for which he won the American Statistical Association's Excellence in Statistical Reporting Award, Felix Salmon described how the use of the Gaussian copula led to massive losses for investors during the 2008 financial crisis. In *Recipe for disaster: the formula that killed Wall Street*, the author explains how, since the 1980s, the Gaussian copula has become the go-to model of financial analysts in order to model and handle dependence across assets. He underlines the role this choice had during the crisis.

In this exercise, we look at two stocks from the CAC40 index in France: **Engie** (engie.csv), a French utility company operating in the electricity market as well as natural gas, nuclear, renewable energy and petroleum and **Veolia** (veolia.csv), another French energy services provider. We would like to test the Gaussian copula assumption on real negative log returns data, for a period ranging from 2005 to 2022.

(a) Read in the data. Then, plot the raw prices (*Adj.Close*) of both stocks. Discuss the stationarity assumption.

(b) First, compute and plot the negative log returns in independent plots. Discuss the dates of occurrence of extreme events. Then, produce a scatterplot of the negative log returns of both stocks and discuss their (visual) dependence.

(c) Plot the pseudo-samples obtained from estimating the marginal distributions by their empirical CDFs.

(d) Fit a Gaussian, a t, a Gumbel and a Clayton copula to the data. Select the copula model using an information criterion, such as the AIC.

(e) Compute the theoretical tail dependence coefficients for the coulae fitted in (d), using the parameters obtained above. Hint: use the `BiCopPar2TailDep()` function from the `VineCopula` package.

(f) Compute the non-parametric estimators of the tail dependence coefficients obtained from the data. Are these estimates close to the theoretical ones obtained with the best model selected in (d)?