Ca’ Foscari
University
of Venice

Single Cycle Degree programme

in Environmental Sciences, Global change and Sustainability

Final Thesis

# Climate response to the 1815 Mount Tambora Eruption: Analysis of a Multi-model Simulation Ensemble

**Supervisor**
Ch. Prof. Davide Zanchettin
**Assistant supervisor**
Ch. Prof. Angelo Rubino

**Graduand**
Margaux Marchant
Matriculation Number 893720

**Academic Year**
2022 / 2023

# Abstract

This Thesis explores the climatic impacts of the 1815 Mount Tambora eruption, an event that significantly altered global climate patterns, notably causing the "year without summer" in Europe the following year. The study uses an ensemble of simulations of the Tambora eruption with three climate models (MPI-ESM1-2.LR, CanESM5, and MIROC-ES2L), each generating nine realizations, to discern whether the diversity of climate responses in the multi-model ensemble is determined more by model specificities than by the initial conditions, here determined by El Niño Southern Oscillation (ENSO). The Thesis work is structured as follows. First, a review is made of the state-of-the-art knowledge about the Mount Tambora eruption and its consequences on the global climate, and information about the models and the Tambora experiment is acquired. Following this, data from the multi-model ensemble is acquired and pre-processed, to homogenize the spatial structure of the data. Then, the study delves into the description and application of mathematical and statistical tools for data analysis, including principal component and cluster analysis. The analysis is conducted on seasonal gridded anomalies of near-surface air temperature for winter and summer of 1816. The results reveal the diversity of post-eruption anomalies within each model and across different models, suggesting a relevant contribution by noise and initial condition dependencies. Model specificities remain substantial even when excluding the noisy components, indicating that multiple response mechanisms exist, and that these are differently represented in the considered models. The Thesis ends with a discussion of the results and an outlook on possible future developments of the work.

# 1. Introduction

The release of aerosol precursors in the atmosphere by volcanic eruptions is an important cause of climate variability, primarily because of the temporary enhancement of the stratospheric aerosol layer preventing solar radiation from reaching the earth's surface, thus resulting in a negative radiative forcing of climate. This phenomenon tends to induce a strong global cooling of the earth’s surface temperature after large volcanic eruptions, as occurred with the Mount Tambora eruption in 1815. As a result, the Mount Tambora eruption has served as a reference for studying the climatic response following volcanic eruptions. (Raible et al., 2016; Timmreck 2012). Despite its fame, the understanding of the climate response to this event shows uncertainties due to gaps in the knowledge of the volcanic forcing intensity and the mechanism of climate responses (Zanchettin et al., 2019).

Even before focusing on the climate response, already our knowledge of the volcanic forcing produced by the Tambora eruption shows uncertainties. The amount of sulfur injected into the atmosphere is now well estimated (Toohey, 2017) thanks to material preserved in natural records such as ice cores (Sigl et al., 2015). However, the volcanic forcing derived from it is uncertain due to our still partial understanding of aerosol microphysical processes and necessary simplifications in their implementation in numerical models, which complicates the description of the spatial and temporal evolution of volcanic aerosols in the atmosphere after an eruption. Aerosol climate models respond to large volcanic eruptions based on input parameters such as the quantity of $SO_2$ injected into the atmosphere (Timmreck, 2012). When using dynamic aerosol climate models, the models calculate the spatio-temporal evolution of the aerosol cloud based on these injections, resulting in varying aerosol optical properties and volcanic forcing across different models (Zanchettin et al., 2019; Marshall et al., 2018). However, in approaches like VolMIP[1], aerosols are not explicitly calculated; instead, predefined aerosol optical properties are imposed to avoid the computationally expensive process of simulating aerosol microphysics.

In 2019, Zanchettin et al. studied the role of forcing uncertainties and initial‑condition unknowns on the climate response to volcanic eruptions, focusing on simulations forced by the Mount Tambora eruption of 1815. The study was based on an ensemble of simulations using the same model with different volcanic forcing, sampling the range of reconstructed magnitude estimates for the Tambora eruption, and encompassing as much as possible different initial conditions, revealing that in summer, the influence of forcing uncertainties tends to overshadow the spread caused by initial conditions due to the pronounced direct radiative impact. Conversely, during winter, the effect of initial conditions predominates. The uncertainties about what was the forcing of Tambora, what determined the model response and why model outputs differ, remain. The figure 1 displays the differences in responses from simulations from the same model but different volcanic forcing.

---

[1] VolMIP : Model Intercomparison Project on the climatic response to Volcanic forcing

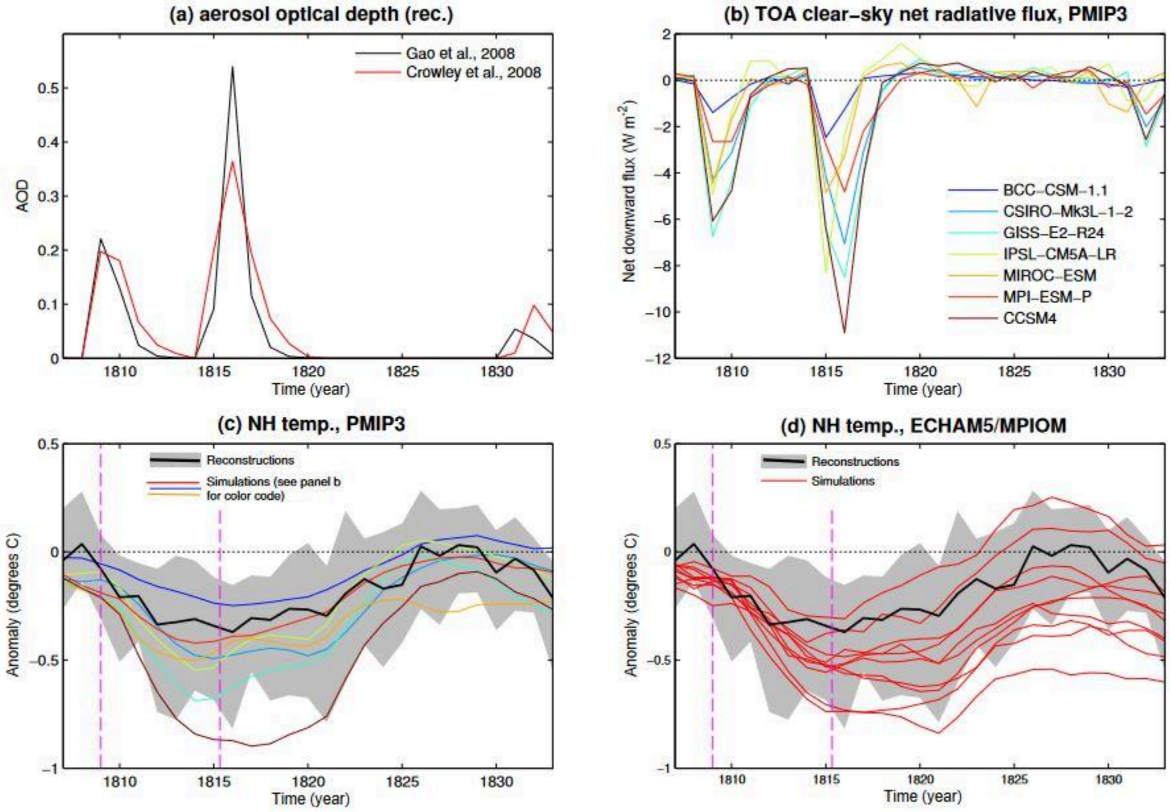


Figure 1 : Uncertainties in volcanic forcing and climate model responses to eruptions in the early 1800s (a). Two estimates of annual-average global aerosol optical depth (AOD) at 550 nm; (b) Net clear-sky radiative flux anomalies for a multi-model ensemble of last-millennium simulations (PMIP3; see: Braconnot et al., 2012); (c) comparison between simulated and reconstructed Northern Hemisphere average summer temperature anomalies (relative to 1799–1808) of different models; (d) same as (c) but single-model ensemble. Adapted from Zanchettin et al. (2016).

This thesis aims to assess how consistent the post-Tambora climate anomalies are across different models and if the diversity of climate response of the multi-model ensemble is determined more by the model specificities than by the initial conditions, here identified based on the state of the El Niño Southern Oscillation.

The thesis is structured into five chapters. The first chapter provides a comprehensive state of the art, offering a thorough review of the existing literature. The second chapter delves into a detailed description of the data employed in this study. In the third chapter, the methodology used in the research is described. The fourth chapter presents and analyzes the various results obtained. Finally, the fifth chapter concludes with a discussion concerning the key findings of this thesis and exploring potential possibilities for future research.

## 2. State of the art

Mount Tambora is an active stratovolcano located on Sumbawa Island in Indonesia (8°15′S 118°0′E; figure 2), reaching 2850 meters (4300 meters before the eruption) (Pfister et al., 2018). The volcano is situated in the active inner volcanic arc of the SundaBanda (Self, 1984). The volcano was dormant for centuries until signs of an impending eruption appeared in 1812, marked by rumbles and dark cloud formations. The main eruption (10-11 of April 1815) resulted from the gradual cooling of magma in a closed chamber between 1.5 and 4.5 kilometers deep. The eruption, with a pressure of 4-5 kbar and a chamber temperature of 700-850 °C, led to the collapse of the volcano roof, releasing 33 km3 of magma, believed to be homogeneous trachyandesite (Foden, 1986; Raible et al., 2016). Sounds of the explosion were heard from more than 2500 km away. The region was immersed into total darkness for two days, the first sun rays were not seen until April the 12th, two days after the eruption. Tsunamis had been induced due to pyroclastic density currents shoving into the water, with some waves reaching four meters high. The tambora volcano continued to rumble until 1818.

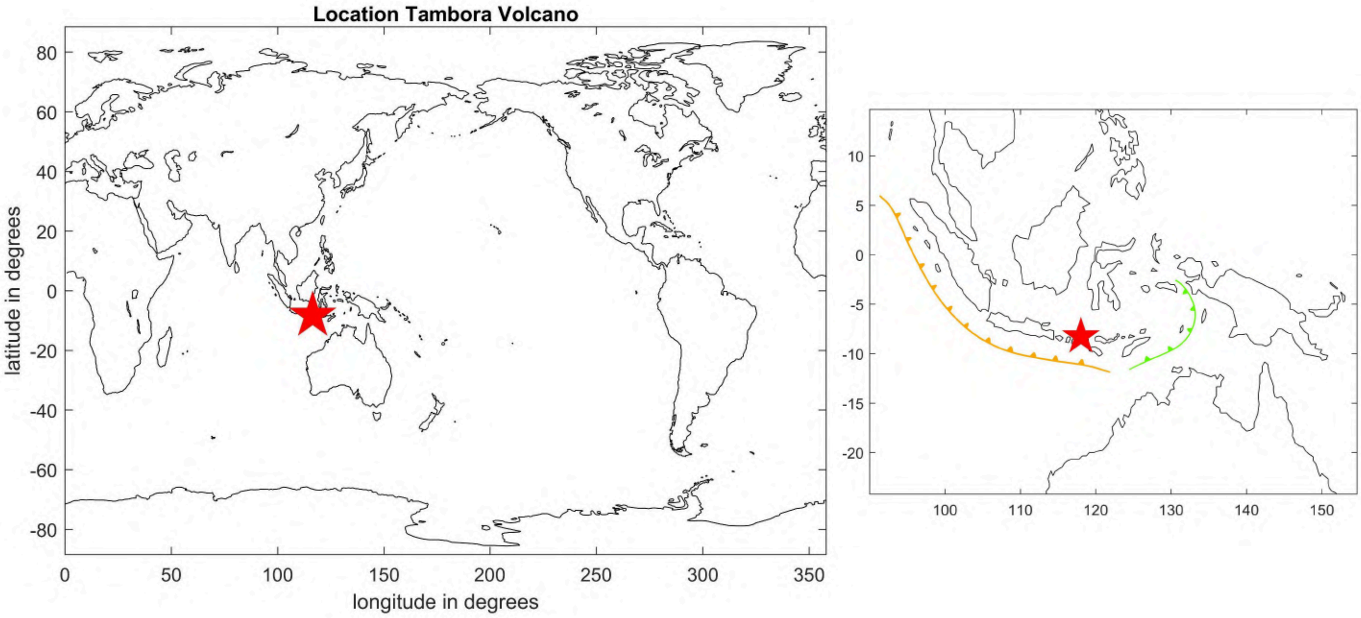


Figure 2: Location Mount Tambora, represented by the red star. The same map layout as the left panel is used later to present the results. The Sunda volcanic arc is represented in orange, while the Banda volcanic arc is represented in green.

The Mount Tambora eruption was a Caldera-forming event, associated with large and long-lived chambers (Suhendro et al., 2021). Caldera-forming eruptions are characterized by the spread of multiple hazardous pyroclastic density currents, as observed in the case of Mount Tambora, where two such currents occurred. Pyroclastic density currents consist of solid particles and gas, displaying hazardous behaviors and significantly impacting the local area (Sulpizio et al., 2014; Cole et al., 2015). This eruption is known as one of the largest volcanic events in the past millennia (Self et al., 1984). It unleashed approximately 140 gigatons of magma, propelling plumes up to 43 kilometers into the stratosphere, resulting in over 70,000 direct fatalities in the surrounding area. The impact extended to agriculture and food supply, causing after-eruption famines in Bali and Lombok. Water was contaminated because of the ashes, causing multiple diseases within the local populations . Worldwile, the eruption is also thought to be the cause of the proliferation of diseases such as Cholera in the Bengal region or typhus in Europe and the eastern Mediterranean region (Oppenheimer, 2003; De Jong Boers, 1995).

Beyond its immediate consequences, the eruption was believed to have resulted in the injection of around 60 megatons of sulfur into the atmosphere during the eruption creating a natural barrier to radiative forcing (Oppenheimer, 2003). The eVolv2k database provides new estimates of stratospheric sulfur injection derived from ice-core data, indicating that the Tambora eruption released approximately 30 megatons of sulfur into the atmosphere (Toohey et al., 2017). The sulfur, in the form of sulfur dioxide ($SO_2$) and hydrogen sulfide ($H_2S$) that oxidizes within days into $SO_2$, reacts with water vapor to form sulfate aerosols (Self et al., 2005). These aerosols reflect incoming solar radiation back to space, inducing a cooling effect on the Earth's surface. The spread of sulfur dioxide circled the tropics within two to four weeks, with sulfate gases oxidizing into sulfate aerosols after about eight weeks. The aerosols were therefore transported towards the poles, taking two to five years for the particles that made it to the poles (Brönnimann et al., 2016). After a large eruption, particles inside the first 15 kilometers of the atmosphere disappeared within a few weeks (Luterbacher et al., 2015), while the particles forming the radiative barrier above 15 kilometers can last a maximum of two to three years, more frequently between one to two years (Self et al., 2005). Because of its location, the eruption had a worldwide impact, indeed, due to the atmospheric circulation volcanic eruptions located in the tropics impact both the North hemisphere and South hemisphere (Hagen, 2023).

El Niño is a significant warming of the tropical Pacific Ocean that happens every three to seven years. It is part of a climate phenomenon known as the El Niño Southern Oscillation (ENSO), marked by an interannual variability in air pressure over the Tropical Pacific. ENSO has two main phases: a positive phase, which leads to warmer conditions called El Niño, and a negative phase, resulting in cooler conditions known as La Niña. This cycle of warming and cooling has a notable impact on global weather patterns (Wang et al., 2017; Trenberth, 1996).

The North Atlantic Oscillation (NAO) index is defined by the contrast in surface sea-level pressure between the Subtropical High and the Subpolar Low. This index exhibits two phases, each associated with distinct temperature variations. Positive phases of the NAO typically correlate with warmer temperatures in the Eastern United States and northern Europe, yet colder temperatures in Greenland, and frequently in southern Europe and the Middle East. Conversely, during a negative phase of the NAO, the temperature patterns are reversed (figure 3). The NAO demonstrates a pattern of variability over the course of decades, showing significant changes that can last longer than ten years (Woollings et al., 2015; NOAA Climate.gov).

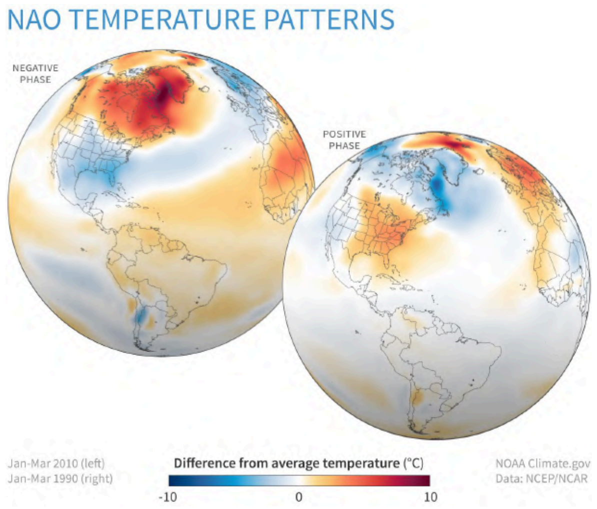


Figure 3: NAO temperature patterns, negative and positive phases. Image from the National Oceanic and Atmospheric Administration (NOAA).

In the following year, 1816, known in Europe as the "year without summer," global temperatures decreased by 0.4 to 0.7 °C (Stothers, 1984). Despite this global temperature decrease, the changes in temperatures during the summer of 1816 were not homogeneously distributed around the Earth's surface. Some locations, such as Western Europe and North America, experienced a decrease in temperature, while Eastern Europe, Western Russia, and Eastern Scandinavia showed no changes or a weak increase in temperatures. Tree ring records indicated a cooling of the summer of 1816 in most of the Arctic, Northern Europe, Eastern North America, and Asia (Raible et al., 2016). In Europe, the general temperature drop had consequences on harvests and water supplies, delaying crop growth, and preventing many plants from reaching maturity (Luterbacher, 2015).

The winter of 1816 was characterized in the Northern Hemisphere by a phenomenon that is currently known as post-eruption winter warming, attributed to the North Atlantic Oscillation (NAO) typically developing a positive phase after a strong volcanic eruption. Positive phase of the NAO leads to warm anomalies over Northern Europe in winter (Wunderlich et al., 2017). According to (Shindell et al., 2004), winter responses to large tropical eruptions involve warming in Northern Eurasia, while negative temperature anomalies cover Northern Africa and the Middle East. The ENSO response to volcanic eruptions is stronger in summer than winter, a response quantified in the study on another large eruption, the Pinatubo Volcano in 1991 (Predybaylo et al., 2017; Christiansen, 2008). In 2022, Zanchettin et al. demonstrated that the post-eruption signal after the 1991 Pinatubo eruption is related to the initial state of ENSO.

Climate models are projections of climate variability through time, they can be used to represent the future, especially in the climate change context that we are currently experiencing, but they can also be used to reproduce past events, generally past climates. Climate models are based on grid decomposition of the earth system, this grid is most of the time three dimensional with the latitude and longitude and the vertical representation. Climate models are based on physical and biogeochemical processes representing the system desired, such as the atmosphere and ocean systems. These models can also be coupled creating atmosphere/ocean models or earth system models if more modules are added. Climate model calculations are based on inputs, representing different forces describing the climate state such as the initial state of ENSO and NAO or by the introduction of elements inducing events ($SO_2$ injection in our case). Climate models are not always reliable, a lot of bias is induced due to the large scale of climate models, the lack in knowledge and simplification of physical and chemical processes and the estimation of past initial conditions. Biases also emerge as a result of the ocean-atmosphere coupling.

# 3. Data description

## 3.1 volc-long-eq experiment

The volc-long-eq experiment is an experiment of the “Model Intercomparison Project on the climatic response to volcanic forcing” (VolMIP, Zanchettin et al., 2016): where volcanic forcing from an idealized Tambora eruption, applied in terms of aerosol optical properties, is imposed on various global ocean-atmosphere coupled/earth system models under controlled initial conditions from an non-altered climate, to explore model diversity in the climate response to the eruption. Currently, three models have participated in the experiment (further described in the following section). For each model, nine realizations[2] are performed with different initial conditions including warm, cold, and neutral phases of the El Niño-Southern Oscillation (ENSO), and strong, weak, and neutral phases of the Atlantic Meridional Overturning Circulation (AMOC). The aerosol optical properties used for the volcanic forcing are provided by the Easy Volcanic Aerosol (EVA) tool (Toohey et al., 2016). The volc-long-eq experiment has background climate conditions set to pre-industrial values, and sampled from a control experiment (piControl: preindustrial control); they should be representative of the Earth’s climate around the year 1850, with uncertainty linked to the unknown states of ENSO and AMOC. These forcing constraints included in setting the preindustrial conditions (set to be constant at their 1850 values) are: atmospheric carbon dioxide concentration, well-mixed greenhouse gas atmospheric concentration excluding carbon dioxide, aerosols (methane, nitrous oxide and sulfates), ozone concentrations, water vapor concentrations, land use, and solar irradiance. As initial conditions for the different realizations of the experiment are taken from the piControl, paired anomalies (volc-long-eq(s,t) minus piControl(s,t), where ‘s’ indicates the spatial dimension and ‘t’ the time dimension) can be calculated, that is differences between the climate state including the volcanic eruption and that without volcanic forcing, with the same amount of unperturbed internal variability (ENSO, AMOC) in both datasets.

---

[2] Outputs of model, results of the model under chosen conditions (here the ENSO and AMOC values)

## 3.2 Models

Three models[3] were used, each contributing with nine realizations :

MPI-ESM1-2.LR

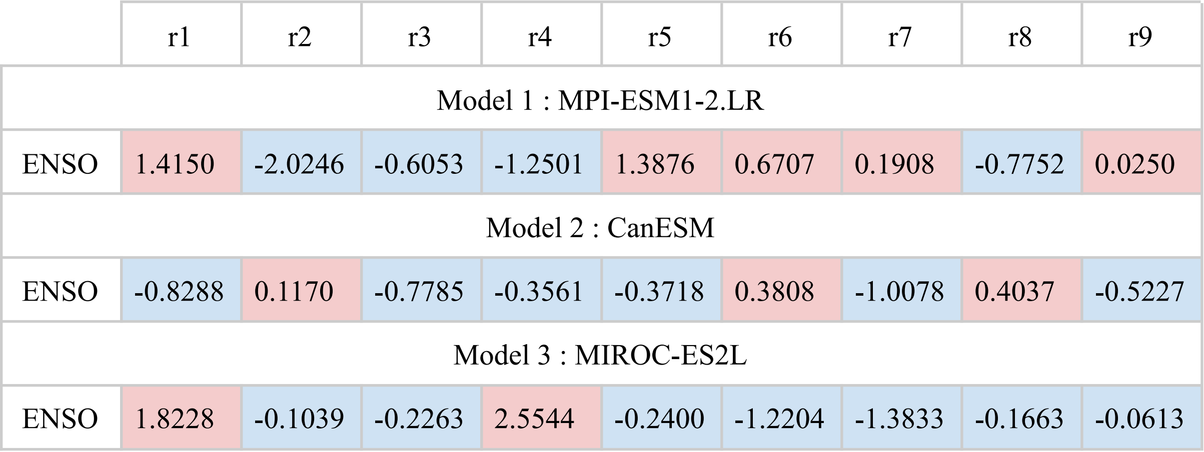
The Max-Planck-Institute Earth System Model is a coupled Earth system model (ocean, atmosphere and land modules) with an atmospheric model resolution of 1.9° x 1.9° and 47 vertical levels and an ocean model resolution of 1.5° with 40 vertical levels (Mauritsen et al., 2019). The Equilibrium Climate Sensitivity[4] (ECS) is estimated as 2.83 K.

CanESM5

The fifth version of the Canadian Earth System Model is an Earth system model with an atmospheric resolution of 2.8° x 2.8° and 40 vertical levels and an ocean model resolution of 1° x 1° and 45 vertical levels (Swart et al., 2019). The Equilibrium Climate Sensitivity (ECS) is estimated as 2.83 K.

MIROC-ES2L

The Model for Interdisciplinary Research on Climate, Earth System version 2 for Long-term simulations is a coupled ocean/atmosphere model with an atmospheric resolution of 2.8° x 2.8° and 40 vertical levels and an ocean model resolution of 1°(longitude) x 0.5-1°(latitude) and 62 vertical levels. The resolution for the ocean model depends on the latitude, thinner closer to the equator and larger towards the poles (Hajima et al., 2020). The Equilibrium Climate Sensitivity (ECS) is fixed as 2.7 K. The three models include prescribed ozone, meaning that the ozone value is predetermined and not calculated by the model.



| | r1 | r2 | r3 | r4 | r5 | r6 | r7 | r8 | r9 |
|---|---|---|---|---|---|---|---|---|---|
| Model 1 : MPI-ESM1-2.LR | | | | | | | | | |
| ENSO | 1.4150 | -2.0246 | -0.6053 | -1.2501 | 1.3876 | 0.6707 | 0.1908 | -0.7752 | 0.0250 |
| Model 2 : CanESM | | | | | | | | | |
| ENSO | -0.8288 | 0.1170 | -0.7785 | -0.3561 | -0.3718 | 0.3808 | -1.0078 | 0.4037 | -0.5227 |
| Model 3 : MIROC-ES2L | | | | | | | | | |
| ENSO | 1.8228 | -0.1039 | -0.2263 | 2.5544 | -0.2400 | -1.2204 | -1.3833 | -0.1663 | -0.0613 |

Figure 4: Recapitulative table of the ENSO initial states used in the volc-long-eq experiment. In blue are the negative values, and in red are the positive values.

---

[3] Models' descriptions come from the article written by Zanchettin and al, in 2022

[4] Measure used in climate science to quantify how the Earth's surface temperature responds to a doubling of the atmospheric concentration of carbon dioxide (Knutti et al., 2017).

As previously mentioned, the experimental protocol specified that the ENSO values provided as input should correspond to three cold, three neutral, and three warm phases. MPI-ESM1-2.LR and CanESM5 models adhere to this protocol by presenting at least 3 positive and three negative ENSO values. Conversely, the MIROC-ES2L model falls short, displaying only two positive values of ENSO. Consequently, these ENSO values, utilized as initial conditions, do not respect the experiment protocol (figure 4).

All the data used in the analysis are stored in NetCDF files, where NetCDF stands for Network Common Data Form. This format is widely employed for storing model outputs. The usefulness of the NetCDF format lies in its ability to handle multidimensional data and grant easy access to portions of the data, making it suitable for representing a range of variables. In our files, each file incorporates multiple variables, including time, latitude, longitude, and near-surface air temperature. The NetCDF format also facilitates the inclusion of detailed information about the data, such as units, dimensions, and the data source. This descriptive capability increases the understanding and interpretation of the stored information.

# 4. Methods

## 4.1 Download of the data

The data comes from the Coupled Model Intercomparison Project Phase 6 (CMIP6). This project aims to enhance our understanding of past, present, and future climate changes resulting from natural, unforced variability or in response to changes in radiative forcing in a multi-model framework. A significant objective of CMIP involves the dissemination of multi-model output to the public in a standardized format.

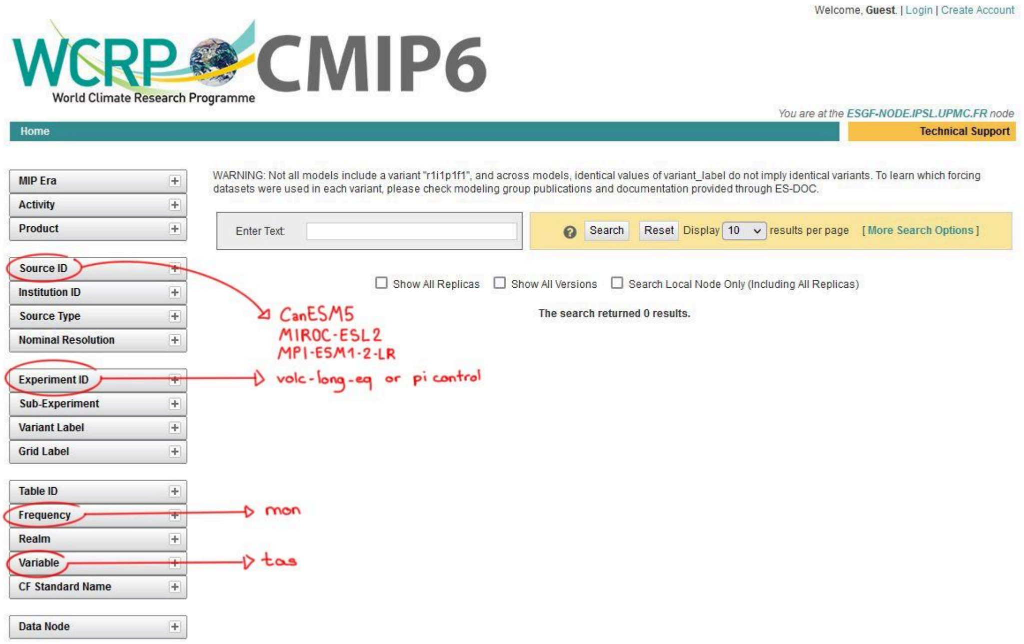
The data used in this thesis comes from the World Climate Research Programme (WCRP) (https://esgf-node.ipsl.upmc.fr/projects/cmip6-ipsl/). The WCRP provides an interface for research to find the needed data (figure 5). To access the three models under the volc-long-eq experiment, a few selections have to be made.

For the realizations with volcanic forcing:

- Variable : tas, corresponding to the near-surface air temperature
- Frequency : mon, here monthly dataset
- experiment ID : volc-long-eq
- Source ID : CanESM5, MIROC-ESL2 and MPI-ESM1-2.LR

For the piControl simulations:

- Variable : tas, corresponding to the near-surface air temperature
- Frequency : mon, here monthly dataset
- experiment ID : piControl
- Source ID : CanESM5, MIROC-ESL2 and MPI-ESM1-2.LR



WARNING: Not all models include a variant "r1i1p1f1", and across models, identical values of variant_label do not imply identical variants. To learn which forcing datasets were used in each variant, please check modeling group publications and documentation provided through ES-DOC.

The search returned 0 results.

Figure 5: Visual interface of the World Climate Research Programme.

Some of the data downloaded had to be merged together. That was the case for some piControl simulations, where some simulations were divided into different files with different range dates. To merge them into one seamless dataset, I used the Climate Data Operator (CDO[5]). Because the thesis was performed on Windows, the cygwin interface had to be installed in order to run CDO.

## 4.2 Anomaly Calculation

For each model, there is a corresponding piControl simulation, and for each realization with each model there is a temporal section of the piControl simulation representing the climate evolution without the Mount Tambora eruption of 1815. The piControl simulation serves as a baseline, and we utilize it to compute near-surface temperature anomalies for each realization during the summer (June/July/August) and winter (January/February/March) of the year 1816. The anomaly of the near-surface temperature is calculated for each simulation of each model by subtracting the piControl to the realization.

$$Anomaly_{i,j} = simulation_{i,j} - piControl_{i}$$

---

[5] From the Max-Planck Institut

With $i = 1, 2 \; or \; 3$ corresponding to model and $j = 1, 2, 3, 4, 5, 6, 7, 8 \; or \; 9$ corresponding to the realization of the model.

I made a function called "anomaly_function.m" (Annexe 18) that creates a 3D matrix. In this matrix, the ‘x’ and ‘y’ represent latitude and longitude, and the ‘z’ component includes all 27 realizations. Since not all models have the same longitude-latitude grid, I performed an interpolation of all model output to the same grid : MPI-ESM 1-2.LR; so that all the anomalies have the same grid. I repeated this process for both winter and summer 1816.

## 4.3 Empirical Orthogonal Functions

Empirical Orthogonal Function (EOF) analysis or Principal Component Analysis (PCA), is a mathematical tool used to decompose signals. This tool is particularly used within climate data to isolate spatial patterns that contribute to the data variability. EOF analysis is used to find components of data that capture most of the total variability through a linear combination of the original variables (Hannachi, 2004). EOF analysis is generally used to analyze spatial data through time, but not only. It can also be used for spatial data for different models with the member/realizations as the 3rd dimension, allowing meteorological patterns that are then interpreted in the light of known phenomena (Hannachi et al., 2007; Wallace et al., 2002). Here this method will be applied to the gridded near-surface temperature variable, as a spatially resolved variable with multiple ensemble realizations.

The areal representativity of individual grid points in the employed global gridded dataset at a resolution of 1° latitude x 1° longitude is not homogeneous, with smaller grid areas as latitude increases towards the poles. This difference in distribution will have a significant impact on the computed EOFs. To solve this problem each data point must be weighted by the cosine of its latitude.

The following description of the functioning of empirical orthogonal functions comes from the article (Hannachi, 2004).

$$A = \begin{pmatrix} \cos(lat1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \cos(latn) \end{pmatrix}$$

$A$ is a diagonal matrix , with on its diagonal the cosine of the different latitudes. This matrix $A$, has to be multiplied to the anomaly matrix (called $X$) to calculate the weighted anomaly :

$$X_{weighted} = X \times A$$

After the weighted anomaly is calculated (henceforth the weighted anomaly will be called $X$), the covariance matrix has to be calculated. This matrix called $R$ contains the covariance between each pair of elements of the matrix :

$$R = X^T \times X$$

The EOFs analysis aims to find the linear combination of all the grid points that explains the maximum variance. To find the linear combination the following matrix has to be solved :

$$R \times C = C \times D$$

$D$ is a diagonal matrix containing the eigenvalues $\lambda_i$ of $R$. The column vectors ($C_i$) of $C$ are the eigenvectors of $R$ corresponding to the eigenvalues. For each eigenvalue, we find the corresponding eigenvector. The eigenvectors are the EOFs, and each of them can be looked at as a map. The eigenvectors are ordered regarding the eigenvalues, the first eigenvector is associated with the highest eigenvalue, the second eigenvector is associated with the second highest eigenvalue, etc. The first component represents the direction in the dataset with the most substantial variability, and accounts for the largest portion of the total variance. As you move to the 2nd component, the 3rd component, and so on, each subsequent principal component captures the remaining variance in decreasing order. Meaning that the first component should capture better the data variability than the second, etc.

An important information output from the EOFs analysis is the Principal Components (PC) matrix.

$$PCs = \begin{pmatrix} PC_{11} & PC_{12} & \cdots & PC_{1n} \\ PC_{21} & PC_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ PC_{2n} & \cdots & \cdots & PC_{nn} \end{pmatrix}$$

The PCs matrix is composed of lines that represent the different components/EOFs from 1 to n (1 to 27 in our case) and columns that represent the different realizations. $PC_{11}$ is the correlation between the 1st component (row 1) and the 1st realization (column 1), $PC_{12}$ corresponds to the correlation between the 1st component and the 2nd realization, etc. The PCs matrix will also be referred as loadings in the result section.

This method will be applied to the anomaly matrix calculated previously. To apply the EOFs analysis I used a Toolbox created by Zelun Wu (https://github.com/zelunwu/eof).

## 4.4 Clustering Analysis

A clustering analysis is a machine learning tool, aiming to classify data based on their similarity (Garcia-Dias et al., 2020). The criteria of similarity depends on the type of clustering analysis implemented, in this thesis a k-mean cluster analysis is performed. The k-mean approach is one of the most used clustering analyses where 'k' represents the number of desired clusters, in our case three to match the number of models. The k-mean approach identifies 'k' points, called centers, to minimize the mean squared distance between each point and its nearest center (Kanungo et al., 2000).

A k-mean approach clustering analysis is applied to the different components of the EOFs analysis. The objective is to assess if the blind clustering identifies the realizations from the different models as distinguished. 27 clustering analyses were performed, starting by clustering just the 1st component, then the 1st and 2nd components, then the 1st, the 2nd and the 3rd component, etc, until finally having all 27 components clustered together. The more components added to the clustering the more local features and noise are considered.

After applying the clustering analysis, the number of realizations from each model per cluster is determined, giving a 3 x 3 matrix.

$$\begin{bmatrix} & model\,1 & model\,2 & model\,3 \\ Cluster\,3 & 3 & 3 & 3 \\ Cluster\,2 & 3 & 3 & 3 \\ Cluster\,1 & 3 & 3 & 3 \end{bmatrix}$$

If all models are equal and sample initial conditions well, and if the data was forced to be clustered in three clusters then the number of realizations per model in each cluster should be three, as displayed in the matrix above.

The root mean square (rms) of the clustering analysis minus the theoretical matrix above is calculated for each clustering analysis performed. The minimum rms value corresponds to the case of the results being as close as possible to the theoretical matrix, hence perfect indistinguishable models. If adding components in the clustering leads to an increase in the rms, it means that the observed clusters are gradually moving away from the theoretical distribution. If noise is considered random, it should follow the theoretical distribution.

As a last index the calculation of the Chi-squared ($\chi 2$) test will be performed. The Chi-squared test is a statistical tool used to assess the goodness of fit of observed data to theoretical data. The Chi-squared test involves a null hypothesis, $H_0$: there is no significant difference between the observed data and the theoretical data. The alternative hypothesis is, $H_A$: There is a significant difference between the observed and theoretical data. The critical value has to be found, for this the confidence interval has been settled at 0.05 and the degree of freedom is the number of categories -1, here 8, the critical value for this is 15.057. If the

Chi-squared value is above this threshold the null hypothesis can be rejected. The Chi-Squared statistics was calculated with the following formula:

$$X^2 = \sum \frac{(O - 3)^2}{3}$$

With O the Observed matrix.

# 5. Results

## 5.1 EOF analysis

### 5.1.1 EOF1 and EOF2

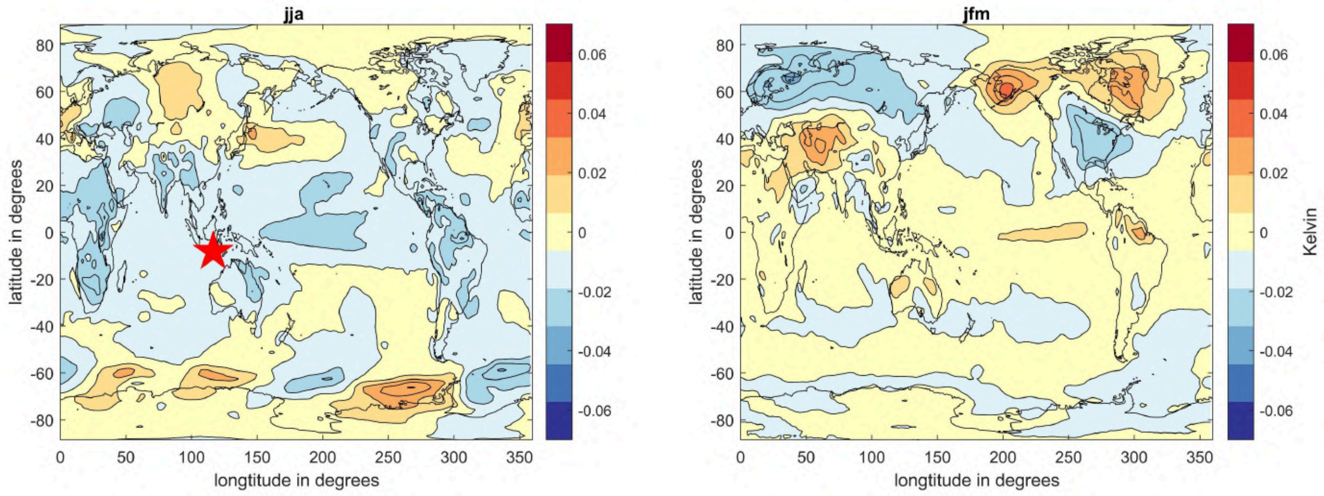


Figure 6: First EOF of the ensemble of temperature anomalies for summer (left panel) and winter (right panel). The red star represents the Mount Tambora location.

The analysis began by calculating anomaly matrices for the summer and winter of 1816 (anomalies for the three models are in the Annexes 1 to 6), followed by conducting an EOF analysis on summer and winter temperatures. The results for the first components of the EOF analyses are displayed in figure 6. The two EOF have respectively an explained variance of 12.9 % for summer and 18.2 % for winter.

Figure 6 illustrates the 1$^{st}$ component of the EOF analysis for near-surface air temperature for the whole globe for the boreal summer (left panel) and winter (right panel) anomalies of 1816. For summer, the pattern reveals a remarkable spatio-temporal heterogeneity, including a general cooling over the African continent and warming in Europe and North Asia. However, the most noticeable anomalies are found in the southern hemisphere, particularly with a circumglobal zonal wavelike pattern around Antarctica. These types of planetary-scale anomalies are a known feature of atmospheric variability during wintertime, when planetary waves are typically at their strongest during the course of the year. In this case, their plausible origin is linked with sea-ice dynamics. The extent and amplitude of the temperature anomalies associated with this wavelike signal in the southern hemisphere are large enough that they likely dominate the ensemble variance, hence likely to critically influence the outcome of the EOF analysis. Since many features of post-eruption variability concern the Northern Hemisphere (e.g., post-eruption winter warming) and the Tropics (e.g., ENSO) and as I am interested in contrasting ensemble responses in summer (when the anomalies are considered to be dominated by radiative cooling) and winter (when dynamical responses predominate), predominate), I decided to remove the southern extratropics from the following analysis, and repeated the EOF calculation for the restricted domain, northward of 30°S latitude.

The second right panel displays the 1st component of the EOF analysis for near-surface air temperature during the winter of 1816. This EOF reveals warm anomalies extending over Scandinavia and northern Asia/Siberia, while the southern part of Asia and the Mediterranean region experience cooling. This pattern, together with the north-south dipole of anomalies over eastern North America, with Canada cooling and the United States and Mexico warming corresponds to the typical imprint of a positive phase of the NAO. Noticeably, a positive phase of the NAO and the associated warming over Scandinavia are a known feature of post-eruption climate variability during the first winters after a strong tropical eruption. The sign of the loadings/PC of EOF1 (Annexe 1) indicates that 44% (11% of MPI-ESM 1-2.LR, 56 % of CanESM5 and 67% of MIROC-ES2L) of the multi-model realizations produce the post-eruption winter warming over Scandinavia (strong positive loading), whereas 56% produce a post-eruption winter cooling (strong negative loading). This result thus suggests that the mechanism behind the post-eruption winter warming, centered around the NAO response, is not always active in the considered models or may be overwhelmed by other mechanisms.

Acknowledging the potential bias introduced by the Antarctic signal and aiming to focus on patterns primarily in the northern hemisphere, the thesis will adopt a modified approach. Moving forward, the EOF analysis will be applied only to the region from the North Pole to 30° S latitude. This adjustment aims to mitigate the influence of the Antarctic signal while retaining essential modes like the El Niño Southern Oscillation in the analysis.

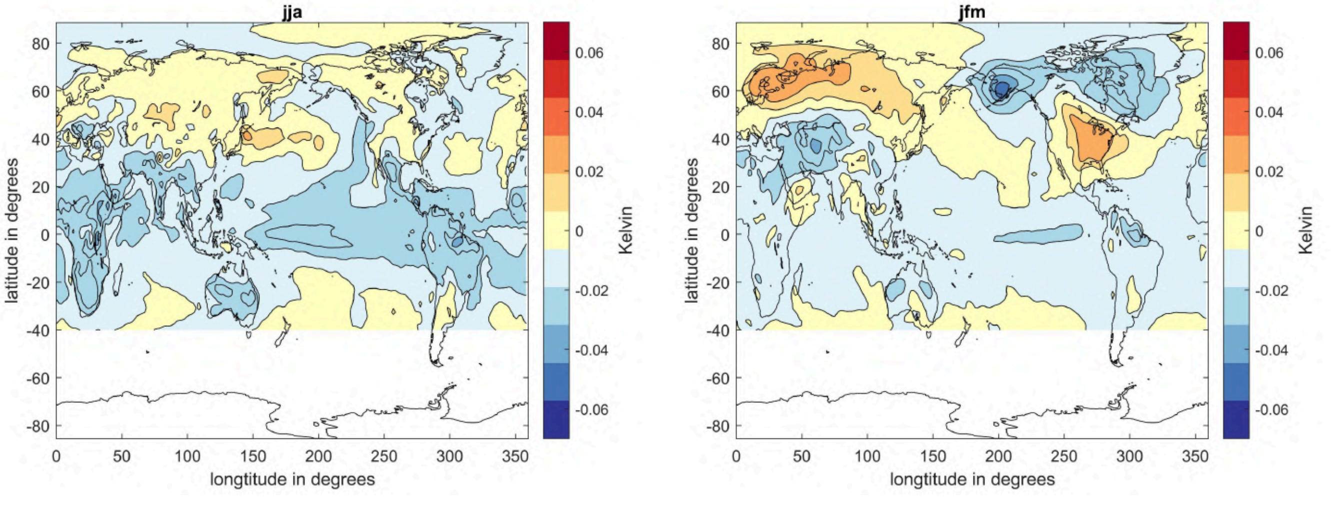


Figure 7: First EOF of the ensemble of temperature anomalies for summer (left panel) and winter (right panel).

The EOF analysis was applied to the summer and winter 1816 anomalies, focusing on the region from 90°N to 30°S latitude, excluding the Antarctic signal. The explained variance is 15.2 % for summer and 19 % for winter. The winter EOF does not change from the EOF calculated from 90°N to 90°S (figure 7), as it was expected due to the higher variance of winter than summer. For winter the first EOF is strongly associated with the NAO pattern. with 44% of realizations (89% MPI-ESM1-2.LR model, 44% of the CanESM5 model and 33% of the MIROC-ES2L) showing a post-eruption winter warming over Scandinavia (large positive loading) and 56% a post-eruption cooling (large negative loading). Analyzing the PCs matrix reveals that the MPI-ESM1-2.LR generally exhibits a positive NAO phase. The CanESM5 and MIROC-ES2L models show nuanced correlations, generally leaning towards a negative NAO phase. Therefore, whether a simulation - or by extension a model - produces or not the post-eruption winter warming is the main discriminant aspect that characterizes post-eruption winter anomalies, confirming the pivotal role of this known phenomenon of natural climate variability.

As expected, the EOF pattern of summer temperature anomalies changes substantially when the analysis is performed without the Antarctic signal. 37% of the realizations (88% of the MPIE-ESM1-2.LR model, 22% of the CanESM5 model and 22% of the MIROC-ES2L model) show a warming of the northern hemisphere, especially Europe and Asia, while the Southern Hemisphere is dominated by a general cooling. Additionally, there is a broad cooling over the tropics reflecting the expected imprint of radiative cooling (the volcanic aerosol scatters more solar radiation back to space) and the Equatorial Pacific shows strong cooling, despite the coldest anomalies over the Central Pacific. This spatial variability in the Pacific cooling may suggest that EOF1 embeds to some relevant extent the response of ENSO to the volcanic eruption. However, the location of the coldest anomalies in the Central Pacific does not correspond to the region where ENSO anomalies are typically observed in post-eruption periods, which are more in the classical cold-tongue region in the Eastern Equatorial Pacific. This supports the interpretation that EOF1 mostly captures the direct radiative response represented by the meridional asymmetry between tropical and extra tropical anomalies.

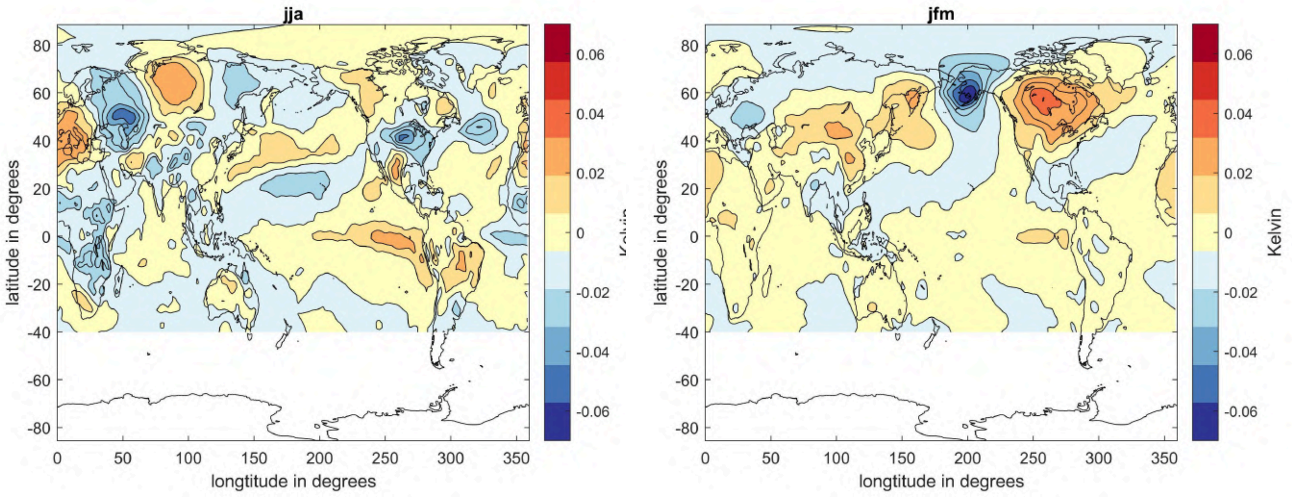


Figure 8: Second EOF of the ensemble of temperature anomalies for summer (left panel) and winter (right panel).

In the summer (figure 8, left panel), the second EOF reveals a robust signature in the Eastern Equatorial Pacific, which is a core region for El Niño/La Niña events. The explained variance for the 2nd EOF of summer is 8.3 %, the EOF shows more local features than the 1st EOF. The sign of the loadings/PC of EOF2 (Annexe 14) indicates that 52% (56% of MPI-ESM 1-2.LR, 44 % of CanESM5 and 56% of MIROC-ES2L) of the multi-model realizations produce a warming of the Central Pacific that corresponds to the habitual ENSO anomalies' location, with a warming of North Asia, a cooling of Western Russia and a European warming. While 48% of the realizations show a general cooling of the European region, that could correspond to the European summer cooling, well known as the "year without summer". The 2nd EOF of winter (right panel), with an explained variance of 11.8 %, shows a warming of North America and South/East Asia with a slight cooling of Siberia and west Russia, and a large cooling around Alaska for 48% (positive loading) of the realizations (44% of MPI-ESM 1-2.LR, 67% of CanESM5 and 33% of MIROC-ES2L).

Overall the 1st and 2nd EOFs are showing interesting patterns such as the winter warming for the 1st EOF of winter and strong ENSO signal for summer in the 2nd EOF. The NAO signal is particularly visible for the 1st winter EOF, with a divergence between the models for the phase: MPI-ESM1-2.LR shows a positive NAO phase while CanESM5 and MIROC-ES2L display a negative NAO pattern. The explained variance decreases more rapidly for summer, revealing more local features than large scale patterns.

### 5.1.2 EOF3 and EOF4

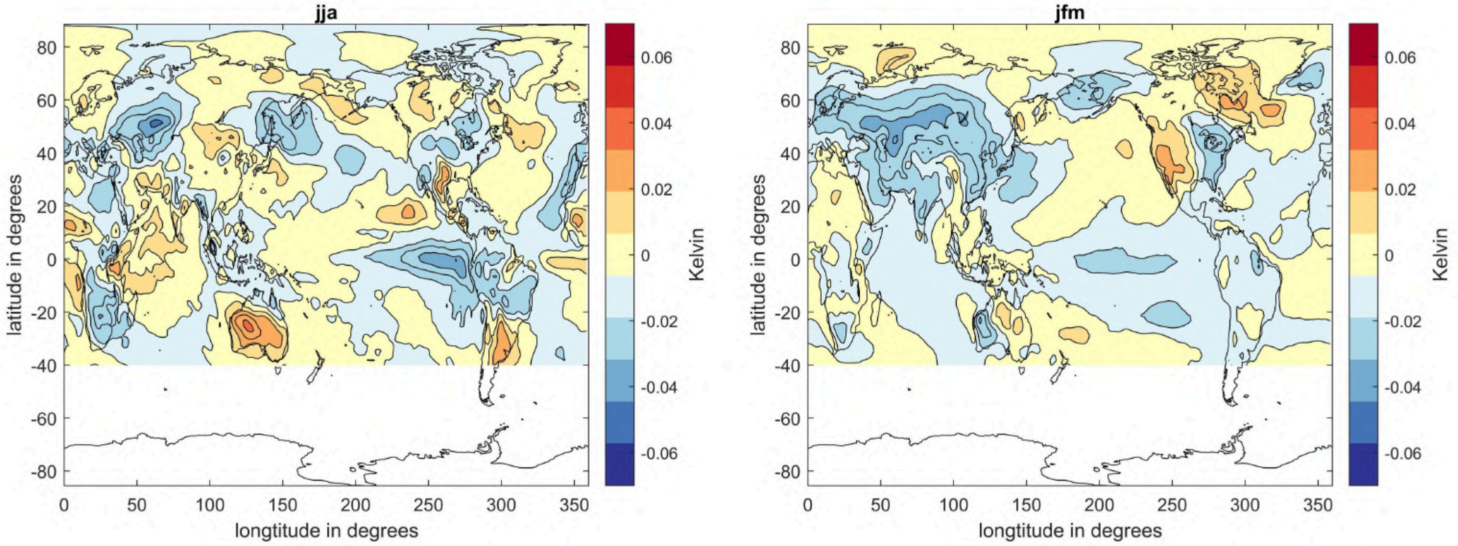


Figure 9: Third EOF of the ensemble of temperature anomalies for summer (left panel) and winter (right panel).

The 3rd EOF of summer in figure 9 (left panel), with an explained variance of 7.2%, shows as the 2nd EOF more local features than the 1st EOF. Less large-scale climate patterns are visible, except a clear cooling of the central east pacific, a critical region for ENSO, for 56% of the realizations. The CanESM5 model clearly shows a preferred La Niña; MPI-ESM1-2.LR preferentially exhibits an El Niño (67%); MIROC-ES2L has equal occurrence of a La Niña or an El Niño (56% and 44%, respectively).

For winter (right panel) the explained variance is 9.5%. The 3rd EOF still displays an apparent large-scale pattern, with a general cooling of the Asian and European continents, while the west American coast manifests a warming extending to the northernmost part of North America, and the east American coast a cooling for 56% of the realizations. In particular, this pattern is seen in 44% of the MPI-ESM 1-2.LR, 44% of the CanESM5 and 89% of the MIROC-ES2L realizations (meaning that 8 out of 9 realizations have positive loadings to the 3rd component).

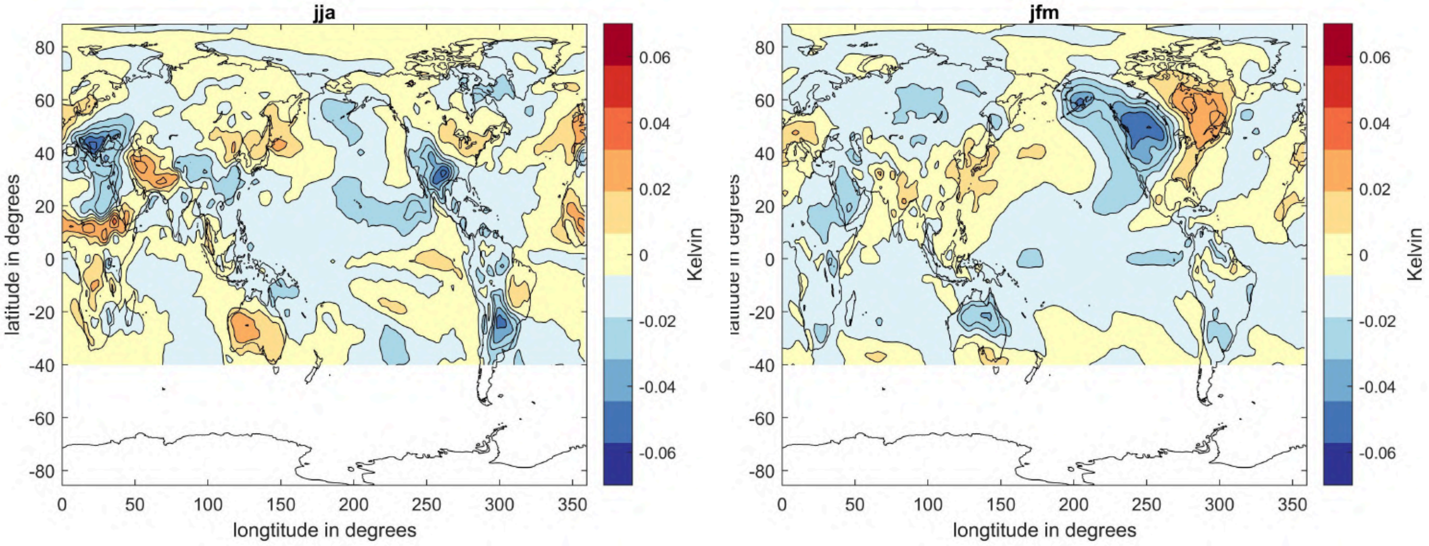


Figure 10: Fourth EOF of the ensemble of temperature anomalies for summer (left panel) and winter (right panel).

The left panel of figure 10, displaying the 4th EOF for the summer with an explained variance of 6.8%, demonstrates various regional areas exhibit cooling and warming but no more large-scale recognisable climate patterns. On another hand, the winter 4th EOF (right panel, explained variance of 8.0%) produces a strong cooling of the western part of North America while the eastern part demonstrates a warming, with a slight cooling of north Asia and a slight warming of the Europe area for 44% of the realizations. 78% of the CanESM5 simulations display negative loadings to the 4th winter EOF, associated with a warming of the North American west coast and a cooling of the North American east coast, with a slight warming of north Asia and a slight cooling of the Europe area.

Overall the 3rd and 4th EOF for summer show more and more local features as the explained variance decreases. The explained variance of winter decreases slower with the number of EOF than summer, and even if the local features are apparent, strong large-scale anomalies remain still visible. The rest of the EOFs are displayed in the Annexes 7, 8 and 9 for summer and the Annexes 10, 11 and 12 for winter. Graphics of the explained variances are also displayed in the Annexe 13.

## 5.2 Cluster analysis

### 5.2.1 Summer

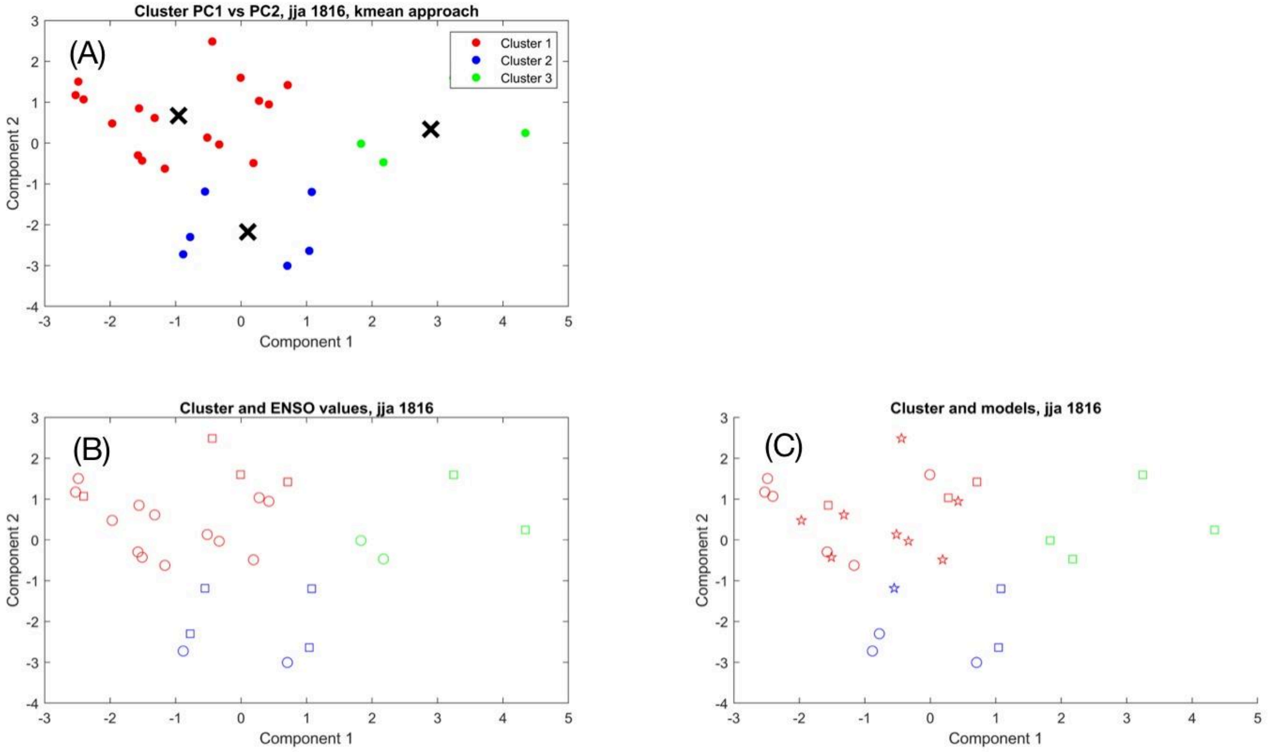


Figure 11: Cluster Analysis summer 1816 (A), linked with the initial conditions of ENSO (B), and the models belonging (C). Red, blue, and green colors correspond to the clusters. For the (B) panel, the round corresponds to negative ENSO values, and the square corresponds to positive ENSO values. For the (C) panel, the square represents the MPI-ESM1-2.LR model, the round represents the CanESM5 model, and the stars represent the MIROC-ES2L model.

Figure 11 illustrates the results of the clustering analysis conducted on the 1st and 2nd components for summer. For summer, the clusters do not have an equal number of realizations (A). Specifically, the first cluster (red) comprises 17 realizations, the second cluster (blue) has six realizations, and the last cluster (green) includes four realizations. If the models were indistinguishable and initial conditions were sampled uniformly across models, as per protocol, the clustering would have led to three clusters of equal size, each with nine realizations and ideally with three realizations from each model, following the theoretical distribution.

The (B) panel of Figure 11 indicates that the initial state of ENSO does not affect the clustering: each cluster contains roughly equal proportions of warm and cold states of ENSO. In contrast, the (C) panel of Figure 10 reveals intriguing patterns. A distinct green cluster exclusively comprises anomalies from the MPI-ESM1-2.LR model, while eight out of nine anomalies from the MIROC-ES2L model are concentrated in the red cluster. The CanESM5 model is represented in both the red and blue clusters. This result indicates a link between the models and the structures of the clusters. The convergence of each model into different clusters suggests that the models are not equal.

### 5.2.2 Winter

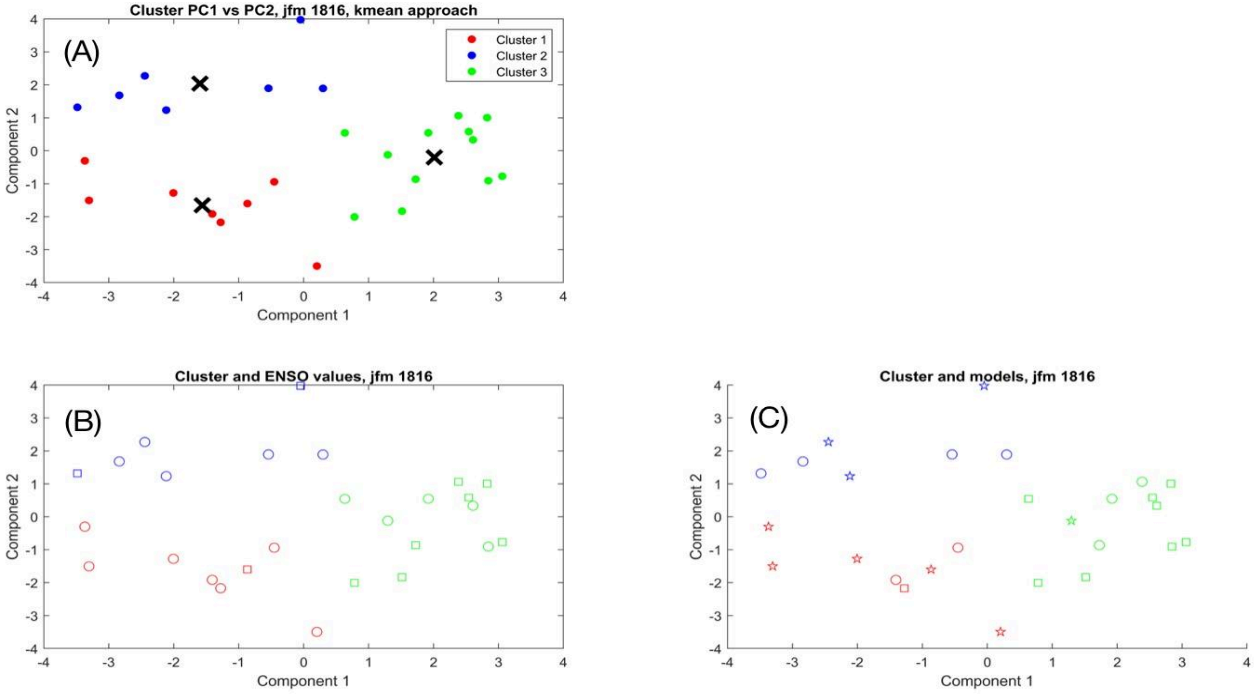


Figure 12: Cluster Analysis winter 1816 (A), linked with the initial conditions of ENSO (B), and the models belonging (C). Red, blue, and green colors correspond to the clusters. For the (B) panel, the round corresponds to negative ENSO values, and the square corresponds to positive ENSO values. For the (C) panel, the square represents the MPI-ESM1-2.LR model, the round represents the CanESM5 model, and the stars represent the MIROC-ES2L model.

Figure 12 illustrates the results of the clustering analysis conducted on the 1st and 2nd components for winter (A). Notably on panel (B), the green cluster predominantly consists of positive ENSO values, with seven out of twelve realizations displaying a positive ENSO signal. Conversely, the red cluster is primarily composed of negative ENSO values, with seven out of eight realizations exhibiting a negative ENSO signal. Similarly, the blue cluster is characterized by five out of seven realizations representing negative ENSO values. These observations suggest a clear association between the clusters and the ENSO sign, with each cluster displaying a tendency toward either positive or negative ENSO values. In the panel (C), the distribution of models within different clusters is presented. The MPI-ESM1-2.LR model is predominantly associated with the green cluster, with eight out of twelve realizations concentrated in this cluster. On the other hand, the CanESM5 model is represented across all three clusters, and similarly, the MIROC-ES2L model is distributed across the clusters, although a majority of its realizations are located in the red cluster. These patterns reveal preferential associations between each model and specific clusters, with the MPI-ESM1-2.LR model exhibiting a strong presence in the green cluster, while the CanESM5 and MIROC-ES2L models show more dispersed distributions across multiple clusters. This indicates that the model MPI-ESM1-2.LR is significantly different from the two other models.

Overall both summer and winter models are showing some preferential associations, especially MPI-ESM 1-2.LR. For the initial conditions, summer clusters are not affected by initial conditions, while for winter clusters, preferential associations are observed in connection with the initial states of ENSO. Winter clusters are more sensitive to the initial ENSO states than summer clusters.

### 5.2.3 Root mean square and Chi-square test

The subsequent phase of the analysis focuses on the outcomes derived from the calculated clustering with addition of a new EOF/component everytime. Then, the root mean square (rms) is calculated for each clustering analysis, resulting in 27 rms values for the summer and 27 rms values for the winter. The graphical representation of these results is presented in figure 13.

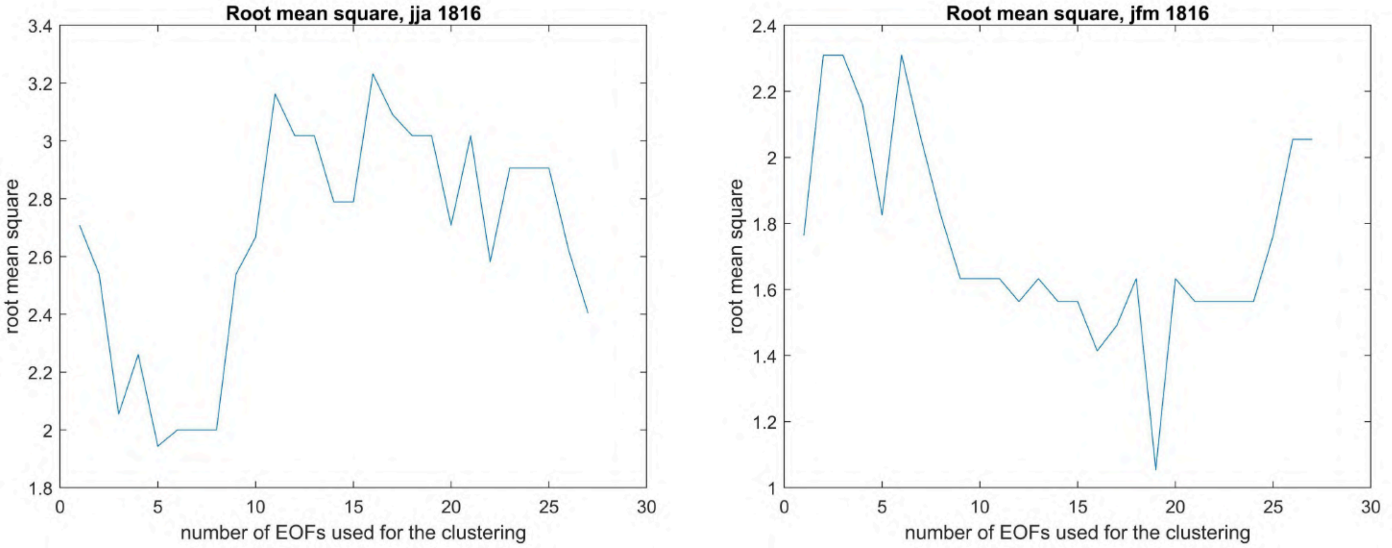


Figure 13: Root mean square calculation of the clusters' structures for summer (left panel) and winter (right panel).

The left panel in figure 13 displays the variation of the rms value with the number of components employed in the clustering analysis for the summer. Across the series of clusterings, starting with the inclusion of only the first component and progressing to the eighth clustering involving eight components, there is a discernible decline in rms values. This trend indicates that the calculated clusters are converging towards the theoretical distribution describing model exchangeability, that is where simulations from each of the three models are equally distributed across clusters (3 simulations per cluster). The minimum of the rms values occurs at the clustering involving five components. This implies that when the fifth component is incorporated, the observed values align most closely with the theoretical distribution. This means that if we consider the overall variability expressed by the first 5-7 components, the models are mostly indistinguishable. However, after the eighth component is included in the clustering, there is a noticeable increase in rms values, reflective that the introduction of additional components (typically more noisy) in the clustering process is impacting the adherence to the theoretical distribution. This change could be attributed to a change in the importance of the different mechanisms of the different models or highlight regional specificities that determine the clustering. The elevated rms values observed in the initial clustering (with one component) could be attributed to differences of the models.

The right panel in figure 13 illustrates analog results for winter. Elevated rms values are found from the 1st clustering (1 component) to the 7th clustering (7 components). These high values suggest a departure from the theoretical distribution, that is an imbalance in the representation of models or initial conditions within the clusters. These differences between the realizations could reflect different dynamics (deterministic components) or some random (noise). Because they are the first components, the noise should not have an impact. For instance, winter warming is observed over Eurasia in the first EOF, while the third EOF also exhibits a similar signature. This component characterizes a distinct 'flavor' of winter warming, indicating slight differences across realizations generated by the same underlying mechanism (deterministic). This implies that, even though there is a shared factor or mechanism causing winter warming, variations or nuances in how this warming is manifested exist across different realizations. Following the 7th clustering and extending to the 24th clustering, there is a consistent decline in the rms values, reaching a minimum at the 19th clustering. This decline implies that, during this phase, the observed values closely align with the theoretical distribution, suggesting a more balanced representation of models and initial conditions in the clusters. However, starting from the 25th clustering, there is a general increase in the rms values, which indicates the departure from the theoretical distribution, which could indicate changes in the model representation .

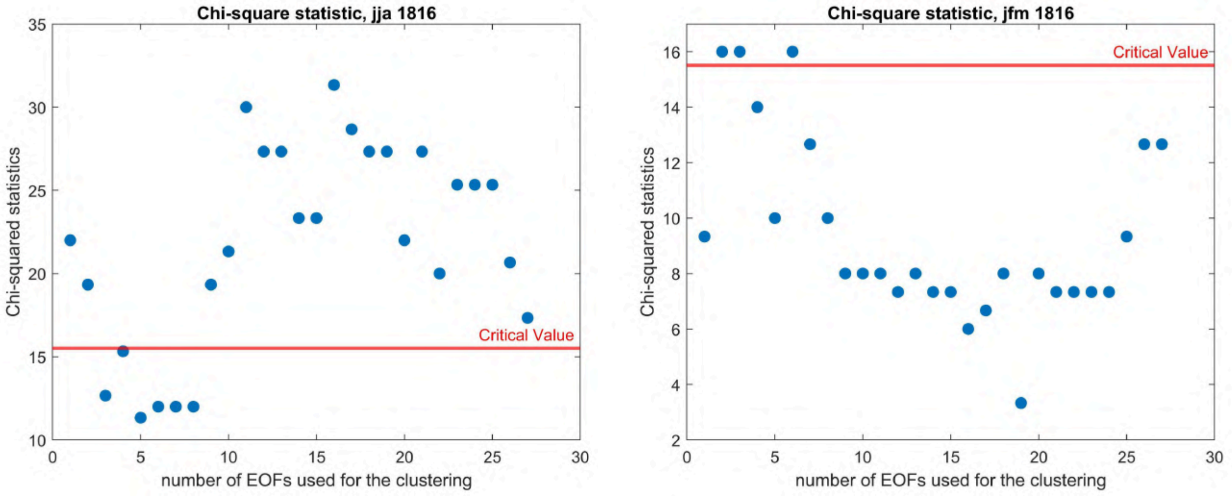


Figure 14: Representation of the Chi-squared statistics with a degree of freedom of 8 (number of categories -1) in function of the number of components in the clustering analysis for summer (left panel) and winter (right panel).

The visualization (figure 14) aims to display the Chi-squared test results evolving as the number of components increases. In the case of summer (left panel), a majority of the values surpassed the critical threshold, leading to the rejection of the null hypothesis[6]. This rejection indicates a significant disparity between the observed data and the theoretical distribution, hence the models and/or the initial condition dependencies are not indistinguishable across the ensemble. Notably, clusterings involving three to eight components exhibit Chi-squared statistics below the critical value, suggesting insufficient evidence to reject the null hypothesis. This implies that the observed data aligns reasonably well with the theoretical distribution, that is, for that portion of variability we can consider the ensemble as made of indistinguishable realizations in terms of model and initial condition dependencies. The deviation from the theoretical distribution can have many causes, the first one could be a robust mechanism across the models represented by a convergence of the results into one of the clusters. The models could also be different enough that the clustering algorithm separates each model (9 realizations) into one particular cluster. Another reason could be the sampling of the initial conditions and initial condition dependencies of the response, if the initial conditions are not well sampled each cluster will not contain 3 realizations from each model. Noise should not be able to deviate results, because pure noise should be clustered as the theoretical distribution. The first two clustering analyses are not well aligned with the theoretical distribution according to their $\chi 2$ values. By looking at the structure of the two first clustering analyses (Annexe 16) for the summer, we can observe a convergence of the results into one cluster : 15 out of 27 realizations for the first clustering analysis and 17 out of 27 for the second. For the first clustering analysis CanESM5 and MIROC-ES2L realizations are converging in the first cluster (7 out of 9 realizations each) and MPI-ESM1-2.LR in the 2nd cluster (6 out of 9 realizations). For the second clustering CanESM5 and MIROC-ES2L are still mainly represented in the first cluster while MPI-ESM1-2.LR is dispersed between the clusters. These first results indicate a similar response between CanESM5 and MIROC-ES2L, while the MPI-ESM1-2.LR model seems to diverge from them. From the 3rd to the 8th clustering analysis, the $\chi 2$ statistics are under the threshold for significance: indicating indistinguishable models. After the 8th clustering analysis, when the $\chi 2$ statistics are above the critical value, a convergence of most the CanESM5 realizations into the 1st cluster is observable until the 15th clustering analysis, while MPI-ESM1-2.LR realizations are more evenly distributed over the three clusters and MIROC-ES2L is almost equally distributed between the two first clusters. After the 15th clustering analysis a change is visible in the structures with the convergence of the realizations of the MIROC-ES2L model into the 2nd cluster, while the MPI-ESM1-2.LR is mainly represented in the 3rd cluster and CanESM5 realizations in the 1st cluster. This convergence of each model in a different cluster highlights that the models are significantly different.

---

[6] There is no significant difference between the observed data and the theoretical data. Models are equivalent and initial conditions are consistently sampled.

The right panel (winter) illustrates that a majority of the Chi-squared statistics fall below the critical value. However, three clusterings, those with 2, 3, and 6 components, surpass the critical values. The elevation above critical values results in the rejection of the null hypothesis, which is a significant difference between the observed data and the theoretical distribution. By looking at these three clusters' structures (Annexe 17) we can observe that the clusters with three and four components demonstrate a convergence of the MPI-ESM1-2.LR model into the 3rd cluster (eight out of nine realizations), with a predominance of the MIROC-ES2L in the 2nd cluster. The clustering analysis with six components shows a convergence of the MIROC-ES2L into the 1st cluster (eight out nine realizations).

There are notable differences in how the two seasons are related to different models. The summer clustering does not align well with the theoretical distribution, suggesting that the initial conditions are either not sampled properly or that they do not affect the response (domination of the deterministic component) or the models are not exchangeable. In contrast, winter clustering appears to be more in line with the theoretical distribution, indicating more comparable models and an importance of the initial conditions.

## 6. Discussion

The results foster discussion on several theoretical and practical aspects of our investigation, spanning from the overall realism of the simulated responses to strong volcanic eruptions, hence the capability of current climate models to realistically reproduce the climatic consequences of historical events, to the consistency across models and the sources of model diversity.

The results demonstrate that EOFs (that is, a mathematical deconstruction of the signal) can separate known climate patterns, hence they allow for a physical interpretation of the results. For instance, summer is known to be sensitive to direct radiative forcing, which is captured by the first EOF; Dynamical changes are apparently captured by the 2nd EOF, most noticeably the early development of an El Nino (Pausata et al., 2023). For summer, the explained variance drops fast with the order of the EOF, showing more local features after the 2nd component. In contrast, winter anomalies show a clear winter warming of the European region in the first EOF, a pattern which is strongly related to a positive NAO phase. This confirms that winter regional responses are strongly modulated by dynamical (i.e., circulation) changes rather than being mostly driven by direct radiative effects. The EOF are manifesting different loadings depending on the models, the 1st EOFs for both summer and winter have positive loadings for MPI-ESM1-2.LR while CanESM5 and MIROC-ES2L are showing negative loadings of the EOFs. The fact that MPI-ESM1-2.LR yields more robustly a post-eruption winter warming does not necessarily imply that it simulates more realistically the mechanisms of post-eruption climate variability compared to the other models. First of all, there is not just one mechanism. The core mechanism proposed in the past was in brief: tropical volcanic eruptions yields enhancement of the stratospheric aerosol layer in the tropics, since aerosol absorb infrared and near-infrared radiation, this leads to a warming of the tropical stratosphere, enhancement of the stratospheric polar vortex by thermal wind balance and then this propagates into the troposphere and is seen as positive NAO, with the circulation anomalies causing the typical meridional asymmetry over Eurasia. But, now we know that the polar vortex enhancement may be due to changes in planetary wave activity by surface cooling (so not just thermal wind response), and the ENSO response therefore matters as well. Then, we have just a few observations and not so many reconstructions, so maybe we don’t have the full spectrum of post-eruption variability, that instead the many simulations in VolMIP can reveal.

This thesis provided a simple metric to assess whether, under the assumption of a well implemented protocol of the volc-long-eq experiment, the models are indistinguishable among each other. The metric is based on the rms calculation between the observed clustering analysis and the theory (models are equal and initial conditions well sampled). The smaller rms values the closer the observed clusters are to the theory. These results are completed with a Chi-square test stating a null hypothesis: there is no significant difference between the observed data and the theoretical data, meaning that the models are indistinguishable. A rejection of this null hypothesis will imply a deviation from the theory. This deviation from the theoretical distribution can have many causes, the first one could be a robust mechanism

across the models making them converge  into one unique cluster. The models could also be different leading to a convergence of each model (9 realizations) to one particular cluster. Another reason could be the sampling of the initial conditions and initial condition dependencies of the response, if the initial conditions are not well sampled each cluster will not contain 3 realizations from each model. Noise should not be able to deviate results, because pure noise should be cluster has the theoretical distribution.
For the summer, the first two rms values are higher than the rest and the Chi-square test led to the rejection of the null hypothesis, by looking at the clusters' structures a robust mechanism is demonstrate across CanESM5 and MIROC-ESL , while MPI-ESM1-2.LR differs from them. The following drop in rms values and the acceptance of the alternative hypothesis indicates indistinguishable models. An increase in the rms values is then visible with a chi-square test rejecting the null hypothesis. The clusters’ structures indicate that models differ  between each other with first a convergence of the CanESM5 realizations in the 1st cluster, then, after the 15th clustering analysis the three models are converging into three different clusters, demonstrating that models are not indistinguishable. The winter rms calculations are generally lower than summer, indicating a general better fit of the winter clustering to the theory than summer. The winter chi-square test indicated that only 3 clustering analyses are rejecting the null hypothesis (3rd, 4th and 6th clustering analysis). By looking at the clusters’ structures, the convergence of the MPI-ESM1-2.LR model into the 3rd cluster is visible for the 3rd and 4th clustering analysis, diverging from the other 2 models. The 6th clustering analysis displays the convergence of the MIROC-ES2L realizations into the 1st cluster.  Summer and winter are both showing first clustering analysis indicating disparities in the models, models are not indistinguishable. Then winter shows indistinguishable models while summers demonstrate models differing from each other.

The theory is based on the well sampling of the initial conditions. Divergence from the theory could be due to the bias from the MIROC-ES2L initial conditions sampling that does not respect the volc-long-eq experiment protocol. By protocol, the initial conditions should encompass, for each model, three negative, three neutral and three positive ENSO phases; MIROC-ES2L, however, only displays 2 positive ENSO phases. Another limitation of this study is that we are working with autocorrelated data without treatment of the dependencies. EOF analysis aims to identify spatial patterns in a dataset. Ignoring spatial dependencies can lead to misleading spatial structures in the identified EOF modes, as the analysis may capture autocorrelation patterns rather than the true variability. Spatial autocorrelation may complicate the separation of signals and noise in EOF analysis. Without considering dependencies, the identified patterns may include both meteorological signals and autocorrelation-induced noise.

This study is the first one done on the volc-long-eq experiment, opening a large area for future research. This work could be expanded by adding other variables in the EOF analysis, such as the precipitation levels or sea-level pressure. This would allow insight into whether disparities in the models and realizations extend to variables other than the near-surface temperature. This could also allow us to see other signals less visible or not visible with the temperature. Here, the initial conditions used were the ENSO values, adding other modes such as NAO could complete the work and see if the climate responses are more sensitive to certain initial conditions rather than others. This would require more realizations than the 9 per model used here for a satisfactory sampling (see, for instance, the 27 realizations used for the volcano-pinatubo-full experiment that consider both, ENSO and NAO states for the initial conditions).

The VolMIP project proposes different experiments such as the volc-pinatubo-full, used to represent the 1991 Pinatubo eruption in the Philippines. Comparing the results obtained from different eruptions enables a comparative analysis. This can help identify patterns, variations, or unique characteristics in simulated post-eruption volcanic climates, contributing to a better understanding of the underlying processes. Applying this method to a different eruption would allow us to assess its generalizability. If the method proves effective across various volcanic events, it suggests that it may have broader applicability and could be considered a robust technique. Using other volcanic eruptions could also address the initial sampling problem, by applying the method to realizations that satisfy the experiment protocol. All these future possible applications could help to access the sources of models' diversities.

This study showed us that disparities across the models are a huge source of uncertainties in the modeling of climate response to large volcanic eruptions and that perceiving the influence of initial conditions on climate predictions poses significant challenges. If we had to predict today, a future-like-Tambora eruption, the uncertainties in the climate responses would be difficult to assess, unless the initial conditions are sufficiently accounted for. First, due to the disparities of the models, but also due to the uncertainty of future climate state in a climate change context (IPPC, 2023). The variations among different models introduce complexities in accurately forecasting the climatic impacts of such eruptions. Additionally, the evolving nature of our climate amplifies the challenge.

## 7. Conclusion

This study was driven by the desire to enhance our comprehension of the climate variability induced by strong volcanic eruptions. The aim was to gain a deeper understanding of the factors contributing to the diversity in simulated responses, both across different models and within simulations produced by the same model. To this purpose, I used an ensemble of simulations from the volc-long-eq experiment of the "VolMIP" initiative, conducted with three models (MPI-ESM1-2.LR, CanESM5 and MIROC-ES2L). The idea was, first, to combine all simulated data in a multi-model ensemble; then use Empirical Orthogonal Functions (EOFs) to deconstruct the associated patterns and remove the noisy components; finally, to determine whether the groups obtained by application of a blind clustering algorithm on the multi-model EOFs corresponds to the individual model ensembles (i.e., the models are significantly different) or, alternatively, the models are indistinguishable and diversity in the multi-model ensemble is determined by other factors, such as initial conditions. To assess this objectively, I utilized a metric based on root mean square (rms) calculations and Chi-square tests to assess model indistinguishability.

The results, focused on gridded Tropical and Northern Hemisphere near-surface air temperature, showed the diversity of the post eruption anomalies within each model and across different models, suggesting multiple response mechanisms and a relevant contribution of initial conditions. The results also showed different seasonal behaviors, as summer responses showed sensitivity to models' disparities (clusters corresponding more to the initial model ensembles) while winter was impacted by initial conditions and showed less inter-model differences (clusters corresponding less to the initial model ensembles).

This is the first study based on the output of the volc-long-eq experiment. The main limitation of this study is linked to the small size of the ensemble, as only 9 simulations per model were available. As discussed in this Thesis, more simulations would allow us to pinpoint with more confidence the range of post-eruption responses and explore how response mechanisms to volcanic eruptions operate under different initial conditions. This seems especially relevant for winter, where the first EOF describes the well known post-eruption winter warming. Why in certain simulations this phenomenon of volcanically-forced variability does not occur remains to be explored.

This study should foster more participation in the volc-long-eq or similar experiments towards better understanding model diversity. A natural follow up to this thesis, based on the available data, is to extend the analysis on a multivariate context, for instance including precipitation or pressure/circulation changes. The method for assessing multi-model diversity proposed here could also be applied to other data, especially projected climate anomalies and volcanic experiment under realistic conditions instead of the idealized conditions of volc-long-eq.

# Acknowledgments

I would like to express my appreciation to the individuals listed below, whose invaluable contributions were pivotal to the successful completion of this thesis.

I am deeply thankful to my supervisor, Davide Zanchettin, for his guidance, support, and invaluable insights throughout the entire research process. His expertise and encouragement have been crucial in shaping this work, complemented by the valuable contribution of my co-supervisor, Angelo Rubino.

Special shoutout to my dear friend Nathan, who has consistently been a reliable source of advice throughout this thesis journey. Your insightful perspectives and motivational talks have played a significant role.

I am infinitely thankful to my greatest friend Ella, for consistently being there for me, not only during this thesis work but throughout my entire life.

Finally, I would like to express my deepest gratitude and love to my parents, sister and cousin. Thomas, Sophie, Mathilde and Milly, for their unwavering support and love.

# Glossary

| | |
|---|---|
| AMOC | Atlantic meridional overturning circulation |
| AO | Arctic Oscillation |
| CDO | Climate Data Operator |
| CMIP6 | Coupled Model Intercomparison Project Phase 6 |
| EOFs | Empirical orthogonal functions |
| ENSO | El Niño Southern Oscillation |
| ECS | Equilibrium Climate Sensitivity |
| NAM | Northern Annular Mode |
| NAO | North Atlantic Oscillation |
| NetCDF | Network Common Data Form |
| PCA | Principal Component Analysis |
| PDC | Pyroclastic Density Current |
| PiControl | Simulation of the climate state in 1816 with preindustrial states, simulation run without volcanic forcing. |
| Realization | Simulation of one model for a specific initial state of ENSO. |
| Rms | Root mean square |
| $SO_2$ | Sulfur dioxide |
| volc-long-eq | Experiment of the “Model Intercomparison Project on the climatic response to volcanic forcing, where volcanic forcing from an idealized Tambora eruption are applied in terms of aerosol optical properties on various global ocean-atmosphere coupled models or earth system models under controlled initial conditions. |
| VolMIP | Model Intercomparison Project on the climatic response to Volcanic forcing |
| $\chi^2$ | Chi-square |

# References

Brönnimann, S., & Krämer, D. (2016). Tambora and the " Year Without a Summer " of 1816. A perspective on earth and human systems science (Vol. 90). Geographica Bernensia.

Cole, P. D., Neri, A., & Baxter, P. J. (2015). Hazards from pyroclastic density currents. In *The encyclopedia of volcanoes* (pp. 943-956). Academic Press.

Christiansen, B. (2008). Volcanic eruptions, large-scale modes in the Northern Hemisphere, and the El Niño–Southern Oscillation. *Journal of Climate*, *21*(5), 910-922.

Danabasoglu, G., Yeager, S. G., Kim, W. M., Behrens, E., Bentsen, M., Bi, D., ... & Yashayaev, I. (2016). North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability. *Ocean Modelling*, *97*, 65-90.

De Jong Boers, B. (1995). Mount Tambora in 1815: A volcanic eruption in Indonesia and its aftermath. *Indonesia*, 37-60.

Foden, J. (1986). The petrology of Tambora volcano, Indonesia: a model for the 1815 eruption. *Journal of Volcanology and Geothermal Research*, *27*(1-2), 1-41.

Garcia-Dias, R., Vieira, S., Pinaya, W. H. L., & Mechelli, A. (2020). Clustering analysis. In *machine learning* (pp. 227-247). Academic Press.

Hagen, M., & Azevedo, A. (2023). Influence of Volcanic Activity on Weather and Climate Changes. *Atmospheric and Climate Sciences*, *13*(2), 138-158.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., ... & Kawamiya, M. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, *13*(5), 2197-2244.

Hannachi, A., & O'Neill, A. (2001). Atmospheric multiple equilibria and non‐Gaussian behaviour in model simulations. *Quarterly Journal of the Royal Meteorological Society*, *127*(573), 939-958.

Hannachi, A. (2004). A primer for EOF analysis of climate data. *Department of Meteorology, University of Reading*, *1*(29), 3-3.

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *27*(9), 1119-1152.

IPCC, 2023: *Climate Change 2023: Synthesis Report.* Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, doi: 10.59327/IPCC/AR6-9789291691647

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C., Silverman, R., & Wu, A. Y. (2000, May). The analysis of a simple k-means clustering algorithm. In *Proceedings of the sixteenth annual symposium on Computational geometry* (pp. 100-109).

Knutti, R., Rugenstein, M. A., & Hegerl, G. C. (2017). Beyond equilibrium climate sensitivity. *Nature Geoscience*, *10*(10), 727-736.

LeGrande, A. N., Tsigaridis, K., & Bauer, S. E. (2016). Role of atmospheric chemistry in the climate impacts of stratospheric volcanic injections. *Nature Geoscience*, *9*(9), 652-655.

Luterbacher, J., & Pfister, C. (2015). The year without a summer. *Nature Geoscience*, *8*(4), 246-248.

Marshall, L., Schmidt, A., Toohey, M., Carslaw, K. S., Mann, G. W., Sigl, M., ... & Tummon, F. (2018). Multi-model comparison of the volcanic sulfate deposition from the 1815 eruption of Mt. Tambora. *Atmospheric Chemistry and Physics*, *18*(3), 2307-2328.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... & Roeckner, E. (2019). Developments in the MPI‐M Earth System Model version 1.2 (MPI‐ESM1. 2) and its response to increasing CO2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998-1038.

McCormick, M., & Veiga, R. E. (1992). SAGE II measurements of early Pinatubo aerosols. *Geophysical Research Letters*, *19*(2), 155-158.

Oppenheimer, C. (2003). Climatic, environmental and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. *Progress in physical geography*, *27*(2), 230-259.

Pausata, F. S., Zhao, Y., Zanchettin, D., Caballero, R., & Battisti, D. S. (2023). Revisiting the mechanisms of ENSO response to tropical volcanic eruptions. *Geophysical Research Letters*, *50*(3), e2022GL102183.

Pfister, C., & White, S. (2018). A year without a summer, 1816. *The Palgrave Handbook of Climate History*, 551-561.

Predybaylo, E., Stenchikov, G. L., Wittenberg, A. T., & Zeng, F. (2017). Impacts of a Pinatubo‐size volcanic eruption on ENSO. *Journal of Geophysical Research: Atmospheres*, *122*(2), 925-947.

Raible, C. C., Brönnimann, S., Auchmann, R., Brohan, P., Frölicher, T. L., Graf, H. F., ... & Wegmann, M. (2016). Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects. *Wiley Interdisciplinary Reviews: Climate Change*, *7*(4), 569-589.

Self, S., Rampino, M. R., Newton, M. S., & Wolff, J. A. (1984). Volcanological study of the great Tambora eruption of 1815. *Geology*, *12*(11), 659-663.

Self, S., Marti, J., & Ernst, G. G. (2005). Effects of volcanic eruptions on the atmosphere and climate. *Volcanoes and the Environment*, 152-174.

Shindell, D. T., Schmidt, G. A., Mann, M. E., & Faluvegi, G. (2004). Dynamic winter climate response to large tropical volcanic eruptions since 1600. *Journal of Geophysical Research: Atmospheres*, *109*(D5).

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., ... & Woodruff, T. E. (2015). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, *523*(7562), 543-549.

Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., ... & Masson-Delmotte, V. (2015). Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. *Nature Geoscience*, *8*(10), 784-788.

Stothers, R. B. (1984). The great Tambora eruption in 1815 and its aftermath. *Science*, *224*(4654), 1191-1198.

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., ... & Winter, B. (2019). The Canadian earth system model version 5 (CanESM5. 0.3). *Geoscientific Model Development*, *12*(11), 4823-4873.

Suhendro, I., Toramaru, A., Miyamoto, T., Miyabuchi, Y., & Yamamoto, T. (2021). Magma chamber stratification of the 1815 Tambora caldera-forming eruption. *Bulletin of Volcanology*, *83*, 1-20.

Sulpizio, R., Dellino, P., Doronzo, D. M., & Sarocchi, D. (2014). Pyroclastic density currents: state of the art and perspectives. *Journal of Volcanology and Geothermal Research*, *283*, 36-65.

Timmreck, C. (2012). Modeling the climatic effects of large explosive volcanic eruptions. *Wiley Interdisciplinary Reviews: Climate Change*, *3*(6), 545-564.

Toohey, M., & Sigl, M. (2017). Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE. *Earth System Science Data*, *9*(2), 809-831.

Trenberth, K. E. (1996). El Niño southern oscillation (ENSO). *Sea*.

Wallace, J. M., & Thompson, D. W. (2002). The Pacific center of action of the Northern Hemisphere annular mode: Real or artifact?. *Journal of Climate*, *15*(14), 1987-1991.

Wang, C., Deser, C., Yu, J. Y., DiNezio, P., & Clement, A. (2017). El Niño and southern oscillation (ENSO): a review. *Coral reefs of the eastern tropical Pacific: Persistence and loss in a dynamic environment*, 85-106.

Woollings, T., Franzke, C., Hodson, D. L. R., Dong, B., Barnes, E. A., Raible, C. C., & Pinto, J. G. (2015). Contrasting interannual and multidecadal NAO variability. *Climate Dynamics*, *45*, 539-556.

Wunderlich, F., & Mitchell, D. M. (2017). Revisiting the observed surface climate response to large volcanic eruptions. *Atmospheric Chemistry and Physics*, *17*(1), 485-499.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., ... & Tummon, F. (2016). The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP):

Experimental design and forcing input data for CMIP6. *Geoscientific Model Development*, *9*(8), 2701-2719.

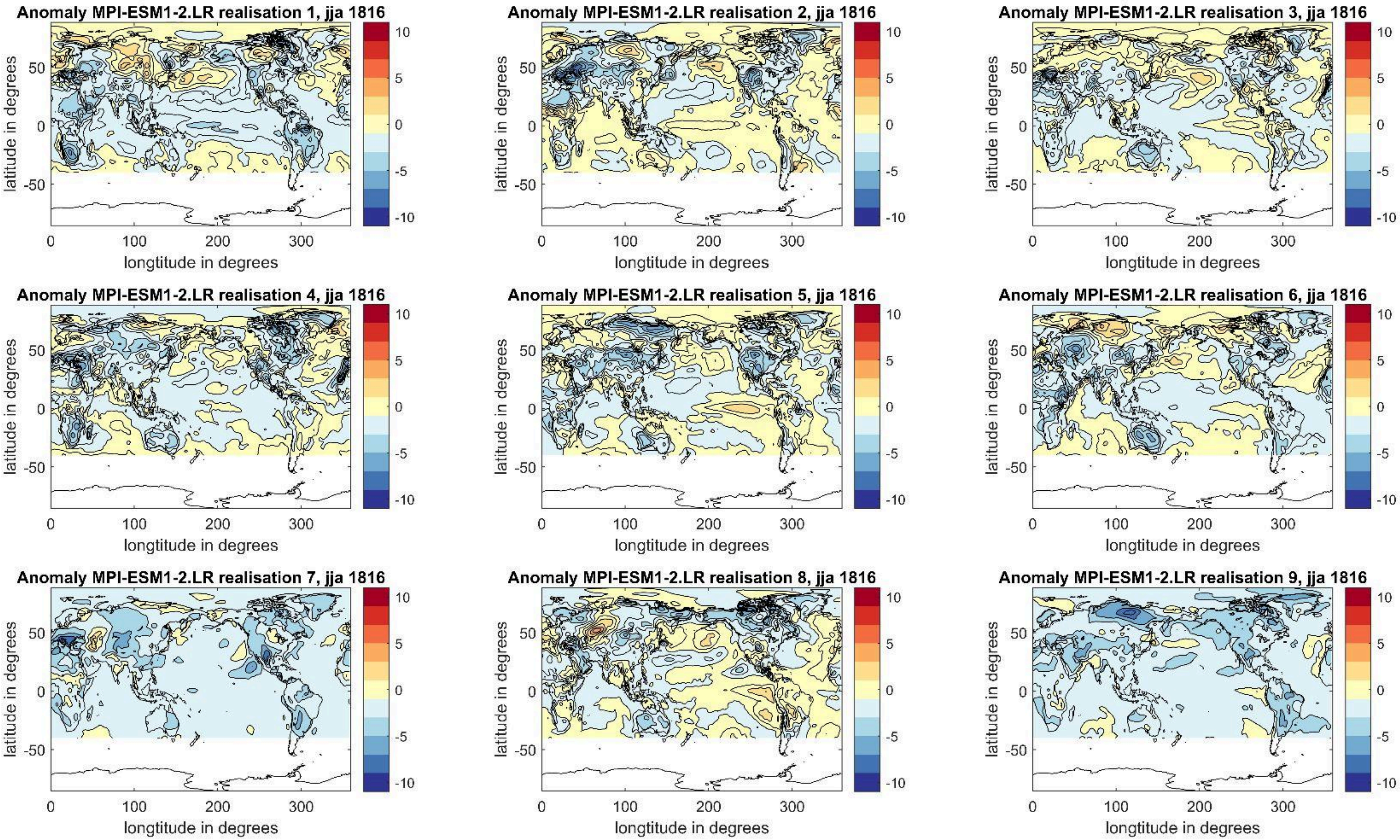
Zanchettin, D., Timmreck, C., Toohey, M., Jungclaus, J. H., Bittner, M., Lorenz, S. J., & Rubino, A. (2019). Clarifying the relative role of forcing uncertainties and initial‐condition unknowns in spreading the climate response to volcanic eruptions. *Geophysical Research Letters*, *46*(3), 1602-1611.

Zanchettin, D., Timmreck, C., Khodri, M., Schmidt, A., Toohey, M., Abe, M., ... & Weierbach, H. (2022). Effects of forcing differences and initial conditions on inter-model agreement in the VolMIP volc-pinatubo-full experiment. *Geoscientific Model Development*, *15*(5), 2265-2292.

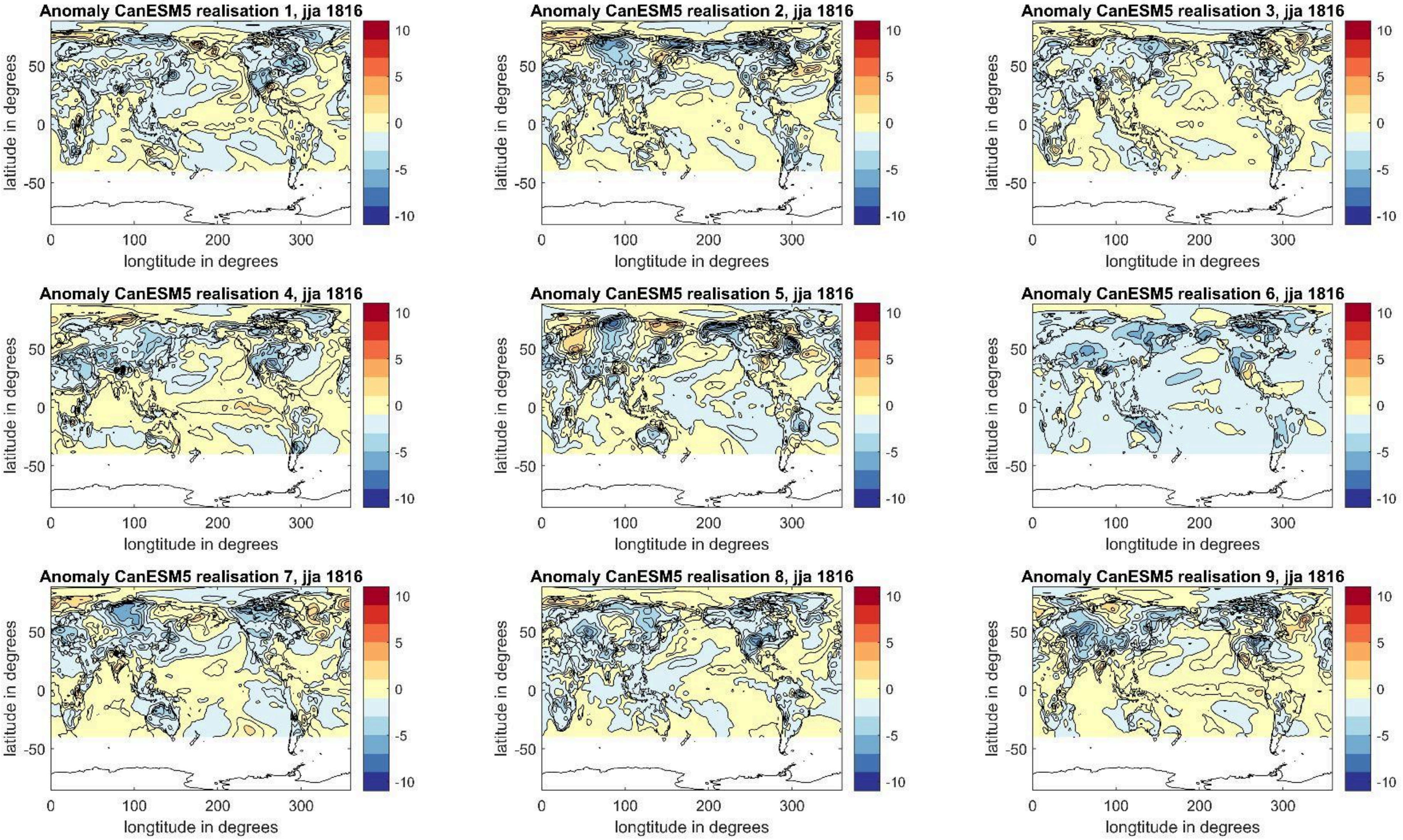
# Annexe

This section is dedicated to extra-materials of the thesis. The annexe is composed of different graphics representing the different near-surface temperature anomalies and the different EOFs. Explained variances, loadings and clustering analysis are also displayed. Finally, the matlab functions created for this project are presented.

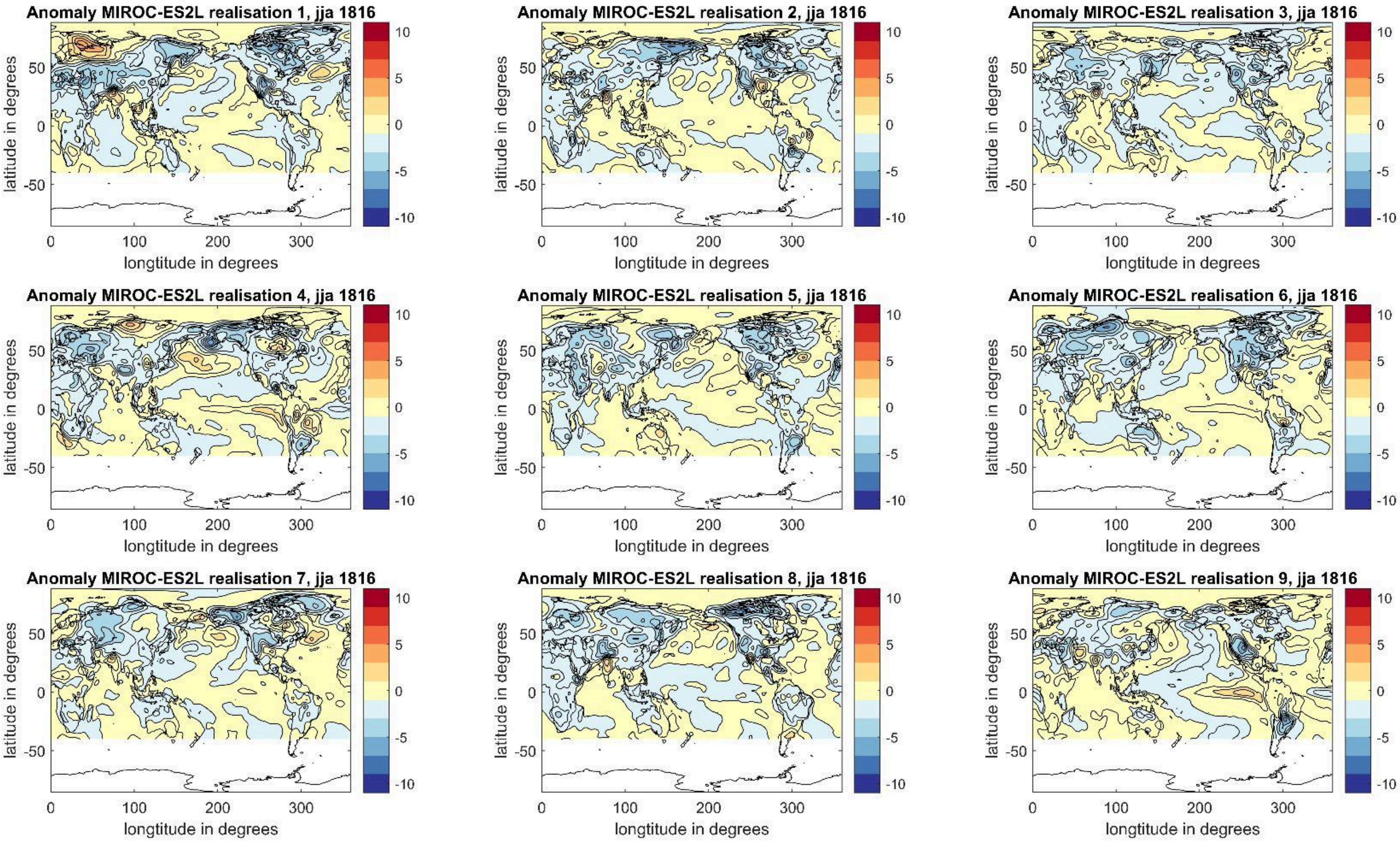
| | |
|---|---|
| Annexe 1 | Anomalies of the near-surface temperature calculated from the 9 realizations of the MPI-ESM1-2.LR for summer, with the anomalies in K |
| Annexe 2 | Anomalies of the near-surface temperature calculated from the 9 realizations of the CanESM5 for summer, with the anomalies in K |
| Annexe 3 | Anomalies of the near-surface temperature calculated from the 9 realizations of the MIROC-ES2L for summer, with the anomalies in K |
| Annexe 4 | Anomalies of the near-surface temperature calculated from the 9 realizations of the MPI-ESM1-2.LR for winter, with the anomalies in K |
| Annexe 5 | Anomalies of the near-surface temperature calculated from the 9 realizations of the CanESM5 for winter, with the anomalies in K |
| Annexe 6 | Anomalies of the near-surface temperature calculated from the 9 realizations of the MIROC-ES2L for winter, with the anomalies in K |
| Annexe 7 | Principal components 1 to 12 of the EOF analysis for summer in K |
| Annexe 8 | Principal components 13 to 24 of the EOF analysis for summer in K |
| Annexe 9 | Principal components 25 to 27 of the EOF analysis for summer in K |
| Annexe 10 | Principal components 1 to 12 of the EOF analysis for winter in K |
| Annexe 11 | Principal components 13 to 24 of the EOF analysis for winter in K |
| Annexe 12 | Principal components 25 to 27 of the EOF analysis for winter in K |
| Annexe 13 | Explained variances of the different EOFs for summer and winter. |
| Annexe 14 | Table of the loadings between the Principal Components and the realizations of each model for summer. |
| Annexe 15 | Table of the loadings between the Principal Components and the realizations of each model for winter. |
| Annexe 16 | Table of the clusters' structures for summer |
| Annexe 17 | Table of the clusters' structures for winter |
| Annexe 18 | Matlab code: Anomaly calculation |
| Annexe 19 | Matlab code: Clustering calculation and identification of the clusters's structures |



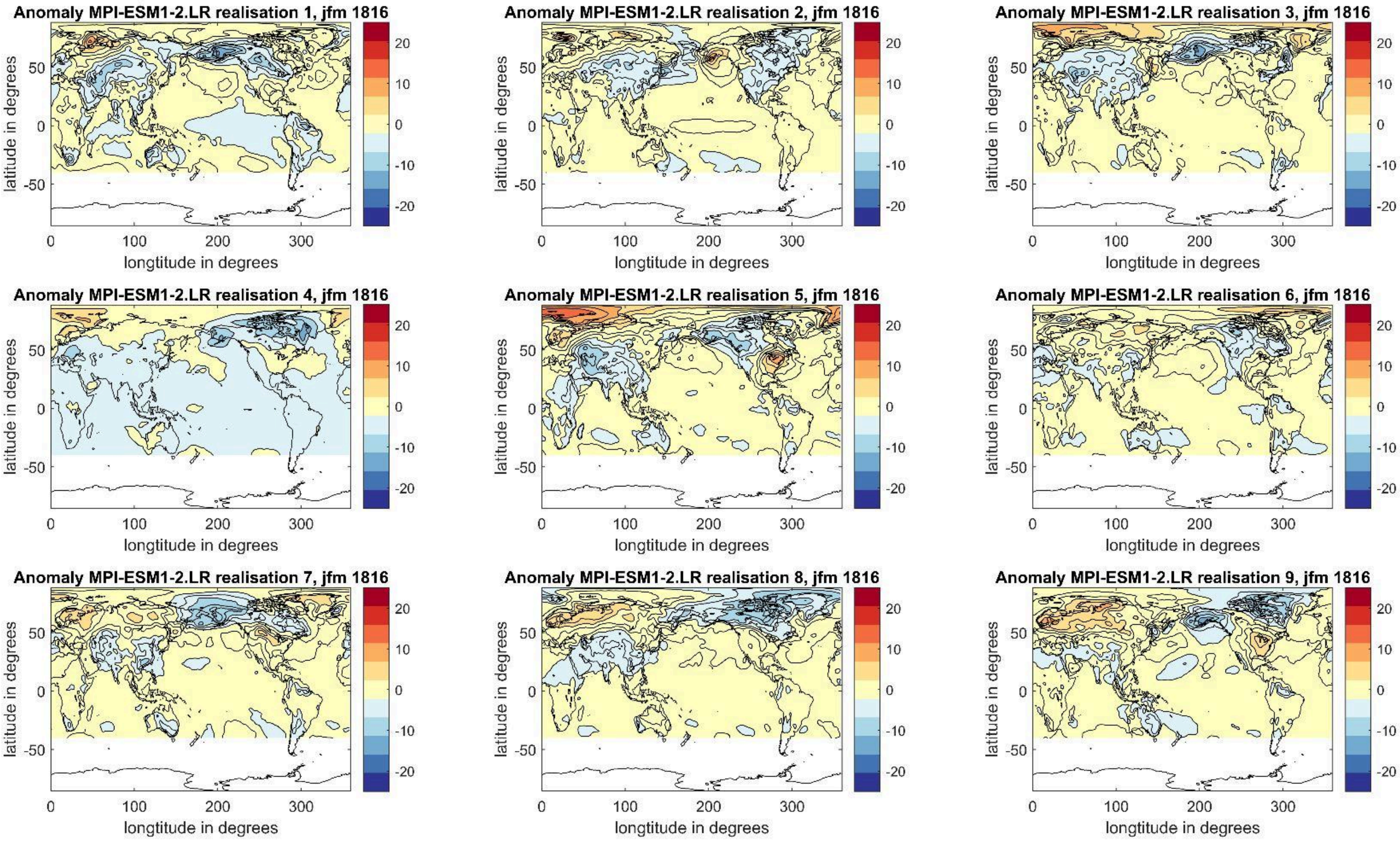
Annexe 1 : Anomalies of the near-surface temperature calculated from the 9 realizations of the MPI-ESM1-2.LR for summer, with the anomalies in K



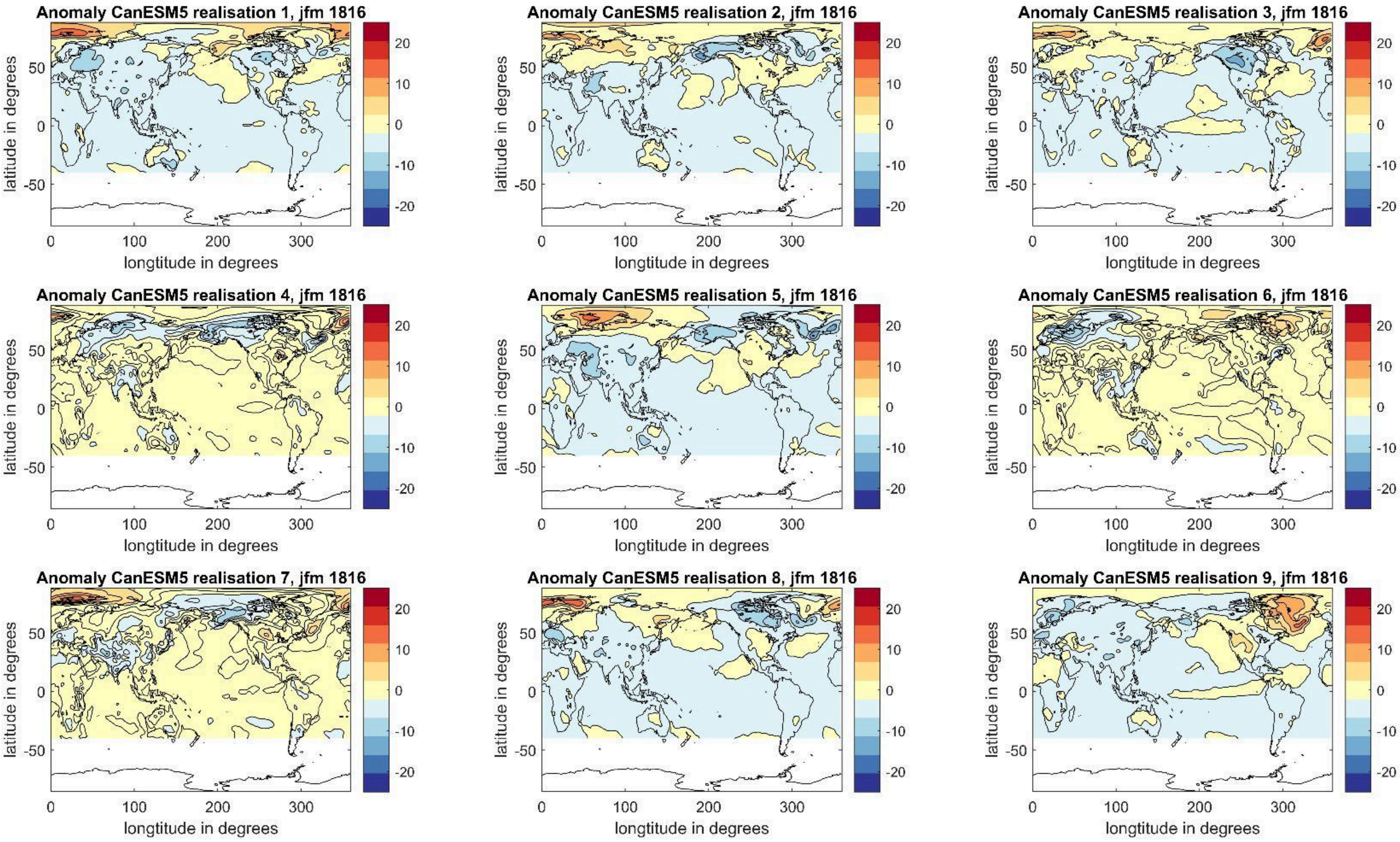
Annexe 2 : Anomalies of the near-surface temperature calculated from the 9 realizations of the CanESM5 for summer, with the anomalies in K



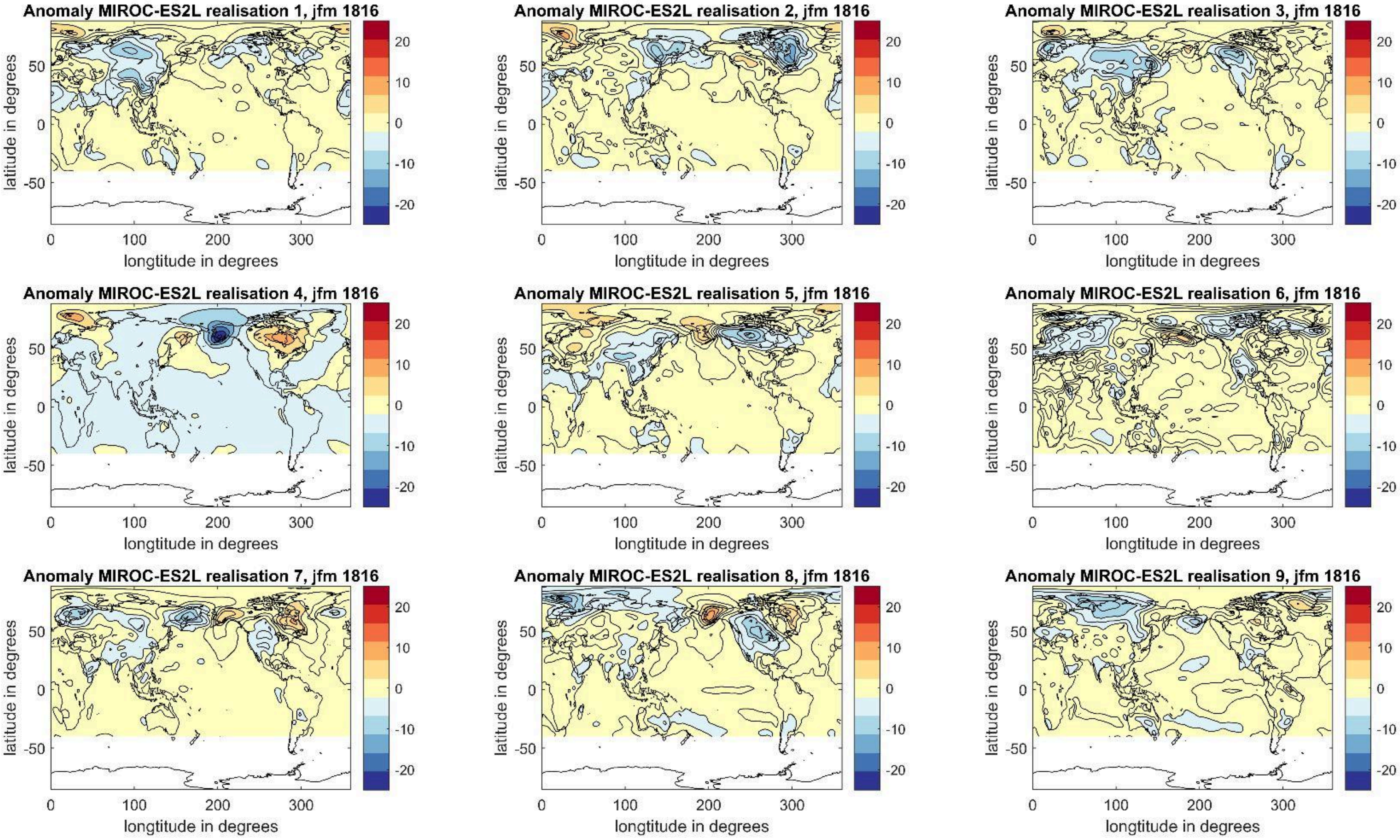
Annexe 3 : Anomalies of the near-surface temperature calculated from the 9 realizations of the MIROC-ES2L for summer, with the anomalies in K



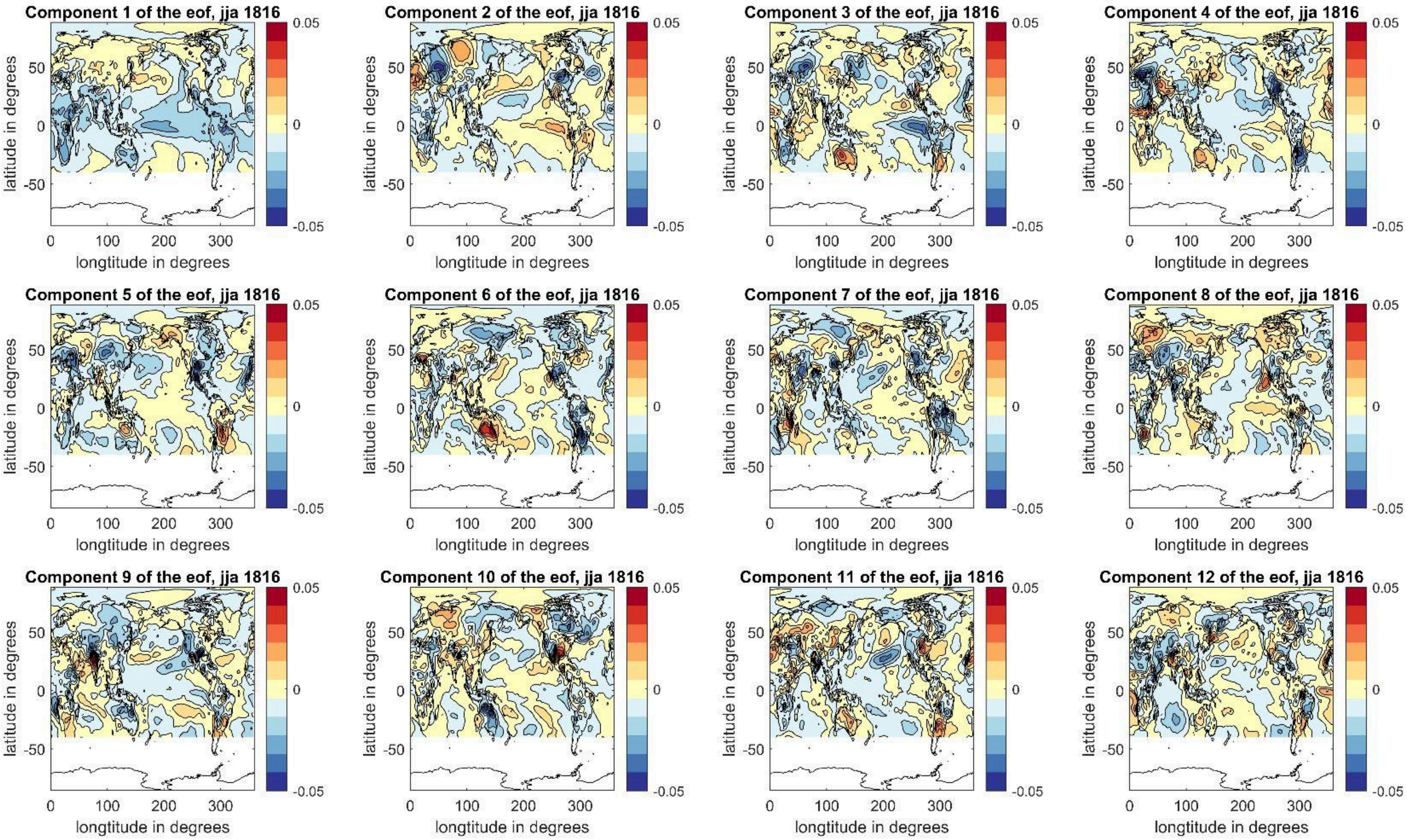
Annexe 4 : Anomalies of the near-surface temperature calculated from the 9 realizations of the MPI-ESM1-2.LR for winter, with the anomalies in K



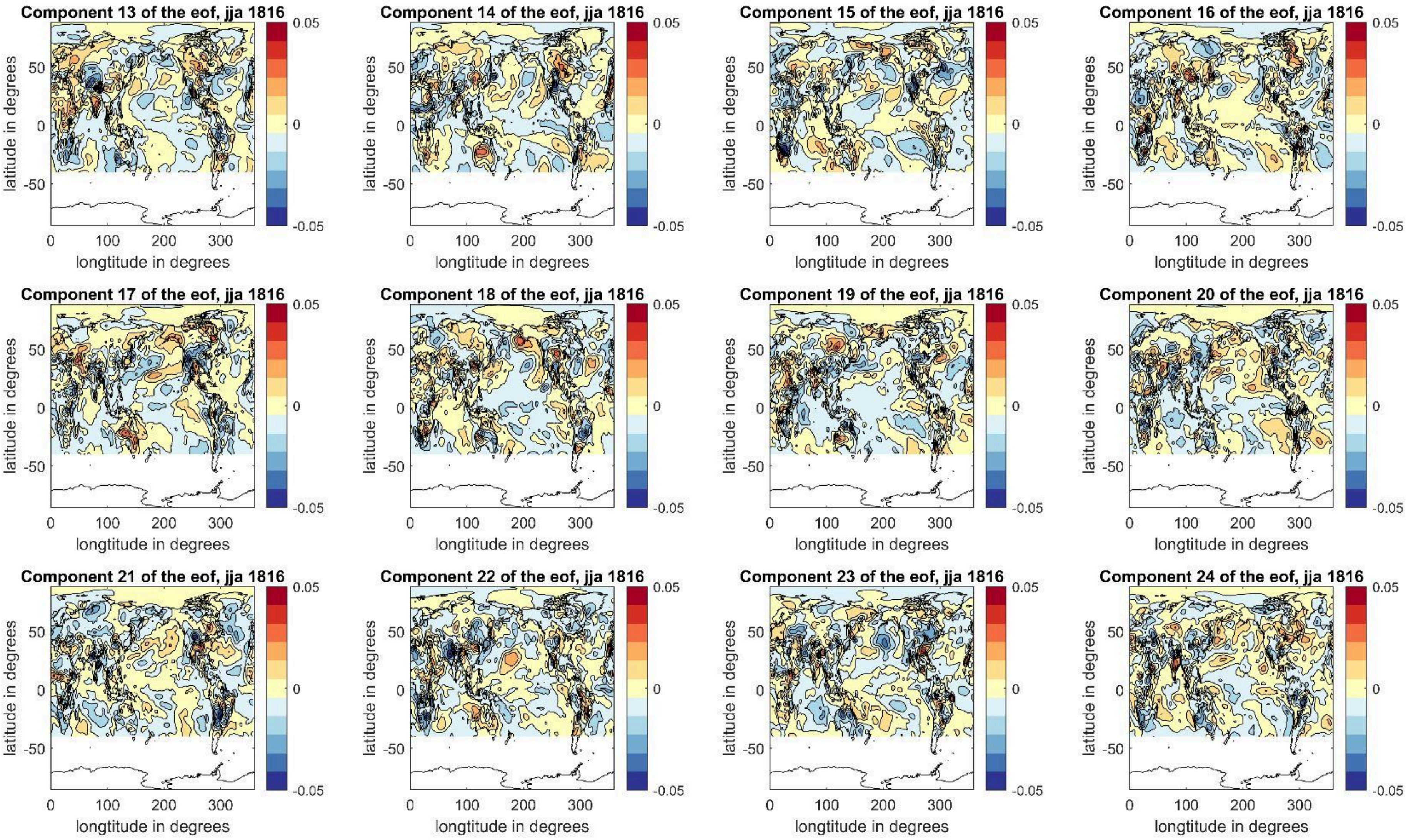
Annexe 5 : Anomalies of the near-surface temperature calculated from the 9 realizations of the CanESM5 for winter, with the anomalies in K



Annexe 6 : Anomalies of the near-surface temperature calculated from the 9 realizations of the MIROC-ES2L for winter, with the anomalies in K



Annexe 7 : Principal components 1 to 12 of the EOF analysis for summer in K


Component 13 of the eof, jja 1816

Component 14 of the eof, jja 1816

Component 15 of the eof, jja 1816

Component 16 of the eof, jja 1816

Component 17 of the eof, jja 1816

Component 18 of the eof, jja 1816

Component 19 of the eof, jja 1816

Component 20 of the eof, jja 1816

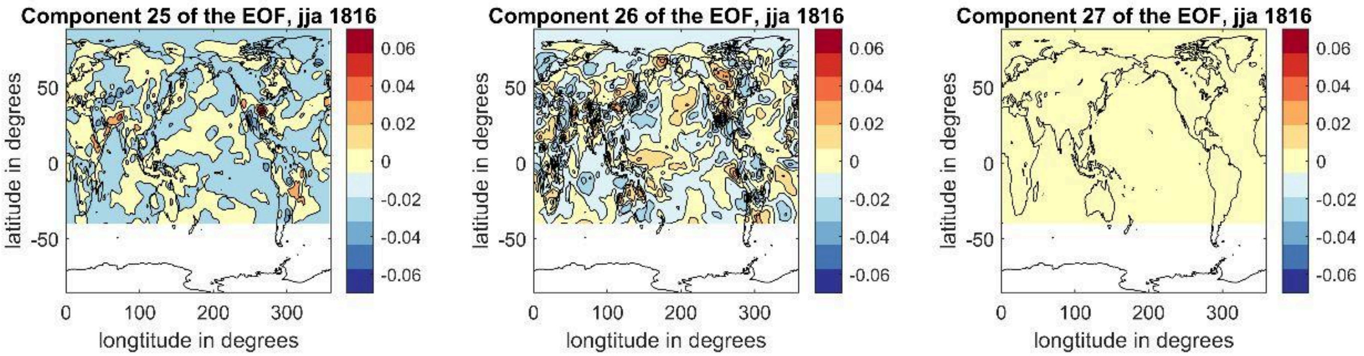
Component 21 of the eof, jja 1816

Component 22 of the eof, jja 1816

Component 23 of the eof, jja 1816

Component 24 of the eof, jja 1816

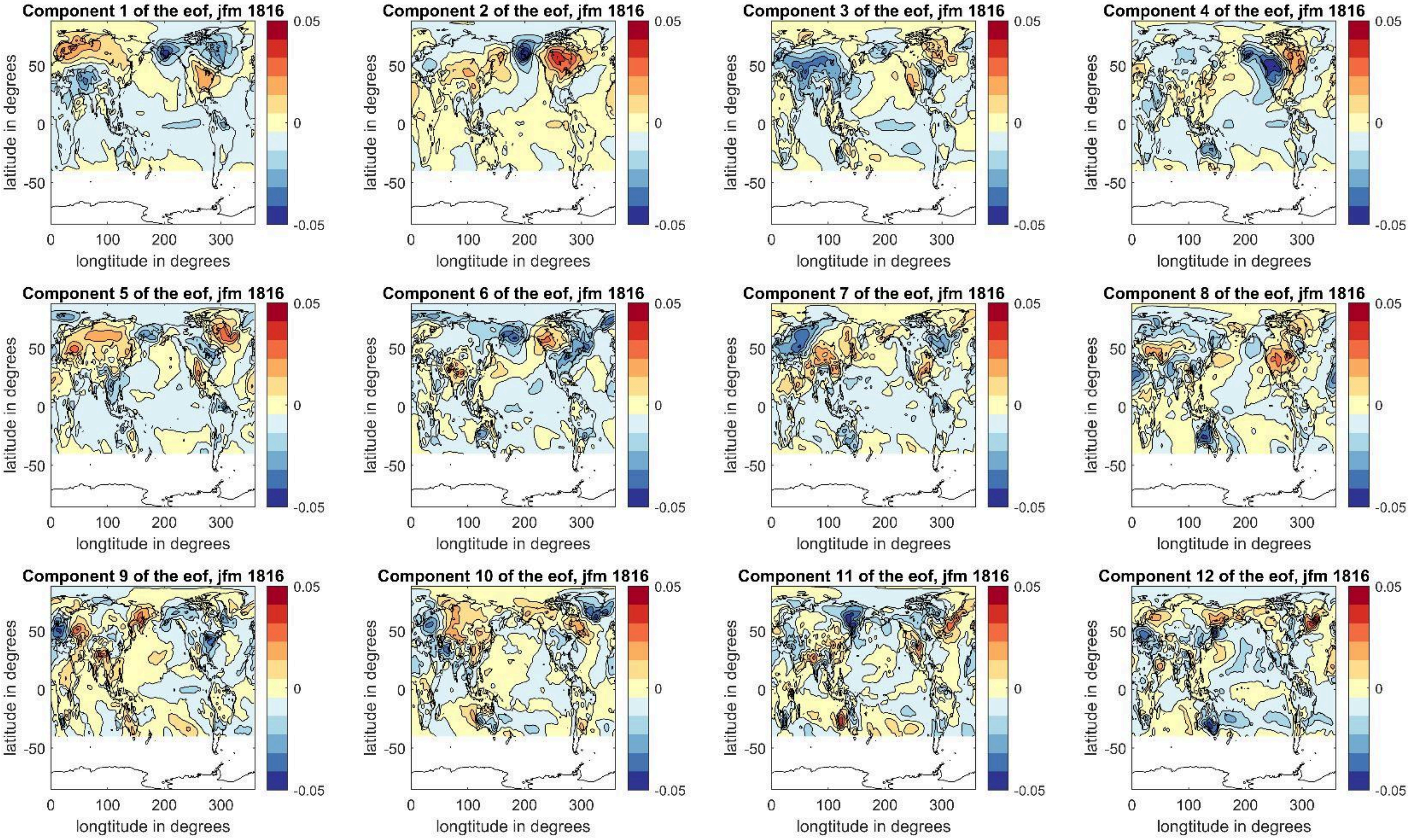
Annexe 8 : Principal components 13 to 24 of the EOF analysis for summer in K


Component 25 of the EOF, jja 1816

Component 26 of the EOF, jja 1816

Component 27 of the EOF, jja 1816

Annexe 9 : Principal components 25 to 27 of the EOF analysis for summer in K



Component 1 of the eof, jfm 1816

Component 2 of the eof, jfm 1816

Component 3 of the eof, jfm 1816

Component 4 of the eof, jfm 1816

Component 5 of the eof, jfm 1816

Component 6 of the eof, jfm 1816

Component 7 of the eof, jfm 1816

Component 8 of the eof, jfm 1816

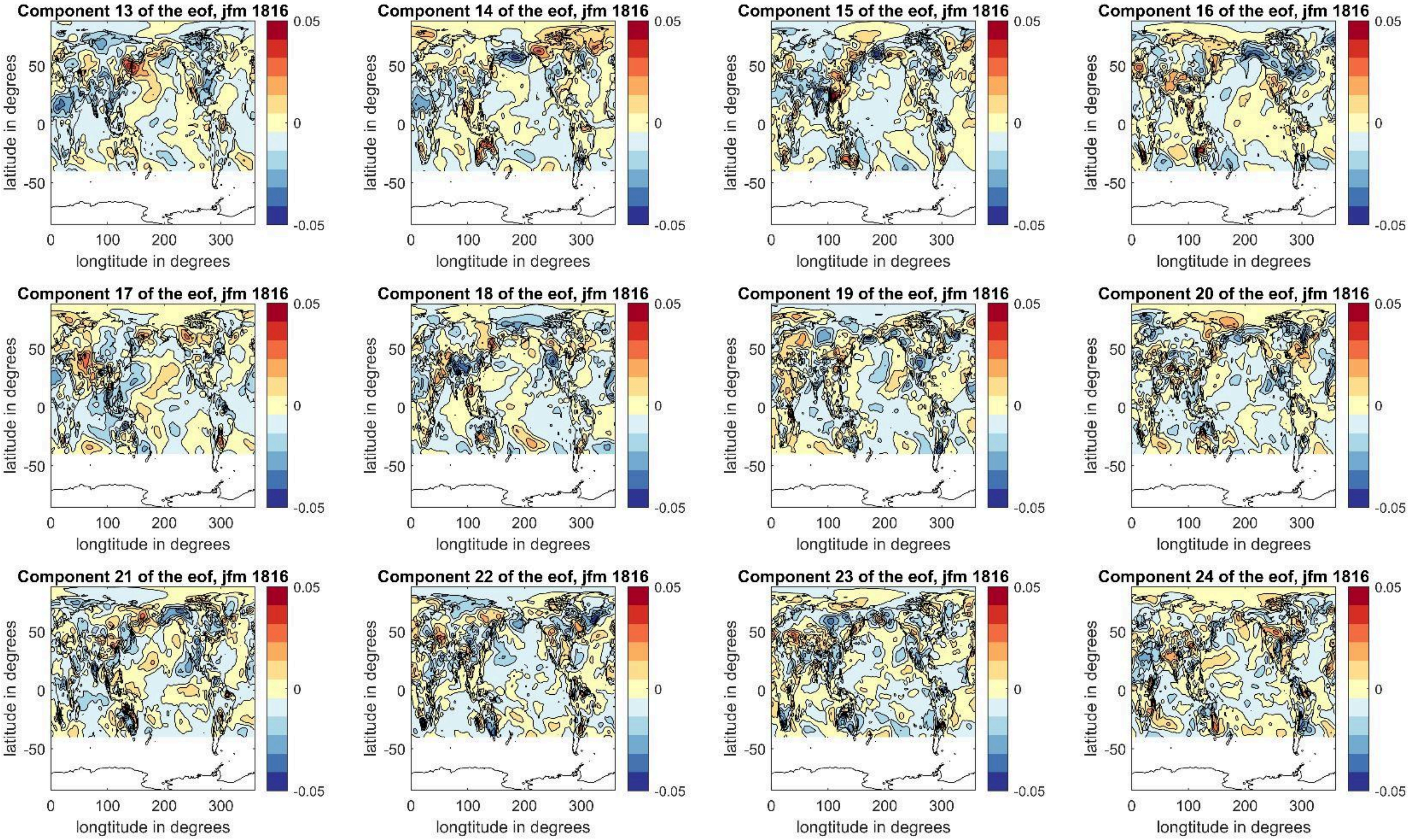
Component 9 of the eof, jfm 1816

Component 10 of the eof, jfm 1816

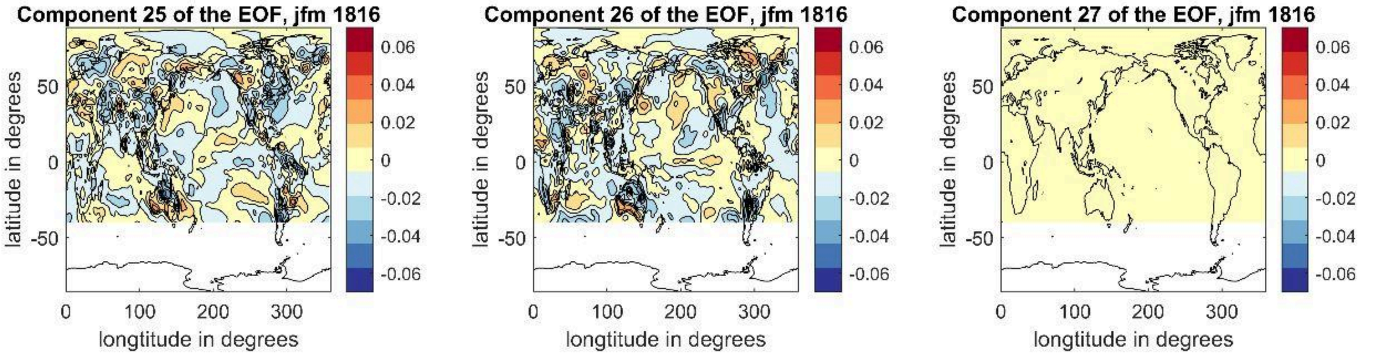
Component 11 of the eof, jfm 1816

Component 12 of the eof, jfm 1816

Annexe 10 : Principal components 1 to 12 of the EOF analysis for winter in K

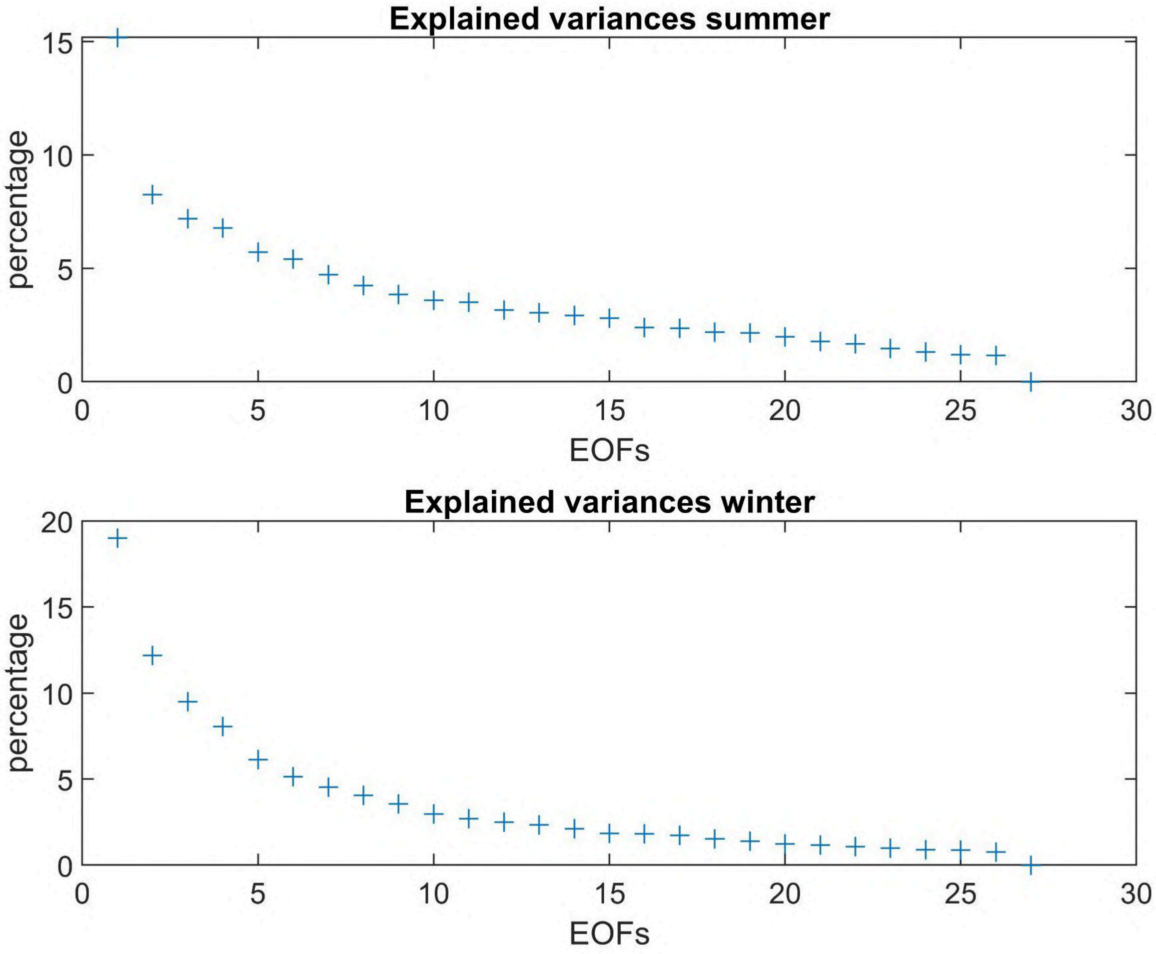


Annexe 11 : Principal components 13 to 24 of the EOF analysis for winter in K

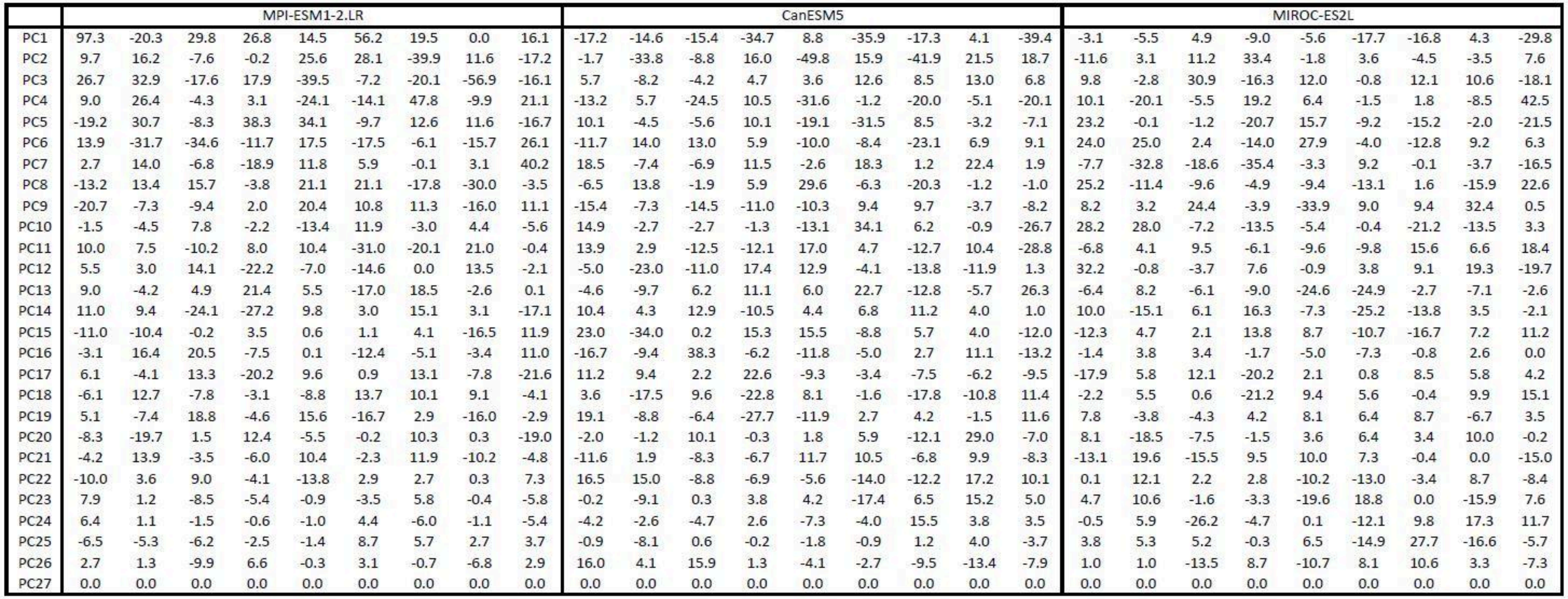


Annexe 12 : Principal components 25 to 27 of the EOF analysis for winter in K

Annexe 13 : Explained variances of the different EOFs for summer and winter.



Annexe 14 : Table of the loadings between the Principal Components and the realizations of each model for summer.

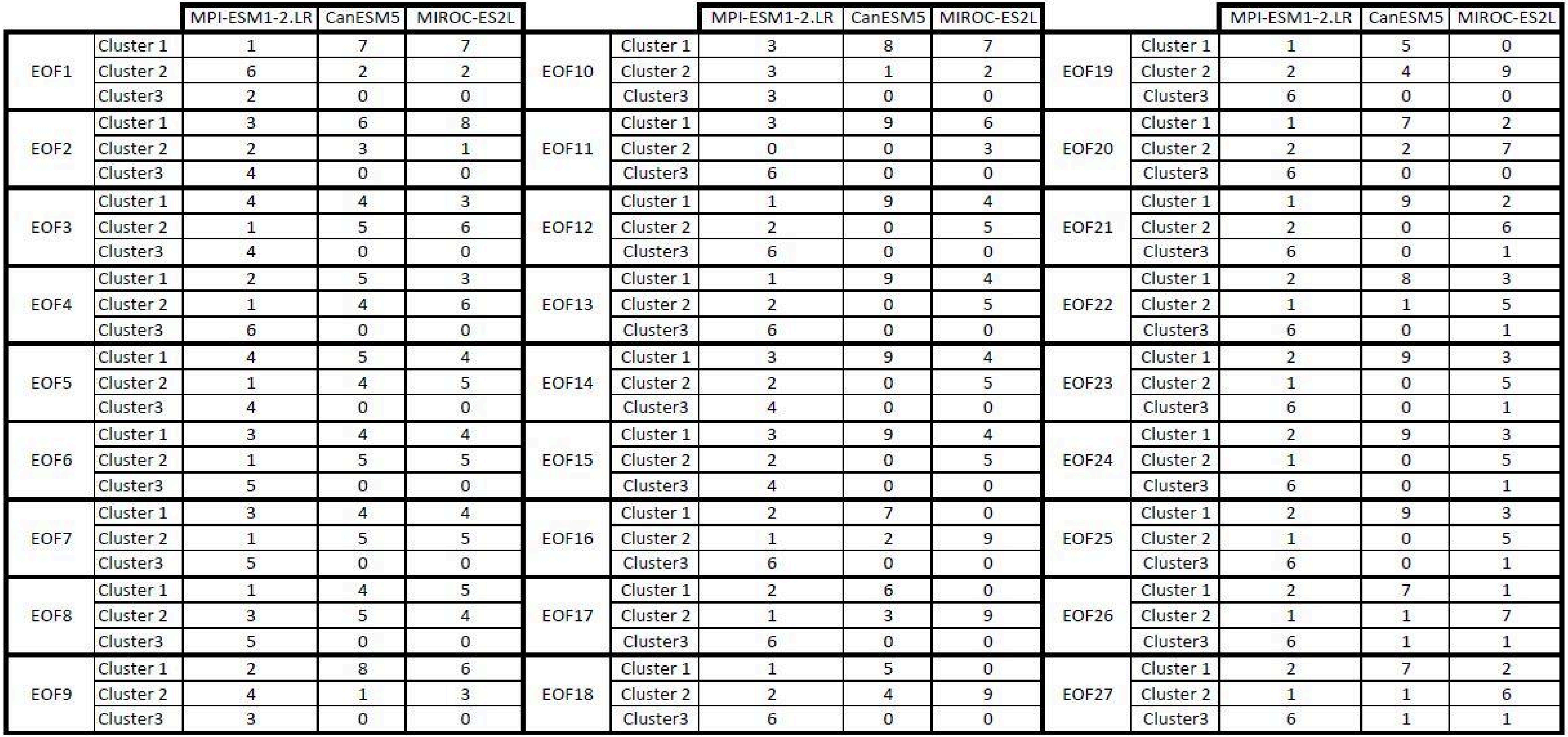


| | MPI-ESM1-2.LR | | | | | | | | | CanESM5 | | | | | | | | | MIROC-ES2L | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PC1 | 97.3 | -20.3 | 29.8 | 26.8 | 14.5 | 56.2 | 19.5 | 0.0 | 16.1 | -17.2 | -14.6 | -15.4 | -34.7 | 8.8 | -35.9 | -17.3 | 4.1 | -39.4 | -3.1 | -5.5 | 4.9 | -9.0 | -5.6 | -17.7 | -16.8 | 4.3 | -29.8 |
| PC2 | 9.7 | 16.2 | -7.6 | -0.2 | 25.6 | 28.1 | -39.9 | 11.6 | -17.2 | -1.7 | -33.8 | -8.8 | 16.0 | -49.8 | 15.9 | -41.9 | 21.5 | 18.7 | -11.6 | 3.1 | 11.2 | 33.4 | -1.8 | 3.6 | -4.5 | -3.5 | 7.6 |
| PC3 | 26.7 | 32.9 | -17.6 | 17.9 | -39.5 | -7.2 | -20.1 | -56.9 | -16.1 | 5.7 | -8.2 | -4.2 | 4.7 | 3.6 | 12.6 | 8.5 | 13.0 | 6.8 | 9.8 | -2.8 | 30.9 | -16.3 | 12.0 | -0.8 | 12.1 | 10.6 | -18.1 |
| PC4 | 9.0 | 26.4 | -4.3 | 3.1 | -24.1 | -14.1 | 47.8 | -9.9 | 21.1 | -13.2 | 5.7 | -24.5 | 10.5 | -31.6 | -1.2 | -20.0 | -5.1 | -20.1 | 10.1 | -20.1 | -5.5 | 19.2 | 6.4 | -1.5 | 1.8 | -8.5 | 42.5 |
| PC5 | -19.2 | 30.7 | -8.3 | 38.3 | 34.1 | -9.7 | 12.6 | 11.6 | -16.7 | 10.1 | -4.5 | -5.6 | 10.1 | -19.1 | -31.5 | 8.5 | -3.2 | -7.1 | 23.2 | -0.1 | -1.2 | -20.7 | 15.7 | -9.2 | -15.2 | -2.0 | -21.5 |
| PC6 | 13.9 | -31.7 | -34.6 | -11.7 | 17.5 | -17.5 | -6.1 | -15.7 | 26.1 | -11.7 | 14.0 | 13.0 | 5.9 | -10.0 | -8.4 | -23.1 | 6.9 | 9.1 | 24.0 | 25.0 | 2.4 | -14.0 | 27.9 | -4.0 | -12.8 | 9.2 | 6.3 |
| PC7 | 2.7 | 14.0 | -6.8 | -18.9 | 11.8 | 5.9 | -0.1 | 3.1 | 40.2 | 18.5 | -7.4 | -6.9 | 11.5 | -2.6 | 18.3 | 1.2 | 22.4 | 1.9 | -7.7 | -32.8 | -18.6 | -35.4 | -3.3 | 9.2 | -0.1 | -3.7 | -16.5 |
| PC8 | -13.2 | 13.4 | 15.7 | -3.8 | 21.1 | 21.1 | -17.8 | -30.0 | -3.5 | -6.5 | 13.8 | -1.9 | 5.9 | 29.6 | -6.3 | -20.3 | -1.2 | -1.0 | 25.2 | -11.4 | -9.6 | -4.9 | -9.4 | -13.1 | 1.6 | -15.9 | 22.6 |
| PC9 | -20.7 | -7.3 | -9.4 | 2.0 | 20.4 | 10.8 | 11.3 | -16.0 | 11.1 | -15.4 | -7.3 | -14.5 | -11.0 | -10.3 | 9.4 | 9.7 | -3.7 | -8.2 | 8.2 | 3.2 | 24.4 | -3.9 | -33.9 | 9.0 | 9.4 | 32.4 | 0.5 |
| PC10 | -1.5 | -4.5 | 7.8 | -2.2 | -13.4 | 11.9 | -3.0 | 4.4 | -5.6 | 14.9 | -2.7 | -2.7 | -1.3 | -13.1 | 34.1 | 6.2 | -0.9 | -26.7 | 28.2 | 28.0 | -7.2 | -13.5 | -5.4 | -0.4 | -21.2 | -13.5 | 3.3 |
| PC11 | 10.0 | 7.5 | -10.2 | 8.0 | 10.4 | -31.0 | -20.1 | 21.0 | -0.4 | 13.9 | 2.9 | -12.5 | -12.1 | 17.0 | 4.7 | -12.7 | 10.4 | -28.8 | -6.8 | 4.1 | 9.5 | -6.1 | -9.6 | -9.8 | 15.6 | 6.6 | 18.4 |
| PC12 | 5.5 | 3.0 | 14.1 | -22.2 | -7.0 | -14.6 | 0.0 | 13.5 | -2.1 | -5.0 | -23.0 | -11.0 | 17.4 | 12.9 | -4.1 | -13.8 | -11.9 | 1.3 | 32.2 | -0.8 | -3.7 | 7.6 | -0.9 | 3.8 | 9.1 | 19.3 | -19.7 |
| PC13 | 9.0 | -4.2 | 4.9 | 21.4 | 5.5 | -17.0 | 18.5 | -2.6 | 0.1 | -4.6 | -9.7 | 6.2 | 11.1 | 6.0 | 22.7 | -12.8 | -5.7 | 26.3 | -6.4 | 8.2 | -6.1 | -9.0 | -24.6 | -24.9 | -2.7 | -7.1 | -2.6 |
| PC14 | 11.0 | 9.4 | -24.1 | -27.2 | 9.8 | 3.0 | 15.1 | 3.1 | -17.1 | 10.4 | 4.3 | 12.9 | -10.5 | 4.4 | 6.8 | 11.2 | 4.0 | 1.0 | 10.0 | -15.1 | 6.1 | 16.3 | -7.3 | -25.2 | -13.8 | 3.5 | -2.1 |
| PC15 | -11.0 | -10.4 | -0.2 | 3.5 | 0.6 | 1.1 | 4.1 | -16.5 | 11.9 | 23.0 | -34.0 | 0.2 | 15.3 | 15.5 | -8.8 | 5.7 | 4.0 | -12.0 | -12.3 | 4.7 | 2.1 | 13.8 | 8.7 | -10.7 | -16.7 | 7.2 | 11.2 |
| PC16 | -3.1 | 16.4 | 20.5 | -7.5 | 0.1 | -12.4 | -5.1 | -3.4 | 11.0 | -16.7 | -9.4 | 38.3 | -6.2 | -11.8 | -5.0 | 2.7 | 11.1 | -13.2 | -1.4 | 3.8 | 3.4 | -1.7 | -5.0 | -7.3 | -0.8 | 2.6 | 0.0 |
| PC17 | 6.1 | -4.1 | 13.3 | -20.2 | 9.6 | 0.9 | 13.1 | -7.8 | -21.6 | 11.2 | 9.4 | 2.2 | 22.6 | -9.3 | -3.4 | -7.5 | -6.2 | -9.5 | -17.9 | 5.8 | 12.1 | -20.2 | 2.1 | 0.8 | 8.5 | 5.8 | 4.2 |
| PC18 | -6.1 | 12.7 | -7.8 | -3.1 | -8.8 | 13.7 | 10.1 | 9.1 | -4.1 | 3.6 | -17.5 | 9.6 | -22.8 | 8.1 | -1.6 | -17.8 | -10.8 | 11.4 | -2.2 | 5.5 | 0.6 | -21.2 | 9.4 | 5.6 | -0.4 | 9.9 | 15.1 |
| PC19 | 5.1 | -7.4 | 18.8 | -4.6 | 15.6 | -16.7 | 2.9 | -16.0 | -2.9 | 19.1 | -8.8 | -6.4 | -27.7 | -11.9 | 2.7 | 4.2 | -1.5 | 11.6 | 7.8 | -3.8 | -4.3 | 4.2 | 8.1 | 6.4 | 8.7 | -6.7 | 3.5 |
| PC20 | -8.3 | -19.7 | 1.5 | 12.4 | -5.5 | -0.2 | 10.3 | 0.3 | -19.0 | -2.0 | -1.2 | 10.1 | -0.3 | 1.8 | 5.9 | -12.1 | 29.0 | -7.0 | 8.1 | -18.5 | -7.5 | -1.5 | 3.6 | 6.4 | 3.4 | 10.0 | -0.2 |
| PC21 | -4.2 | 13.9 | -3.5 | -6.0 | 10.4 | -2.3 | 11.9 | -10.2 | -4.8 | -11.6 | 1.9 | -8.3 | -6.7 | 11.7 | 10.5 | -6.8 | 9.9 | -8.3 | -13.1 | 19.6 | -15.5 | 9.5 | 10.0 | 7.3 | -0.4 | 0.0 | -15.0 |
| PC22 | -10.0 | 3.6 | 9.0 | -4.1 | -13.8 | 2.9 | 2.7 | 0.3 | 7.3 | 16.5 | 15.0 | -8.8 | -6.9 | -5.6 | -14.0 | -12.2 | 17.2 | 10.1 | 0.1 | 12.1 | 2.2 | 2.8 | -10.2 | -13.0 | -3.4 | 8.7 | -8.4 |
| PC23 | 7.9 | 1.2 | -8.5 | -5.4 | -0.9 | -3.5 | 5.8 | -0.4 | -5.8 | -0.2 | -9.1 | 0.3 | 3.8 | 4.2 | -17.4 | 6.5 | 15.2 | 5.0 | 4.7 | 10.6 | -1.6 | -3.3 | -19.6 | 18.8 | 0.0 | -15.9 | 7.6 |
| PC24 | 6.4 | 1.1 | -1.5 | -0.6 | -1.0 | 4.4 | -6.0 | -1.1 | -5.4 | -4.2 | -2.6 | -4.7 | 2.6 | -7.3 | -4.0 | 15.5 | 3.8 | 3.5 | -0.5 | 5.9 | -26.2 | -4.7 | 0.1 | -12.1 | 9.8 | 17.3 | 11.7 |
| PC25 | -6.5 | -5.3 | -6.2 | -2.5 | -1.4 | 8.7 | 5.7 | 2.7 | 3.7 | -0.9 | -8.1 | 0.6 | -0.2 | -1.8 | -0.9 | 1.2 | 4.0 | -3.7 | 3.8 | 5.3 | 5.2 | -0.3 | 6.5 | -14.9 | 27.7 | -16.6 | -5.7 |
| PC26 | 2.7 | 1.3 | -9.9 | 6.6 | -0.3 | 3.1 | -0.7 | -6.8 | 2.9 | 16.0 | 4.1 | 15.9 | 1.3 | -4.1 | -2.7 | -9.5 | -13.4 | -7.9 | 1.0 | 1.0 | -13.5 | 8.7 | -10.7 | 8.1 | 10.6 | 3.3 | -7.3 |
| PC27 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Annexe 15 : Table of the loadings between the Principal Components and the realizations of each model for winter.

| | MPI-ESM1-2.LR | | | | | | | | | CanESM5 | | | | | | | | | MIROC-ES2L | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PC1 | 47.1 | -37.0 | 17.3 | 57.3 | 77.7 | 19.8 | 57.3 | 63.4 | 80.6 | -29.5 | 47.0 | -17.1 | -9.8 | 48.0 | -83.4 | 7.4 | 31.5 | -84.4 | -15.1 | 24.1 | -48.4 | 5.5 | 2.0 | -42.6 | -75.5 | -85.3 | -57.8 |
| PC2 | -34.8 | -57.7 | 15.6 | 11.8 | -22.4 | -33.3 | 15.8 | -22.2 | 31.8 | -40.5 | 20.9 | -26.9 | 37.1 | 14.7 | 28.9 | 36.5 | -21.8 | 48.7 | -29.6 | -6.7 | -32.6 | 122.9 | -74.4 | 17.2 | -8.6 | -39.5 | 49.0 |
| PC3 | 67.1 | 0.1 | 54.0 | -24.1 | 15.9 | -12.2 | 3.8 | -13.0 | -82.0 | 38.0 | -0.1 | -49.5 | -30.1 | 76.1 | -24.1 | 7.5 | -4.7 | 63.3 | 11.4 | -36.3 | 14.9 | 8.4 | -11.8 | -13.0 | 4.3 | -39.5 | -24.4 |
| PC4 | 62.6 | -32.7 | -6.4 | 2.6 | 60.6 | 15.3 | -24.7 | -32.3 | 6.7 | -13.9 | -27.0 | 21.0 | 2.8 | -31.8 | -21.8 | -30.7 | -41.3 | -46.2 | 11.4 | -51.3 | 50.3 | 68.6 | 1.1 | 9.4 | 19.1 | 27.0 | 1.6 |
| PC5 | 38.5 | -18.9 | -6.8 | -33.1 | -8.9 | 34.9 | 49.2 | 34.8 | 23.7 | -27.4 | -4.1 | -1.5 | -14.9 | -17.9 | 21.4 | -17.3 | -7.5 | 49.0 | -48.1 | -26.2 | -42.0 | -2.8 | 2.8 | -8.6 | 27.5 | 45.9 | -41.7 |
| PC6 | 37.4 | 37.1 | 30.3 | -4.0 | -35.8 | 11.3 | 16.7 | 6.5 | -6.0 | -32.8 | -16.6 | -55.3 | 0.3 | 1.2 | -7.9 | -28.4 | -38.2 | -36.1 | 6.7 | 58.9 | -8.7 | 6.6 | -8.7 | 5.2 | 11.2 | 13.5 | 35.5 |
| PC7 | 8.5 | -27.8 | 21.5 | 19.9 | 12.9 | 36.3 | 7.9 | -39.6 | -1.1 | 47.8 | -40.1 | -14.3 | 43.0 | -1.1 | 33.9 | -23.9 | 21.4 | -14.2 | -17.5 | 8.0 | -24.7 | -26.4 | -30.0 | 24.2 | -14.3 | -2.4 | -7.8 |
| PC8 | -3.7 | -59.7 | -28.8 | -33.9 | 1.1 | 22.5 | 11.4 | -17.4 | 21.8 | -4.4 | -5.6 | -33.8 | 7.8 | 15.1 | 15.1 | 13.7 | -15.8 | 10.3 | 50.6 | 19.6 | 16.3 | -16.4 | 36.2 | -24.7 | -0.8 | -7.2 | 10.8 |
| PC9 | 27.0 | -27.1 | 10.9 | 12.8 | -60.7 | 33.6 | -9.8 | -24.5 | -22.1 | -24.3 | 19.5 | 21.6 | -13.0 | -11.1 | -12.3 | 5.1 | 48.7 | -3.8 | 8.5 | 1.2 | 2.1 | 14.2 | 15.5 | 6.5 | -19.9 | 2.7 | -1.3 |
| PC10 | 4.0 | -16.4 | -23.1 | 23.0 | -16.2 | 7.6 | -37.1 | -4.0 | 9.5 | 21.2 | 26.4 | -10.3 | -14.5 | 46.7 | -2.6 | 13.7 | -13.5 | -23.9 | -13.7 | -0.9 | -18.3 | -3.9 | -14.2 | -4.1 | 35.6 | 33.1 | -4.1 |
| PC11 | 28.7 | -15.0 | -3.5 | 27.3 | -15.9 | -19.7 | -5.8 | 2.3 | 5.5 | 14.6 | -12.7 | 10.8 | 24.2 | -25.7 | -31.2 | 32.1 | -31.4 | 29.3 | -11.2 | 28.3 | 13.8 | -16.7 | -5.4 | -6.9 | -8.5 | 6.3 | -13.6 |
| PC12 | -0.4 | 25.0 | -15.0 | 22.5 | -0.8 | 23.6 | 2.0 | -14.8 | 30.4 | 9.0 | -26.9 | -26.4 | -42.8 | -3.6 | -4.6 | 17.6 | 8.2 | 17.4 | -13.3 | -17.2 | 16.2 | 1.5 | -3.9 | -7.2 | -24.7 | 4.7 | 23.6 |
| PC13 | -1.7 | -23.7 | 6.5 | 15.8 | 8.0 | -13.1 | -17.4 | 44.7 | -17.6 | 16.4 | -19.5 | -17.0 | -10.6 | -20.4 | -1.4 | -1.9 | 14.6 | 3.1 | 27.1 | -5.3 | -32.5 | 13.4 | 6.8 | -6.6 | -11.0 | 23.7 | 19.5 |
| PC14 | -13.9 | 2.2 | 20.7 | 6.9 | 10.5 | 22.1 | 13.4 | -20.4 | -10.4 | 16.2 | 15.8 | 16.5 | -15.7 | -25.9 | -3.6 | 2.8 | -10.2 | 0.9 | 7.1 | 11.8 | -14.2 | 5.7 | -12.8 | -52.4 | 23.1 | -4.8 | 8.5 |
| PC15 | 18.4 | 15.2 | 9.8 | -3.0 | -0.4 | 4.6 | -31.2 | 5.5 | 19.2 | -19.3 | 15.7 | -11.7 | 31.0 | -8.3 | 7.3 | 0.0 | 11.9 | 8.2 | 15.9 | -26.3 | 0.5 | -22.3 | -22.5 | -24.7 | 0.3 | -2.4 | 9.0 |
| PC16 | -14.3 | -2.0 | 24.9 | 3.8 | -26.9 | 12.0 | -6.3 | 19.5 | 14.7 | 7.6 | -19.4 | 27.2 | 0.3 | 26.4 | 9.5 | -19.6 | -26.9 | 0.4 | 9.5 | -10.5 | 11.8 | 7.0 | -1.0 | -17.7 | -22.0 | 1.5 | -9.5 |
| PC17 | 5.1 | 23.2 | -4.2 | -0.1 | -2.9 | 0.9 | -26.8 | -16.1 | 19.0 | 11.1 | 0.8 | -8.5 | -2.7 | -11.0 | 7.4 | -8.0 | -5.6 | 13.1 | 18.0 | 12.3 | -26.9 | 25.1 | 20.6 | 6.9 | 0.0 | -15.0 | -35.4 |
| PC18 | -25.1 | 5.3 | 6.1 | 4.0 | -9.7 | 23.5 | 11.8 | 8.5 | -2.8 | -3.8 | -13.5 | -14.5 | 13.6 | -5.2 | -28.0 | 21.4 | -0.5 | -5.9 | 20.5 | -18.5 | 2.3 | 2.9 | -9.4 | 16.8 | 22.4 | -0.6 | -21.6 |
| PC19 | 12.7 | 5.0 | 13.5 | -29.6 | -13.6 | -25.1 | 11.3 | -11.0 | 19.4 | 17.1 | -15.5 | 8.4 | 0.3 | 1.4 | -2.5 | 26.1 | 12.6 | -23.1 | 2.2 | -9.8 | -11.9 | 4.5 | 2.7 | -8.5 | 2.5 | 7.8 | 3.2 |
| PC20 | -2.9 | -11.9 | 27.7 | 18.1 | -1.8 | -15.4 | 5.2 | -13.1 | 17.0 | -5.6 | 9.5 | -0.9 | -14.3 | -5.1 | -3.2 | -8.9 | -12.8 | 5.3 | 6.3 | -17.1 | -8.7 | -22.1 | 17.4 | 19.8 | 11.9 | -4.8 | 10.3 |
| PC21 | 11.3 | 9.5 | -13.5 | -2.1 | -2.2 | 14.0 | 9.6 | 2.7 | -14.1 | 7.6 | 15.9 | 8.5 | 2.2 | -5.2 | 11.3 | 18.1 | -28.7 | -9.0 | 9.6 | -17.2 | -18.9 | -5.2 | 0.0 | 13.4 | -22.0 | -6.1 | 10.4 |
| PC22 | 6.8 | 1.5 | -22.6 | 8.3 | -19.7 | -12.4 | 22.8 | 3.2 | 6.2 | 23.2 | 12.1 | -5.5 | 4.6 | -5.8 | -5.8 | -27.8 | 6.1 | 0.5 | 12.6 | -10.2 | 8.4 | 5.0 | -9.3 | -3.5 | 6.0 | -6.8 | 2.0 |
| PC23 | 6.6 | 5.6 | -9.7 | 29.2 | -2.3 | -12.8 | 11.9 | -5.2 | -11.6 | -19.0 | -18.4 | 1.7 | 3.2 | 7.5 | 24.1 | 9.0 | 1.9 | -6.6 | 6.5 | -6.8 | -3.0 | -0.1 | 7.8 | -15.4 | 10.4 | -4.8 | -9.8 |
| PC24 | 0.8 | -5.2 | 8.2 | 0.1 | -9.7 | 5.2 | -7.8 | 16.7 | -1.4 | 12.8 | 5.3 | -14.7 | 1.7 | -15.1 | 15.9 | 7.9 | -2.3 | -12.5 | -22.1 | -6.1 | 15.5 | 4.3 | 17.1 | -3.7 | 10.0 | -20.4 | -0.6 |
| PC25 | -5.2 | -2.1 | 11.7 | 2.3 | 2.4 | -9.9 | 7.0 | -4.0 | -0.3 | 0.2 | 18.1 | -15.0 | -9.5 | -9.4 | 16.8 | 6.0 | -1.8 | -7.2 | 8.4 | 2.9 | 16.2 | 0.7 | -12.2 | 1.1 | -16.3 | 22.1 | -23.0 |
| PC26 | -10.0 | 6.6 | 2.6 | 4.5 | 0.6 | -1.1 | 5.1 | -11.1 | -6.2 | 1.7 | 7.9 | -11.5 | 22.4 | 5.0 | -11.8 | -4.7 | -2.1 | 1.3 | -12.5 | -10.0 | -4.3 | 6.7 | 24.7 | -12.3 | -11.6 | 17.6 | 2.6 |
| PC27 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Annexe 16 : Table of the clusters’ structures for summer



| | | MPI-ESM1-2.LR | CanESM5 | MIROC-ES2L |
|---|---|---|---|---|
| EOF1 | Cluster 1 | 1 | 7 | 7 |
| | Cluster 2 | 6 | 2 | 2 |
| | Cluster3 | 2 | 0 | 0 |
| EOF2 | Cluster 1 | 3 | 6 | 8 |
| | Cluster 2 | 2 | 3 | 1 |
| | Cluster3 | 4 | 0 | 0 |
| EOF3 | Cluster 1 | 4 | 4 | 3 |
| | Cluster 2 | 1 | 5 | 6 |
| | Cluster3 | 4 | 0 | 0 |
| EOF4 | Cluster 1 | 2 | 5 | 3 |
| | Cluster 2 | 1 | 4 | 6 |
| | Cluster3 | 6 | 0 | 0 |
| EOF5 | Cluster 1 | 4 | 5 | 4 |
| | Cluster 2 | 1 | 4 | 5 |
| | Cluster3 | 4 | 0 | 0 |
| EOF6 | Cluster 1 | 3 | 4 | 4 |
| | Cluster 2 | 1 | 5 | 5 |
| | Cluster3 | 5 | 0 | 0 |
| EOF7 | Cluster 1 | 3 | 4 | 4 |
| | Cluster 2 | 1 | 5 | 5 |
| | Cluster3 | 5 | 0 | 0 |
| EOF8 | Cluster 1 | 1 | 4 | 5 |
| | Cluster 2 | 3 | 5 | 4 |
| | Cluster3 | 5 | 0 | 0 |
| EOF9 | Cluster 1 | 2 | 8 | 6 |
| | Cluster 2 | 4 | 1 | 3 |
| | Cluster3 | 3 | 0 | 0 |
| EOF10 | Cluster 1 | 3 | 8 | 7 |
| | Cluster 2 | 3 | 1 | 2 |
| | Cluster3 | 3 | 0 | 0 |
| EOF11 | Cluster 1 | 3 | 9 | 6 |
| | Cluster 2 | 0 | 0 | 3 |
| | Cluster3 | 6 | 0 | 0 |
| EOF12 | Cluster 1 | 1 | 9 | 4 |
| | Cluster 2 | 2 | 0 | 5 |
| | Cluster3 | 6 | 0 | 0 |
| EOF13 | Cluster 1 | 1 | 9 | 4 |
| | Cluster 2 | 2 | 0 | 5 |
| | Cluster3 | 6 | 0 | 0 |
| EOF14 | Cluster 1 | 3 | 9 | 4 |
| | Cluster 2 | 2 | 0 | 5 |
| | Cluster3 | 4 | 0 | 0 |
| EOF15 | Cluster 1 | 3 | 9 | 4 |
| | Cluster 2 | 2 | 0 | 5 |
| | Cluster3 | 4 | 0 | 0 |
| EOF16 | Cluster 1 | 2 | 7 | 0 |
| | Cluster 2 | 1 | 2 | 9 |
| | Cluster3 | 6 | 0 | 0 |
| EOF17 | Cluster 1 | 2 | 6 | 0 |
| | Cluster 2 | 1 | 3 | 9 |
| | Cluster3 | 6 | 0 | 0 |
| EOF18 | Cluster 1 | 1 | 5 | 0 |
| | Cluster 2 | 2 | 4 | 9 |
| | Cluster3 | 6 | 0 | 0 |
| EOF19 | Cluster 1 | 1 | 5 | 0 |
| | Cluster 2 | 2 | 4 | 9 |
| | Cluster3 | 6 | 0 | 0 |
| EOF20 | Cluster 1 | 1 | 7 | 2 |
| | Cluster 2 | 2 | 2 | 7 |
| | Cluster3 | 6 | 0 | 0 |
| EOF21 | Cluster 1 | 1 | 9 | 2 |
| | Cluster 2 | 2 | 0 | 6 |
| | Cluster3 | 6 | 0 | 1 |
| EOF22 | Cluster 1 | 2 | 8 | 3 |
| | Cluster 2 | 1 | 1 | 5 |
| | Cluster3 | 6 | 0 | 1 |
| EOF23 | Cluster 1 | 2 | 9 | 3 |
| | Cluster 2 | 1 | 0 | 5 |
| | Cluster3 | 6 | 0 | 1 |
| EOF24 | Cluster 1 | 2 | 9 | 3 |
| | Cluster 2 | 1 | 0 | 5 |
| | Cluster3 | 6 | 0 | 1 |
| EOF25 | Cluster 1 | 2 | 9 | 3 |
| | Cluster 2 | 1 | 0 | 5 |
| | Cluster3 | 6 | 0 | 1 |
| EOF26 | Cluster 1 | 2 | 7 | 1 |
| | Cluster 2 | 1 | 1 | 7 |
| | Cluster3 | 6 | 1 | 1 |
| EOF27 | Cluster 1 | 2 | 7 | 2 |
| | Cluster 2 | 1 | 1 | 6 |
| | Cluster3 | 6 | 1 | 1 |

Annexe 17 : Table of the clusters’ structures for winter .

| | | MPI-ESM1-2.LR | CanESM5 | MIROC-ES2L |
|---|---|---|---|---|
| EOF1 | Cluster 1 | 3 | 4 | 3 |
| | Cluster 2 | 0 | 2 | 5 |
| | Cluster3 | 6 | 3 | 1 |
| EOF2 | Cluster 1 | 1 | 2 | 5 |
| | Cluster 2 | 0 | 4 | 3 |
| | Cluster3 | 8 | 3 | 1 |
| EOF3 | Cluster 1 | 0 | 4 | 3 |
| | Cluster 2 | 1 | 2 | 5 |
| | Cluster3 | 8 | 3 | 1 |
| EOF4 | Cluster 1 | 5 | 3 | 2 |
| | Cluster 2 | 1 | 5 | 7 |
| | Cluster3 | 3 | 1 | 0 |
| EOF5 | Cluster 1 | 4 | 4 | 5 |
| | Cluster 2 | 0 | 2 | 4 |
| | Cluster3 | 5 | 3 | 0 |
| EOF6 | Cluster 1 | 2 | 5 | 8 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 3 | 1 | 0 |
| EOF7 | Cluster 1 | 1 | 5 | 7 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 4 | 1 | 1 |
| EOF8 | Cluster 1 | 0 | 4 | 6 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 2 | 2 |
| EOF9 | Cluster 1 | 0 | 4 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 2 | 3 |
| EOF10 | Cluster 1 | 0 | 4 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 2 | 3 |
| EOF11 | Cluster 1 | 0 | 4 | 5 |
| | Cluster 2 | 4 | 2 | 1 |
| | Cluster3 | 5 | 3 | 3 |
| EOF12 | Cluster 1 | 0 | 3 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 3 | 3 |
| EOF13 | Cluster 1 | 0 | 4 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 2 | 3 |
| EOF14 | Cluster 1 | 0 | 4 | 4 |
| | Cluster 2 | 3 | 3 | 2 |
| | Cluster3 | 6 | 2 | 3 |
| EOF15 | Cluster 1 | 0 | 4 | 4 |
| | Cluster 2 | 3 | 3 | 2 |
| | Cluster3 | 6 | 2 | 3 |
| EOF16 | Cluster 1 | 0 | 4 | 4 |
| | Cluster 2 | 4 | 3 | 2 |
| | Cluster3 | 5 | 2 | 3 |
| EOF17 | Cluster 1 | 0 | 4 | 3 |
| | Cluster 2 | 3 | 3 | 3 |
| | Cluster3 | 6 | 2 | 3 |
| EOF18 | Cluster 1 | 0 | 4 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 2 | 3 |
| EOF19 | Cluster 1 | 2 | 4 | 4 |
| | Cluster 2 | 4 | 4 | 2 |
| | Cluster3 | 3 | 1 | 3 |
| EOF20 | Cluster 1 | 0 | 3 | 5 |
| | Cluster 2 | 4 | 4 | 1 |
| | Cluster3 | 5 | 2 | 3 |
| EOF21 | Cluster 1 | 0 | 3 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 3 | 3 |
| EOF22 | Cluster 1 | 0 | 3 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 3 | 3 |
| EOF23 | Cluster 1 | 0 | 3 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 3 | 3 |
| EOF24 | Cluster 1 | 0 | 3 | 5 |
| | Cluster 2 | 4 | 3 | 1 |
| | Cluster3 | 5 | 3 | 3 |
| EOF25 | Cluster 1 | 0 | 4 | 5 |
| | Cluster 2 | 4 | 4 | 1 |
| | Cluster3 | 5 | 1 | 3 |
| EOF26 | Cluster 1 | 1 | 4 | 7 |
| | Cluster 2 | 4 | 4 | 2 |
| | Cluster3 | 4 | 1 | 0 |
| EOF27 | Cluster 1 | 1 | 4 | 7 |
| | Cluster 2 | 4 | 4 | 2 |
| | Cluster3 | 4 | 1 | 0 |

# Annexe 18 : Anomaly Calculation

```
function [Anomaly] = anomaly_function(months,year)
% This function calculates the global temperature anomaly for the 3 months
% chosen for a specific year chosen
% This function uses 3 models : MIROC-ES2L (9 realizations) , MPI-ESM1-2.LR (9
% realizations) and CanESM5 (9 realizations)
% The aim is to do the calculation of the anomaly for each realization and
% to create a matrix 3d with lat*long*27 realizations
% The inputs are the numbers of the corresponding months ex : for jja 1816
% the input is [6,7,8]
% The second input is the year of the study, so 1816 for example.
% model 1 refers to MPI-ESM1-2.LR
% model 2 refers to CanESM5
% model 3 refers to MIROC-ES2L
% Set up longitudes and latitudes
lon_1 =
ncread('MPIESM_2_model/tas_Amon_MPI-ESM1-2-LR_volc-long-eq_r11i1p1f1_gn_181501-183412.nc','lon');
lat_1 =
ncread('MPIESM_2_model/tas_Amon_MPI-ESM1-2-LR_volc-long-eq_r11i1p1f1_gn_181501-183412.nc','lat');
lon_2 = ncread('CanESM5_model/tas_Amon_CanESM5_volc-long-eq_r25i1p2f1_gn_181504-187003.nc','lon');
lat_2 = ncread('CanESM5_model/tas_Amon_CanESM5_volc-long-eq_r25i1p2f1_gn_181504-187003.nc','lat');
lon_3 = ncread('MIROC_model/tas_Amon_MIROC-ES2L_volc-long-eq_r9i1p1f2_gn_181504-183512.nc','lon');
lat_3 = ncread('MIROC_model/tas_Amon_MIROC-ES2L_volc-long-eq_r9i1p1f2_gn_181504-183512.nc','lat');
% Find index for the dates
% Model 1
% All the 9 realizations are starting in 1815/01/01 for MPI ESM
dt =
ncread('MPIESM_2_model/tas_Amon_MPI-ESM1-2-LR_volc-long-eq_r11i1p1f1_gn_181501-183412.nc','time');
% Here days since 1850 01 01 00:00:00
dt_1 = datetime(1850,1,1) + days(dt); %,'Format','yyyy-MM')
[y_1,m_1] = ymd(dt_1);
date_1 = [y_1,m_1];
index_1 = [find(date_1(:,1)== year & date_1(:,2) == months(1)), find(date_1(:,1)== year & date_1(:,2) ==
months(2)),find(date_1(:,1)== year & date_1(:,2) == months(3))];
% index1 is a vector with the indexes of the months of the year for the
% Model 1
% Model 2
% days since 1850 01 01 00:00:00
dt = ncread('CanESM5_model/tas_Amon_CanESM5_volc-long-eq_r1i1p2f1_gn_181504-187003.nc','time');
dt_2 = datetime(1850,1,1) + days(dt); %,'Format','yyyy-MM')
[y_2,m_2] = ymd(dt_2);
date_2 = [y_2,m_2];
index_2 = [find(date_2(:,1)== year & date_2(:,2) == months(1)), find(date_2(:,1)== year & date_2(:,2) ==
months(2)),find(date_2(:,1)== year & date_2(:,2) == months(3))];
% Model 3
% days since 1850 01 01 00:00:00, now modified so from 1815/04/01
dt = ncread('MIROC_model/tas_Amon_MIROC-ES2L_volc-long-eq_r9i1p1f2_gn_181504-183512.nc','time');
dmy=[0; cumsum(365.25/12.*ones(248,1))]; % because modified file, not days since 1815 bus months since
1815
dt_3 = datetime(1815,4,15) + dmy;%,'Format','yyyy-MM') % modified so 1815/04/01
[y_3,m_3] = ymd(dt_3);
date_3 = [y_3,m_3];
index_3 = [find(date_3(:,1)== year & date_3(:,2) == months(1)), find(date_3(:,1)== year & date_3(:,2) ==
months(2)),find(date_3(:,1)== year & date_3(:,2) == months(3))];
% Calculation mean temperature 3 months , for the realization r11-r19 for
% model 1
temp_1 = ones(192,96,9);
```

```
% Download data 1816 for MPE-ISM-1-2-LR
m = 1;
  for y = 11:19
      temp =
ncread(sprintf('MPIESM_model/tas_Amon_MPI-ESM1-2-LR_volc-long-eq_r%di1p1f1_gn_181501.nc',y),'tas');
      temp_1(:,:,m) = mean(temp(:,:,index_1),3);
      m = m +1;
  end
% the result of this loop is a 3d matrix with the mean average the
% year 1816
% Calculation mean temperature 1816 , for the realization r1-r9 MIROC
temp_3 = ones(128,64,9);
% Download data 1816 for MPE-ISM-1-2-LR
n = 1;
  for x = 1:9
      temp =
ncread(sprintf('MIROC_model/tas_Amon_MIROC-ES2L_volc-long-eq_r%di1p1f2_gn_181504-183512.nc',x),'tas')
;
      temp_3(:,:,n) = mean(temp(:,:,index_3),3);
      n = n +1;
  end
% the result of this loop is a 3d matrix with the mean average of the
% year 1816
% Calculation mean temperature 1816 , for the realization r1-r25 CanESM5
temp_2 = ones(128,64,9);
% time started 1815-04
% Download data 1816 for MPE-ISM-1-2-LR
o = 1;
  for z = 1:9
      temp =
ncread(sprintf('CanESM5_model/tas_Amon_CanESM5_volc-long-eq_r%di1p2f1_gn_181504-187003.nc',z),'tas');
      temp_2(:,:,o) = mean(temp(:,:,index_2),3);
      o = o +1;
  end
% Calculation pi control
% picontrol MPI-ESM 1-2-LR
% r11 : 1932 (january 01)
% r12 : 2023
% r13 : 2042
% r14 : 2050
% r15 : 2074
% r16 : 2139
% r17 : 2238
% r18 : 2309
% r19 : 2317
% r11 : jja 1815 : jja 1932
% ...
% r19 : jja 1815 : jja 2317
% Here I have 19 files from 1850 01 to 2229 12
% I'm going to create two files
% days since 1850 01 01 00:00:00
dt_pi_1 = ncread("picontrol_MPIESM2\out_pi_control_MPIESM.nc",'time');
dt_pc_1 = datetime(1850,1,1) + days(dt_pi_1);%,'Format','yyyy-MM')
[y_pc_1,m_pc_1] = ymd(dt_pc_1);
date_pc_1 = [y_pc_1,m_pc_1];
% Loop that create a matrix 9*3, each row is a realization and column 1,
% column 2, column3 are month 1, month 2 and month 3.
pi_temp_1 = ones(9,3);
```

```
m = 1;
for a = [1933,2024,2043,2051,2075,2140,2239,2310,2318] % The corresponding piControl years
pi_temp_1(m,:) = [find(date_pc_1(:,1)== a & date_pc_1(:,2) == months(1)), find(date_pc_1(:,1)== a &
date_pc_1(:,2) == months(2)),find(date_pc_1(:,1)== a & date_pc_1(:,2) == months(3))];
m = m+1;
end
% pi control
pi_1 = ones(192,96,9);
for z = 1:9
   pi_temp = ncread("picontrol_MPIESM2\out_pi_control_MPIESM.nc",'tas');
   pi_1(:,:,z) = mean(pi_temp(:,:,pi_temp_1(z,:)),3);
end
% Calculation anomaly
anomaly_1 = temp_1 - pi_1;
% Calculation pi control + anomaly for r1-r9 MIROC
% r1 : 1850
% r2 : 1860
% ...
% r9 : 1950
% days since 1850 01 01 00:00:00
dt_3 = ncread("picontrol_MIROC\tas_Amon_MIROC-ES2L_piControl_r1i1p1f2_gn_185001-204912.nc",'time');
dt_pc_3 = datetime(1850,1,1) + days(dt_3);%,'Format','yyyy-MM')
[y_pc_3,m_pc_3] = ymd(dt_pc_3);
date_pc_3 = [y_pc_3,m_pc_3];
% Loop that create a matrix 9*3, each row is a realization and column 1,
% column 2, column3 are june july and august
pi_temp_3 = ones(9,3);
p = 1;
for b = [1851,1861,1871,1881,1891,1901,1911,1921,1931] % The corresponding  piControl years
pi_temp_3(p,:) = [find(date_pc_3(:,1)== b & date_pc_3(:,2) == months(1)), find(date_pc_3(:,1)== b &
date_pc_3(:,2) == months(2)),find(date_pc_3(:,1)== b & date_pc_3(:,2) == months(3))];
p = p+1;
end
% pi control
pi_3 = ones(128,64,9);
for z = 1:9
   pi_temp =
ncread("picontrol_MIROC\tas_Amon_MIROC-ES2L_piControl_r1i1p1f2_gn_185001-204912.nc",'tas');
   pi_3(:,:,z) = mean(pi_temp(:,:,pi_temp_3(z,:)),3);
end
% Calculation anomaly
anomaly_3 = temp_3 - pi_3;
% Calculation pi control + anomaly for r1-r9 CanESM5
% r1 : 5551
% r2 : 5556
% r3 : 5574
% r4 : 5621
% r5 : 5642
% r6 : 5786
% r7 : 5789
% r8 : 5803
% r9 : 5844
% days since 1850 01 01 00:00:00
dt_pi_2 = ncread("picontrol_CanESM5\out_pi_control_CanESM5.nc",'time');
dmy=[0; cumsum(365.25/12.*ones(12611,1))]; % because modified file, not days since 1815 bus months since
1815
dt_pc_2 = datetime(5550,1,15) + dmy; %days(dt_pi_CAN);%,'Format','yyyy-MM') % put 15th of january and not
1st otherwise it was confusing between the month
```

```
[y_pc_2,m_pc_2] = ymd(dt_pc_2);
date_pc_2 = [y_pc_2,m_pc_2];
% Loop that create a matrix 9*3, each row is a realization and column 1,
% column 2, column3 are june july and august
pi_temp_2 = ones(9,3);
q = 1;
for c = [5552,5557,5575,5622,5643,5787,5790,5804,5845] % The piControl years corresponding
pi_temp_2(q,:) = [find(date_pc_2(:,1)== c & date_pc_2(:,2) == months(1)), find(date_pc_2(:,1)== c &
date_pc_2(:,2) == months(2)),find(date_pc_2(:,1)== c & date_pc_2(:,2) == months(3))];
q = q+1;
end
% pi control
pi_2 = ones(128,64,9);
for z = 1:9
  pi_temp = ncread("picontrol_CanESM5\out_pi_control_CanESM5.nc",'tas');
  pi_2(:,:,z) = mean(pi_temp(:,:,pi_temp_2(z,:)),3);
end
% Calculation anomaly
anomaly_2 = temp_2 - pi_2;
% Interpolation, that all the matrices have the same lat and long.
% so 192*96
% MPI ESM is already 192*96 but CanESM5 and MIROC are 128*62
[X_1,Y_1] = ndgrid(lon_1,lat_1); % grid that we want 192*96
[X_2,Y_2] = ndgrid(lon_2,lat_2); % grid that we want to change 128*62
[X_3,Y_3] = ndgrid(lon_3,lat_3); % grid that we want to change 128*62
% CanESM5
interp_2 = ones(192,96,9);
for i = 1:9
i_2 = interpn(X_2,Y_2,anomaly_2(:,:,i),X_1,Y_1,'spline');
interp_2(:,:,i) = i_2;
end
% MIROC
interp_3 = ones(192,96,9);
for j = 1:9
i_3 = interpn(X_3,Y_3,anomaly_3(:,:,j),X_1,Y_1,'spline');
interp_3(:,:,j) = i_3;
end
% Creation of 3d matrix with all the realizations
Anomaly = ones(192,96,27);
Anomaly(:,:,1:9) = anomaly_1;
Anomaly(:,:,10:18) = interp_2;
Anomaly(:,:,19:27) = interp_3;
end
```

# Annexe 19 : Clustering calculation and identification of the clusters’s structures

```
function [sum_c] = clustering_eof(pcs)
% This function calculates the clustering for the 27th EOFs
% The inputs are the loadings, called pcs here
% The output is a matrix 3*3*27 with the 27 clustering analysis
% organized in function of the number of realizations of each model per
% cluster
% the matrix columns are : MPI-ESM1-2.LR, CanESM5 and MIROC-ES2L
% the matrix rows are : Cluster 1, Cluster 2 and Cluster 3

norm_pcs = normalize(pcs(:,:));
pc_clust = norm_pcs.';
sum_c = ones(3,3,27);
for e = 1:27
rng(1); % For reproducibility
[idx,C] = kmeans(pc_clust(:,1:e),3); % Calculation of the cluster with k-mean approach
c1 = find(idx == 1);  % The different clusters
c2 = find(idx == 2);
c3 = find(idx == 3);
% Testing for which model for cluster 1
c1m1 = 0;
c1m2 = 0;
c1m3 = 0;
for l = 1 : length(c1)
if c1(l) <= 9 % If index < 9, selects the indexes of the realizations under 9, so the MPI-ESM1-2.LR model
  c1m1 = c1m1 +1;
end
if c1(l) > 9 & c1(l) <= 18 % CanESM5
  c1m2 = c1m2 +1;
end
if c1(l) >= 19 % MIROC-ES2L
  c1m3 = c1m3 +1;
end
end
% Testing for which model for cluster 2
c2m1 = 0;
c2m2 = 0;
c2m3 = 0;
for l = 1 : length(c2)
if c2(l) <= 9
  c2m1 = c2m1 +1;
end
if c2(l) > 9 & c2(l) <= 18
  c2m2 = c2m2 +1;
end
if c2(l) >= 19
  c2m3 = c2m3 +1;
end
end
% Testing for which model for cluster 3
c3m1 = 0;
c3m2 = 0;
c3m3 = 0;
for l = 1 : length(c3)
if c3(l) <= 9
  c3m1 = c3m1 +1;
end
```

```
if c3(l) > 9 & c3(l) <= 18
   c3m2 = c3m2 +1;
end
if c3(l) >= 19
   c3m3 = c3m3 +1;
end
End

% Creation of the matrix with all the clusters
sum_c(1,1,e) = c1m1;
sum_c(2,1,e) = c2m1;
sum_c(3,1,e) = c3m1;
sum_c(1,2,e) = c1m2;
sum_c(2,2,e) = c2m2;
sum_c(3,2,e) = c3m2;
sum_c(1,3,e) = c1m3;
sum_c(2,3,e) = c2m3;
sum_c(3,3,e) = c3m3;
T =
table([c1m1;c2m1;c3m1],[c1m2;c2m2;c3m2],[c1m3;c2m3;c3m3],'VariableNames',{'MPI-ESM1-2.LR','CanESM5','MIROC-
ES2L'},'RowNames',{'Cluster 1','Cluster 2','Cluster 3'});
display(T)
end
end
```