Data Science Lab in Environment

Bettani Elia, matricola 776536

3 September 2019

Project
Analyze and compare two geospatial datasets.
Constraints for the exercise:
- 2 different datasets (e.g. model and observations, n different models, 2 different observational datasets, etc.)
- Perform at least 1 operation in space and 1 operation in time (e.g. aggregation, normalization, resampling, regridding etc.)
- Display data both as function of space and time (time series and maps)
- Use some sort of statistical analysis on the data (arrange the data according to some pdf, hypothesis testing, time series analysis, detection of trends, etc.)

1 Introduction

The World Climate Research Program (WCRP) is an international programme that helps to coordinate climate research all over the world. The aim of the WCRP organization is to facilitates analysis and prediction of Earth system change in order to give support to a wide range of practical applications of direct relevance, benefit and value to society. In particular, the project CMIP (Coupled Model Intercomparison Project) of WCRP started in 1995 and now is in its sixth phase (CMIP6). This project wanted to develop and adopt a common infrastructure for collecting, organizing, and distributing output from models performing common sets of experiments. Connecting to its webpage, CMIP6 provides a free and open source access to stored databases containing the results of different model simulations, concerning various climatic elements. The idea behind the project is to choose two different outputs of model simulations, comparing them and seeing how good each model simulation fits to reality, verifying some simple assumptions that are known to be true in the current state of the art.

2 Datasets

From the open source web platform of CMIP6 two different datasets of simulations that concerned surface temperature were taken into consideration. Both datasets contain time series from January 1850 to December 2014 and cover the whole world. The values contained in both model represent fictitious surface temperatures simulated by

two different models. These models are built taking into account numerous different parameters regarding geosphere and biosphere factors. The first dataset contains the simulations of the model called 'Amon CESM 2', while the second contains the simulations of the model 'Amon MRI-ESM2-0'; to simplify things, we will refer to the first one calling it 'model 1', and to the second one calling it 'model 2'.

3 Types of data

The data are stored in CMIP6 platform in 'nc' file format. Python can open and read these files organizing data into arrays, via suitable libraries. Each array collects the time series of fictitious surface temperatures; so, the inner structure of an nc file allows the compiler to manage three-dimensional data, locating a single observation in space (latitude and longitude), and in time. The granularity of the data analysed was monthly. The datasets analysed, provide simulations at 192 different latitude levels, 288 longitude levels and 1980 (12 months per 165 years) time levels.

4 Workflow

Starting from a fixed set of parameters that affects surface temperatures, models try to replicate and simulate what would happen. Tuning parameters means to create different scenarios. Every model involve an amount of error, even if it is very accurate. Looking back to what they simulate in the past, it is possible to evaluate how reliable and how far from reality a model is. The aim is to consider some statistically significant phenomena observed in reality in the past, and see if models replicate well what was empirically observed.

4.1 Temperatures anomalies are more accentuated at high latitudes far from the equator

Selecting a subset of observation referring to the XXth century, both models point out well this phenomena in the northern hemisphere. In the southern hemisphere, the trend of high and low latitudes are basically the same for both model 1 and model 2.
In the following graphs, in red is shown the trend of temperature anomalies for points in the equatorial band (for both hemisphere); in blue is shown trend of temperature anomalies far from the equator.

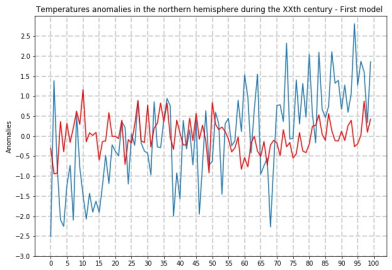

Figure 1: Northern hemisphere, first model

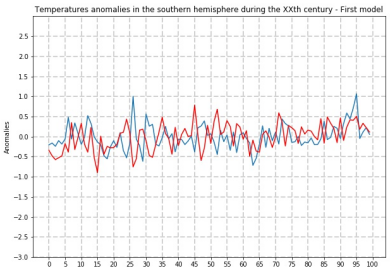

Figure 2: Southern hemisphere, first model

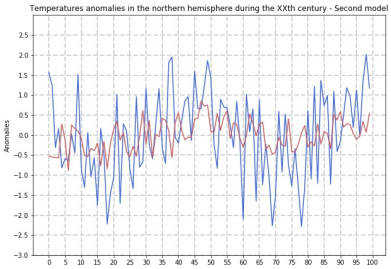

Figure 3: Northern hemisphere, second model

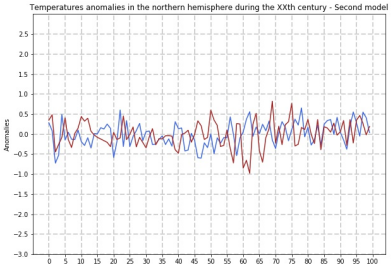

Figure 4: Southern hemisphere, second model

The computation of the anomalies above is built aggregating in the space dimension the observations. If we proceed calculating the mean and the anomalies from a time perspective instead, the situation is heterogeneous for both models and both hemisphere:

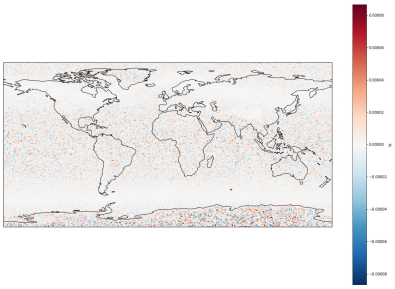

Figure 5: First model

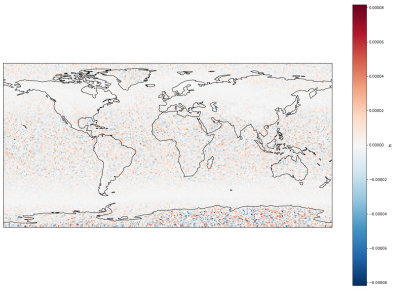

Figure 6: Second model

4.2 The warming is more evident in terrestrial areas than in marine areas

Both model 1 and model 2 replicate well this phenomenon in the subsets of observations considered. The magnitude of average anomalies for terres-

trial stations is significantly higher than in marine stations, as shown in the graphs below; in light blue is possible to find the trend for marine station, while in brown it is highlighted the trend of a subset of terrestrial observations.

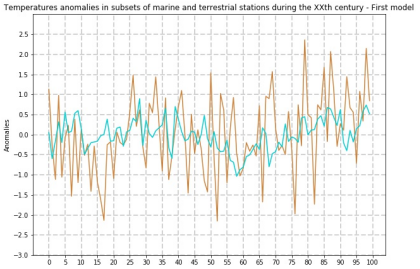

Figure 7: First model

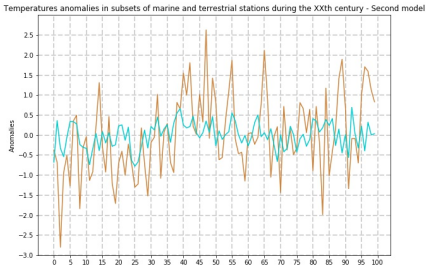

Figure 8: Second model

Focusing on the effective warming of temperatures, it is possible to calculate anomalies of the last 50 years of the XXth century, with the reference of the mean computed for the previous 50 years of the century of interest. It is predictable to find trends shifted below the 0 and with higher slope for the terrestrial subset, but the behaviour of the trend line is peculiar and different (also regarding the comparison between model 1 and 2 on the subset of observations selected).

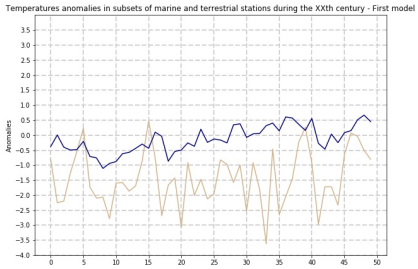

Figure 9: First model

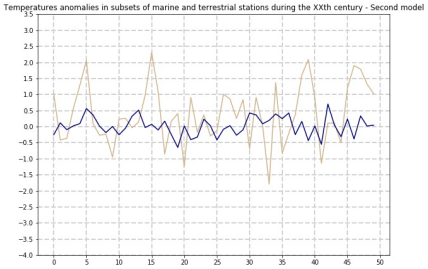

Figure 10: Second model

This could be explained with a distortion in the subset of data analysed due to a selection bias.

4.3 Southern hemisphere (especially antarctic region) suffers mild cooling

Both model anomalies showed to have more accentuated anomalies in the antarctic region comparing to the entire southern hemisphere, confirming what was observed in reality. In the graphs below, the trend of the entire southern hemisphere is highlighted in brown, while the one of the antarctic region is highlighted in light blue.

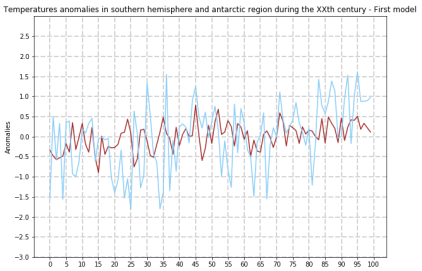

Figure 11: First model

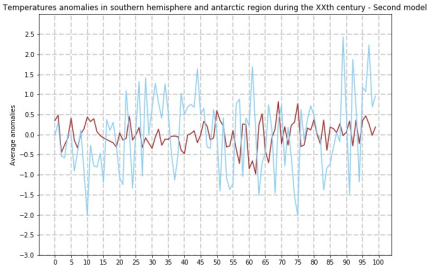

Figure 12: Second model

However, the two model do not replicate the mild cooling observed empirically. In facts, both model 1 and 2 showed a balanced and constant trend along the XXth century, for both areas analysed:

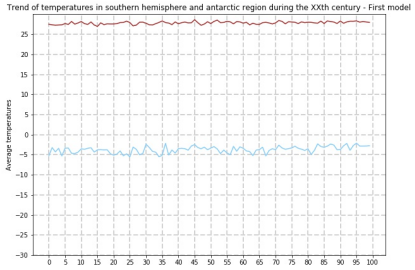

Figure 13: First model

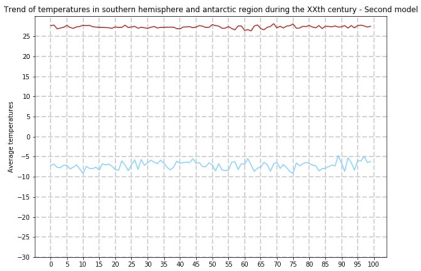

Figure 14: Second model

5 Conclusions

The two model analysed seem to well replicate macro-phenomena as well as inner trends, but not micro-phenomena (not explicit or evident). Unfortunately, it seems that the more the examined phenomenon is hidden, the less the model is able to reproduce this particular trend, as shown in the last example. This remark is logical: no matter how accurate are, models represent always an approximation of a multiform reality.
Finally we need to verify the magnitude of the warming showed in models: in facts at first those model were built in order to replicate and inspect the trend of the overall temperature. For this reason, being able to register the magnitude and the slope of this trend, is the most important requirement to fulfill.
The results are reported in the graphs below .
It is remarkable how the mean temperature per year changes from a model to another (see 1960's for example), but the global balance keep staying very pessimistic; alongside the last century, the average annual temperature is grown by half degree (Celsius).
In the final graph is pretty evident also

how the second model choose a cautious approach. In particular, it is possible to appreciate it looking at the intercepts at the beginning of the trend, fixed under 10.2 in the second model while at 10.8 in the first one. This approach is slightly noticeable in the whole checks made so far, but looking at the global trend this is shown explicitly.

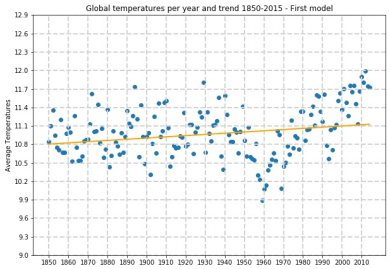

Figure 15: First model

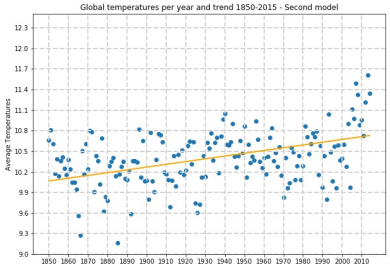

Figure 16: Second model

References

[1] CMIP6 website and open source platform and WCRP website:
https://esgf-node.llnl.gov/projects/cmip6/
https://www.wcrp-climate.org

[2] Dettagli su cambiamento climatico nel secolo scorso:
https://www.wwf.it/il-pianeta/cambiamenti-climatici/
https://www.focus.it/temi/cambiamento-climatico
http://www.nationalgeographic.it/argomento/cambiamento+climatico