The Chimera project
A distributed multi-metric approach for weather forecast ensemble combination

Elvis Cantelli[1], Vitor Marrafon[2], Gustavo Beneduzi[1]
1. Universidade de São Paulo
2. Universidade Federal de Itajubá

The 7th edition of the INFIERI School series, held at USP, August 28 to September 9, 2023

UNIFEI
Universidade Federal de Itajubá - MG

CLIMATEMPO
A StormGeo Company

Introduction

One of the main problems in meteorology is how to combine the different weather forecasts that an ensemble model produces into one single solution. Numerical models that simulate the atmosphere in the entire globe are developed and run every day in various research and commercial centers around the world, assimilating observed data to assimilate into their own physical descriptions.

These forecast models, which include temperature, wind, precipitation and several other variables, often work better under certain atmospheric systems than others, producing a scenario where there is not a model that performs better than the others in a general picture, for instance in the whole region of South America.

The main aim of the project was to develop a composition that is able to produce the best possible weather forecast for the whole continental region, based on the everyday-runs of several weather forecast models available from different centers.

Region classification

The first step was to separate and classify the South America into several pieces with a self-defined climate behavior. For this, we used gridded (0.5° resolution) observed data (max and min daily temperatures, and daily precipitation) from NOAA Climate Prediction Center to compose a daily climatology of the last 7 years. With this, we assumed each pixel as a sample and the timeseries of the temperatures and rain as "dimensions" (features) that compose this climate character to clusterize the pixels into a 2D map of clumps by using the T-SNE (t-distributed stochastic neighbor embedding) algorithm.

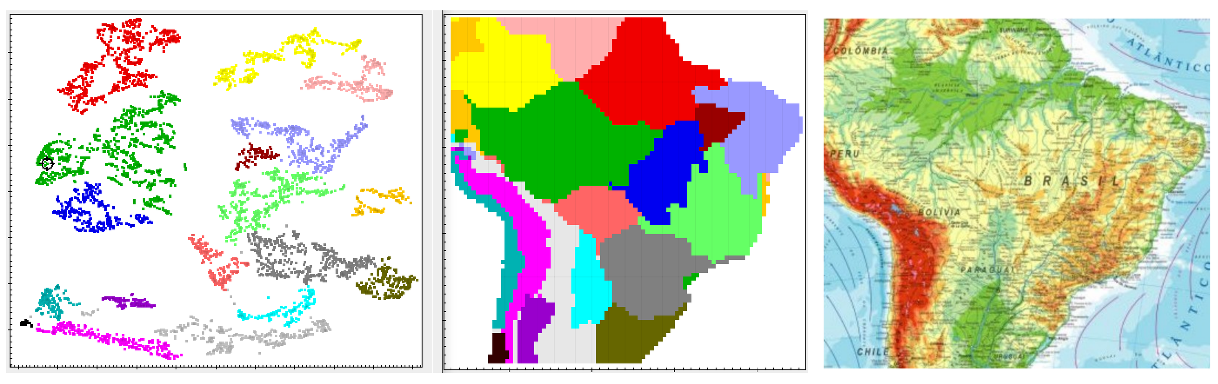

Figure 1: T-SNE manifold (left), correspondent geographic info of the clumps (middle) and landscape map of the brazilian region

As a matter of fact, we were able to reconstruct landscape and biome features only by using climate data. This regionalization aims to analyze the behavior of the models in each type of climate, keeping consistence of regimes that dominate these regions.

Figure 2: Final region configuration in the South America (left) and Brazil (right)

Model evaluation

There are several types of accuracy metrics to evaluate how a model performs when compared to real observed values, and among them are pearson-r, RMSE, R^2, etc. The problem arises from trying to combine these metrics into one to determine the performance taking all of them into account, as errors are not limited, correlations lie between -1 and 1 and so on.

We interpolate grid surface data from a vast list of realtime weather stations to compose our historical observed reference and evaluate the metrics of the model over their daily runs.

Since T-SNE seeks to preserve similarities, we can define a "perfect" model by setting its features as the best metric values (1 for correlations, 0 for errors, etc) and run it together with the models (as the T-SNE samples), so the distance of the models to this "reference model" can be seen as how well the model has performed (Figure 3).

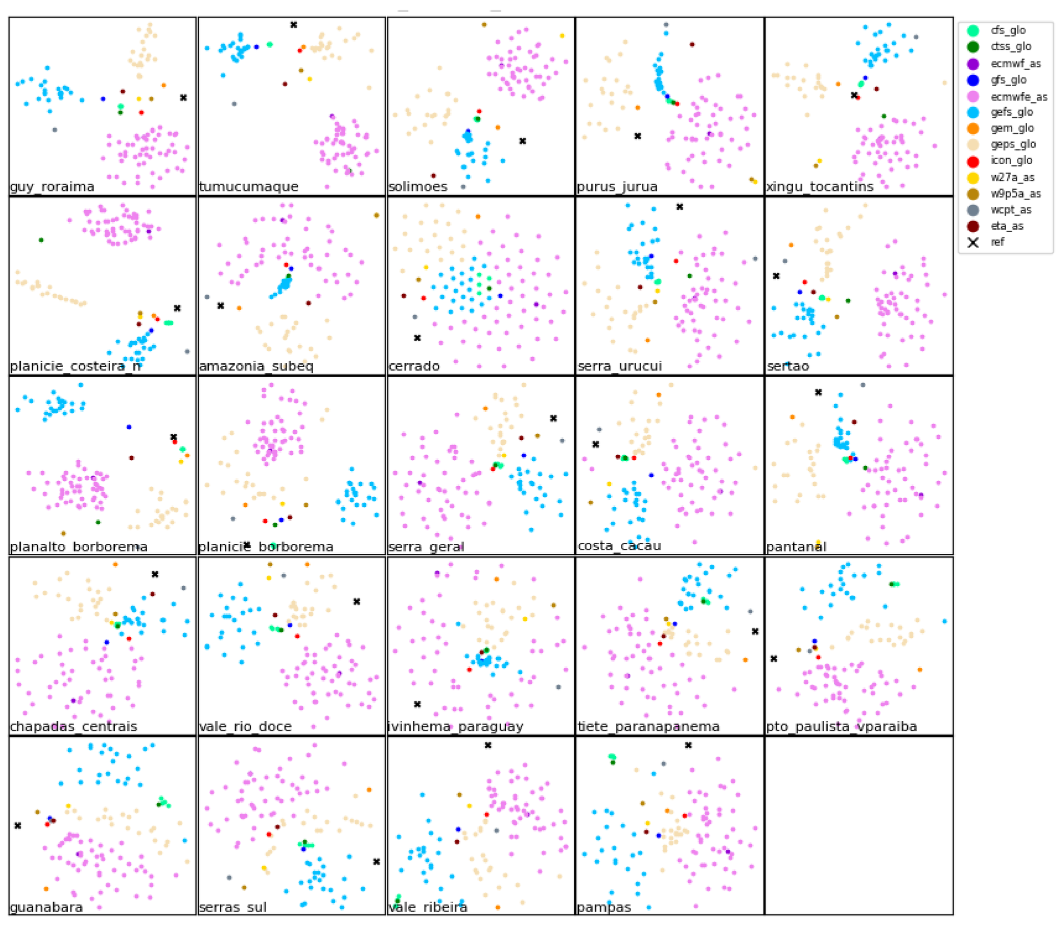

Figure 3: T-SNE maps of the combined model metrics for precipitation in each of the brazilian climate regions. The X mark represents what would be the "perfect model" with the best metrics.

Based on these distances, we compose a rank list with weights for each model to be summed together in a final solution (Figure 4).

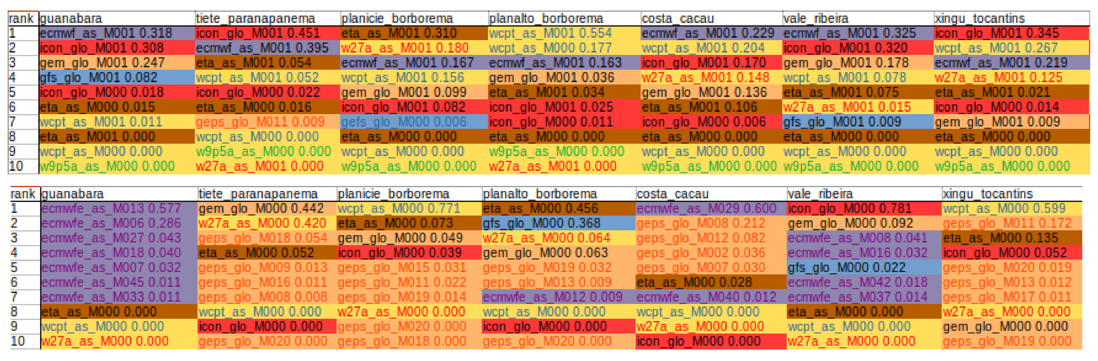

Figure 4: Example of the weight table with the ranked models for each climate region. Upper table contains the temperature rank, and the bottom table contains the precipitation ranks

Temperature bias correction

Gridded models are not able to fully reproduce the weather from seaside cities and locations close to sharp landscape features due to their resolution. Often, seaside city coordinates fall into a pixel that is mostly water, so the temperatures are far from realistic.

To address this problem, we developed 7 different bias correction methods based on observed data, and clustered the bias behavior for each weather forecast model, so we fit the most suitable correction to a given bias pattern. The improvement can be seen in figure 5.

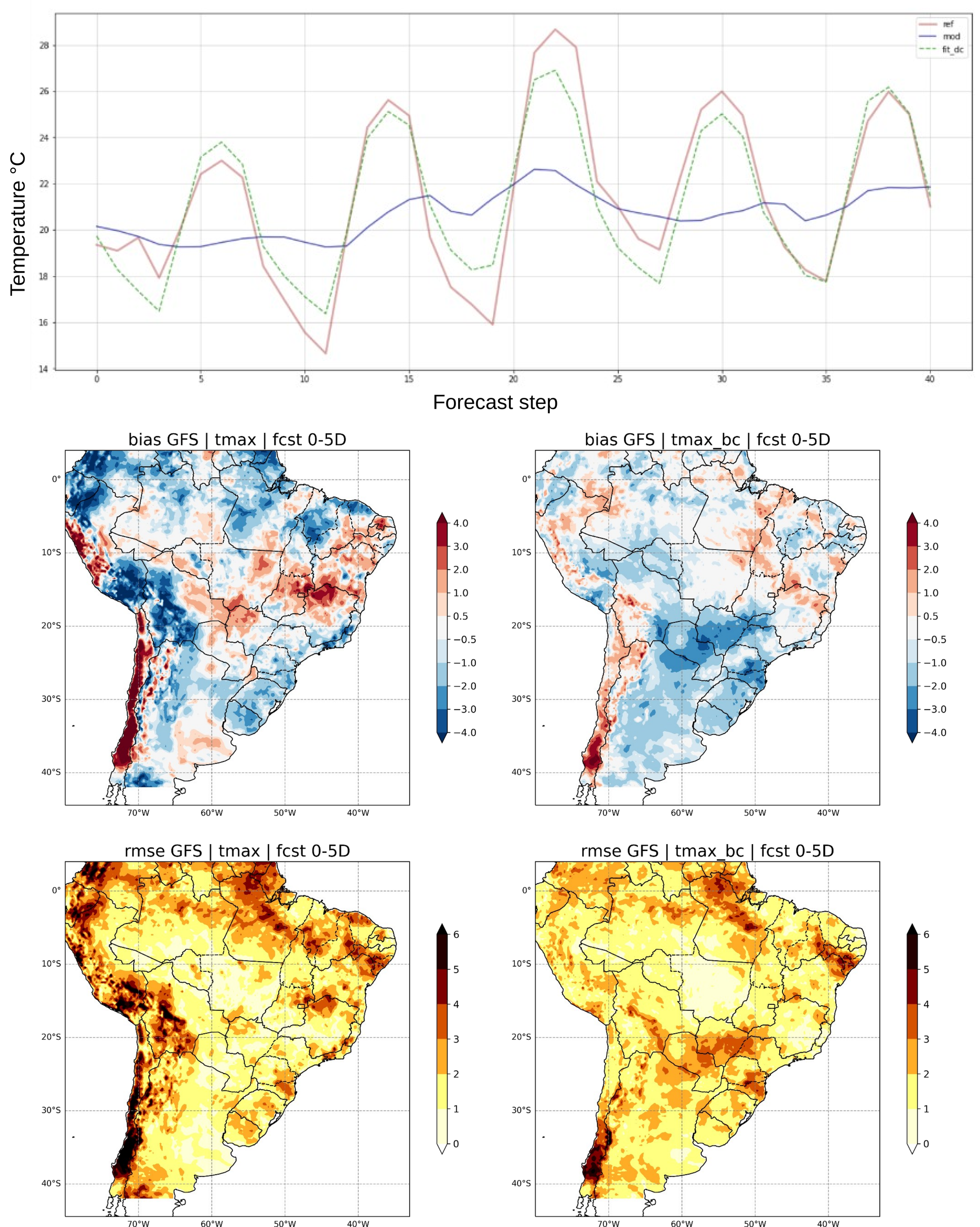

Figure 5: Upper panel shows the temperature timeseries for the Arraial do Cabo – RJ case, where the original GFS model data is the blue line, the observed is the green, and the bias-corrected is the red line. In the bottom panels, we show the overall bias for GFS (left) and the bias-corrected version (right)

Final combination

The final combination extracts the ranked regions for each weather model with shape-based weight maps, use the a digital topography map to bring the values to sea level (in the case of temperatures), and adds them together to finally transform back to the original elevation. For the precipitation, we use a map 3-hour rain probability based on all the models to impose a cutoff below 7%, so we lower the false alarm ratio. The final products are shown in figures 6 and 7.

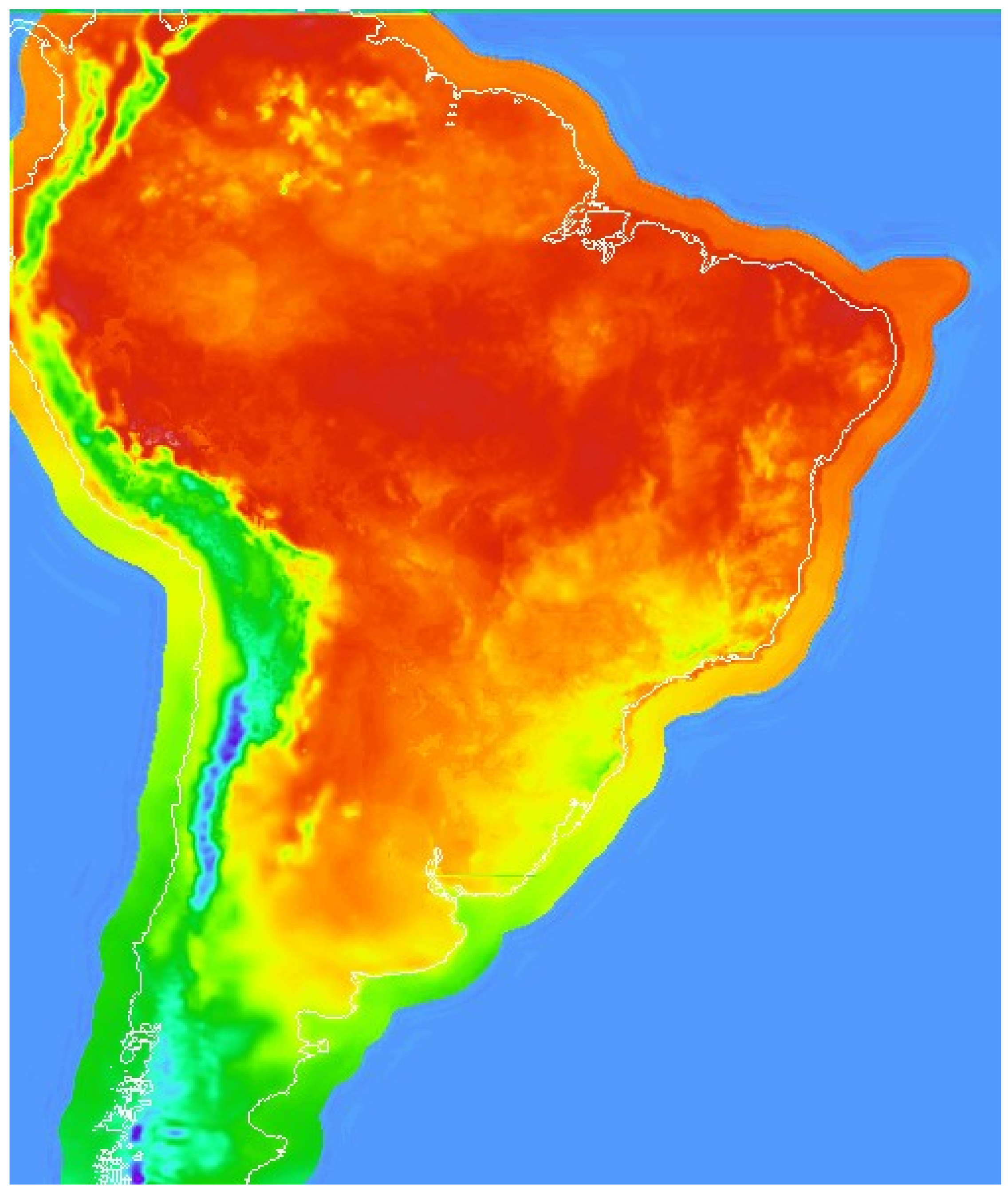

Figure 6: Combined temperature map

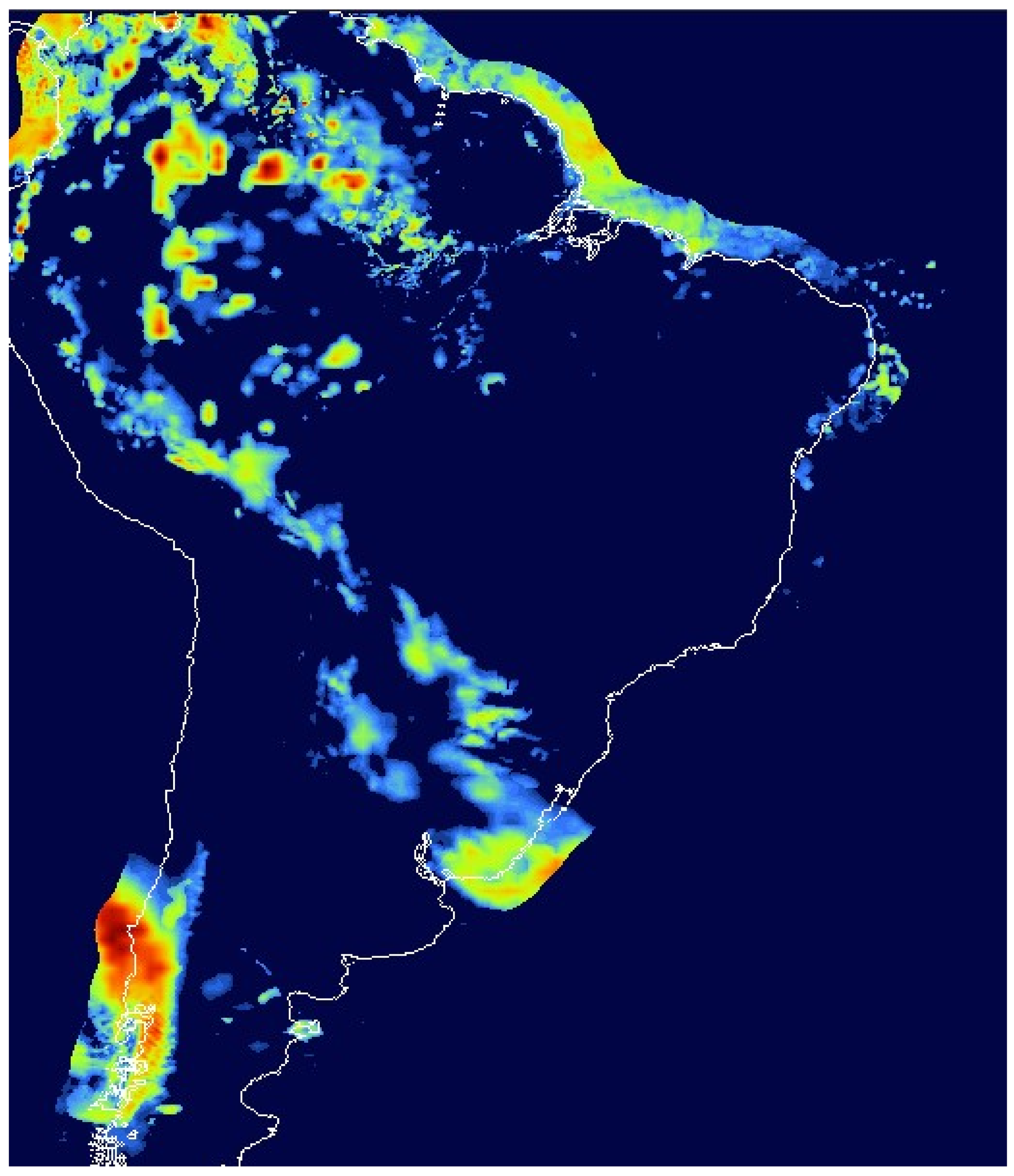

Figure 7: Combined precipitation map

Performance

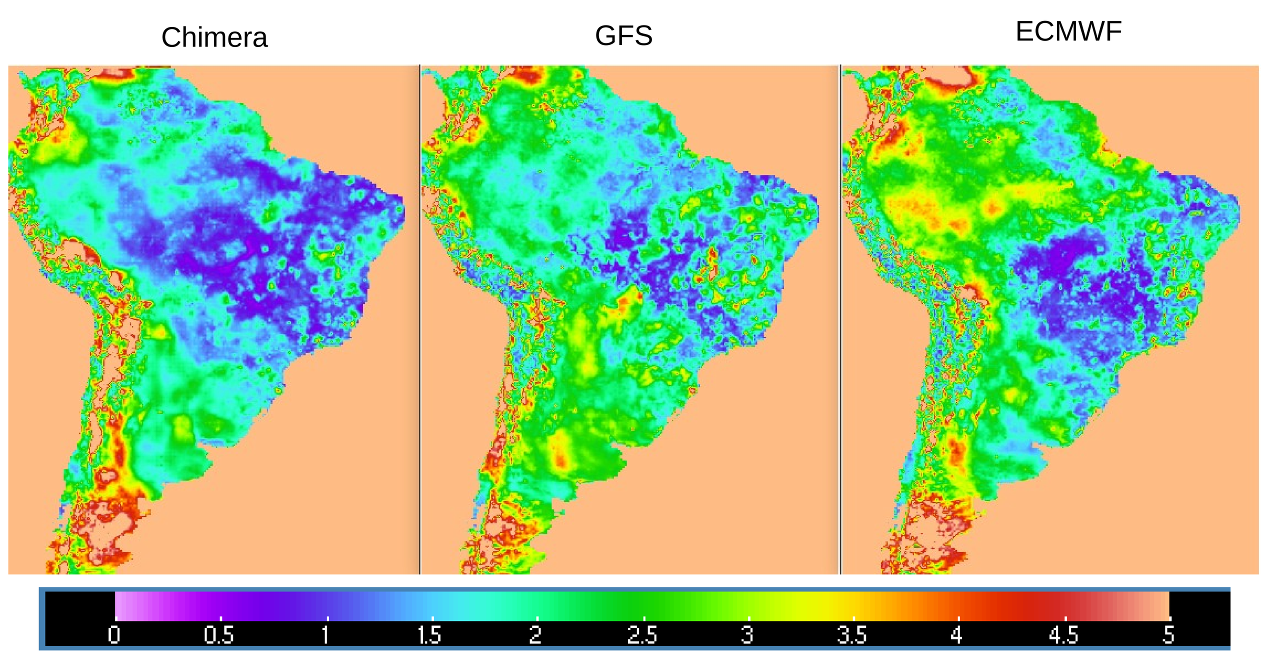

Figure 8: Max temperature RMSE for Chimera, GFS and ECMWF

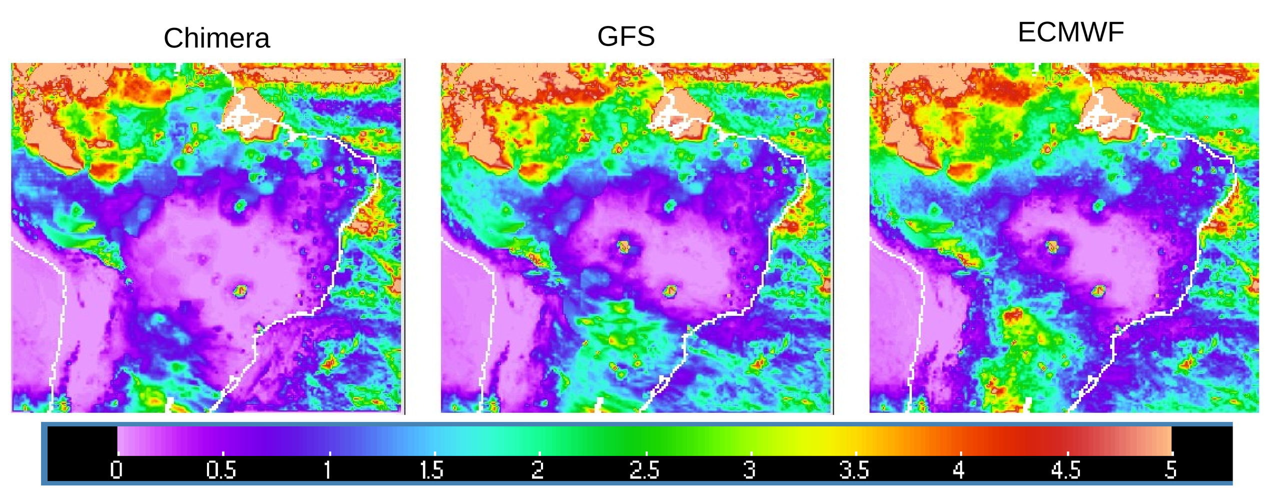

Figure 9: Precipitation RMSE for Chimera, GFS and ECMWF