Assessing several downscaling methods for daily minimum and maximum temperature in a mountainous area. Are we able to statistically simulate a warmer climate in the Pyrenees?

Marc Lemus-Canovas[1] & Swen Brands[2]

[1]Climatology Group, Department of Geography, University of Barcelona, Barcelona, Spain.
[2]MeteoGalicia - Xunta de Galicia, Santiago de Compostela, Spain.

1 Why this study?

The aim of this study is to test the capacity of sevaral perfect prog statistical downscaling (SD) variants to reproduce different statistical climate aspects (*Maraun et al. 2015*) in a mountainous region. This study will allow us to answer the following questions:

- **Can we statistically reproduce the present climate?**
- **Can we reproduce the warmer periods of the current climate** and, therefore, **reproduce a future warming signal**?

2 Downscaling structure

In this study we downscale Maximum (Tx) and minimum (Tn) temperature for the 1981-2015 period. To downscale both variables we used several methods, reanalysis datasets, domains and set of predictors (Fig. 1)

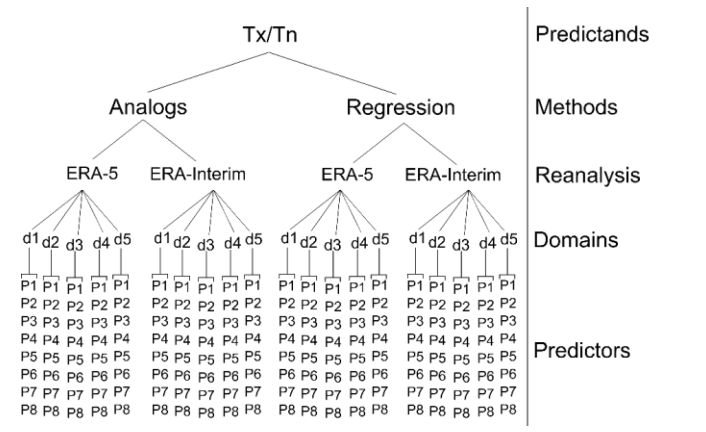

Figure 1. Methodological tree. From Predictands to Predictors.

Regarding the set of predictors, 8 combinations (Table 1) have been used in order to get the best combination to properly downscale the above mentioned statistical aspects. We tried to structure the predictors in: **near-surface** variables (i.e. p1), **middle-high altitude** variables (i.e. p2) and a **mix of both** (i.e. p7).

Predictors	Variables	Predictors	Variables
p1	T2m	p5	SLP + T2m + T850
p2	T850	p6	SLP + T850 + Z500
p3	SLP + T2m	p7	SLP + T2m + T850 + Z500
p4	SLP + T850	p8	SLP + T2m + U850 + V850

Table 1. *Set of predictors.*

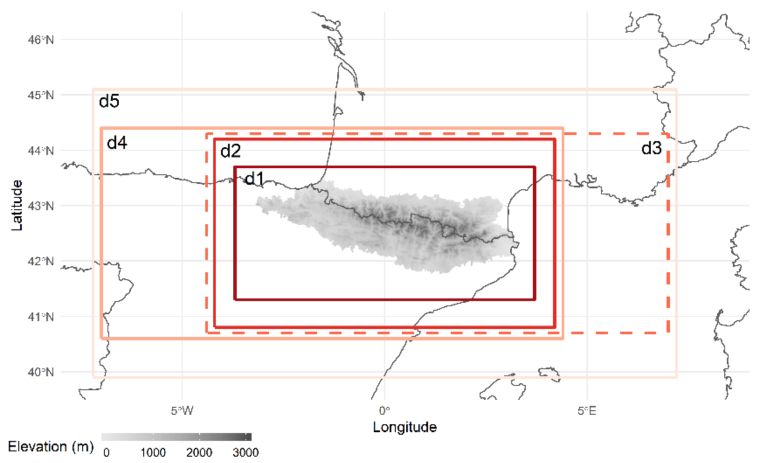

Figure 2. *Study area and set of domains used for each predictor*

3 What and how we want to evaluate?

We use a set of metrics to respond several questions related to the performance of the statistical downscaling. A first group of metrics assess the performance of the **temporal**, **distribution**, **variability** and **trend** aspects; A second group examines the **robustness/stationarity to climate change** and **extreme** conditions. The validation procedure is performed by means of a **K-fold cross validation of 5 folds** containing 7 years of daily data each one. One of this folds contains **the warmest years** in order to test the robustness to climate change when we apply the Warm-test and Extreme-test explained below (Gutiérrez et al. 2013).

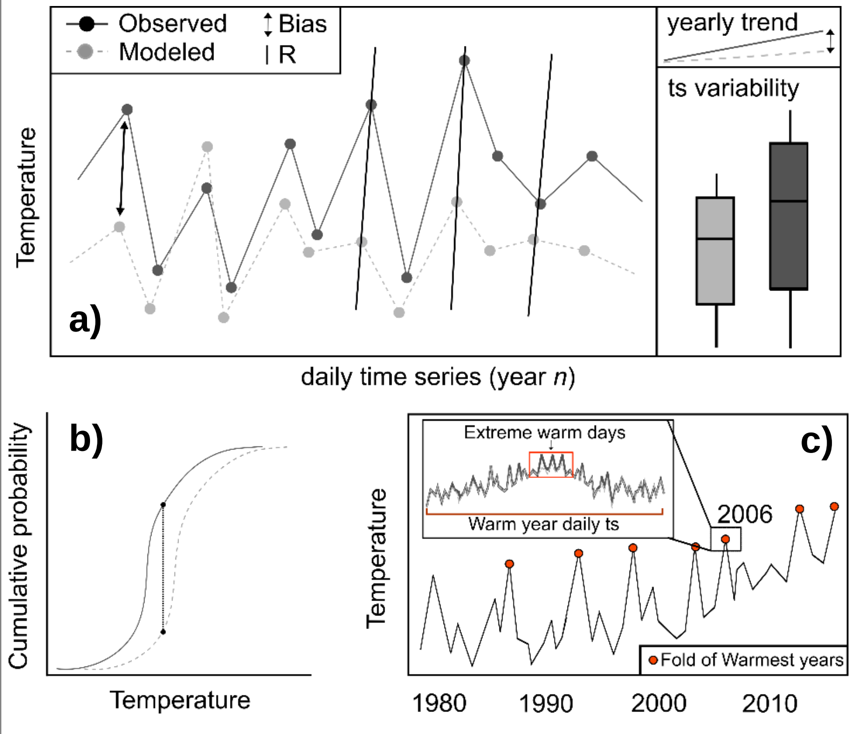

Figure 3 . *Scheme of the statistical aspects asessed.* ***a)*** *centered-left, time series examines temporal Correlation and Bias; top-right, Bias-10yTrend; bottom-right, Bias-CV.* ***b)*** *test differences on distributions (Ks-test).* ***c)*** *warm-test and extreme-warm*

- *Is there a good* ***temporal correlation*** *between observed and modelled?*
  - Evaluation metric: **R**
- *Is there a* ***bias in the magnitud, variability and trend*** *of the modelled data with respect to the observed?*
  - Evaluation metrics: **Bias**, **Bias-CV** (coef. Variation), **Bias-10yTrend** (Sen's slope)
- *Does the prediction reproduce the* ***same statistical distribution*** *as the observed data?*
  - Evaluation metrics: **Kolmogrov-Smirnov Test** (Ks-test)
- *How well can we simulate* ***hot periods and extremes without using them in the training period****?*
  - Evaluation metrics: **Warm-test (Gutierrez et al. 2012); Extreme-test**

Robustness to climate change tests

The **Warm-test** is based on a comparison of biases. The assumption is that the biases (bt) between the mean bias in the warm (bw) fold and the rest of mean biases (bk) of the others folds must be 0. Consequently, in a t-test: H0 ≡ bt = 0 (Pval < 0.05 documents a significant difference of the bias in warm conditions compared to the bias in normal conditions).
The **Extreme-test** is identical to the warm-test but compares the biases of the mean values above the 90th percentile.

4 How well do the downscaling methods perform?

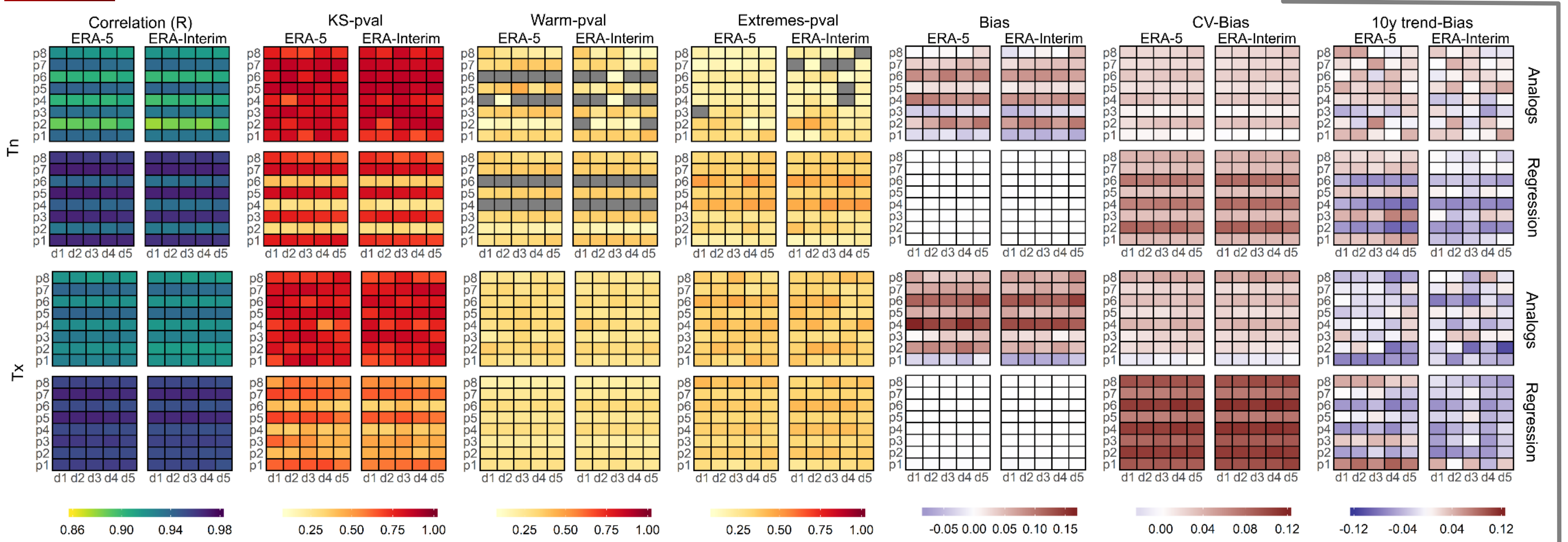

Figure 4. *Heatmaps presenting the performance of the whole combination of predictors and domains for each evaluation metric (columns) and each predictand (left rows) and each method (right rows)*

- **Correlation**: **Regression** performs systematically better than analogs and best results are obtained by combining **surface** and **middle to high-altitude variables**. i.e. **p1, p5, p7**. Both methods perform better for Tx than for Tn.
- **Distribution**: The observed distribution is better reproduced by **analogs** than by regression. In the case of regression, near surface predictors yield best results for this statistical aspect, especially for **Tn**.
- **Warm and extreme test**: The two methods perform similar when reproducing the unconditional mean of the anomalous warm period and, in this aspect, perform **better for Tx** than for Tn. P5 is the best suited predictor combination especially when ERA-5 is considered. The methods capability to reproduce the mean value of the distributions upper tail in the warm period is seemingly larger then for the unconditional mean. This is because the spread of the test distribution fit to the 4 bias values for normal temperature conditions is larger for the upper-tail mean, thereby reducing the power of the test. For Tn regression performs systematically better than analogs.
- **Bias**: Is essentially zero for regression by definition. For analogs, a positive bias is detected for most of predictors.
- **CV-Bias**: For both methods and both predictands, the internal variability is overestimated. A less positive bias Is detected in analogs and for Tn.
- **10y tr-Bias**: The trend over/infra estimation is very slight for both methods and predictands, because in absolute terms is lower than 0.12ºC/10y

5 Mapping the optimal predictor combination...

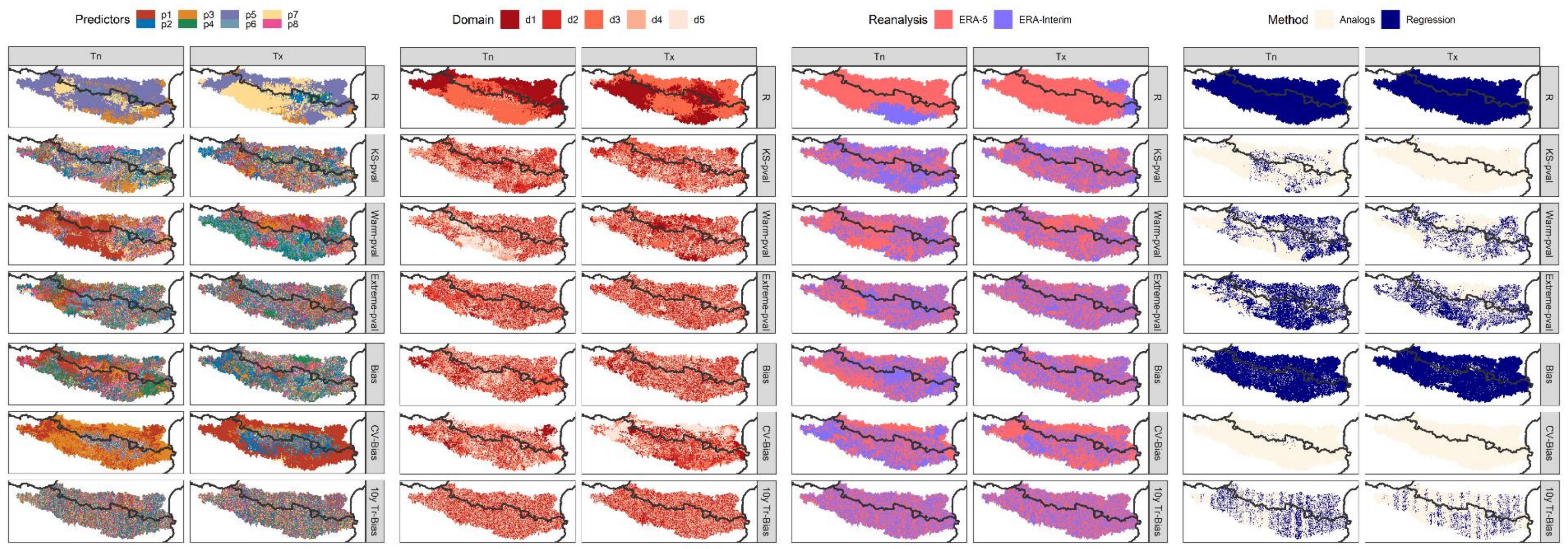

Figure 5. *Best combinations of predictors, domains, reanalysis datasets and methods analysed in the geographical space.*

- **Predictors**: p5, p7 and p3 perform well for **R**; north-south pattern in **Warm-pval** for Tx (north: surface predictors (p1,p3); south (mixed predictors (p4,p6); **Extreme-pval** randomly distributed and non-correlated to Warm-pval results; **CV-bias** clearly depends on elevation (lower areas: p1, p3; higher areas: p2).
- **Domains**: Better R for the **smallest and Mediterranean domains**; Difficult to find geographical patterns for the other metrics.
- **Reanalysis**: **ERA-5** presents **slightly better** results than ERA-Interim, mainly in R, Warm-pval, Extreme-pval and CV-bias
- **Method**: **Analogs** is useful to model statistical **distribution** and **internal variability**. **Regression** performs better for the **temporal correlation** and **Bias**. The **warming** signal is slightly better performed by regression in Tn, the two methods perform similar for Tx.

6 Does the performance depend on elevation?

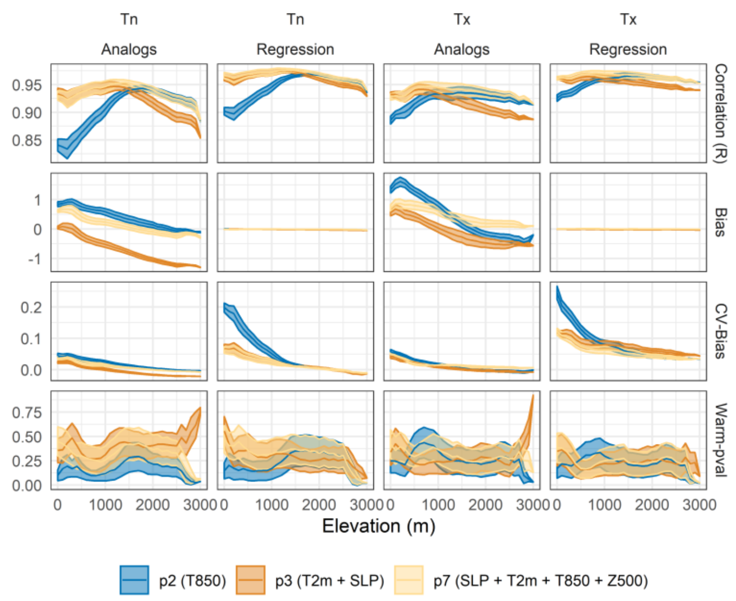

Figure 6. Elevation dependence of different predictor sets and for several evaluation metrics. Results are for the first domain (d1)

- Using **surface variables (p3) improves the result for the lower areas** of the Pyrenees. However, as we increase the altitude, the results are getting worse.
- The **opposite effect is produced by p2 (middle-altitude variable)**, which performs better at intermediate elevations.
- Using a **combined predictor solves the elevation dependence issue**, except for the Warm-pval metric.

Conclusions

1. **Several requirements must be met to assure that PP statistical downscaling models yields reliable results under warmer climate conditions. Fulfilling all of these requirements is a difficult task.**
2. **The decision on which predictor to be used depends on the aim of the study. Which is the best combination for an extremes analysis, variability assessment or future trend analysis?**
3. **This kind of study allows us to check which method is most suitable for use in a warming climate.**
4. **Results indicate that it is straightforward to use predictors on several surface and pressure levels to avoid elevation-dependency in the applied performance metrics.**

Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R., & Herrera, S. (2013). Reassessing statistical downscaling techniques for their robust application under climate change conditions. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-11-00687.1
Maraun, D., Widmann, M., Gutiérrez, J. M., Kotlarski, S., Chandler, R. E., Hertig, E., Wibig, J., Huth, R., & Wilcke, R. A. I. (2015). VALUE: A framework to validate downscaling approaches for climate change studies. *Earth's Future*. https://doi.org/10.1002/2014EF000259

EGU General Assembly 2020

Contact: Marc Lemus-Canovas
mlemus@ub.edu