Proposals of metrics for a more physical evaluation of climate models

M. J. Casado, E. Rodríguez-Camino, A. Pastor, M. C. Sánchez de Cos, J. M. Sánchez-Laulhé, C. Jiménez, and P. Ramos

Agencia Estatal de Meteorología (AEMET), Spain

Sofia, Bulgaria 7-11Sep 2015

WHY AND WHICH MODEL EVALUATION?

- Climate models are the primary tools available for investigating the response of the climate system to various forcings, for making climate predictions on seasonal to decadal time scales and for making projections of future climate over the coming century and beyond (IPCC, 2013). Climate models, which are based on well stablished physical principles, are not perfect, and therefore, a compulsory evaluation of their accuracy is needed before being used for estimating the possible evolution of the Earth´s climate.
- Climate models have been generally evaluated by focusing on their performance on annual, seasonal or monthly means. Nevertheless, **daily scales usually associated with changes at the synoptic scale are likely to be those that most strongly affect human, physical or biological systems** (Perkins et al. 2007).
- As climate models are very complex systems, they have different capabilities and limitations which can be evaluated using a variety of methods and approaches. The model performance metrics are intended to include measures of model performance in presenting mean climate, variability (i.e., ENSO, NAO), and key physical processes (e.g., convection, fluxes).
- Evaluated models able to catch the essential of **synoptic scales –responsible for many meteorological extremes-**,and of the main physical process behind the climate system should be a distinctive feature of future climate services and in particular of the Copernicus Climate Change Service (C3S).
- The main goal of this study is to focus on a better estimation of the **correctness of the coupling between subsystems** of the climate system, on the **proper simulation of weather at synoptic scale** and on the **correct representation of essential variability modes**, for which the following physically based metrics are proposed.

CLIMATE VARIABILITY

Atmospheric dynamics have long been characterized in terms of repeating patterns or cycles. Although the exact timing and magnitude of long-term oscillations in teleconnection patterns is chaotic, pattern statistics do exhibit regular features (Stoner et al., 2009). These patterns are characterized by a quasi-fixed large scale spatial structure and an associated time series that identifies the amplitude and phase evolution of this structure (Wallace and Gutzler, 1981; Barnston and Livezey, 1987).

To evaluate how climate models simulate climate variability, different metrics are proposed checking e.g. positions, amplitudes (Figs. 1 and 2) and phase evolution of the climate variability patterns (Casado and Pastor, 2012).

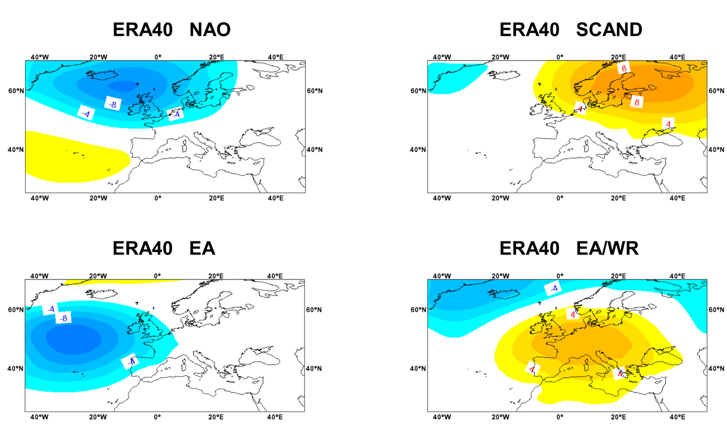

Fig. 1. ERA40 variability modes (Interval 2 hPa). (Casado and Pastor, 2012)

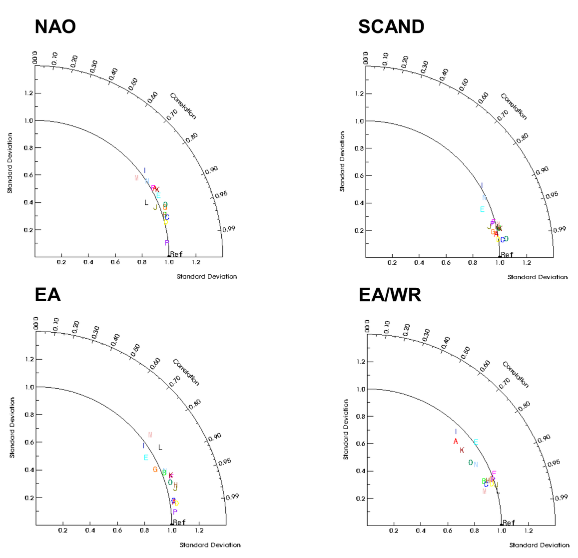

Fig. 2. Taylor diagrams for comparison of spatial patterns between AR4 models (color letters) and ERA40 (black point). (Casado and Pastor, 2012)

CIRCULATION TYPES

The evolution of daily synoptic weather patterns is the main driver of day-to-day weather change. Classifications of circulation regimes at synoptic time scale were introduced as an attempt to link persistent and recurring patterns with synoptic-scale or planetary-scale atmospheric dynamics.

To evaluate how well climate models simulate the daily synoptic patterns, different metrics, are proposed checking e.g. positions and amplitudes of the principal centers of action (Cattiaux et al. (2013), and frequencies, persistences or lifetimes of each circulation type (CT) (Fig. 3) (Pastor and Casado, 2012).

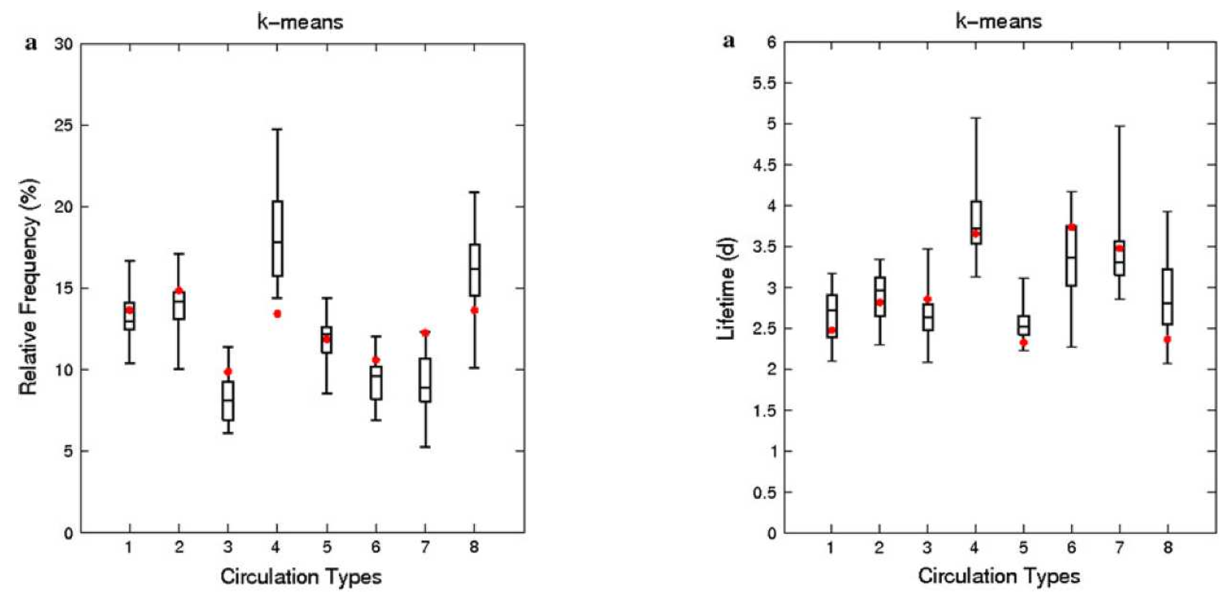

Fig. 3. Box-plots of relative frequencies (%) (left) and the mean lifetime (days) (right) of 8 CTs for 16 AR4 models based on k-means. ERA40 values appear as red dots. (Pastor and Casado, 2012)

FEEDBACK LOOPS

The complex internal feedbacks of the climate system determining its highly non-linear behaviour can either amplify ('positive feedback') or dampen ('negative feedback') the effects of a perturbation in one climate variable. Climate models should be able to simulate the main feedbacks of the system (Flato et al., 2013).

The summer time evolution of the Azores anticyclone and the feedback loop responsible for its dynamics is an example of evaluation of models based on their simulation of main feedback loops (Fig. 4) (Sánchez del Cos et al., 2015).

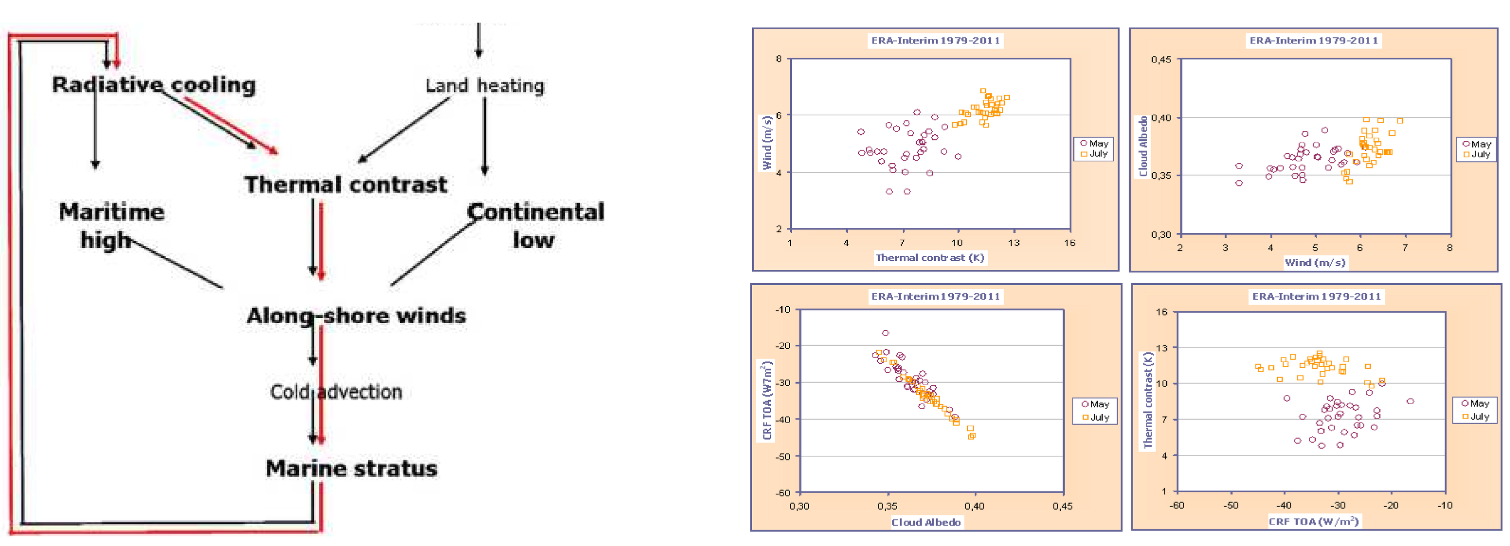

Fig. 4. Schematic diagram showing the feedback associated with the Azores high-pressure system (left) and relationships among pairs of variables over the selected area (ERA-Interim data 1979-2011) for the months of May and July (right). (Sánchez del Cos et al., 2015)

COUPLING BETWEEN COMPONENTS

Land-surface processes and interaction between land-surface and atmosphere are especially relevant for the evaluation of climate models. A novel approach proposed for evaluating regional climate models (RCMs) is based on the comparison of empirical relationships among model outcome variables (Fig.5) (Sánchez et al., 2013).

The similarity of 2D-scattered plots between surface fluxes for RCMs respect to ERA-Interim is estimated using the Hellinger coefficient (Hellinger, 1909).

The Hellinger coefficient was originally designed to estimate the proximity of probability density functions (pdf's). The Hellinger coefficient is defined as:

$$d_{Hell}^{(s)} = \int_R q(x)^s p(x)^{(1-s)} dx$$

where q(x) and p(x) are two pdf's to compare, and s is a parameter (0<s<1). The Hellinger coefficient can be thought of as measure of the ''overlap'' between two distributions. Hellinger coefficient yields information about differences or similarities in relative position, shape and orientation of the pdf's.

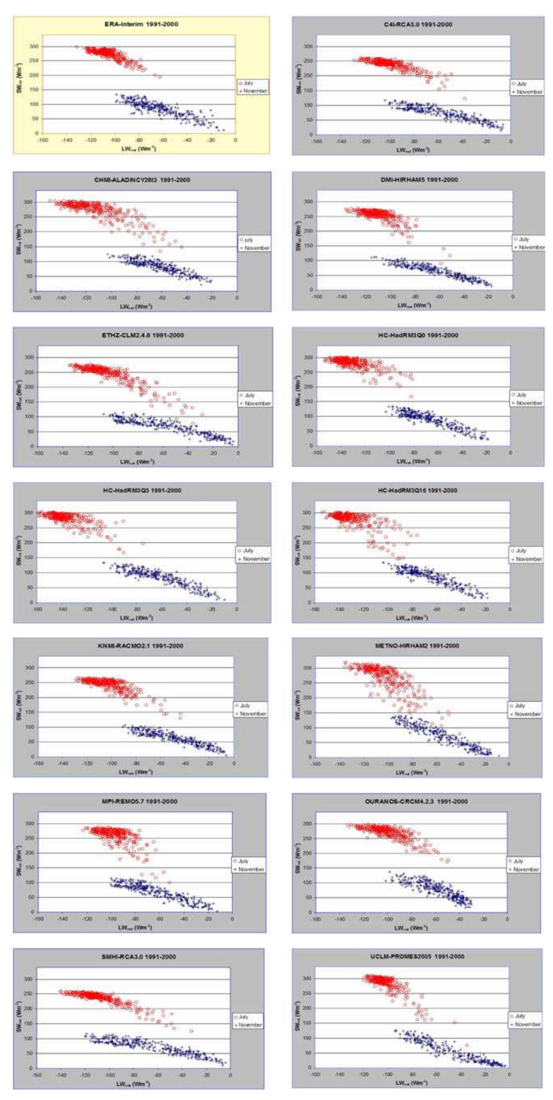

Fig. 5. Scattered plots of Swnet as a function of Lwnet for ERA-Interim and thirteen ENSEMBLES RCMs over the selected area. Red circles and blue crosses correspond to dry (July) and wet (November) seasons, respectively. (Sánchez del Cos et al. 2013).

CONCLUSIONS

- Evaluation of climate models have so far mainly focused on outcome variables (usually temperature and precipitation) disregarding essential aspects as the correctness of the underlying weather simulation. Climate models performance over past and present climate periods should put special emphasis on the introduction of more physically based metrics.
- The models' rankings are highly dependent on the region, variables and metrics selected for the evaluation. Therefore it is advisable the use of as much as possible different evaluation approaches, this would improve our confidence in climate models. Moreover, the choice of the 'best'' model will be strongly dependent on the specific applications designed by users.
- Estimation of the correctness of the coupling between subsystems of the climate system, of the proper simulation of weather at synoptic scale and of the correct representation of essential modes of variability should be incorporated to list of quality control criteria to be met by any climate model selected for C3S.
- As C3S will provide data for a wide variety of sectors –some of them extremely dependent of certain time scales- the correct representation of scales ranging from weather patterns up to the main variability modes affecting Europe climate should be contemplated in the evaluation process of C3S climate models.

REFERENCES

- *Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon Weather Rev 115: 1083–1126*
- *Casado M , Pastor A (2012) Use of variability modes to evaluate AR4 climate models over the Euro-Atlantic region. Clim Dyn 38: 225–337*
- *Cattiaux J, Quesada B, Arakélian A, Codron F, Vautard R and Yiou P (2012) North-Atlantic dynamics and European temperature extremes in the IPSL model: sensitivity to atmospheric resolution. Clim Dyn 40 (9-10): 2293-2310*
- *Flato GJ and co-authors (2013) Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the IPCC*
- *Hellinger E (1909) Neue Begru¨ndung der Theorie quadratischer Formen von unendlich vielen Vera¨nderlichen. J Reine Angew Math 136:210–271*
- *IPCC, Climate Change (2013) The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.*
- Pastor A, Casado M (2012) Use of circulation types classifications to evaluate AR4 climate models over the Euro-Atlantic region. Clim Dyn 39: 2059–207
- *Perkins SE, Pitman AJ, Holbrook NJ, McAneney J (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. J Clim 20:4356–4376*
- *Sanchez de Cos C, Sanchez-Laulhe JM, Jimenez-Alonso C, Sancho-Avila JM, Rodriguez- Camino E (2013) Physically based evaluation of climate models over the Iberian Peninsula. Clim Dyn 40: 1969-1984*
- *Sanchez de Cos C, Sanchez-Laulhe JM, Jimenez-Alonso C, Rodriguez- Camino E (2015) On summer subtropical highs feedback for evaluating climate models (In review)*
- *Stoner K, Hayhoe K, Wuebbles DJ (2009) Assessing general circulation model simulations of atmospheric teleconnection patterns. J Clim 22(16): 4348–4372*
- *Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the Northern hemisphere winter. Mon Weather Rev 109:784–812*