Scale-separation diagnostics and the Symmetric Bounded Efficiency for the inter-comparison of gridded products with different spatial resolutions

B. Casati [1] (barbara.casati@ec.gc.ca), C. Lussana [2], A. Crespi [3]

1= Meteorological Research Division, Environment and Climate Change Canada, Dorval (QC), Canada
2 = Division for Climate Services, the Norwegian Meteorological Institute, Oslo, Norway
3 = Center for Climate Change and Transformation, Eurac Research, Bolzano, Italy

Scale-separation methods are a class of spatial verification methods (Gilleland et al., 2010) which enable i) the comparison of the scale structure of gridded products, such as reanalyses, forecasts from numerical models, and gridded observations; ii) the assessment of bias, error and skill on different scales; and iii) the analysis of the scale dependency of forecast predictability. In this study we apply novel scale-separation diagnostics and introduce the Symmetric and Bounded Efficiency for the comparison of precipitation reanalyses with different spatial resolutions. The COSMO-REA6 reanalysis (6km resolution) is compared against the ERA5 reanalysis control member (HRES, 31km resolution) and 10 ensemble members (EDA, 62km resolution). 24-hour accumulated precipitation fields are decomposed into the sum of components on different spatial scales by using a 2D Haar wavelet filter. The separate scale components are then compared by using the continuous verification statistics listed in Table 1.

$En_X^2 = \mu_{X^2} = \sigma_X^2 + \mu_X^2;$	$En_Y^2 = \mu_{Y^2} = \sigma_Y^2 + \mu_Y^2$	$MSE = (\mu_Y - \mu_X)^2 + \sigma_Y^2 + \sigma_X^2 - 2\sigma_Y\sigma_X r_{Y,X}$	
$NB_\mu = \frac{\mu_Y - \mu_X}{\lvert\mu_Y\rvert + \lvert\mu_X\rvert};$	$NB_\sigma = \frac{\sigma_Y - \sigma_X}{\sigma_Y + \sigma_X}$	$NSE = 1 - \frac{MSE}{\sigma_X^2};$	$SS_{rand} = 1 - \frac{MSE}{(\mu_Y - \mu_X)^2 + \sigma_Y^2 + \sigma_X^2}$
$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_Y}{\sigma_X} - 1\right)^2 + \left(\frac{\mu_Y}{\mu_X} - 1\right)^2}$		$SBE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_Y - \sigma_X}{\sigma_Y + \sigma_X}\right)^2 + \left(\frac{\mu_Y - \mu_X}{\lvert\mu_Y\rvert + \lvert\mu_X\rvert}\right)^2}$	

Table 1: scale-separation diagnostics, where *X* and *Y* indicate the gridded observation and forecast compared, μ the mean, σ the standard deviation, *r* the correlation, following conventional statistical notation.

The scale-separation diagnostics are illustrated in Figure 1. The energy (En^2) and its square-root (En, Fig.1a) are proportional to magnitude and number of precipitation features on each scale. Comparison of the energies informs on biases on different scales and on the scale structure. All products show that the dominant precipitation features have scales ranging between 200-400km. As expected, the COSMO-REA6 (in grey) has larger energy, much more small-to-medium scale features, than the ERA5 products. The ERA5 control member (in red) has slightly more small-scale details than the lower resolution ensemble members (in blue). The energies on the different scales are compared by the Normalized Bias (NB), which is the difference between the two energies divided by their sum. For the wavelet components (which have null spatial average $\mu_X = 0$) the energy En_X^2 is the variance σ_X^2, whereas for the largest scale component (which is a constant field with $\sigma_X^2 = 0$) the En_X is the field spatial mean μ_X. Hence, the NB comparing square-root energies is the NB comparing the scale component standard deviations and the field spatial averages (Table 1). Normalized (and bounded) statistics facilitate the comparison and aggregation of verification results, e.g. for sites with different climatologies. The square-root energy NB is bounded (it ranges between -1 and 1) and symmetric (its absolute value does not change if we swap the two products compared): these two key properties enable the definition of the Symmetric and Bounded Efficiency.

While assessing the Mean Squared Error (MSE) and correlation on different scales (Fig.1b,c), the correlation separates the performance of the ERA5 products, whereas the MSE does not. In fact, the MSE highly depends on the variability of the products compared, both observation and forecast variabilities (Table 1), so that higher resolution forecasts (which have higher variability) tend to be penalized more heavily with respect to coarser/smoother forecasts. The reduction of variance, also known as Nash Sutcliffe Efficiency (NSE), is the MSE skill score against sample climatology ($MSE_{clim} = \sigma_X^2$), and it was introduced to reduce the effect of the variability while assessing the forecast performance. However, the NSE compensates solely for the observed variability (the MSE is normalized by the observed variance only). Then, the MSE shortcoming when comparing forecasts with different resolutions is still not addressed by the NSE: the higher resolution/noisier forecast is still penalized more heavily compared to the coarser/smoother one, and also the NSE cannot separate the performance of the two ERA5 products (Fig.1d). The MSE skill score against random chance (SS_{rand}, Table 1) normalizes the MSE by the variances of both observation and forecast products (the forecast variability is factored in), and is therefore capable of separating the performance of the ERA5 products (Fig.1e). The SS_{rand} allows a fairer comparison of products with different resolutions and enables the assessment of the added value of increasing resolution.

The Kling and Gupta Efficiency (KGE) combines in a single performance measure the comparison of forecast and observation variances and averages, and the two products correlation. The KGE also allows a fair comparison of forecasts with different resolutions, and is capable of separating the performance of the ERA5 products (Fig.8 of Casati et al., 2023). However, the KGE compares variances and averages by their ratio, which renders the KGE asymmetric and unbounded (the score can attain very large negative values e.g. when comparing a smooth forecast versus a noisy observation). The Symmetric and Bounded Efficiency (SBE) is then defined as the KGE, but the variances and averages are compared by the NB (Table 1). The SBE is then bounded and capable of separating the performance of the ERA5 products (Fig.1f). Both SS_{rand} and SBE are symmetric, hence they are invariant with-respect-to the definition of observation and forecast in the comparison of the two products, whereas NSE and KGE are not symmetric, and their values (as well as the no-skill to skill transition scale) change if the order of comparison of the two products is swapped.

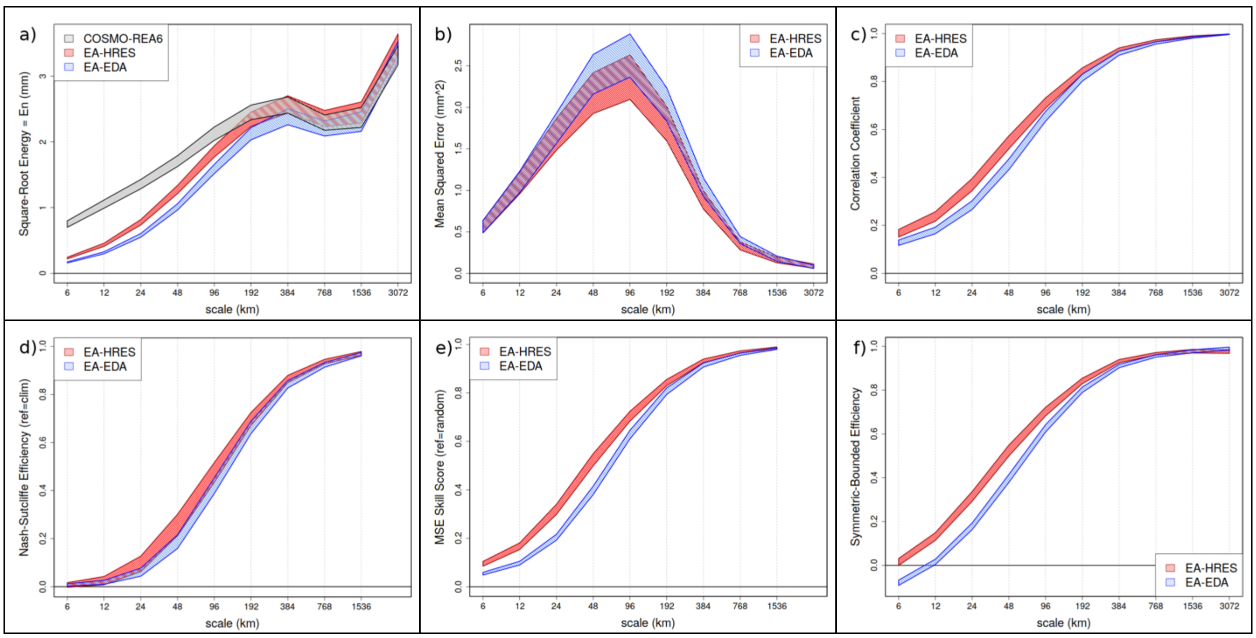

Figure 1: scale-separation diagnostics comparing the ERA5 products against COSMO-REA6. The shading shows the bootstrap 90% confidence intervals of the aggregated statistics on 50 intense precipitation case studies.

The scale-separation diagnostics in Figure 1 consistently show that the reanalyses are in strong agreement on large scales, and exhibit weaker agreement on (less predictable) small scales. The stronger agreement of HRES with COSMO-REA6 (compared to EDA) is due to a more similar representation of the variability at small-to-medium scales, as well as a better linear dependence (correlation). On the largest scale, on the other hand, HRES is slightly underperforming EDA due to over-forecast bias.

References

Casati, B., C. Lussana, and A. Crespi (2023) Scale-separation diagnostics and the Symmetric Bounded Efficiency for the inter-comparison of precipitation reanalyses. International Journal of Climatology, 43, 2287-2304.

Gilleland, E., Ahijevych, D. A., Brown, B. G. and Ebert, E. E. (2010) Verifying forecasts spatially. Bulletin of the American Meteorological Society, 91, 1365–1376.