An intensity-scale skill score to assess the added value of enhanced resolution

B. Casati, A. Glazer, J. Milbrandt, V.Fortin Meteorological Research Division, Environment Canada

Corresponding author: barbara.casati@ec.gc.ca

Motivation: enhanced resolution NWP systems produce precipitation forecast fields enriched of realistic small-scale details. Traditional verification approaches however fail to detect the added value of the enhanced resolution, possibly due to the higher variability and small timing and location displacements.

Aim: develop a scale-separation verification approach which enables to compare the performance (bias, error and skill) of coarse versus high resolution forecasts. We wish to extract the added value of the enhanced resolution!

Data: we test the new scale-separation verification on the MesoVICT (http://www.ral.ucar.edu/projects/icp/) case studies. We verify the Environment Canada GEM (Global Environmental Multi-scale) model, run in three configurations:

- GEMglb = global, ~ 33km
- GEMreg = LAM regional, ~ 15km
- GEMhres = LAM high-res, ~ 2.5km

against the VERA analysis ~ 8km.

Fig.s 2 to 7 show verification statistics evaluated for the MesoVICT core case study illustrated in Fig. 1.

Fig. 3. MSE and energy of the scale components of the binary images

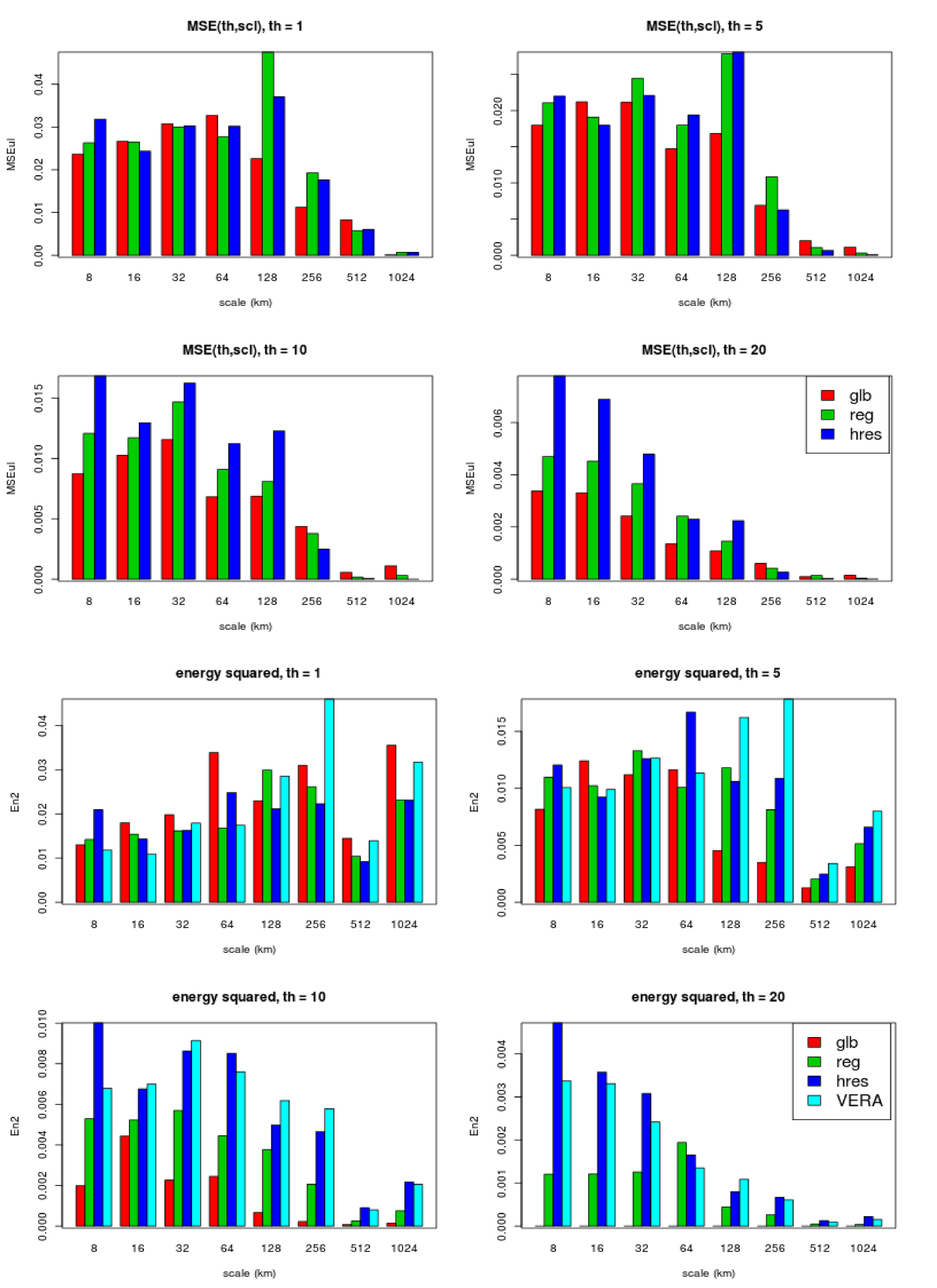


Fig. 6. *MSEu,l,random* and Intensity-Scale skill score evaluated as in 11.

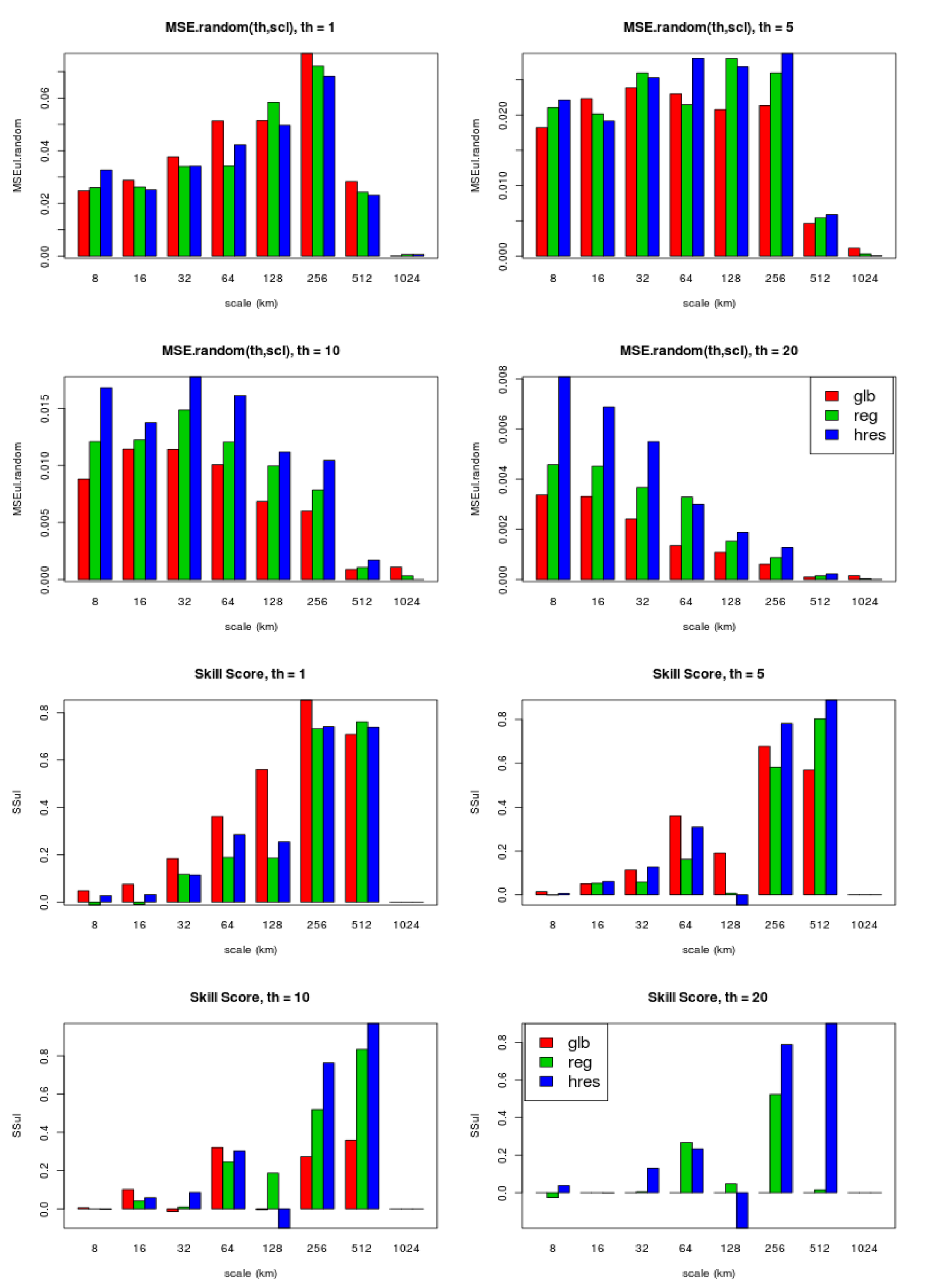


Fig. 1. MesoVICT core case study

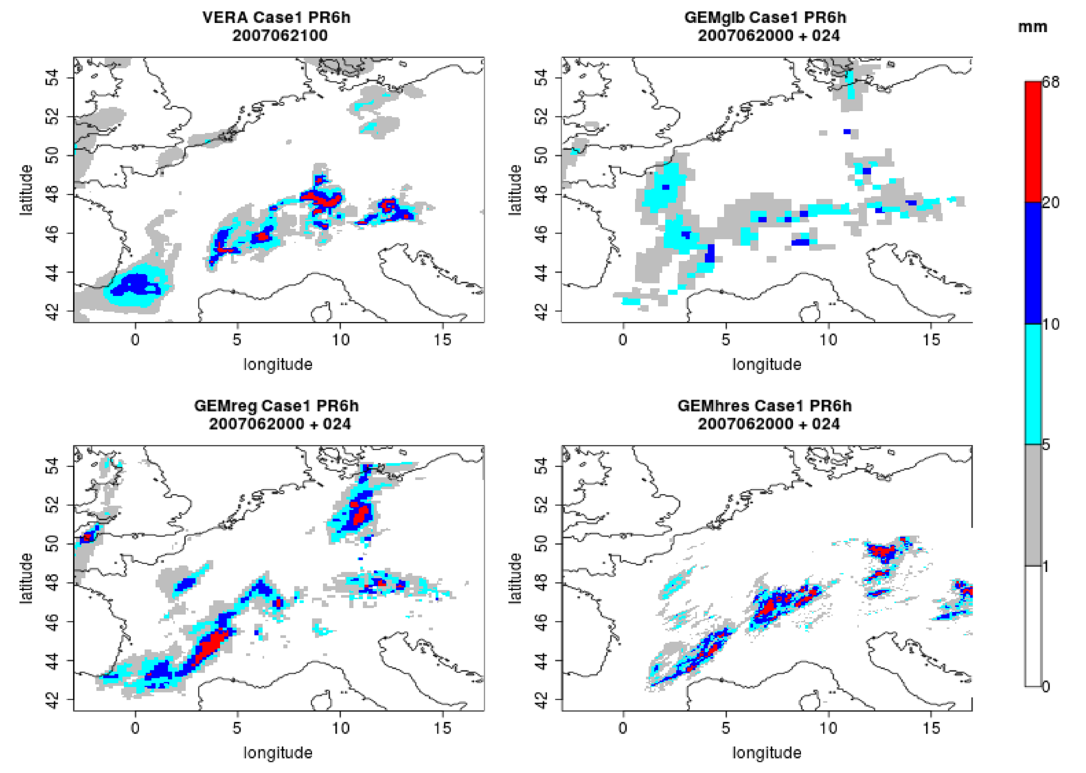


Fig. 2. The wavelet-based scale-separation MSE skill score

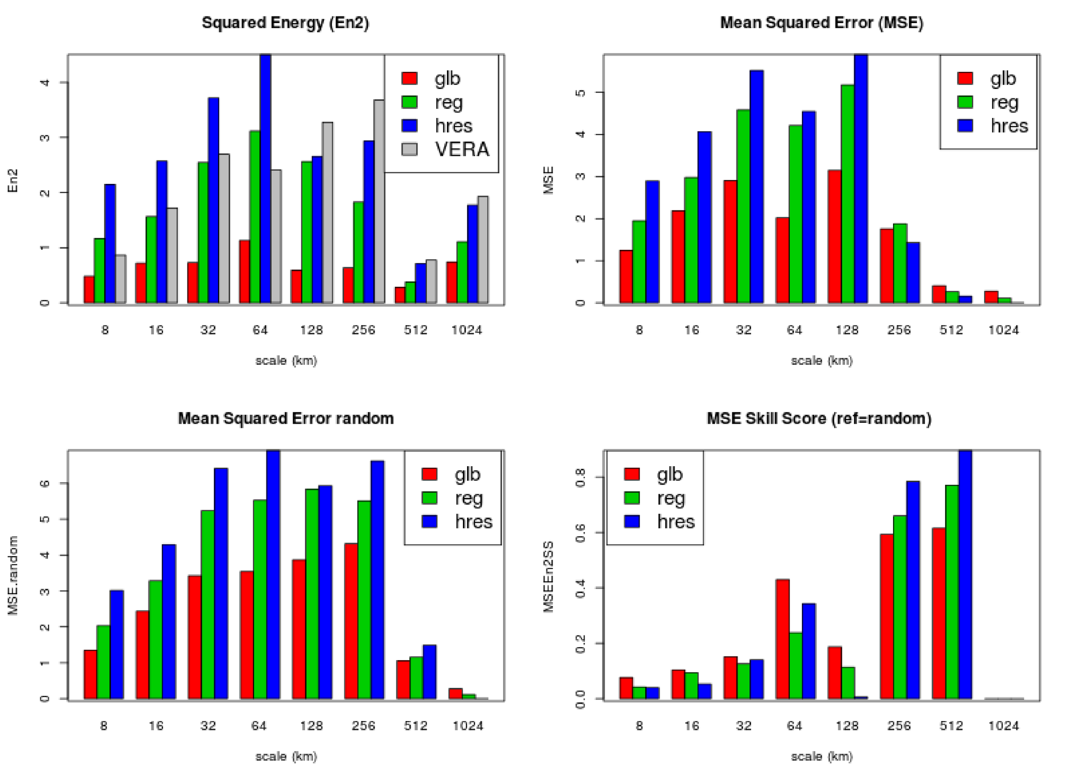


Fig. 4. Contingency table statistics

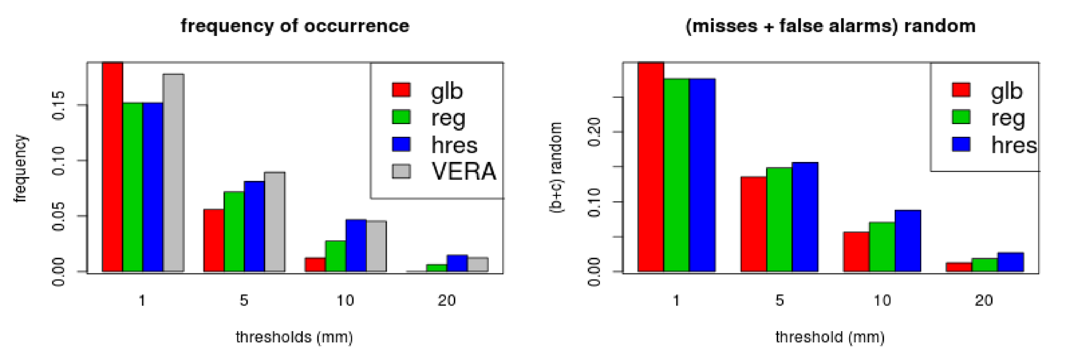


Fig. 5. Intensity-Scale skill score

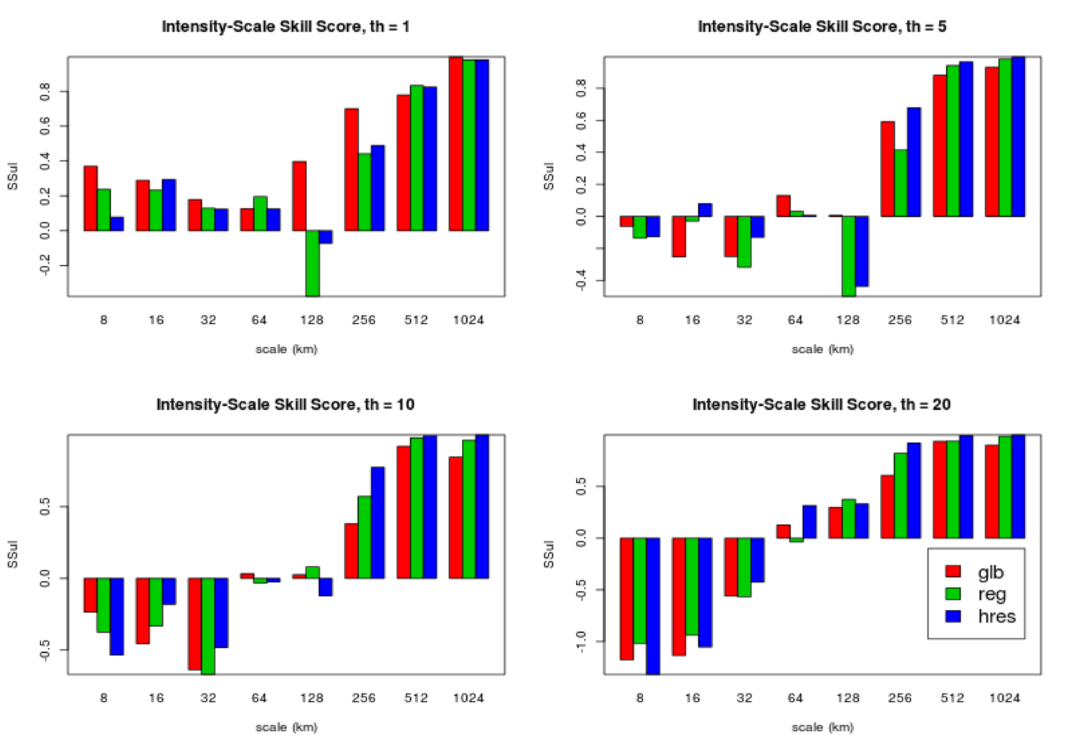
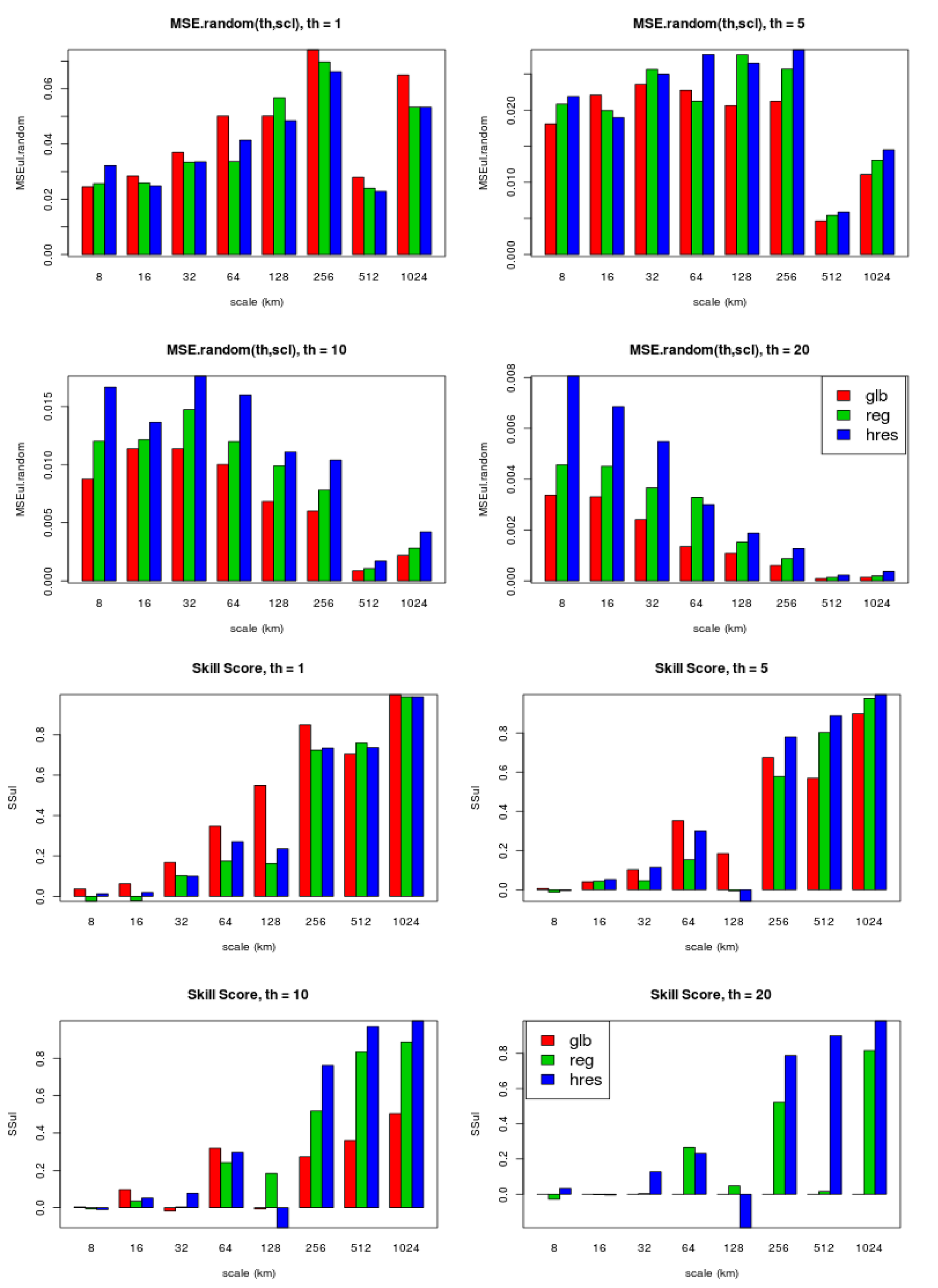


Fig. 7. *MSEu,l,random* and Intensity-Scale skill score evaluated as in 12.



The wavelet-based scale-separation MSE skill score

1. Forecast (*Y*) and obs (*X*) are decomposed into the sum of components on different scales by using 2D discrete Haar wavelet transforms:

$$X = \sum_{j=1}^{J} W_j^m(X) + W_J^f(X)$$

2. The Mean Squared Error (*MSE*) and the forecast and obs energy (*En*^{2}) are then evaluated for each scale component:

$$\mathrm{MSE}(Y,X) = \overline{(Y-X)^2}; \quad En^2(X) = \overline{X^2}$$

3. Note that:

$$\mathrm{MSE}(Y,X) = \overline{(Y-X)}^2 + \sigma_Y^2 + \sigma_X^2 - 2\sigma_Y\sigma_X r_{Y,X}$$

$$\overline{W_j^m(X)} = 0 \quad \Rightarrow \quad \sigma^2_{W_j^m(X)} = En^2(W_j^m(X))$$

$$\overline{W_J^f(X)} = \overline{X} \quad \Rightarrow \quad \sigma^2_{W_J^f(X)} = 0$$

4. Therefore, the MSE for random forecast on each scale components is given by:

$$\mathrm{MSE}_{random}(W_j^m(Y), W_j^m(X)) = En^2(W_j^m(Y)) + En^2(W_j^m(X))$$

$$\mathrm{MSE}_{random}(W_J^f(Y), W_J^f(X)) = \overline{(Y-X)}^2$$

5. The MSE skill score (reference=random) is finally evaluated on each scale component as:

$$\mathrm{SS} = 1 - \mathrm{MSE}/\mathrm{MSE}_{random}$$

Note 1: the reference forecast accounts for the forecast variability, therefore the skill score is suitable for comparing forecasts with different resolutions.

Note 2: *MSErandom* is distributed across the scales in proportion to the energy (i.e. number of events and magnitude of the signal on each scale).

The Intensity-Scale skill score

Casati,Ross,Stephenson (2004) Met.Apps. **11**
Casati (2010) Wea&For **25**

6. Intensity: Forecast and obs are transformed into binary 0/1 images (*Yu,Xu*) by thresholding the precipitation intensity.

7. Scale: The binary forecast and obs are decomposed into the sum of components on different scales by using discrete Haar wavelet transforms.

8. For each threshold *u* and spatial scale *l*, the Mean Squared Error (*MSEu,l*) and energy (*En*^{2}*u,l*) of the scale components of the binary images are evaluated (Fig. 3).

9. In virtue of the thresholding, the IS statistics are related to the contingency table entries (Fig. 4):

$$\mathrm{MSE}_u = \frac{b+c}{n}; \quad En^2(Y_u) = \frac{a+b}{n}; \quad En^2(X_u) = \frac{a+c}{n}$$

$$\mathrm{MSE}_{u,random} = En^2(Y_u)\,(1-En^2(X_u)) + En^2(X_u)\,(1-En^2(Y_u))$$

10. The Intensity-Scale skill score (Fig. 5) is calculated as:

$$\mathrm{IS}_{u,l} = 1 - \mathrm{MSE}_{u,l}/\mathrm{MSE}_{u,l,random}$$

where *MSEu,l,random* is the *MSEu,random* equipartitioned across the scales.

Issue: in real practice, the MSE for a random forecast is not equipartitioned across the scales (it should be proportional to the number of events on each scale).

Proposal: two re-formulation of the *MSEu,l,random* and Intensity-Scale skill score are tested:

11. *MSEu,l,random* is re-defined as in 4:

$$\mathrm{MSE}_{u,l,random} = En^2_{u,l}(Y) + En^2_{u,l}(X) \quad \text{for } W_l^m(\cdot)$$

$$\mathrm{MSE}_{u,L,random} = \overline{(Y_u - X_u)}^2 \quad \text{for } W_L^f(\cdot)$$

12. *MSEu,l,random* is re-defined as in 9:

$$\mathrm{MSE}_{u,l,random} = En^2_{u,l}(Y)(1-En^2_{u,l}(X)) + En^2_{u,l}(X)(1-En^2_{u,l}(Y))$$

the Intensity-Scale skill score is then re-evaluated as in 10.

Results: the two re-formulation of the Intensity-Scale skill score leads to very similar results (compare Fig.s 6 and 7). We favor definition 11 for its statistical meaning and flexibility (it can be applied to thresholded as well as to continuous fields).

Conclusions

The wavelet-based scale-separation MSE skill score and scale-separation statistics defined in equations 1-5 are informative on the forecast bias, error and skill on different scales, and are suitable for comparing models with different resolutions. These can be evaluated also on thresholded binary fields to attain informations on the forecast performance for specific intensity events.