# Downscaling of ECMWF seasonal integrations by RegCM

**Čedo Branković and Mirta Patarčić**

*Croatian Meteorological and Hydrological Service (DHMZ)*
*Grič 3, 10000 Zagreb, Croatia*
*cedo.brankovic@cirus.dhz.hr*

**ABSTRACT**

The 50-km RegCM was used to dynamically downscale ECMWF operational and experimental seasonal integrations over the southern Europe/Mediterranean area. For operational seasonal forecasts, a major disadvantage was found to be a relative scarcity of ECMWF archived data available for downscaling: too few upper-air levels at 12-hour frequency. Also, downscaling of the whole ensemble (40 members) may not always be practical; this implies designing an algorithm to select sub-ensembles. For experimental integrations (from the ENSEMBLES project), the input for downscaling was much better, but the size of ensembles (9 members) could be considered too small. When compared against gridded CRU data, an overall better representation of T2m and precipitation was obtained in RegCM. Although objective skill scores indicate no clear winner, RegCM tends to be better for higher thresholds (severe events) and over the mountains.

## 1. Introduction

Seasonal forecasts produced by global models cannot give sufficient details required at local spatial and temporal scales. This is particularly crucial for surface fields that are closely linked to orography, like precipitation or surface temperature. Fig.1 illustrates this for the 11-year (1991-2001) JAS (July to September) average: ensemble mean precipitation from ECMWF experimental seasonal forecasts at approximately 1.875° resolution is compared with CRU[1] data on 0.5×0.5° grid. Although the observed 11-year averages for Zagreb (dark blue star, 2.9 mmday$^{-1}$) and Split (1.5 mmday$^{-1}$) fit reasonably well within the ECMWF field, the global model is generally drier than CRU, in particular over the Alps and surrounding areas, and lacks the structure associated with orography.

One way to increase the resolution of the fields of interest is by performing dynamical downscaling, whereby the output of a global model is used to force a regional model. Fig.2 shows what difference in orography might be expected when going from the ECMWF seasonal model resolution (1.875°) to the 50-km RegCM; the maximum difference reaches about 800 m over the Alps. Also, major orographic features of the central-south Europe downscaling domain become clearly visible at a higher resolution.

We apply the Regional Climate Model version 3 (RegCM3) to dynamically downscale ECMWF seasonal integrations. RegCM is a regional model that is specifically designed to be used in the climate community worldwide (e.g. Pal et al. 2007). In the following, the results of dynamical downscaling of ECMWF operational seasonal forecasts are first shown, and then we discuss the results of downscaling of ECMWF experimental seasonal integrations, which are a part of the EU project ENSEMBLES.

---

[1] Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, U.K.

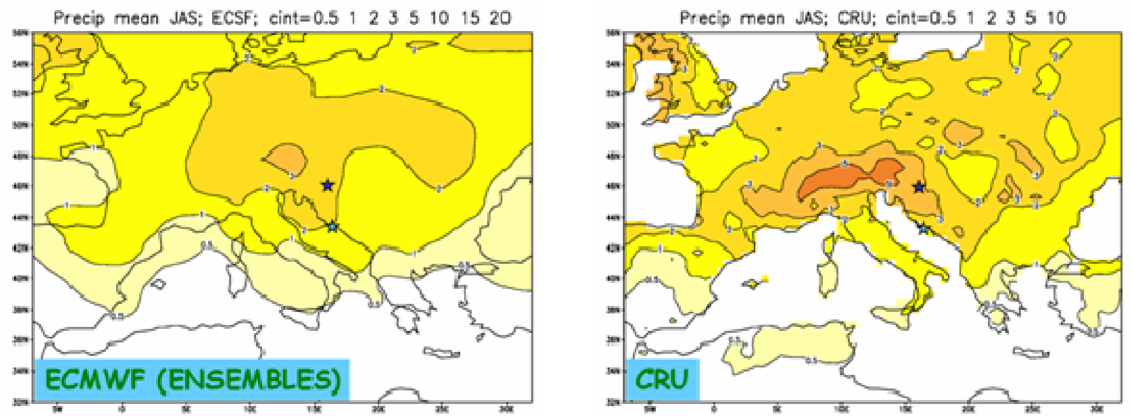


*Figure 1: The 11-year JAS precipitation averages for ECMWF experimental seasonal ensembles (left) and CRU data (right). Contouring 0.5, 1, 2, 3, 5,10 mmday$^{-1}$. The stars indicate the locations of Zagreb (dark blue) and Split (light blue), respectively.*

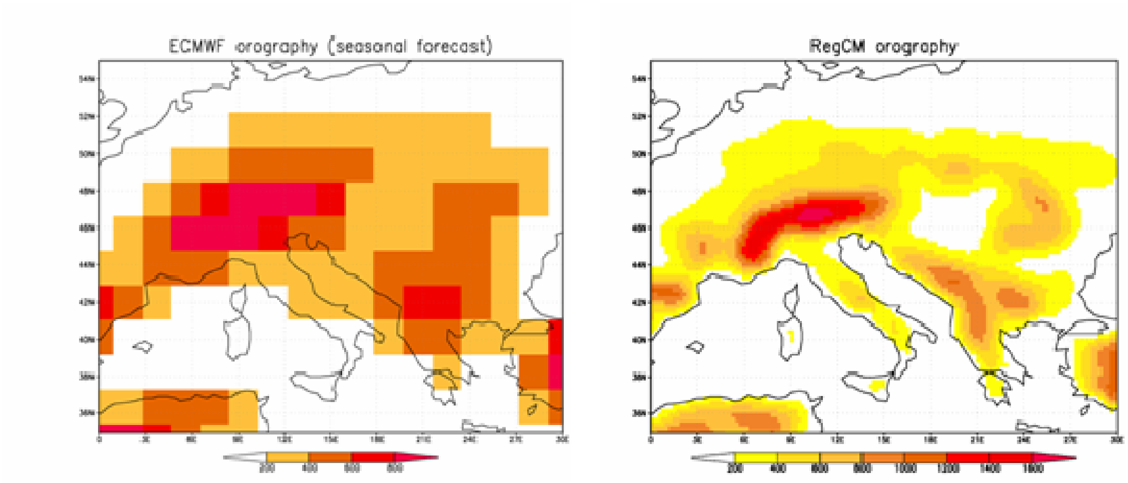


*Figure 2: Orography in ECMWF seasonal model (left) and RegCM used for dynamical downscaling (right). Contours every 200 m.*

## 2. Downscaling of ECMWF operational seasonal forecasts

Dynamical downscaling was applied to operational forecasts for only one season, JJA 2003. Input fields for RegCM (initial and boundary conditions) were taken from ECMWF seasonal forecast archive. Only fields at six standard pressure levels (1000, 925, 850, 700, 500, 200 hPa) with the 12-hour frequency were available. Thus, at the time when the experiments were made, the ECMWF archive contained relatively poor input for dynamical downscaling. To reduce computational requirements for downscaling, a subset of seasonal forecasts was made by applying clustering technique to the full operational ensemble of 40 members. Based on the 500 hPa geopotential height, the most populated clusters for the 1$^{st}$ and 2$^{nd}$ months in the JJA 2003 season were defined, and the common members from these most populated clusters were selected. The resulting sub-ensemble contained 12 seasonal forecasts to be downscaled.

The RegCM domain covered central and southern Europe and the northern Mediterranean, and the model horizontal resolution was 50 km. The model top was at 200 hPa with 14 levels in the vertical. Fig.3 (top panels) shows the JJA 2003 precipitation averaged over the 12-member ECMWF sub-ensemble and RegCM ensemble. Seasonal averages for Zagreb and Split are also indicated. Despite having more structure in the precipitation field, RegCM clearly overestimates the observed seasonal averages over Croatia, i.e. the station data fit better with the ECMWF sub-ensemble. When compared with the Global Precipitation Climatology

Project, GPCP[2], field (Huffman et al. 2001; Fig.3, bottom panel), a better agreement of ECMWF sub-ensemble than RegCM ensemble average precipitation with this verification data is seen.

The insufficiently resolved forcing fields used to generate RegCM initial and lateral boundary conditions (which also adversely affected the definition of the regional model vertical resolution) might be considered as the main contributor to model's relatively poor performance. In addition, to this "external" cause, model deficiencies, like a general tendency to overestimate total precipitation in summer (see e.g. Branković et al. 2004) may have also played a role in such a deficient result seen in Fig.3.

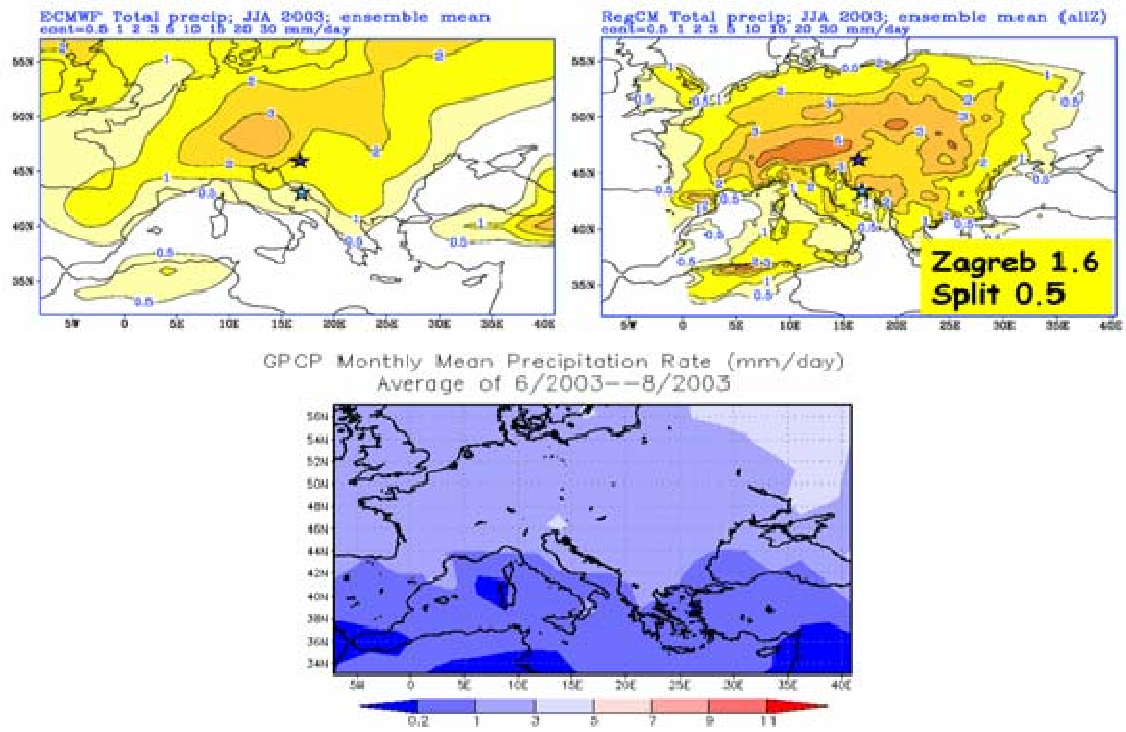


*Figure 3: JJA 2003 sub-ensemble mean precipitation for ECMWF operational seasonal forecast (top left), for RegCM ensemble (top right) and the GPCP verifying data (bottom panel).*

The results for upper-air fields, on the other hand, indicate no advantage of either model (not shown). This implies that the scarcity of the forcing data might have had a major detrimental effect on the regional model's physics. Nevertheless, despite being severely handicapped during the integrations, RegCM proved to be exceptionally robust.

## 3. Downscaling of ECMWF experimental seasonal integrations

ECMWF experimental seasonal integrations, part of the EU project ENSEMBLES, were run with same global model as for operational seasonal forecasts discussed in section 2: $T_L$95-L40. For the 11-year period, from 1991 to 2001, 6-month integrations were run twice a year (from May and November initial data) with 9-member ensembles. The data for the regional model initial and boundary conditions were available from ECMWF model levels with 6-hour frequency. Thus, the input for dynamical downscaling was much better than that for downscaling of operational forecasts described in section 2. In RegCM, the same horizontal resolution was used as before (50 km), but the number of vertical levels was increased to 18, and the model top was at 100 hPa. For the 11-year period, downscaling has been carried out for the two seasons only, winter (JFM) and summer (JAS). For verification purposes, both CRU and ECMWF Re-analysis (ERA-40) data were used.

[2] http://cics.umd.edu/~yin/GPCP/main.html

## 3.1. Area averaged skill measures

For upper-air fields, no major improvement in regional model skill scores is seen when compared with those from the global model (Fig.4). The skill score considered is anomaly correlation coefficient (ACC) computed with respect to ERA-40. When ACCs for temperature at 850 hPa (T850) are positive, only the winter of 1996 benefits from dynamical downscaling (Fig.4, left). For positive ACCs in summer, the global model always outperforms the regional model by a relatively narrow margin (Fig.4, right). The improvement by RegCM is found mostly for negative ACC values, and the benefit of downscaling in these cases is therefore questionable.

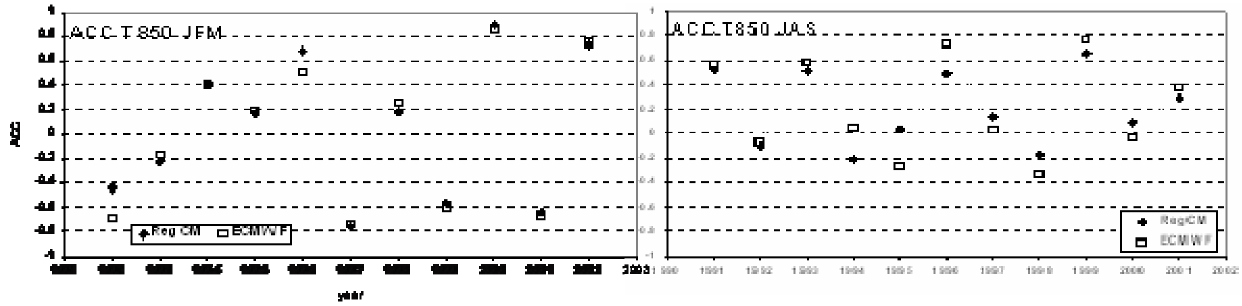


*Figure 4: T850 ensemble-mean anomaly correlation coefficients (ACC) over the RegCM integration domain for winter (left) and summer (right).*

Fig.5 shows another measure of performance of global and regional ensembles. For 2m temperature (T2m), area-mean anomalies are displayed for both models and both verifying analysis (CRU and ERA-40). Vertical bars denote the range of anomalies within ensembles, i.e. the highest and smallest anomalies from individual ensemble members. To some extent they indicate ensemble spread for individual years.

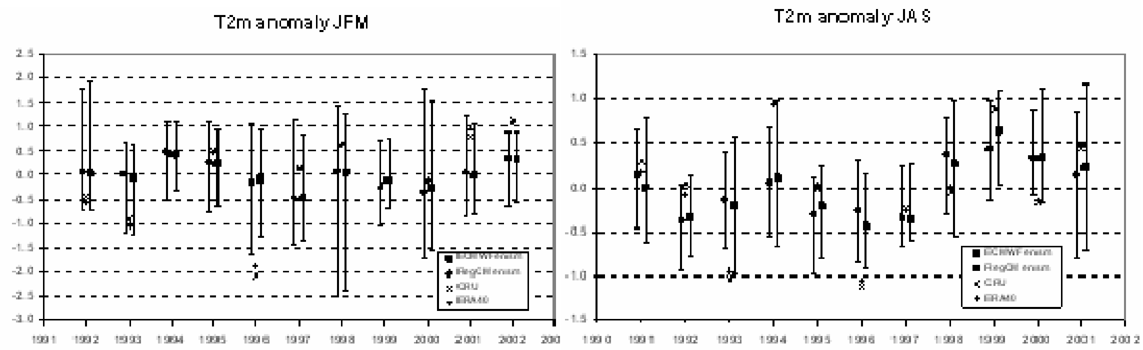


*Figure 5: T2m area-averaged anomalies for winter (left) and summer (right) over the RegCM integration domain. Solid circles and squares are ensemble means, and crosses verifying analyses. Vertical bars denote extreme values within ensembles.*

For the given 11-year period, the interannual variation of observed T2m anomalies, as revealed by × and + crosses in Fig.5, varies between +1.0° and -2.0° in winter (Fig.5, left) and between +1.0° and -1.0° in summer (Fig.5, right). For both models, the variability of ensemble averages (solid circles and squares) is smoother than observed, as it might be expected, but broadly follows observed anomalies. The vertical bars reveal that only for few years the observed anomaly is outside the range of ensemble members, thus indicating overestimation (JAS 1993 and both seasons in 1996) or underestimation (JAS 1994 and JFM 2002) of the whole ensemble. The year 1996 was characterised by extremely cold anomalies in the 11-year period. A relatively close positioning of ensemble averages and similar extension of vertical bars imply no major differences between the global and regional models. However, the area averaging may have concealed the differences that arise from small scales and therefore a spatial distribution of anomalies (or any other skill measure) should be examined.

### 3.2. Spatial distribution of skill (accuracy) measures

Verification of predicted T2m and precipitation against CRU and ERA-40 data is presented and discussed in terms of categorical forecasts. Following Wilks (1995), for verification of categorical forecasts the definition of contingency table elements is normally required, which in turn are based on the occurrence/non-occurrence frequencies of a discrete variable. The simplest approach is adopted here: either an event will or will not occur (the so-called binary forecast). The contingency table is then of the 2×2 form and contains four forecast/event pairs:

| | | Observed | |
|---|---|---|---|
| | | Yes | No |
| Forecast | Yes | ***a*** | ***b*** |
| | No | ***c*** | ***d*** |

In the above table, ***a*** denotes the frequency of the event that was correctly predicted, ***b*** is the frequency of the event that was forecast but did not occur, ***c*** denotes the frequency of the event that was observed but was not predicted, and ***d*** is the frequency of correctly predicted non-occurrence. The total number ***n*** of all forecast/event pairs is $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c} + \boldsymbol{d}$.

When $\boldsymbol{b} = \boldsymbol{c} = 0$, forecasts are perfectly accurate. If not so, various accuracy measures can be defined. The most basic and relatively simple ones are the hit rate, the false alarm ratio and the threat score. They are defined by the following expressions:

Hit rate: $$\boldsymbol{HR} = \frac{\boldsymbol{a}+\boldsymbol{d}}{\boldsymbol{n}}$$

False alarm ratio: $$\boldsymbol{FAR} = \frac{\boldsymbol{b}}{\boldsymbol{a}+\boldsymbol{b}}$$

Threat score: $$\boldsymbol{TS} = \frac{\boldsymbol{a}}{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}}$$

***HR*** ranges from 0 to 1; the latter is the best possible outcome. Sometime ***HR*** is dominated by the most frequent category and, as an overall measure of accuracy, could be misleading (for example, in the rain/no-rain events non-occurrences ***d*** could be much larger than ***a***, $\boldsymbol{d} \gg \boldsymbol{a}$). Therefore, a careful choice of thresholds for a given parameter is desirable. Alternatively, ***TS*** can be used because it gives more "weight" to the cases when the event to be forecast (occurrence) is less frequent than the non-occurrence. Thus, ***TS*** provides better information on real accuracy of the event in question (in particular, for rare events) because correct forecasts of non-occurrences are removed from consideration (Wilks 1995). The range of ***TS*** is between 0 and 1; the latter again is the best score. The range of ***FAR*** is also between 0 and 1, but its orientation is opposite to ***HR***, i.e. the best ***FAR*** is 0. In a sense ***FAR*** complements ***HR***.

### *3.2.1. Near-surface temperature (T2m)*

For categorical forecasts only a single event can be assessed at a time. Therefore, in order to be able to gauge model performance, various thresholds for precipitation and T2m were defined and examined. They, of course, should be related to climatology for a given season (there is no point assessing, for example, the T2m forecast accuracy for negative thresholds in summer!). Fig.6 (left) compares the hit rate for the JAS threshold T2m > +20°C for the global (top) and the regional model (bottom) against CRU verification dataset. The two distinct large regions, one in the north and one in the south of the domain, where ***HR*** = 1 can be seen. In the south, the hit rate is primarily determined by successful forecasts of T2m; in the north, the best hit rate is achieved because the non-occurrence of the T2m > +20°C was successfully predicted. Between these two regions there is a "transition" region where the hit rate assumes values smaller than 1. This area broadly coincides with the extent of the +20°C isotherm in verification data (not shown), and the variation in the hit rate corresponds to the displacement of the isotherm in ensemble members with respect to the verification.

The hit rate is improved with RegCM in parts of western and eastern Europe, and over the mountainous regions: the western Alps, the Balkans, the Apennines and the Pyrenees. For the mountains, the improvement could be, at least partly, ascribed to a better-resolved high orography in RegCM (cf. Fig.2). This essentially implies that the regional model predicts cooler temperatures (in this case, more non-occurrences of T2m > +20°C) than the global model over the mountainous regions, which is confirmed by verification data (not shown). The improvement in the eastern and western parts of the domain is not related to the representation of orography (these are relatively flat lowlands), but rather to fewer occurrences (fewer ensemble members) of the +20°C isotherm positioning northward of its observed value in RegCM than in ECMWF model.

Fig.6 (left) also indicates some areas where the hit rate in RegCM is worse than the global model (yellow regions in Fig. 6 bottom left panel). These are mainly lowlands to the south of 45°N where T2m could be quite high in summer. This deterioration of the RegCM skill is related to a non-negligible model systematic error: over much of the integration domain RegCM underestimates observed T2m (not shown).

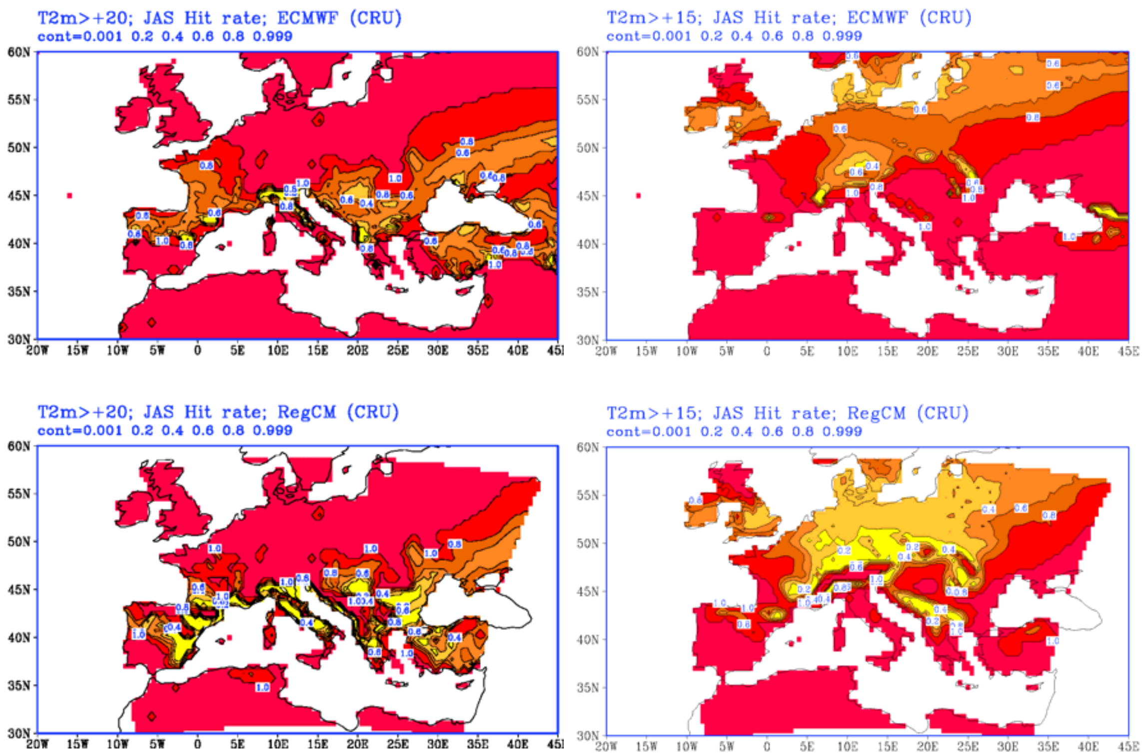


*Figure 6: The JAS hit rate for the threshold T2m > +20°C (left) and for the threshold T2m > +15°C (right) for the ECMWF model (top) and RegCM (bottom). Contouring interval every 0.2.*

The adverse impact of the RegCM systematic error in T2m (erroneous cooling) is better demonstrated for a lower threshold, T2m > +15°C. A relatively large area of central Europe (the northern portion of the domain) with the hit rate below 1 is seen in both models (Fig.6, right). In this region, observed T2m values are relatively close to the +15°C threshold, and for the ECMWF ensembles, small variations in T2m affect the model skill. For RegCM, the erroneous cooling contributes to largely reduced hit rate, because only small proportion of ensemble members actually predicted T2m to be above the +15°C threshold.

The difference between the hit rate and the threat score, for the threshold T2m > +15°C, could be inferred by comparing Fig.6 (right) with Fig.7. Low values of ***TS*** in Fig.7 indicate areas where (correctly predicted) non-occurrences of T2m, for the given threshold, have been excluded from consideration. In RegCM they are mostly found over the mountains of central Europe (the Alps, the Carpathian mountains). To the north of the mountains there is a relative large area with ***TS*** close to zero (Fig.7 right). Again, due to erroneous cooling, only a small fraction of RegCM forecasts predicted a correct occurrence for the threshold considered.

As mentioned above, the false alarm ratio to a certain extent complements the hit rate (or the threat score); it could be expected that when the hit rate is small, the false alarm ratio is large and *vice versa*. For the summer threshold T2m > +20°C (Fig.8), a relative improvement attained by RegCM (***FAR*** = 0) with respect to the ECMWF model is clearly indicated. For central and southern Europe, such an improvement is visible not only over the mountains, but also over other regions that are less affected by a better representation of orography. Only a small proportion, in the bordering area of the domain and in few isolated spots, remains relatively unchanged. Some improvement for RegCM is also seen for the threshold T2m > +15°C, however, not as widespread as in Fig.8 (not shown).

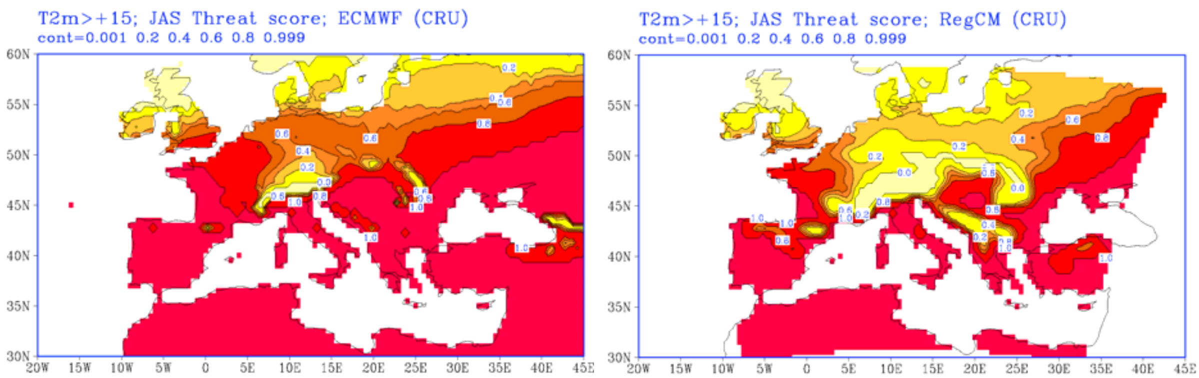


*Figure 7: The JAS threat score for the threshold T2m > +15°C for the ECMWF model (left) and RegCM (right). Contouring interval every 0.2.*

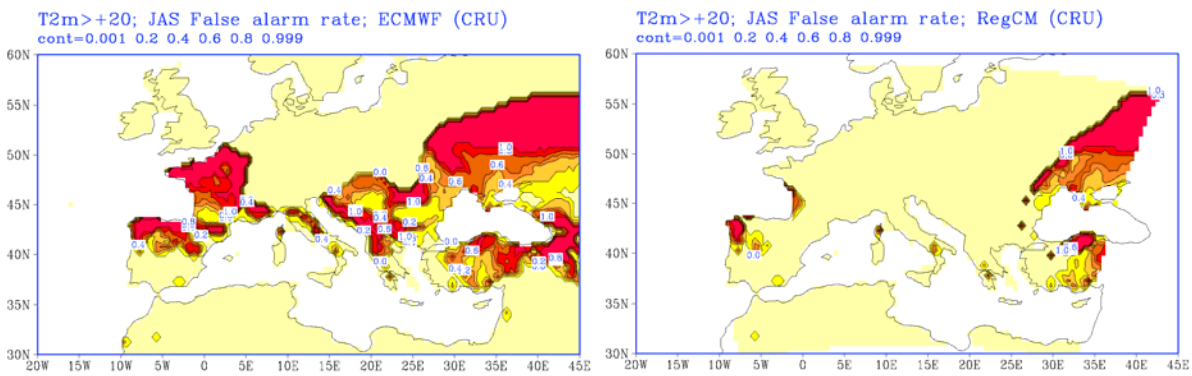


*Figure 8: The JAS false alarm ratio for the threshold T2m > +20°C for the ECMWF model (left) and RegCM (right). Contouring interval every 0.2.*

The above consideration for the JAS T2m thresholds (that are close to severe events) points out to a possible deleterious impact of the RegCM model systematic errors on some skill scores that are based on direct model output. On the other hand, as the results for FAR show, taking into consideration some other skill scores will help to get better understanding of model performance. Clearly, the assessment of model skill would benefit from some sort of statistical post-processing of model data.

The same accuracy measures as shown and discussed in Figs. 6 to 8 were computed for both models with respect to ERA-40 verifying data. Although the results broadly agree with those above, there are some differences, mainly in details. They can be summarized as the following, without showing figures. When compared to Fig.6, the hit rate computed with respect to ERA-40 is improved for ECMWF model, but is slightly worse for RegCM (it becomes zero over much of the northern Italy and over Greece). The threat score for the global model is improved in the Alpine region and to the south of 45°N, whereas almost no differences could be seen for RegCM. The false alarm ratio in ECMWF model has become better (zero) over Italy and Greece, but not much changed elsewhere. For RegCM, the isolated spots seen in Fig.8 (right) have now vanished. Overall, it seems that ERA-40 brings some improvements in accuracy measures for ECMWF model and a slight deterioration for RegCM. It can be assumed that such a result is related to a relative closeness in horizontal resolutions between ECMWF model and ERA-40.

For the winter season (JFM), we discuss only the threshold T2m < 0°C because, in the region of interest (central Europe), near-surface temperature in ECMWF model drops slightly below zero only in the Alpine area. Therefore, the comparison of the two models for some other, lower, winter threshold (severe events) makes no sense. The threat score in RegCM is improved when compared to the ECMWF model over high mountains (the Alps and the Carpathian Mountains) and over the mountains of the Balkan Peninsula (Fig.9). Over western Europe, ***TS*** is zero indicating a correct predictions of non-occurrences for the threshold considered (this also points out to mild winters in the 11-year period). Over the Balkans, more ensemble members than in ECMWF predict T2m below zero. This improvement in RegCM is genuine because, unlike in summer, the temperature systematic error is very small here (not shown) and does not affect adversely the accuracy of the model.

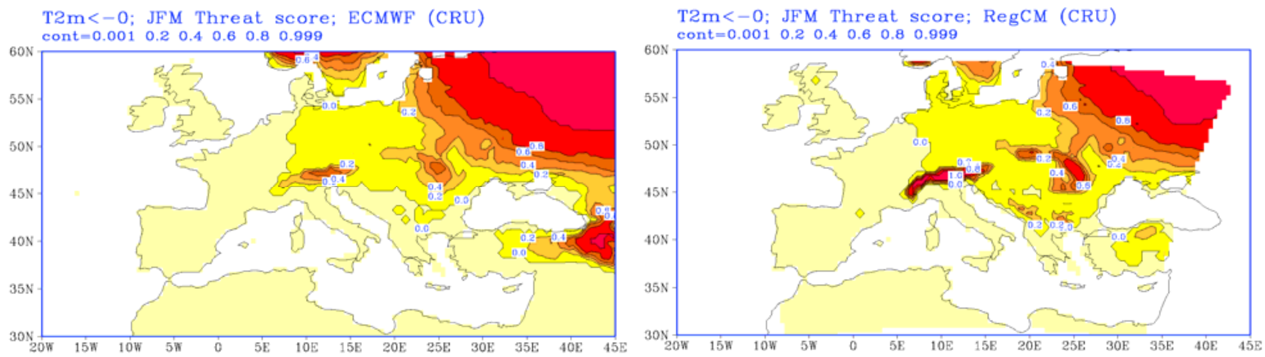


*Figure 9: The JFM threat score for the threshold T2m < 0°C for the ECMWF model (left) and RegCM (right). Contouring interval every 0.2.*

### *3.2.2. Precipitation*

The two distinct types of precipitation in central Europe are well established in summer and winter seasons. In winter, the large-scale precipitation is dominant, mostly related to large weather systems - lows and fronts. In summer, these systems also influence precipitation, however, to a lesser degree than in winter. In addition to the large-scale precipitation, the dominant summer type is the local-scale precipitation (often showery), closely related to variations in orography. The influence of orography on precipitation is not felt only due to changes in heights, but it also depends on prevailing winds with respect to the orientation of mountains ridges. Over central and southern Europe there is, on average, more precipitation in summer (cf. Fig.1) than in winter (not shown). For Croatia, the (climatological) division is in the north-south direction. In summer, larger amounts are recorded in the continental north than in the maritime south. For example, for the given 11-year period. Zagreb had on average 270 mm of rain in JAS, whereas in Split there was only about a half of that amount, 140 mm. In the corresponding winters the opposite was recorded, i.e. the north was drier than the south: in Zagreb 120 mm fell on average in JFM and in Split about 160 mm (shown as stars in Fig.10).

The assessment of accuracy measures for precipitation is made with CRU data only, since this parameter is not available in ERA-40. Because of the distinct annual cycle, Fig.10 shows the threat score for the two different thresholds: for JAS total precipitation, TPR > 2 mmday$^{-1}$ and for JFM, TPR > 1.5 mmday$^{-1}$.

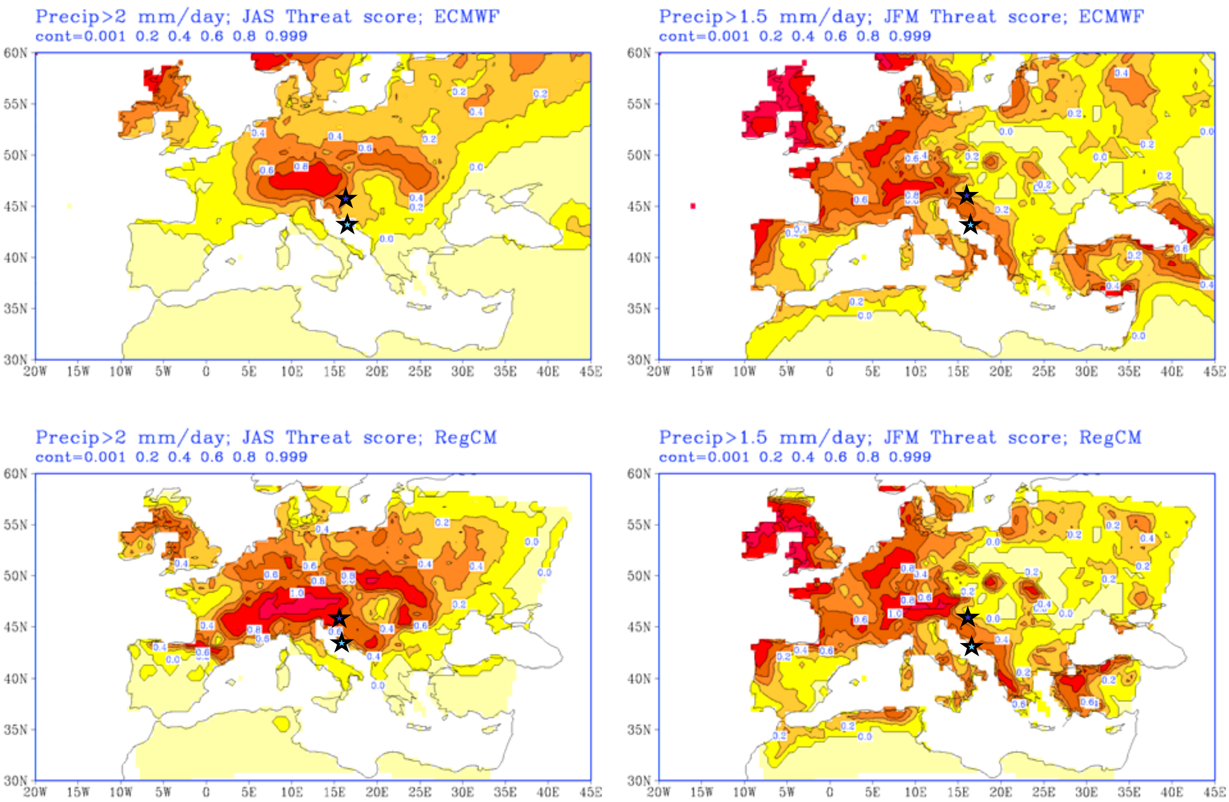


*Figure 10: The threat score for the threshold TPR > 2 mmday$^{-1}$ in JAS (left) and for the threshold TPR > 1.5 mmday$^{-1}$ in JFM (right) for ECMWF model (top) and RegCM (bottom).*

In JAS (Fig.10, left), a clear improvement could be seen with RegCM over much of central Europe, in particular over the mountains. Over a large portion of the Alps and the northern Carpathian Mountains, ***TS*** in RegCM attained the value of 1, and over the western Europe and the western Balkans it has largely increased values when compared with the ECMWF model. The improvement of ***TS*** seen in RegCM is not only associated with high orography but also with a better resolved (less high) orography (for example, the Massif Central in France). In addition, an overall better representation (in terms of detailed structure) of summer precipitation in RegCM irrespective of orography also plays an important role. It is clear from Fig.1 that the global model underestimates precipitation over much of European area, in particular in the west. The RegCM systematic errors in precipitation affect to a lesser degree the results of skill scores than for near-surface temperature. The low ***TS*** values to the south of approximately 42°N correspond to correctly predicted non-occurrences in both models, i.e. no precipitation larger than 2 mmday$^{-1}$ was observed in these latitudes.

In JFM (Fig.10, right), the RegCM model is better than ECMWF over the central Alps, western Turkey and the Atlas Mountains and in some (isolated) spots over the Carpathian Mountains. Somewhat lower scores in the central-northern France are due to both models overestimating precipitation there (not shown). Such an overestimation is also found in eastern Europe, between the Baltic and the Black Seas, resulting in smaller ***TS*** values. However, it seems that the RegCM systematic error in JFM (overestimation of precipitation in most of Europe north of 40°N, except the highest Alps, not shown) does not severely affect the results for ***TS***.

## 4. Conclusions

Dynamical downscaling with the RegCM model was applied to ECMWF operational and experimental seasonal forecasts. Downscaling mostly affects surface fields related to orography (near-surface temperature, precipitation) by introducing and/or improving detailed structure due to better resolved small spatial scales. We analysed the spatial distribution of accuracy measures for some surface fields, because this enables to identify areas where RegCM adds value to global forecasts. There is almost no or very little improvement in upper-air downscaled fields.

For ECMWF operational seasonal forecasts, RegCM initial and boundary conditions were affected by a poor availability of the global model data, and therefore, dynamical downscaling has not resulted in intended improvement. In addition, applying a relatively simple method to select a sub-ensemble for downscaling, might have also affected the results of the regional model.

For experimental seasonal integrations, RegCM, on average, improves T2m and precipitation forecasts over the mountains of central and southern Europe, in particular for higher thresholds examined (severe events). However, it must be emphasised that part of the improvement seen in the RegCM results was attained because model's systematic errors acted in the direction that favoured higher skill scores. For example, for T2m a relatively large erroneous cooling over central Europe improved the model hit rate over high mountains (the Alps, the Carpathian Mountains) by predicting accurately more non-occurrences for high thresholds (e.g. T2m > +20°C) than in the global model. The same systematic error has a deleterious effect on this skill measure over some (warm) lowland areas. For precipitation, the improvement of accuracy measures with RegCM is genuine because, for this parameter, the model systematic error is relatively small. This improvement is seen not only over the mountains of the integration domain, but also in other areas with relatively inconspicuous orography.

By definition, due to a higher horizontal resolution employed, dynamical downscaling by a regional model will always be able to produce better resolved details in forecast fields, in particular for surface and near-surface parameters and those related to orography. However, if the global model results are not very successful, the boundary forcing may adversely affect dynamical downscaling and therefore no (major) improvement can be expected. In the case of good global forecasts, dynamical downscaling can bring an improvement only if a regional model is free of relatively large systematic biases.

### References

Branković, Č., M. Patarčić and L.Srnec, 2004: Seasonal dynamical downscaling with ERA-40 data: A sensitivity study. *Croatian Meteor. J.*, 39, 15-39.

Huffman, G.J., R.F. Adler, M.M. Morrissey, S. Curtis, R. Joyce, B. McGavock and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multi-satellite observations. *J. Hydrometeor.*, **2**, 36-50.

Pal, J.S., F. Giorgi, X. Bi, N. Elguindi, F. Solmon, X. Gao, S.A. Rauscher, R. Francisco, A. Zakey, J. Winter, M. Ashfaq, F.S. Syed, J.L. Bell, N.S. Diffenbaugh, J.Karmacharya, A. Konare, D. Martinez, R.P. da Rocha, L.C. Sloan and A.L. Steiner, 2007: Regional climate modelling for the developing world, *Bull. Amer. Meteor. Soc.,* **88**, 1395-1409.

Wilks, D.S., 1995: *Statistical methods in the atmospheric sciences*. Academic Press, San Diego.