IDŐJÁRÁS
Quarterly Journal of the Hungarian Meteorological Service
Vol. 112, No. 3–4, July–December 2008, pp. 179–190

Analysis of heavy precipitation for France using high resolution ALADIN RCM simulations

Michel Déqué and **Samuel Somot**

Météo-France CNRM, CNRS/GAME,
42 avenue Coriolis, 31057 Toulouse, France; E-mail: deque@meteo.fr

(Manuscript received in final form October 24, 2008)

Abstract—A high resolution version of ALADIN over France is analyzed here in 30-year ERA-40-driven simulations. It is demonstrated that a resolution of 12 km improves some features of the mean precipitation field with respect to the same version at 50 km resolution. This version improves also the representation of precipitation upper quantiles in summer by producing less high precipitation rates.

Key-words: extreme events, precipitation, regional climate modeling, horizontal resolution, France

1. Introduction

Successful reproduction of large-scale features of climate, such as pole-to-equator gradient or location of main subtropical anticyclones from a set of equations, is a great challenge of numerical modeling. The average public, however, is little concerned. Similarly, predicting a global warming of 2 °C if carbon dioxide concentration doubles is little informative for the scientific impact community. The simulation of features at the scale of human activities, like the geographical distribution of precipitation over a country, is much more interesting for decision makers.

Regional climate modeling is thus an important issue. Climate simulation with a purely local model is not possible, except in idealized experiments, since the atmosphere has no borders. The first climate models were global or at least hemispheric (*Kasahara* and *Washington*, 1967; *Manabe et al.*, 1965). With increasing computer resources in the last decades, modelers have refined more and more the resolution of the general circulation models (GCMs). Increasing horizontal resolution corresponds to two categories of needs:

1) allowing a more accurate discretization of the Navier-Stokes equations which cannot be solved analytically; this also applies for vertical or time resolution: the higher the resolution, the closer the discretized solution to the exact solution,
2) enabling a better representation of surface forcing and, as a consequence, the meso- and synoptic scales.

In case 1) the minimum resolution seems to be 300 km, but further increases modify only to a limited extent the large-scale atmospheric circulation (*Sperber et al.*, 1994). This explains why during a quarter of century, GCMs have been run at that resolution. In case 2) there is no mesh size limit, since the scales of the forcing extend up to 1 km (mountain) and even to 1 mm (canopy).

Limited area climate models appeared some 25 years after the GCMs (*Giorgi*, 1990). An additional reason for the need of higher resolution in the 1990s was the availability of atmospheric reanalyses at typically 100 km resolution (ERA-40, *Uppala et al.*, 2005). Indeed, one can use regional climate models (RCMs) at high resolution as intelligent "interpolators" to produce data in area where observations are scarce or doubtful. Another recent and practical reason for higher horizontal resolution is the need by hydrologists for climate simulations taking into account surface elevation in the best possible agreement with their drainage basin. In a pioneer study, *Christensen et al*. (1998) showed that a LAM at high resolution (20 km) was able to produce intense local precipitation events which were not grid point storms.

The ALADIN model is a limited area version of the ARPEGE-IFS forecast system. It has been used in numerical weather prediction by a wide community since the 1990s (*Bubnova et al*., 1995) and more recently in regional climate modeling (*Radu et al.*, 2008; *Farda et al.*, 2008). In the framework of the CECILIA European project presented in this special issue, ALADIN has been used at 10 km resolution over three domains, Czech Republic, Hungary, and Bulgaria. In the present study, we analyze a similar experiment with ALADIN covering a domain over France, with special emphasis on extreme precipitation. Although this study is not formally part of the CECILIA project, it has permitted to explore and extend the capacities of ALADIN to simulate climate at high resolution.

In this paper we want to investigate whether the ARPEGE-climate physical parameterizations, which has been developed for spatial scales between 50 km and 300 km, are still valid at higher resolution. In addition, we wish to verify if the increase in resolution can produce some details, which make the model closer to observation. We do not analyze the temperature field because:

- due to the small size of the integration domain, temperature is largely controlled by advection from the lateral boundary conditions,
- the small-scale details are easy to mimic by simple surface elevation correction (the maps of mean screen level temperature present strong similarities with the pattern of the orography map).

Simulating precipitation is a harder challenge, and the present paper will address this field, with a special attention to heavy daily precipitation.

The paper is organized as follows: in Section 2, we present the experiments and validation data. In Section 3, we validate seasonal mean precipitation. In Section 4, we examine the probability distribution function of daily precipitation, and we focus on its tail in Section 5. Then we conclude in Section 6.

2. Experiments and data

ALADIN is a limited area model issued from the ARPEGE-IFS software developed at Météo-France and ECMWF. Proper references have been given in the previous section. More specifically, here we use version 4.6 of the climate model which is based on cycle 24 of ARPEGE-IFS. Two model discretizations are used here: one with 50 km resolution, the other with 12 km resolution. The two integration domains have the same center, which is also the center of the Lambert projection (2°E, 47°N). The relaxation-free area is shown in *Fig. 1* (solid line for the 12 km resolution, dashed line for the 50 km resolution). The higher resolution domain is larger than the lower resolution one, because the relaxation zone is narrower (8 grid points on each side in both cases).

Fig. 1. Integration domain of the two ALADIN versions where no Davies relaxation is applied (i.e., without the 8 rows at each border): FR12 (solid) and FR50 (dashed).

The total number of grid points is 50 by 50 for the 50 km version, referred to as FR50 in the following and 150 by 150 for the 12 km version, referred to as FR12 in the following. There are 31 vertical levels in both cases, and the time step is 15 min for FR50 and 10 min for FR12. Except for mesh size and time step, the two versions have exactly the same physical parameterizations described in *Radu et al.* (2008).

A 41-year simulation has been carried out with each model (1960–2000). The 6-hourly lateral forcing is ERA-40 reanalysis (*Uppala et al.*, 2005). We examine here only 30 years (1971–2000) of both simulations. We have at disposal a high resolution (8 km grid) reanalysis over France since 1971 (*Quintana-Segui et al.*, 2008). Such a dataset, named SAFRAN, is essential to evaluate the added value of 12 km resolution versus 50 km resolution (*Déqué* and *Somot*, 2008). It has been obtained by optimal interpolation of all climatological stations over France, so the true resolution (distance between two actual observations) is rather 30 km on average. In a numerical model, the true resolution is also coarser than the mesh size, because of numerical diffusion. We have also used raw observations from three rain-gauge series to validate the SAFRAN series (see Section 4).

In order to make an easy comparison between three different resolutions, first we have selected the land grid points of FR50, which are located in the French metropolitan territory. Then, for FR12 and SAFRAN, each land grid point is associated to the nearest neighbor in the FR50 restricted grid. All FR50 grid points which have less than 12 neighbors in FR12 or less than 20 neighbors in SAFRAN are eliminated. Finally, we retain, for each FR50 remaining point 12 neighbors for FR12 and 20 neighbors with SAFRAN. We have thus 166 boxes containing each 1 point of FR50 grid, 12 points of FR12 grid, and 20 points of SAFRAN grid. These numbers of 12 and 20 are a trade-off to keep the maximum number of boxes with the maximum of points inside: if we had imposed more points in the boxes, we would have got less boxes. This selection avoids sampling artifacts in coastal or border regions by ensuring equi-populated boxes. Indeed, the variance of box-averages decreases with the size of the box, and we could get higher extremes with small-populated boxes.

3. Seasonal mean precipitation

Let us first consider 30-year averages for each season. *Fig. 2* shows the precipitation distribution over France in summer (JJA) for SAFRAN, FR50, and FR12. The model is too wet, but reproduces the main geographical contrasts, and increasing resolution does not increase precipitation (a feature often observed in ARPEGE GCM). The model precipitation rates are too large over high mountains (Alps and Pyrenees). One can see that the pattern of Massif Central precipitation (just south to the center of France) is improved by higher resolution (unrealistic double maximum in FR50).

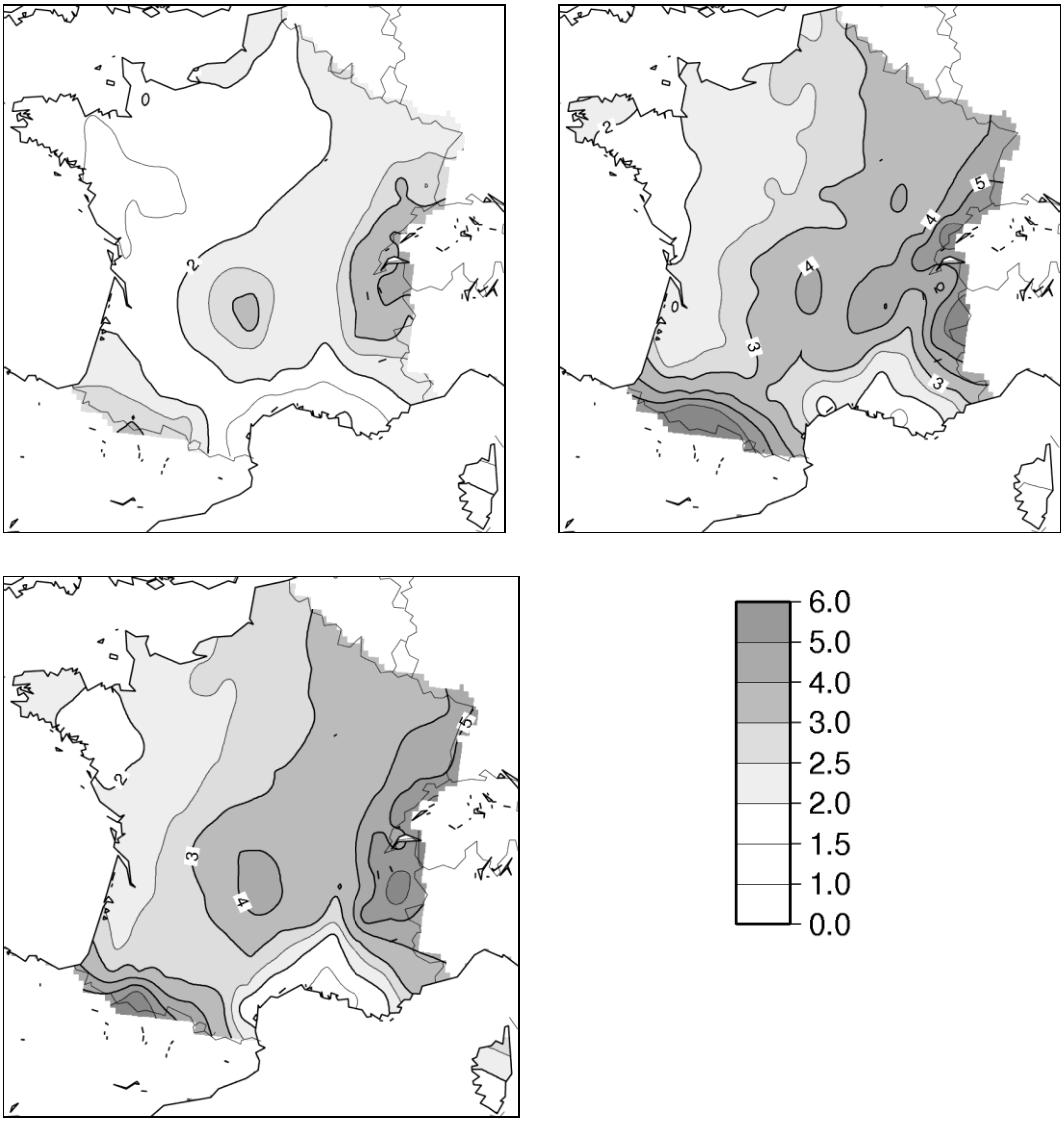

Fig. 2. Mean summer precipitation for SAFRAN (a), FR50 (b) and FR12 (c). Contours: 1, 1.5, 2, 2.5, 3, 4, 5 and 6 mm/day.

To make the comparison more accurate, precipitation is averaged in each box (see previous section for definition of the 166 boxes), and root mean square error (RMSE) is calculated for each season with respect to SAFRAN. *Table 1* shows the results. In each season, the higher resolution is superior to the lower one. The improvement is not spectacular, but systematic.

Table 1. Root mean square error of precipitation (mm/day) over France for seasonal means with the two ALADINs versions calculated on 50 km × 50 km averages.

	DJF	**MAM**	**JJA**	**SON**
FR50	0.72	1.37	1.32	0.76
FR12	0.65	1.24	1.16	0.59

4. *Precipitation probability density functions*

In order to evaluate the precipitation probability density function (pdf), quantile-quantile (q-q) diagrams of ALADIN and SAFRAN versus observation are plotted for three French cities. Paris is representative of the climate of the northern half of France. In the southern half, due to the presence of Massif Central mountains and Mediterranean sea, two climates are observed: Bordeaux is representative of the oceanic climate, whereas Marseilles is representative of the Mediterranean climate. In fact, Bordeaux and Paris climates are not so different, but there is a latitudinal and continental gradient at a time between the two cities. The location of the three cities is indicated in *Fig. 1*.

We consider here daily precipitation over 30 years. Each calendar season is analyzed separately. The quantile resolution is 1/1000 (probabilities between 0.1 and 99.9%). Three methods have been used to evaluate the quantiles in the boxes.

- Method 1 consists of averaging precipitation in each of the 3 boxes closest to the 3 cities, as we did in last section for RMSE over whole France. This method favors the extremes in FR50, because no spatial averaging is performed before sorting daily precipitation.
- Method 2 consists of computing quantiles for each model grid point, whatever the resolution. Then the quantiles are averaged for all grid points of a given box. This allows a fairer comparison between two resolutions, because no preliminary damping of high resolution data is applied before sorting.
- Method 3 consists of sorting the daily precipitation data of all points inside a box (pooling). This allows a better sampling of higher quantiles in high resolution data.

For FR50, the three methods are identical, because there is only one grid point per box. This is also the case with observations: we have a single time series for each city. In order to better capture the differences between the three methods, let us use the following notations:
R_{ijk} is precipitation for box i (i = 1,3 for the closest box to Paris, Bordeaux, and Marseilles), grid point j (j = 1,20 for SAFRAN grid, j = 1,12 for FR12 grid, and j = 1 for FRA50 grid), and day k (k takes about 2700 values depending on the season). A^x means averaging for index x ($x = j$ or k) and S^x means sorting with respect to index x.

- Method 1 computes $S^k(Aj(R_{ijk}))$.
- Method 2 computes $A^j(Sk(R_{ijk}))$.
- Method 3 computes $S^{jk}(R_{ijk})$.

From statistical point of view, methods 2 and 3 evaluate the same quantity, provided that each box is homogeneous. Method 3 is more accurate, in particular for extreme values, whereas Method 2 allows to calculate a confidence interval with some hypotheses. For our three cities, Methods 2 and 3 give similar results, whereas Method 1 provides systematically less precipitation rates, because of spacial averaging of wet and dry grid points.

Fig. 3 shows the q-q plots for DJF and JJA obtained by Method 3. For the three locations, SAFRAN provides a pdf in good agreement with raw station data (except for Paris JJA beyond 99% quantile). This indicates that the SAFRAN interpolation algorithm is not detrimental to high precipitation events. FR12 provides systematically less precipitation than FR50, whatever the quantile. Both models are generally below the diagonal of the diagram, indicating less precipitation than observed, except for summer precipitation below 5 mm/day, where both models exaggerate the quantities.

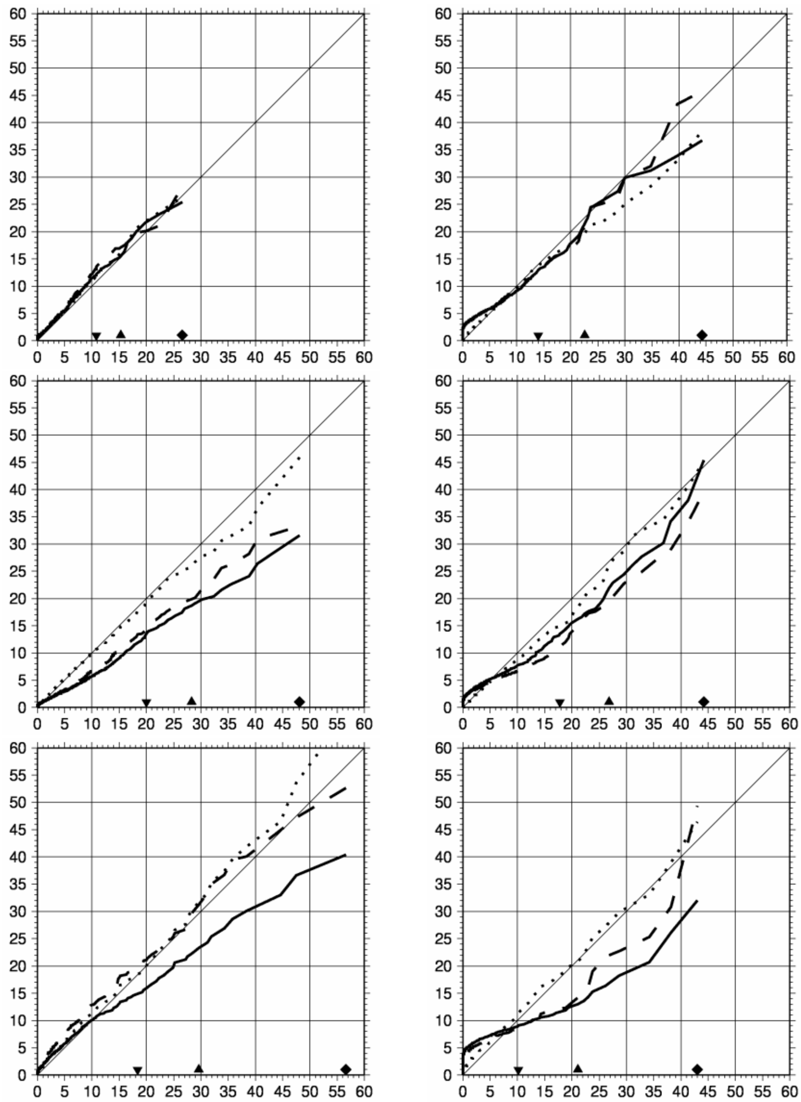

Fig. 3. Quantile-quantile plots of winter (left) and summer (right) daily precipitation (mm/day). Observed quantiles are sorted along the x-axis and model/analysis quantiles are sorted along the y-axis. Solid line corresponds to FR12, dashed line to FR50, and dotted line to SAFRAN. Top panels correspond to Paris, medium panels to Bordeaux, and bottom panels to Marseilles. The 99.9%, 99%, and 97.5% observed quantiles are indicated along the x-axis by a diamond, a triangle, and an inverted triangle, respectively.

From this first analysis, it appears that we can rely upon SAFRAN data for precipitation pdf, that FR12 and FR50 behave similarly for Paris and Bordeaux, but FR12 is worse than FR50 for Marseilles.

5. Heavy precipitation events

In order to make a more systematic comparison over France without inflating the paper with maps, we must use a quality criterion. The ranked probability score (*RPS*, *Epstein*, 1969) is well adapted to pdf comparison for variables like precipitation. It is widely used in probability forecast evaluation. It is simply based on the euclidean distance between the cumulated density functions (cdf) for a set of thresholds t_i, $i = 1, n$:

$$RPS = \sum_{i=1}^{n} \left[\mathrm{Prob}\left(MOD < t_i\right) - \mathrm{Prob}\left(REF < t_i\right) \right]^2 , \tag{1}$$

where *MOD* is FR12 or FR50 daily precipitation and *REF* is SAFRAN precipitation. Generally, the thresholds t_i are chosen to sample the pdf in a homogeneous way by taking the quantiles of a climatological distribution. Here we want to measure the fit of the tail of the distribution, so we selected the quantiles above 95% from the SAFRAN pdf ($n = 49$). The *RPS* is a dimensionless quantity and can be averaged for France, as the *MSE* in Section 3. However, a squared difference between two quantities close to each other (order of magnitude 0.95–1.00 each) produces small values, and we multiplied this score by 10^4 to handle quantities with order of magnitude 1. With this scaling, an *RPS* of 1 corresponds to a departure of 0.01 (e.g., 98% versus 99%) between ALADIN and SAFRAN cdf. *Table 2* shows *RPS* for each season. The 3 aggregation methods have been used to calculate the pdf. Methods 2 and 3 give similar *RPS*, which are smaller than those of Method 1. However, the conclusion is the same, whatever the method: FR12 has a better pdf tail than FR50, except in winter.

Table 2. Ranked probability scores for the tail of precipitation pdf (dimensionless units multiplied by 10^4) over France per season and with the 3 aggregation methods

	Method 1		**Method 2**		**Method 3**	
	FR50	**FR12**	**FR50**	**FR12**	**FR50**	**FR12**
DJF	1.88	2.50	1.65	2.34	1.56	2.16
MAM	3.15	1.44	2.18	1.13	2.04	1.16
JJA	1.73	0.66	1.39	0.64	1.35	0.66
SON	1.37	1.06	1.22	1.03	1.20	1.00

Fig. 4 helps to localize the weaknesses of the model pdf. It shows the ratio of the 99.9% quantiles between ALADIN and SAFRAN. The return period corresponding to this quantile is about 10 years, because one year contains about 90 days per season, so 10 years corresponds to 1000 values and thus to a probability of 0.1%. In winter, the behaviors of FR12 and FR50 are similar: the model overestimates the return value in the center of the country and underestimates it in the rest of the country. Contrary to the RPS results (which take into account all quantiles between 95% and 99.9%), there is no evidence that FR12 is worse than FR50. In particular, the underestimation by 30% in the southeast is improved by increasing resolution. In summer, the superiority of FR12 is obvious. FR50 produces return values in excess by 50% with respect to SAFRAN in some part of the country. However, FR12 is still too weak in the Mediterranean region.

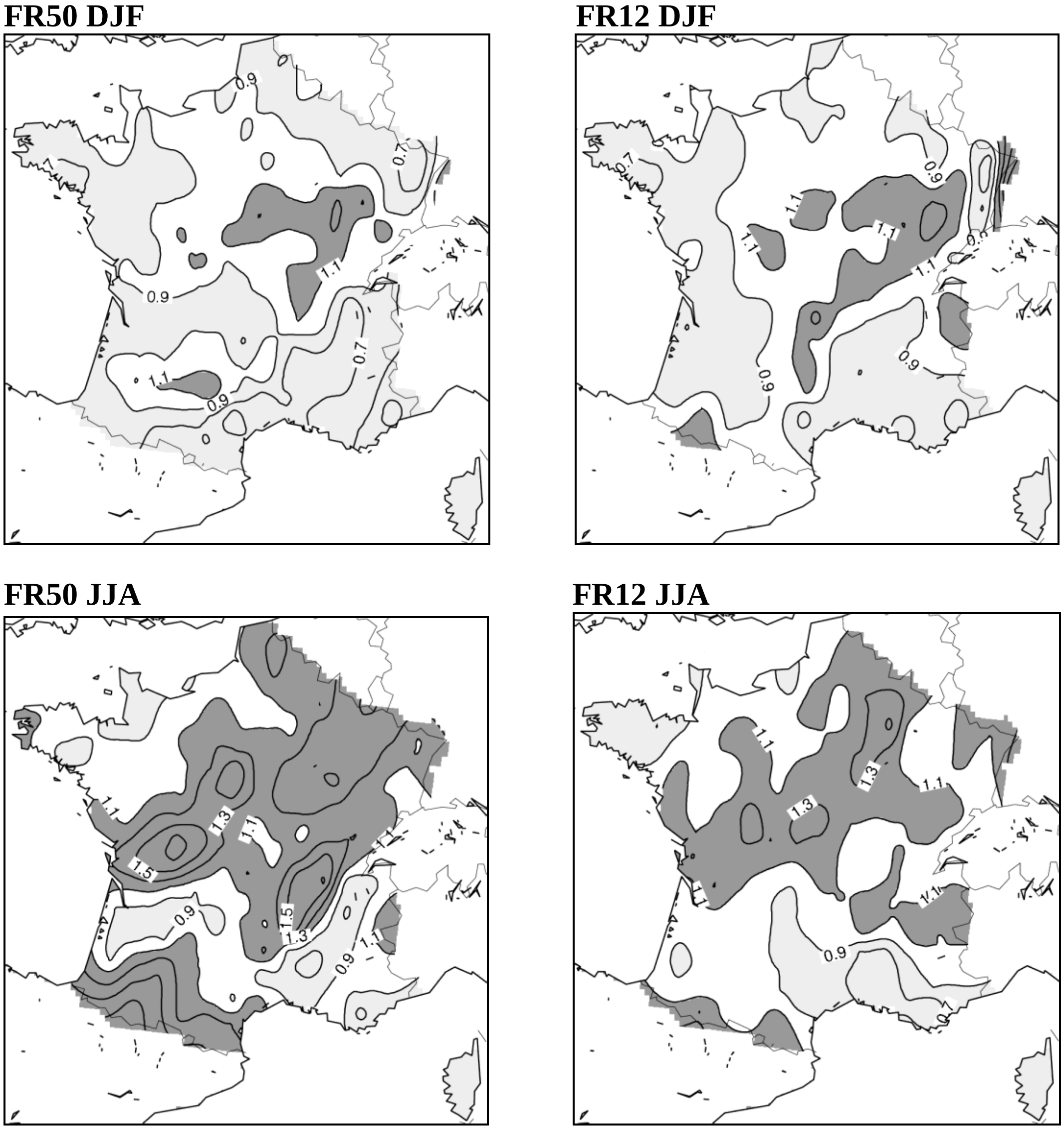

Fig. 4. Ratio of the 99.9% quantiles FRA50 over SAFRAN (left) and FRA12 over SAFRAN (right) in winter (top) and summer (bottom). The contour interval is 0.2, values over 1.1 are dense shaded, and values below 0.9 are light shaded.

We also computed the absolute maximum of precipitation of individual grid points by pooling all seasons (not shown). This 30-year maximum is about 50 mm/day over most of France for SAFRAN and ALADIN, but in the Cevennes region (southeast), it reaches 260, 144, and 108 mm/day for SAFRAN, FR12 and FR50, respectively. The highest events occur in autumn. The accurate representation of this kind of event is out of scope of ALADIN-climate.

6. *Conclusion*

The above results show that ALADIN is able to simulate the seasonal mean precipitation field and the tail of the probability distribution function (from one month to ten years return periods). For most quality criteria, the 12 km resolution version of ALADIN is superior to the 50 km one. The relatively small size of the domain (about 1200 km for the relaxation-free area) is not too detrimental to the simulation over the country, since no spurious strong precipitation stripes have been found over France. However, this small size does not allow the RCM to produce its own climate: we are here in a typical numerical downscaling exercise.

In the validation process of extreme precipitations, we used an alternative strategy by not considering a model grid point as a spatial average of observations, but by considering that a sample of grid points represents statistically a sample of observations. SAFRAN data are not really a series of observations, but the result of optimal interpolations. Using raw rain-gauge data over France is more complex, because the density is one observation per 10 km in some areas and one observation per 100 km in other areas. In *Déqué* (2007), such a comparison is performed with a 50 km resolution RCM.

Two important points are not addressed in this study. The first one is the space and time distributions of the highest precipitation extremes. As shown in previous section, the distribution is not uniform over France and the highest events are observed in the southeastern part in Autumn. The ALADIN-climate model, which is based on hydrostatic equations (in our climate version) with a full parameterization of convection, is not able to reproduce them quantitatively. To our knowledge, only 2-1 km resolution non-hydrostatic models like Meso-NH (*Ducrocq et al.*, 2008) can produce them explicitly. However, a recent study performed in the French project CYPRIM suggests that coarser resolution models like ARPEGE or ALADIN can produce strong rainy events (with respect to model quantiles) at a reasonable frequency. When the non-hydrostatic model Meso-NH is driven by ARPEGE during these events, it produces large amounts of rainfall (above 200 mm/day). This shows that one can use ARPEGE or ALADIN precipitation as a proxy of severe events with a proper statistical postprocessing (*Déqué*, 2007).

The second point is the added value of high resolution. The fact, that the probability distribution is improved in FR12, is not a proof that the spatial distribution of the 12–50 km scale structures is improved. In *Déqué* and *Somot* (2008) we show, that the spatial correlation of FR12 versus SAFRAN inside each 50 km box is significantly better than the correlation of FR50 interpolated onto the FR12 grid.

The perspective of this study is to extend the extreme precipitation analysis to CECILIA domains (Czech Republic, Hungary, Bulgaria, and Romania), where high resolution daily precipitation observations will be available soon. We plan to run the same scenarios (IPCC-A1B 2021–2050 and 2071–2100) with FR12 as in CECILIA. PRUDENCE simulations (*Christensen* and *Christensen*, 2007) at 50 km resolution suggest that the upper quantiles should increase in a warmer climate. As we have illustrated here that 12 km resolution ALADIN has a generally better representation of precipitation extremes than FR50, it is worth checking that the upper quantiles increase as a response of global warming is still valid at 12 km resolution.

Ackowledgements**—**This work was partly supported by European Commission FP-6 project CECILIA (GOCE-CT-2006-037005) and by French program ANR-MEDUP. The authors would like to thank *A. Horányi* for organizing the excellent workshop in Budapest, 2008. They are grateful to the two anonymous referees who suggested large improvements with respect to the first extended abstract.

References

Bubnova, R., Hello, G., Bénard, P., and *Geleyn, J.F.*, 1995: Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of ARPEGE/ALADIN NWP system. *Mon. Weather Rev. 123*, 515-535.

Christensen, O.B., Christensen, J.H., Machenhauer, B., and *Botzet, M.*, 1998: Very high-resolution regional climate simulations over Scandinavia-present climate. *J. Climate 11*, 3204–3229.

Christensen, J.H. and *Christensen, O.B.*, 2007: A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic Change 81*, 7-30.

Déqué, M., 2007: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global Planet. Change 57*, 16-26.

Déqué, M. and *Somot, S.*, 2008: Added value of high resolution for ALADIN Regional Climate Model. *Research Activities in Atmospheric and Oceanic Modelling 38*, 5.2-5.2.

Ducrocq, V., Nuissier, O., Ricard, D., Lebeaupin, C., and *Thouvenin, T.*, 2008: A numerical study of three catastrophic precipitating events over Western Mediterranean region (Southern France). Part II: Mesoscale triggering and stationarity factors. *Q. J. Roy. Meteor. Soc. 134*, 131-145.

Epstein, E.S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteorol. 8*, 985-987.

Farda, A., Déqué, M., Somot, S., Horányi, A., Spiridonov, V., and *Toth, H.*, 2008: The ALADIN model as regional climate model for Central and Eastern Europe. *Stud. Geophys. Geod.* (under revision).

Giorgi, F., 1990: Simulation of regional climate using a limited area model nested in a general circulation model. *J. Climate 3*, 941-963.

Kasahara, A. and *Washington, W.M.,* 1967: NCAR global general circulation model. *Mon. Weather Rev. 95*, 389-402.

Manabe, S., *Smagorinsky, J.*, and *Strickler, R.F.*, 1965: Simulated climatology of a general circulation model with a hydrological cycle. *Mon. Weather Rev. 93,* 769-798.

Quintana-Segui, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., and *Morel, S.,* 2008: Analysis of near-surface atmospheric variables: validation of the SAFRAN analysis over France. *J. Appl. Meteorol. Clim. 47*, 92–107.

Radu, R., Somot, S., and *Déqué, M.,* 2008: Spectral nudging in a spectral regional climate model. *Tellus 60A,* 898-910. doi: 10.1111/j.1600-0870.2008.00341.x

Sperber, K.R., Hameed, S., Potter, G.L., and *Boyle, J.S.,* 1994: Simulation of the northern summer monsoon in the ECMWF model: Sensitivity to horizontal resolution. *Mon. Weather Rev. 122,* 2461-2481.

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., Mcnally, A.P., Mahfouf, J.F., Morcrette, J.J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and *Woollen, J*., 2005: The ERA-40 re-analysis. *Q. J. Roy. Meteor. Soc. 131*, 2961-3012.