



Dynamical and statistical downscaling of the French Mediterranean climate: uncertainty assessment

M. Vrac¹, P. Drobinski², A. Merlo^{1,2}, M. Herrmann³, C. Lavaysse², L. Li², and S. Somot⁴

¹Institut Pierre Simon Laplace/Laboratoire des Sciences du Climat et l'Environnement, CNRS/CEA/UVSQ, UMR8212, CNRS/INSU – Laboratoire des sciences du climat et de l'environnement (LSCE), Gif-sur-Yvette, France

²Institut Pierre Simon Laplace/Laboratoire de Météorologie Dynamique, CNRS/Ecole Polytechnique/ENS/UPMC, Palaiseau, France

³Laboratoire d'Études en Géophysique et Océanographie Spatiales, CNRS/CNES/IRD/OMP, Toulouse, France

⁴Centre National de Recherches Météorologiques, Météo-France/CNRS, Toulouse, France

Correspondence to: M. Vrac (mathieu.vrac@lscce.ipsl.fr)

Received: 5 July 2011 – Revised: 27 June 2012 – Accepted: 11 July 2012 – Published: 6 September 2012

Abstract. ERA-40 reanalyses, and simulations from three regional climate models (RCMs) (ALADIN, LMDZ, and WRF) and from one statistical downscaling model (CDF-t) are used to evaluate the uncertainty in downscaling of wind, temperature, and rainfall cumulative distribution functions (CDFs) for eight stations in the French Mediterranean basin over 1991–2000. The uncertainty is quantified using the Cramer-von Mises score (CvM) to measure the “distance” between the simulated and observed CDFs. The ability of the three RCMs and CDF-t to simulate the “climate” variability is quantified with the explained variance, variance ratio and extreme occurrence. The study shows that despite their differences, the three RCMs display very similar performance. In terms of global distributions (i.e. CvM), all models perform better than ERA-40 for both seasons and variables. However, looking at variance criteria, RCMs are not always much better than ERA-40 reanalyses, whereas CDF-t produces accurate results when applied to ERA-40. In a second step, a combined statistical/dynamical downscaling approach has been used, consisting in applying CDF-t to the RCM outputs. It shows that CDF-t applied to the RCM outputs does not necessarily produce better results than those from CDF-t directly applied to ERA-40. It also shows that CDF-t applied to RCMs generally improves the downscaled CDFs and that the “additional” added value of CDF-t applied to the RCMs is independent of the performance of the RCMs in terms of CvM, explained variance, variance ratio and extreme occurrence.

1 Introduction

Climate varies across a wide range of temporal and spatial scales. Yet, climate modelling has long been approached using global general circulation models (GCM) that can resolve only the broader scales of atmospheric circulations (around 100–200 km grid resolution). Large-scale climate determines the environment for mesoscale and microscale processes that govern the weather and local climate. Resolving such interactions is needed for an improved understanding of the weather and local climate and of how climate both influences and is influenced by human activities.

Hence, there is a need to develop tools for downscaling GCM predictions to generate finer scale projections of local climatologies. Downscaling is the process of deriving regional climate information based on large-scale climate conditions. Both dynamical and statistical downscaling methods have been used extensively in the last decade to produce regional climate (see, e.g. Laprise, 2008; Maraun et al., 2010; or Rummukainen, 2010). Statistical downscaling models (SDM) consist in obtaining high-resolution climate data by deriving statistical relationships between observed small-scale variables (often station level) and larger scale variables (e.g. GCM), using either weather typing (e.g. Huth, 2001; Boé and Terray, 2008), regression models (through linear – e.g. Wilby et al., 2002; Busuioc et al., 2008 – or non-linear models – e.g. Cannon and Whitfield, 2002; Salameh et al., 2009; Ghosh and Mujumdar, 2008), or stochastic weather generators (e.g. Wilks and Wilby, 1999; Yang et al., 2005;

Carreau and Vrac, 2011). Statistical downscaling may be used whenever suitable small-scale observed data are available to derive the statistical relationships. Dynamical downscaling consists in driving a regional climate model (RCM) by a GCM over an area of interest since decreasing grid spacing generally improves the realism of the results (e.g. Mass et al., 2002; Sotillo et al., 2005; Ruti et al., 2007; Herrmann et al., 2011). Different inter-comparisons of downscaling techniques have been performed between RCMs only (e.g. Frei et al., 2003, 2006; Déqué et al., 2007), between SDMs only (e.g. Harpham and Wilby, 2005; Raju and Mujumbar, 2011), or including both approaches. In this latter category, various SDMs have been compared with RCMs but also applied to single RCMs (e.g. Busuioc et al., 2006; Quintana-Seguí et al., 2010) to evaluate the potential added value of applying a statistical model to one given RCM. However, most of the studies, working on one or several RCMs and/or SDMs, do not combine the two approaches but only compare the respective quality of the downscaled data (e.g. Haylock et al., 2006; Schmidli et al., 2006). In the present study, one goal consists in evaluating and comparing the potential added value of applying a statistical model to different RCMs. Moreover, most of those inter-comparisons have generally focused on one climate variable only, such as temperature (e.g. Spak et al., 2007) or more generally precipitation (e.g. Schmidli et al., 2007). In this article, we quantify the uncertainty of statistical and dynamical downscaling of three climate variables: wind, temperature and precipitation. Three different RCMs are employed: ALADIN, LMDZ and WRF. The multi RCMs approach allows here to ensure the robustness of the RCM uncertainty assessment. Moreover, the SDM evaluated here is named the “cumulative distribution function transform” (CDF-t) approach. CDF-t has originally been developed for wind downscaling (Michelangeli et al., 2009) but recently applied to temperature and precipitation (Vigaud et al., 2012; Lavaysse et al., 2012). This method aims at modelling local-scale statistical characteristics using a probabilistic downscaling model. While most of the classical statistical downscaling models generally directly provide local-scale values (e.g. Maraun et al., 2001 for a recent review), probabilistic downscaling models link the cumulative distribution function (CDF) of a large-scale variable with the CDF of the same variable at a much smaller scale, and allow to downscale –value of a combined or to correct– CDFs from which local-scale data can be generated. This study corresponds to the first time that the CDF-t approach is compared with and even applied to RCMs. More specifically, here, we address the specific following issues:

- evaluation of the quality of the downscaling models with respect to three meteorological variables (wind, temperature and precipitation),
- evaluation of the ability of the various downscaling techniques to simulate meteorological extremes,

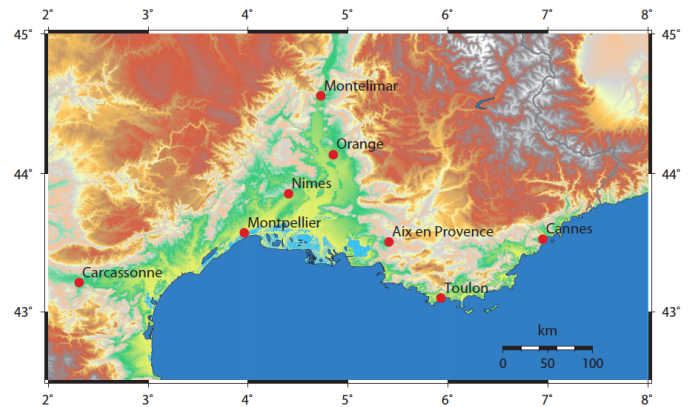


Fig. 1. Map of the studied region (French Mediterranean basin). The red dots indicate the locations of the surface weather stations from which wind speed, temperature and rainfall data are collected.

- evaluation of added value of a combined statistical/dynamical downscaling approach.

We focus here on the French Mediterranean basin (Fig. 1) which is a key spot of occurrences of meteorological extremes, such as windstorms (e.g. Drobinski et al., 2005; Guénard et al., 2006; Lebeaupin Brossier and Drobinski, 2009), heavy precipitation (e.g. Ducrocq et al., 2008; Lebeaupin et al., 2006; Beaulant et al., 2011; or Vrac and Yiou, 2010) and heat waves and droughts (e.g. Schär et al., 2004; Fischer et al., 2007; Vautard et al., 2009).

We investigate wind, temperature and precipitation distributions over the last twenty years (1981–2000) at eight widely distributed meteorological station locations (Salameh et al., 2009; Lavaysse et al., 2012). Uncertainty assessment of the various downscaling techniques with a special attention to meteorological extreme events is a key objective of the MEDUP project (Forecast and projection in climate scenario of Mediterranean intense events: uncertainties and propagation on environment) in the frame of which this study has been conducted.

In the following, Sect. 2 details the datasets, the statistical method and the RCMs used in this study. Section 3 quantifies the uncertainty associated with statistical and dynamical downscaling regarding wind speed, temperature and precipitation with a special focus on their extremes. Section 4 quantifies the possible added value of combining dynamical and statistical downscaling technique and evaluates uncertainty and error propagation at the various stages of the overall downscaling procedure. Finally, Sect. 5 concludes this study and suggests some perspectives.

2 Methodology

2.1 Station and large-scale data

2.1.1 Meteorological surface data

As “station” data, we use the 10-m wind speed (in m s^{-1}), 2-m temperature (in K) and rainfall (in mm day^{-1}) daily data provided by the SAFRAN analysis system (Le Moigne, 2002) at locations where weather stations are localized. We selected stations spread over Southern France (Fig. 1) allowing the sampling of the various sub-climatic regions as in Salameh et al. (2009) and Lavaysse et al. (2012), and use the data provided by SAFRAN in order to avoid “holes” in the data collected by the surface weather stations. The SAFRAN analysis system had been initially designed to provide atmospheric forcing data in mountainous areas for avalanche hazard forecasting (Durand et al., 1993, 1999). The avalanche version of SAFRAN has recently been used to develop a long-term meteorological reanalysis over the French Alps (Durand et al., 2009). This system has been extended over the whole country and modified in order to feed macroscale soil-vegetation-atmosphere transfer models (Le Moigne, 2002). A detailed description of SAFRAN, its validation and its application over France is given by Quintana-Seguí et al. (2008). The number of stations used in the analysis evolved with time (Vidal et al., 2010). Continuously increasing from 3000 to 4000 for precipitation between the late 1950s and present, the increase was much sharper for temperature and wind speed, with a jump from 500 to 4000 and from 500 to 2000, respectively, between the late 1980s and the late 1990s (no significant change since then). We kept temperature and rainfall data over a 20-yr period between 1981 and 2000 (only non zero observed precipitation). Because of the important automation of the anemometers during the 1980s, which significantly modified the measurement of the wind speed, we kept wind speed data over a 10-yr period between 1991 and 2000. In the following, although data come from SAFRAN gridcells that can sometimes be slightly different from real observed time series (Quintana-Seguí et al., 2008), we will refer to “station” data for simplicity. The main characteristics of the eight stations used in this study are given in Table 1 and their locations are indicated by red dots in Fig. 1.

2.1.2 Large-scale data: ERA-40 reanalyses

The European Centre for Medium-Range Weather Forecasts (ECMWF) has released reanalysed datasets for the time frame 1957–2002 (Simmons and Gibson, 2000). The ERA-40 model has a resolution corresponding to a T159 spectral truncation with 60 vertical levels from 1000 to 0.1 hPa. Data are reported on a $1.125^\circ \times 1.125^\circ$ grid every 6 h (00:00, 06:00, 12:00, and 18:00 UTC). In this study, we use the data covering the 1981–2000 period. To use large-scale data onto stations, a simple bi-linear interpolation from the four nearest

grid points of the model are used. This method can result in diffusive solutions, possibly destroying sharp gradients that may be present in variables such as precipitation. The potential consequences of such an interpolation are beyond the scope of this article and are therefore not investigated in the present study.

2.2 The CDF-t method

The statistical downscaling used in this study is the “Cumulative Distribution Function-transform” (CDF-t) method developed by Michelangeli et al. (2009). This approach aims at relating the cumulative distribution function (CDF) of a climate variable (e.g. wind) at a large scale (e.g. from GCM or reanalysis data) to the CDF of this variable at a local scale (e.g. at a station). CDF-t can be seen as a variant of the quantile-quantile correction method (e.g. Wood et al., 2002, 2004, or Déqué, 2007) that can use either non-parametric (as in Déqué, 2007) or parametric (as in Shabalova et al., 2003 or Piani et al., 2010) correspondences, between predictors and predictands quantiles, in order to derive local-scale CDFs (i.e. at the stations) based on the evolutions of the large-scale CDFs between calibration and target (i.e. future or evaluation) period. Although the CDF-t and quantile-quantile methods have a similar philosophy, CDF-t takes into account the change in the large-scale CDF from the historic to the future time period, while quantile-quantile projects the simulated large-scale values onto the historic CDF to compute and match quantiles. In the CDF-t approach, a mathematical transformation T is applied to the large-scale CDF to define a new CDF as close as possible to the CDF measured at the station. Let F_{Gh} and F_{Sh} define respectively the CDFs of the variable of interest from the GCM (subscript G in the following) and from a given station (subscript S) over a historical calibration period (subscript h). We assume that the transformation T allows to go from F_{Gh} to F_{Sh} :

$$T(F_{Gh}(x)) = F_{Sh}(x). \quad (1)$$

Replacing x by $F_{Gh}^{-1}(u)$, where u is any probability in $[0, 1]$, we obtain

$$T(u) = F_{Sh}(F_{Gh}^{-1}(u)), \quad (2)$$

which provides a simple definition of T . Assuming T is stationary in time, the transformation can be applied to F_{Gf} , the large-scale CDF of the climate variable over a validation or future period f , to generate F_{Sf} , the CDF at the station location for the same period f :

$$T(F_{Gf}(x)) = F_{Sf}(x), \quad (3)$$

which is equivalent to

$$F_{Sf}(x) = F_{Sh}(F_{Gh}^{-1}(F_{Gf}(x))). \quad (4)$$

Although the philosophies of the CDF-t and quantile-quantile approaches are relatively close in working both with

Table 1. Main characteristics of the eight stations over the 1981–2000 period for rainfall and temperature and over the 1991–2000 period for wind intensity.

Station	Altitude (m)	Mean annual prec. amount (mm)	Mean winter temperature (°C)	Mean summer temperature (°C)	Mean winter wind intensity (m s ⁻¹)	Mean summer wind intensity (m s ⁻¹)
Cannes	146	878.3	8.7	20.5	1.8	1.5
Nimes	98	711.5	8.1	21.1	3.1	2.9
Orange	98	691.6	6.9	20.5	2.1	2.0
Toulon	198	919.2	7.1	18.8	3.1	2.7
Carcassonne	297	710.9	7.6	19.3	4.0	3.5
Aix-en-Provence	220	629.9	6.6	19.4	1.9	1.8
Montélimar	284	894.6	6.7	20.0	2.7	2.2
Montpellier	7	619.4	8.7	20.9	3.7	3.4

CDFs, they are based on two different angles to correct either quantiles or probability values. The main underlying hypothesis of the CDF-t method is that, although the RCM is not able to predict correctly the CDF of a variable at a local scale, the evolution of this CDF from a time period to another is coherent and makes sense. Therefore, the CDF-t method translates this evolution to the local-scale CDF to determine F_{Sf} . Hence, CDF-t can be considered as a bias-correction method. However, in a purely practical way, CDF-t is used here to correct relatively large-scale statistical distributions to model local-scale ones. In that sense, it performs a change of spatial scale and could be seen as a downscaling approach, although not in the classical sense. The stationary assumption of the statistical model is recurrent whatever the SDM. This assumption has to be made to apply any SDM to climate conditions (e.g. future or past conditions) different from those of the calibration period (e.g. Quintana-Seguí et al., 2010; Maraun et al., 2010). However, CDF-t estimates the change in the local-scale distribution before generating climate (temperature, wind and precipitation) values. In other words, although the transformation T is assumed to be stationary, the statistical properties of the statistically downscaled data are not stationary and are able to evolve with climate change captured at large-scale.

Once F_{Sf} has been determined from Eq. (4), a quantile-quantile (QQ) approach is performed between F_{Gf} and F_{Sf} to generate local-scale time series. While in the “classical” approach (e.g. Déqué, 2007), QQ is directly applied between F_{Gh} and F_{Sh} ; the CDF-t approach generates quantile values through a QQ performed between F_{Gf} (and not F_{Gh}) and F_{Sf} (and not F_{Sh}). This allows to generate local-scale values according to F_{Sf} in chronological agreement with “future” large-scale simulations.

In this study, CDF-t is supplied either by the bi-linear interpolated ERA-40 data onto the stations, as explained previously, or by RCM outputs in the combined dynamical/statistical approach. Note that interpolation of RCM results for comparison to point station data must be performed carefully since spatial characteristics could be

significantly interrupted. As in Déqué (2007) or Michelangeli et al. (2009), a “constant correction” method is applied whenever the large-scale CDFs are off the range of the local-scale ones. The percentage of data for which the “constant correction” is applied is generally quite low whatever the season and the model (ERA-40, ALADIN, LMDz) used as input in CDF-t, for temperature (less than 1 %) and wind (~ 2 %).

However, for precipitation, while CDF-t applied to ERA-40 or ALADIN data performs this “constant correction” for less than 1 % of the data (for both winter and summer seasons), CDF-t applied to LMDz needs a “constant correction” on about 30 % of the data. Hence, except these high percentages for CDF-t applied to LMDz rainfall, the global percentage of values for which a “constant correction” is applied is quite low, even for the rainfall variable from ERA-40 and ALADIN. Indeed, to reduce the potential gap between the large- and local-scale CDFs, a “shift” is applied to F_{Gh} (the historical large-scale CDF) to make the large- and local-scale CDFs have the same first quantile (i.e. first x value), before applying CDF-t. Hence, in Eqs. (1)–(4), F_{Gh} and F_{Gf} correspond to CDFs computed from the “shifted” data (with the same shift defined for F_{Gh}). This technique is detailed (with an additional “inflation”, not performed here) in Kallache et al. (2011).

Note that, by construction, if the CDF of the large-scale variable (i.e. obtained from reanalyses or a given RCM) is stationary between the calibration and target (future) periods, then the estimate of F_{Sf} will be F_{Sh} . That is, the local-scale CDF is also supposed to be stationary. In other words, if no change occurs (neither at large nor at local scales), CDF-t will be a perfect downscaling model. This property is (or clearly should be) a fundamental requirement of any statistical downscaling model (SDM) or bias correction method. If this were not the case, we could not have much confidence in the statistical approach used and its further application in any context would be meaningless. In the following, all computations for this CDF downscaling have been made through the “CDF-t” R package (freely available on www.r-project.org/) with empirical (i.e. data driven) CDFs.

More theoretical and technical details, as well as first validations and comparisons can be found in Michelangeli et al. (2009), while applications of CDF-t are provided in Vignaud et al. (2011) and Lavaysse et al. (2012) to local projections of precipitation and temperature over India and the French Mediterranean, respectively, or in Oetli et al. (2011) to correct RCMs simulations as inputs of a crop yield model. While CDF-t is directly applied to all values of temperature and wind variables, as mentioned in Sect. 2.1.1, only non zero precipitation data are supplied to CDF-t for precipitation, hence linking large- and local-scale CDFs of strictly positive rainfall values (see Lavaysse et al., 2012 for further details).

2.3 Regional climate models

To evaluate the robustness of the uncertainty evaluation in dynamical downscaling, a multi-model approach is used with three different RCMs, run with different dynamical cores, numerical schemes and physical parameterizations. The RCMs are ALADIN, LMDZ and WRF.

In this study, the version 5 of the limited-area atmosphere regional climate model ALADIN-Climate is used without spectral nudging with a 50 km horizontal resolution (see Radu et al., 2008, Déqué and Somot, 2008, and Farda et al., 2010 for version 4 description, and Colin et al., 2010 for version 5 description). ALADIN-Climate shares the same dynamical core as the cycle 32 of its weather forecast counterpart ALADIN and the same physical package as the version 5 of the GCM ARPEGE-Climate (Déqué, 2010). ALADIN-Climate is a bi-spectral RCM with a semi-implicit semi-lagrangian advection scheme. Its configuration includes a 11-point wide bi-periodization zone in addition to the more classical 8 points relaxation zone. This so-called extension zone allows the computation of the fast-Fourier transforms for the spectral-to-grid point space computation. More details can be found in Farda et al. (2010). The planetary boundary layer turbulence physics is based on Louis (1979) and the interpolation of the wind speed from the first layer of the model (about 30 m) to the 10 m-height follows Geleyn (1988). Version 4 of ALADIN-Climate was used for the European ENSEMBLES project in which it was inter-compared with the state-of-the art of the European RCMs at 50 and 25 km (Christensen et al., 2008; Sanchez-Gomez et al., 2008; Christensen et al., 2010). The simulation covers the 1958–2001 ERA-40 period but we use those simulations over 1981–2000. In ALADIN-Climate, nudging is applied on temperature, wind vorticity, wind divergence and logarithm of the surface pressure. The maximum e-folding time depends on the variables (6 h for the vorticity, 24 h for the logarithm of the surface pressure, the specific humidity and the temperature, 48 h for the divergence) following the setting of Guldborg et al. (2005). The maximum e-folding time is reached above 700 hPa and for scales larger than 1280 km. The nudging is linearly decreasing between 700 and 850 hPa

in altitude and between 1280 and 640 km for the horizontal scales. The atmospheric boundary layer and the scales not represented in ERA-40 are not nudged (Herrmann et al., 2011).

The LMDZ model used in this intercomparison study is the regional version of a global atmospheric general circulation model, as described in Li (1999) and Hourdin et al. (2006). It is a grid-point model with the possibility of making a regional enhancement of spatial resolution. The current model has globally 240×180 points in longitude and latitude respectively. But the grid points are not equally distributed on the Earth. The spatial resolution is about 35 km in a rectangular covering an extended area of the Mediterranean and Europe (15° W/ 45° E and 20° N/ 60° N). A similar utilization of LMDZ in studying regional climate extremes around the Mediterranean has been reported in Goubanova and Li (2007). In this work, LMDZ is used as a classical limited area model (such as ALADIN and WRF). The whole Earth, except our interested domain over the Mediterranean and Europe, is nudged to ERA-40 through relaxation of both atmospheric temperature and winds. In LMDZ, a very strong indiscriminate nudging is applied over the whole globe except over the Mediterranean region, with a relaxation time set to half an hour to ensure a close relation with ERA-40 at synoptic scale. No nudging is applied over the Mediterranean domain.

The last RCM is the limited-area model WRF (Weather Research and Forecasting), version 3.1. The WRF model is developed by the National Center for Atmospheric Research (NCAR) (Skamarock et al., 2008). The model solves the non-hydrostatic equations of motion in terrain following sigma coordinates. One single domain is used with a horizontal mesh size of 50 km and a domain centered at 43.02° N and 15.45° W, covering an area of $4900 \text{ km} \times 3150 \text{ km}$. The geographical data are from 5 min resolution USGS (United States Geophysical Survey) data. In the vertical, 28 unevenly spaced η levels are used and a complete set of physics parameterizations is used (Kain and Fritsch, 1990, 1993 for the cumulus parametrization; Janjic, 2002 for the Eta surface layer scheme; Janjic, 1990, 2002 for the Mellor Yamada Janjic PBL scheme; Mlawer et al., 1997 and Dudhia, 1989 for the longwave and shortwave radiation schemes, respectively). The WRF simulation has only been performed on the 1991–2000 evaluation period. To avoid unrealistic departure from the driving fields, indiscriminate nudging is applied as boundary and central nudging with a coefficient of $5 \times 10^{-5} \text{ s}^{-1}$ (6 h) for temperature, humidity and velocity components above the planetary boundary layer (Salameh et al., 2010). Although the WRF model has not been run on the full period (calibration + evaluation) but only on the evaluation period, it allows to increase the number of RCMs studied. Hence, CDF-t cannot be applied to WRF for comparison. However, as the evaluations of all the tested models (CDF-t and RCMs including WRF) are made in the following on the exact same period, the inclusion of the WRF model is useful

to have a more complete assessment of the performances and of the variability of the RCMs simulations.

3 Uncertainty assessment of statistical and dynamical downscaling

The uncertainty evaluation is performed using two different time periods: one for the calibration of CDF-t and one for projections and evaluations. As indicated in Sect. 2.1.1, the observed temperature and rainfall data cover a 20 yr period, whereas wind speed only spans 10 yr. These periods have been split into two periods of equal length. Calibration is performed per season (winter and summer) in 1981–1990 for temperature and precipitation and 1991–1996 for wind speed; projections and evaluations are performed per season on 1991–2000 (temperature and rainfall) and 1996–2000 (wind) periods, based on daily values.

The evaluation of SDMs (any, not only CDF-t) still is a challenging task. Ideally, the best we could do is to calibrate SDMs over a given time period and evaluate them on another time period with climate conditions very different from those of the calibration period. As we do not dispose of reliable observed data for a climate “very” different from the one used for calibration, the least that we can (should) do is to cut into two parts the time series that we have at our disposal. At least, this allows to evaluate the SDMs on data that were not used during the calibration process. In other words, although the calibration periods (only 10- and 5-yr time periods) may not cover the “climatology” in the region, we hope that they contain sufficiently varied situations to capture the main relationships between large- and local-scale data within the calibrated SDM (here, CDF-t) to be applied in another climate context. In that context, to verify whether or not the CDFs of the calibration period and those of the evaluation period are significantly different at 95 % (i.e. $\alpha = 0.05$), the Kolmogorov-Smirnov (KS) test (Darling, 1957) has been applied between those CDFs. The KS statistic is the *supremum* of the absolute differences between F and F_{ds} : $KS = \sup_x |F_{ds}(x) - F(x)|$. The results showed that the two periods are associated with significantly ($\alpha = 0.05$) different CDFs for a vast majority of CDFs (87.5 %, 80 % and 67.5 % of the temperature, wind and rain CDFs, respectively).

The uncertainty quantification, which is a major objective of the MEDUP project, is performed by estimating the “distance” between the downscaled CDF and the observed CDF over the evaluation period. This is quantified over each station using the Cramer-von Mises (CvM) score. Indeed, the CvM can be seen as a measure of the distance between two CDFs (Darling, 1957) and has already been used in previous downscaling evaluation (e.g. Michelangeli et al., 2009). If $F(x)$ is the empirical CDF of the observed data over the evaluation period (i.e. the CDF to be retrieved), and $F_{ds}(x)$ the downscaled CDF, the CvM statistics is defined as the integrated squared difference between F and F_{ds} :

$$CvM = \int_{-\infty}^{\infty} |F_{ds}(x) - F(x)|^2 dx. \quad (5)$$

In the following, the CvM score is computed to compare the observed CDF with (i) the CDF of the interpolated large-scale fields from ERA-40 reanalyses, (ii) the CDF obtained from the CDF-t statistical downscaling method and (iii) the CDF of the interpolated RCM (for each of the three RCMs). The Cramer-von Mises (CvM) test is a goodness-of-fit test. Consequently, for each station, variable and model, the probability to reject or accept the null hypothesis allows an assessment of the degree of confidence into the downscaling method. This can be associated with the notion of uncertainty, hence justifying the title of the article and of the section. Moreover, the following figures display results through box-and-whisker plots (the so-called boxplots): those allow visualizing the spread of the quality of the different downscaling approaches and provide information on the global uncertainty associated with a given model for a given variable over the eight stations altogether. In addition to CvM, the Kolmogorov-Smirnov (KS) score was also used to quantify the distance between observed and modelled CDFs (Darling, 1957). However, the obtained results are similar to those of CvM. Hence, no assessment based on the KS scores are presented in the following.

Figure 2 shows boxplots of the different CvM scores obtained over the evaluation period (1996–2000 for wind, and 1991–2000 for temperature and rain) for winter (November to March) and summer (May to September), and for each of the five models (ERA-40, CDF-t, ALADIN, LMDz, and WRF) over the eight stations, and for the three variables.

The dashed lines indicate theoretical CvM values under which the CDFs of the simulations are not statistically significantly different at 95 % from the CDFs of the observations. The values of CvM are given as a function of the downscaling model (ALADIN, LMDZ, WRF and CDF-t) and the driving large-scale fields from ERA-40 reanalyses. The CvM variability represented by the box in the box-plot, quantifies the spatial variability over the eight weather stations. However, since there is no signal of any spatial pattern, the CvM values will not be shown as a function of the weather station locations. From Fig. 2, the inter-model variability of the CvM scores is slightly higher for wind speed than for temperature, and slightly higher for temperature than for precipitation. As expected, ERA-40 reanalyses display high CvM values for all three variables indicating the poor quality of the CDFs of the ERA-40 reanalyses with respect to the local observations for wind speed, temperature and precipitation. Looking at the median of the boxplots, it is worth noticing that, with CvM as a diagnostics of the quality of the downscaled variable, CDFs of wind speed and temperature from the RCMs are not better simulated than rainfall CDFs. Indeed, while CDF-t seems predominantly below the thresholds (i.e. not significantly

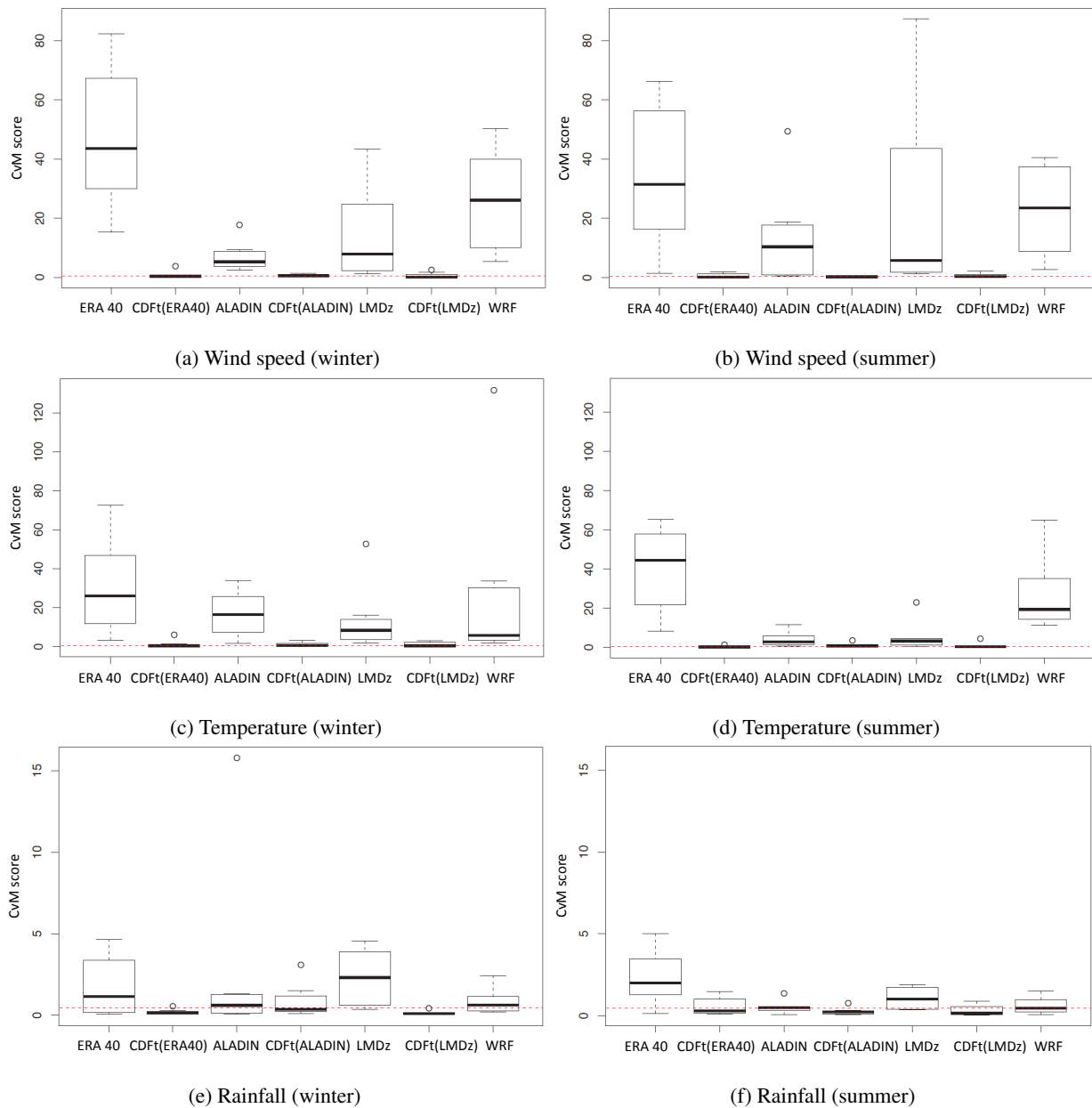


Fig. 2. Boxplots of Cramer-von Mises (CvM) scores computed over the evaluation period (1996–2000 for wind, and 1991–2000 for temperature and rain) in winter (left column) and summer (right column). The dashed lines indicate the threshold CvM values under which the modelled CDFs (from ERA-40, RCMs or CDF-t) are not statistically significantly different at 95 % from the CDF of the observations.

different from the observations), the three RCMs have their boxplots higher than the theoretical thresholds for both seasons (first two rows of Fig. 2). For rainfall, CvM scores from CDF-t are still below the thresholds (Fig. 2e and f). However, this is also true for ALADIN and WRF, even though the variability of their CvM statistics is larger (except for summer in Fig. 2f, when ALADIN boxplot is smaller than CDF-t boxplot). For both seasons, CvM values for LMDZ rainfall are higher than the thresholds, meaning that rainfall CDFs significantly differ from the observations. Interestingly, WRF and

ALADIN provide lower CvM scores for precipitation than for the other variables and all models perform better than ERA-40 for both seasons and variables (except LMDZ winter rainfall). Surprisingly, in general, the RCMs seem to produce better downscaled precipitation than temperature and wind speed, at least using the CvM metrics.

Although CvM scores provide quantitative information regarding the quality of the local-scale CDFs, a complementary and more commonly used diagnostics to evaluate the quality of the local-scale CDFs is the variance. In the following,

although it would certainly be possible to compute the different criteria directly with the obtained CDFs (i.e. without generating values), it is generally easier to generate values and then to compute the criteria. Indeed, the CDF-t method employed in this article corresponds to the empirical version, i.e. working with empirical CDF. Hence, we do not get an analytic form of the CDF. Moreover, even with a parametric formulation of the various CDFs involved in this approach, the local-scale CDFs obtained are not necessarily trivial and their properties (variance, etc.) can be difficult to determine.

Figures 3 and 4 (left columns) show the boxplots of the variance explained by each of the five models (ERA-40, CDF-t, ALADIN, LMDz, and WRF) in winter (Fig. 3) and summer (Fig. 4) over the evaluation period and for the eight weather stations. The explained variance (%ev) expressed in percent is defined as

$$\%ev = \frac{\sum_{i=1}^n (S_i - \bar{O})^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \times 100, \quad (6)$$

where n is the number of days of the period of evaluation, S_i is the simulated value for day i , O_i is the observed value at day i , and \bar{O} is the mean of the observations for the period. The quantity %ev allows to characterize the variability of the simulated data with respect to the mean of the observations. Therefore, this criterion provides a quantitative information about the quality of the variability of the simulations. However, to be correctly interpreted, the variance has to be computed from data normally distributed. Indeed, although the variance can be empirically computed for almost any sample, whatever the real underlying associated distribution, the interpretation of the variance is only correct based on the assumption that the data are Gaussian. This is usually considered the case for temperature and wind (quantile-quantile normal plots confirmed this for our data, not shown) but precipitation data are asymmetric and are usually considered as Log-normally distributed (e.g. Das, 1956; Cho et al., 2004), meaning that the logarithm of strictly positive precipitation values is normally distributed. Hence, in the following, if %ev is directly calculated on the temperature and wind intensity data, it is computed on the log-values of the positive precipitation data. In Figs. 3 and 4 (left columns), the closer %ev is to 100 % (dashed line), the better the variability of the simulations. Note that, in those figures, the range varies from one variable to another.

For wind speed (Figs. 3a and 4a), %ev is much larger than 100 % for ERA-40 and WRF, whereas it is close to 100 % for CDF-t, ALADIN, and LMDZ. For temperature (Figs. 3c and 4c), ERA-40 and ALADIN in winter, and ERA-40, ALADIN and WRF in summer display %ev higher than 100 %. The explained variance %ev is closer to 100 % for CDF-t, LMDz and WRF in winter, and CDF-t and LMDZ in summer. If the latter models are well centered around 100 % on

average over the eight stations, a station-to-station comparison shows differences especially in winter (Fig. 3a). For the rainfall (Figs. 3e and 4e), the differences between seasons are more pronounced. The explained variance %ev displays large values for all RCMs and around 100 % for ERA-40 and CDF-t in winter (Fig. 3e). In summer, while %ev is too low for ERA-40 and ALADIN, it is close to 100 % for CDF-t and WRF, and too large for LMDZ (Fig. 4e). Interestingly, ERA-40 does not overestimate the explained variance as it does for wind speed and temperature. In winter, dynamical downscaling of rainfall with all three RCMs generally degrades the explained variance with respect to the driving field ERA-40.

However, as the percentage of explained variance is by definition calculated with respect to the mean of the observations (i.e. \bar{O} in Eq. 6), it does not reflect whether or not the variance of the simulations (i.e. with respect to their own mean, say \bar{S}) is similar to the variance of the observations. To assess the variability of the simulations with respect to their own mean, the ratio of variances (%rv) is computed (in percentage to ease comparisons with %ev) for each model and station:

$$\%rv = \frac{\sum_{i=1}^n (S_i - \bar{S})^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \times 100 = \frac{\sigma_S^2}{\sigma_O^2} \times 100, \quad (7)$$

where \bar{S} is the mean of the simulated data, and σ_S^2 and σ_O^2 are the estimated variances of the simulations and observations, respectively. This ratio allows to see if the variance of the simulations is close to the variance of the observations. It can also be seen as the percentage of explained variance when the mean of the simulations is exactly the mean of the observations. In other words, the combined evaluation of those two criteria also provides information on the bias since if %ev and %rv are close to each other, it means that the bias is small. Also, if %rv is close to (resp., far from) 100 % and %ev is not, it means that the bias is relatively large (resp., small).

The boxplots of %rv are displayed in Figs. 3 and 4 (right column) for winter and summer, respectively. Globally, %rv is similar to %ev (left columns of Figs. 3 and 4). This means that the high values of %ev are essentially due to differences between the variance of the simulated data and the variance of the observations, rather than differences between the mean values. Nevertheless, some differences with %ev are visible. For wind speed, ERA-40 and WRF display boxplots for %rv smaller than for %ev, meaning that the large values of %ev can be attributed to differences in mean values. Nevertheless, still regarding wind speed, one can note the large values of %ev and %rv for WRF. Indeed, WRF produces too strong surface winds. This generates high %ev because the WRF wind speed mean value is higher than for other RCMs. Moreover, this generates a high %rv since strong winds induce strong surface stress and wind variance (i.e. in Eq. 7)

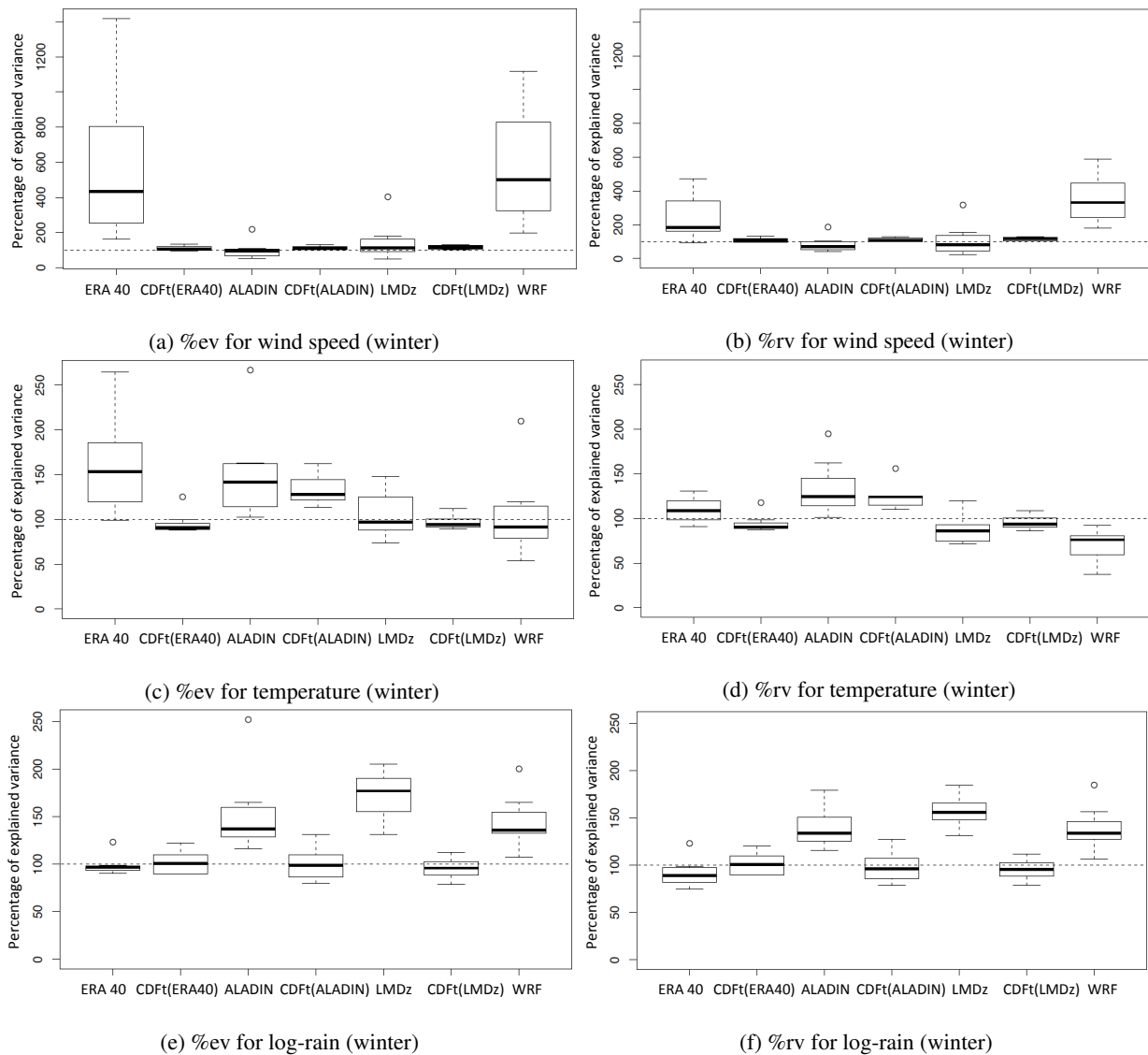


Fig. 3. Boxplots of percentages of explained variance (left column) and ratios (sim/obs) of variances in % (right column) computed over the evaluation period (1996–2000 for wind, and 1991–2000 for temperature and rain) in winter. For rain, the variances are calculated on the log-values of rainfall data (see text for details). The closer to the dashed line at 100 %, the better the variability of the simulations.

is linearly related to the surface stress (e.g. Drobinski et al., 2004).

To evaluate how those biases influence the representation of the “extreme” values, Fig. 5 provides for each model the boxplot (representing the spatial variability between the eight stations) of the percentages of simulated values that are higher than the 95th percentile of the observations in winter (left column) and summer (right column) over the evaluation period. A “perfect” model should give 5% of such values. For the rainfall variable, those percentages are given conditionally on positive rain intensity only. The question we are trying to answer here is: If we know that it rains, what is the global probability to get a simulated data that is higher than the observed 95th percentile. This corresponds

to a conditional probability (or percentage) that allows us to compare SAFRAN and model (RCM or ERA-40) data in a proper way. Indeed, as the frequency of wet days is different from SAFRAN and from the models, it would not be consistent to look at this percentage for all values (i.e. including 0’s). Looking at rainy days only makes that SAFRAN and model data are compared in comparable situations.

On average, ERA-40 overestimates the occurrence of wind speed extremes and underestimates the occurrence of temperature and rainfall extremes. Except for WRF, dynamical downscaling improves the occurrence of wind speed extreme. Occurrence of winter temperature extremes are better predicted when dynamical downscaling is applied. This is true whatever the RCM; although, WRF shows the highest

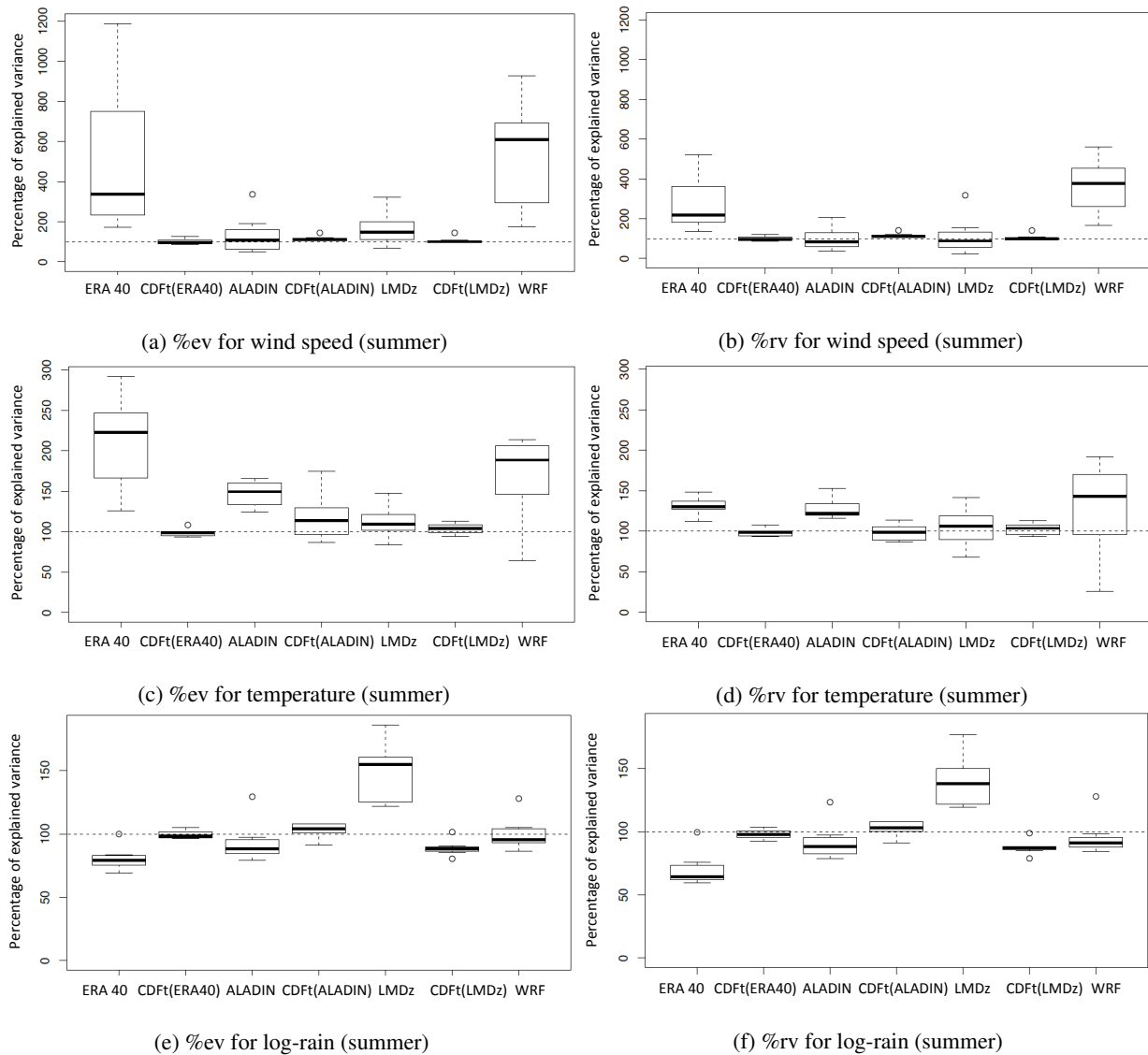


Fig. 4. Same as Fig. 3 but for summer.

spatial variability of those percentages between the stations. Summer extreme temperatures are however strongly overestimated with ALADIN and WRF, the latter still showing the highest spatial variability between stations. The results are different for extreme rainfall. All RCMs overestimate the occurrence of rainfall extremes in winter, while for summer, LMDz still has too many rainfall extremes and WRF too few on average. The same results are obtained when considering higher quantiles (e.g. 99th, not shown). Interestingly, for wind speed and precipitation, the general patterns of the box-plots for climate extreme occurrence are very similar to those for %ev and %rv for both seasons (Figs. 3 and 4). This result indicates that the deficiencies of the RCMs to reproduce accurately the variance of the various climate variables affect in a same way the prediction of extreme occurrence.

One major finding of this section is thus that all RCMs, despite their differences in terms of dynamical core, numerical schemes and physical parameterizations, display very similar performance, which are not necessarily “much” better than ERA-40 reanalyses in terms of variance criteria (although this is true on average for all stations, this conclusion does not stand on station-to-station basis). Surface temperature and wind observations are assimilated in ERA-40 reanalyses. Therefore, it is logical that RCMs do not necessarily perform much better than ERA-40 at locations where observations were assimilated. As precipitation is not assimilated but the product of physical parameterization, ERA-40 rainfall is usually considered as strongly biased and incorrect. Hence, RCMs precipitation are usually better than those from ERA-40. However, this is not the case for the variance criteria,

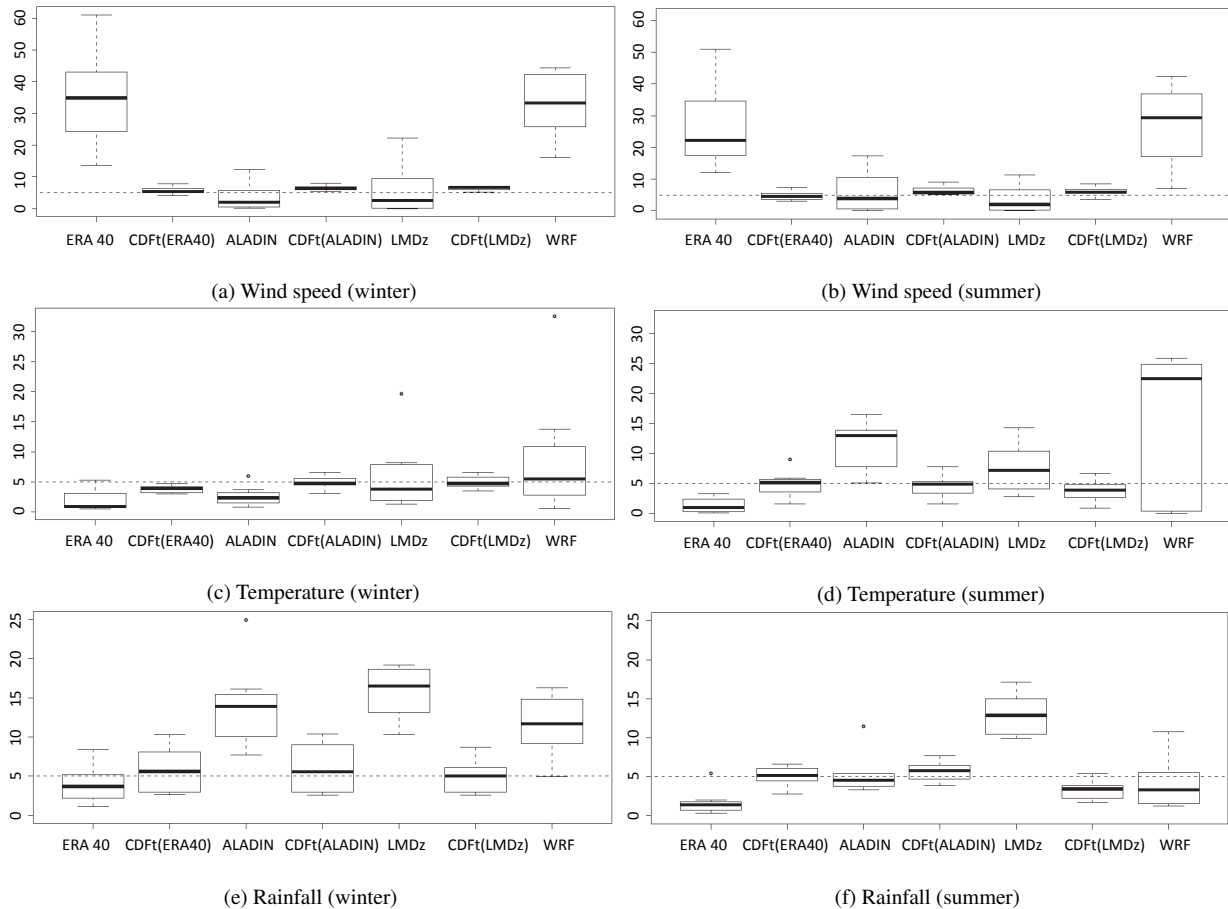


Fig. 5. Boxplots of percentage of values higher than the observed 95th percentile for the evaluation period (1996–2000 for wind, and 1991–2000 for temperature and rain) in winter (left column) and summer (right column). The closer to the dashed line at 5 %, the better the simulation of the extremes.

especially in winter (Fig. 3). Indeed, ERA-40 data assimilate a number of observations over the whole domain and even if they do not assimilate precipitation, they assimilate pressure fields that have a significant influence on precipitation. Moreover, even if the resolution of RCM is higher, they do not necessarily assimilate these observations for this domain: in that case, they can be less competitive. This is also to be contrasted by studies showing a frequent added value of high-resolution simulations, in particular for extremes of wind or rain (e.g. Ruti et al., 2007 or Herrmann et al., 2011).

4 Combined statistical/dynamical downscaling

In this section, we quantify the potential “added value” of combining statistical and dynamical downscaling approaches. It consists in applying CDF-t to the RCM downscaled fields instead of the ERA-40 large-scale fields. The relevant questions addressed in this section are the following: Can the CDF-t approach be used for bias correction of RCM downscaled data as performed similarly in the

GCM community? Does combining statistical and dynamical downscaling techniques improve the overall downscaling performance?

Figures 2, 3, 4 and 5 also display the results for the combined statistical/dynamical downscaling approach. CDF-t applied to ALADIN and LMDz are referred to as CDF-t(ALADIN) and CDF-t(LMDz), respectively. For data availability reasons, WRF simulations have not been downscaled further with CDF-t.

For both seasons, the improvement in terms of CvM diagnostics is clear for all three variables for CDF-t(LMDz). It is also true for wind speed with CDF-t(ALADIN). However, with ALADIN, the gain is smaller for temperature and rainfall. Note also that, based on the results from the height stations, the spatial range of the CvM scores from CDF-t(ALADIN) is similar to that of ALADIN (Fig. 2e). Similar results are found for %ev and %rv (Figs. 3 and 4): the combined statistical/dynamical approach displays improved results for wind speed, temperature and rainfall with respect to dynamical downscaling only. The gain is however smaller in terms of %ev and %rv when CDF-t is applied to ALADIN

outputs. Note also that the spatial variability is reduced for all variables and RCMs when CDF-t is applied. Finally, the use of the combined statistical/dynamical downscaling approach never degrades the results quantified in terms of CvM, %ev and %rv regarding dynamical downscaling only.

For wind speed, temperature and rainfall extremes (Fig. 5), combined statistical/dynamical downscaling generally provides better results (i.e. a percentage of values over the 95th observed percentile closer to 5%) than with dynamical downscaling only. However, for winter temperature extremes, ALADIN gives better score than CDF-t(ALADIN). For summer temperature extremes, CDF-t(ALADIN) only brings a gain on average since the spatial variability of the percentages of extremes, indicated by the size of the box, is much larger than with ALADIN alone.

Concerning the propagation of uncertainty with a combined statistical/dynamical downscaling, we can observe that the “additional” added value of CDF-t applied to RCMs outputs is independent of the quality of the field downscaled with the RCMs. For instance, while ALADIN displays better CvM results for wind speed than LMDz, CDF-t applied to those RCMs produces similar CvM scores (Fig. 2a). This is also true for %ev and %rv (Fig. 3e and f). Note also that some “bad” %ev and %rv values from the combined downscaling are due to stationary or non-stationary properties of the observed data between calibration and evaluation periods while “input” (ERA-40 or RCMs) data show opposite properties. For example, many winter time series of temperature from SAFRAN clearly indicate non-stationary evolutions of their CDFs, while those winter time series of temperature from ALADIN show stationary properties. CDF-t is nevertheless driven by those inputs: If their distribution does not evolve while observations do (or the opposite), it is clear that CDF-t projections will be misled. In other words, the quality of CDF-t strongly depends on the quality of its inputs.

The evaluations tools employed up to now in this study characterize “temporal” CDFs and variances. The CDF-t approach is not designed to correct the spatial correlation since it is applied location per location, i.e. in a univariate context and we did not expect CDF-t to improve the spatial variability of the RCMs. To verify this point, for each variable, the correlation has been calculated for every pair of stations and the resulting values were plotted in function of the distance separating the stations. The results (not shown) confirm that, as expected, CDF-t does not improve the spatial variability of the RCMs simulations since the spatial correlations of the combined statistical/dynamical downscaling results are very close to those of the initial RCMs data. However, the good point is that it does not deteriorate the spatial correlation either. Also, if the spatial correlation between any two stations is not improved by CDF-t (seen for example through variogram analyses, not shown), the spatial structure of statistical properties (such as means or quantiles of given probabilities) is improved since CDF-t makes the distributions of the RCM

data (and so their marginal statistical properties) more similar to those of the observations.

One major finding is thus, that despite the added value of dynamical downscaling to retrieve finer scale meteorological patterns and improve wind speed, temperature and precipitation distributions with respect to ERA-40, the use of a combined statistical/dynamical downscaling approach does not necessarily degrade too much nor necessarily improve the results when compared to those produced by CDF-t directly applied to ERA-40 reanalyses. In other words, refining “too much” the CDFs provided to CDF-t does not necessarily imply improved downscaled CDFs. Indeed, CDF-t applied to ERA-40 reanalyses already provides quite satisfactorily results in term of CvM “distance”, variances, and simulation of extremes. The same conclusion does not necessarily hold for other large-scale data, such as GCM outputs, where a spatial refinement through RCMs may improve the quality of the CDF-t downscaling.

5 Conclusions

In this study, ERA-40 reanalyses, and simulations from three regional climate models (ALADIN, LMDZ, and WRF) and from one statistical downscaling model (CDF-t) are used to evaluate the uncertainty in downscaling wind speed, temperature, and rainfall for eight stations in the French Mediterranean basin from 1991–2000. The Cramer-von Mises score (CvM) is employed to measure the “distance” between those distributions, and the ability of the three regional climate models and CDF-t to simulate the “climate” variability is quantified with the explained variance, variance ratio and extreme occurrence. The main general conclusions are as follows:

- Despite their differences, the three RCMs display very similar performance.
- In terms of global distributions (i.e. CvM), all models perform better than ERA-40 for both seasons and variables.
- However, looking at variance criteria, RCMs are surprisingly not always “much” better than ERA-40 reanalyses.
- CDF-t shows relatively good results for all tested criteria when applied to ERA-40.

More specifically, concerning the comparison RCM vs. CDF-t, we can also conclude that, when applied to ERA-40 (or GCM) data, RCMs tend to improve the spatial correlations while CDF-t does not change much the spatial dependence. Hence, if a specific study focus on the spatial structure, RCMs should be favoured over CDF-t. However, this may be different with a different SDM, such as a multi-sites statistical model (e.g. Harpham and Wilby, 2005;

Frost et al., 2011). On the other hand, if marginal (i.e. one dimension) distributions are of interest, the CDF-t approach is certainly more adapted than RCMs. The results concerning the extremes simulated by the three RCMs driven by ERA-40 are contrasted and depend on the variable studied, the season and the RCM itself. CDF-t shows more stability in reproducing correctly statistical properties of high values (of temperature, wind intensity and rainfall). In general, the discrepancies between observed data and CDF-t results are smaller than the discrepancies between observed data and RCM results. This is due to the fact that CDF-t (and potentially any SDM) is constructed to provide outputs that look like the observed climatology (at least over the calibration period). This feature is obviously not in RCMs that are based on physical processes.

In a second step, a combined statistical/dynamical downscaling approach has been used, consisting in applying CDF-t to the regional climate model outputs (only ALADIN and LMDZ for data availability reason). It shows that, based on the criteria used in this study:

- CDF-t applied to the RCM outputs does not necessarily produce better results than those from CDF-t directly applied to the ERA-40 reanalyses.
- CDF-t applied to the RCMs generally improves the downscaled CDFs and the “additional” added value of CDF-t applied to the RCMs is independent of the performance of the RCMs in terms of CvM, explained variance, variance ratio and extreme occurrence.

Note also that CDF-t applied to RCM outputs generally provides a better spatial structure (e.g. in terms of spatial correlation) than CDF-t applied to ERA-40 (not shown). Indeed, RCMs reproduce better the spatial variability than ERA-40 while CDF-t is not designed to correct the spatial correlations of the simulations.

Moreover, although the calibration and evaluation periods have been tested as significantly different through KS tests for a majority of stations, the CDF-t approach shows good stability and results. However, the periods selected (based on availability and quality of the data) are relatively short, especially for precipitation, and can potentially misrepresent the tail of the distributions, and so the CDF transformations, due to few extremes. This question of the length of data period for calibration and evaluation of any SDM is a very important and recurrent question that is left for future work.

Finally, this study confirmed some intuitive expected results but also showed more surprising results. The natural follow-up is to apply these methods to GCM simulations with coarser resolution than ERA-40 to test the robustness of the results of the present study obtained on the combined statistical/dynamical downscaling approach, and apply it to future GCM projections. Note however that representing correct distributions, variances or extreme properties in present

climate is not a guarantee of correctly representing the evolution of those properties in a climate change scenario.

Another comparison framework could also have been used. Indeed, the (dynamically and statistically) down-scaled data have a spatial density equal to or lower than $50 \times 50 \text{ km}^2$, while the station density is larger than that ($\sim 65 \times 65 \text{ km}^2$). We therefore have less than one observation per grid cell. To solve this, we could have used a very dense network of stations and aggregated them to the spatial resolution of the RCM outputs for comparisons. Hence, we would have one “aggregated” observation per grid cell, which would facilitate the comparisons. However, in the present study, the comparisons were performed based on cumulative distribution functions (CDFs) calculated over 10 yr from mean daily values. This temporal aggregation allows avoiding the high frequency fluctuations and makes data more comparable to RCM outputs. Also, the comparisons are made in terms of distributions and statistical properties, which reduce the potential spatial inconsistencies between RCM outputs and observations. However, this alternative comparison approach (i.e. using aggregated observations) will be made in a future work.

A relevant perspective is also to get use of gridded surface observations (such as daily ECA&D reanalyses, monthly CRU data or even the whole SAFRAN reanalyses database) to produce spatially resolved downscaled field. This is, of course, an evident product for dynamical downscaling but a less frequently used approach for statistical downscaling, although some studies took advantage of such gridded data (e.g. Déqué, 2007 or Quintana-Seguí et al., 2010).

Acknowledgements. We are thankful to E. Martin for providing the SAFRAN reanalysis data. This research has received funding from the ANR-MEDUP, ANR-McSIM and ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie) through contract ADEME 0705C0038, and from GIS “Climat Environnement Société” through the MORCE-MED project.

Edited by: G. Molinie

Reviewed by: S. Hwang and three anonymous referees



The publication of this article is financed by CNRS-INSU.

References

- Beaulant, A.-L., Joly, B., Nuissier, O., Somot, S., Ducrocq, V., Joly, A., Sevault, F., Déqué, M., and Ricard, D.: Statistico-dynamical downscaling for Mediterranean heavy precipitation, *Q. J. Roy. Astron. Soc.*, doi:10.1002/qj.796, 2011.
- Busuioc, A., Giorgi, F., Bi, X., and Ionita, M.: Comparison of regional climate model and statistical downscaling simulations of different winter precipitation change scenarios over Romania, *Theor. Appl. Climatol.* 86, 101–123, doi:10.1007/s00704-005-0210-8, 2006.
- Busuioc, A., Tomozeiu, R., and Cacciamani, C.: Statistical downscaling model based on canonical correlation analysis for winter extreme precipitation events in the Emilia-Romania region, *Int. J. Clim.*, 28, 449–464, 2008.
- Boé, J. and Terray, L.: A Weather-Type Approach to Analyzing Winter Precipitation in France: Twentieth-Century Trends and the Role of Anthropogenic Forcing, *J. Climate*, 21, 3118, doi:10.1175/2007JCLI1796.1, 2008.
- Cannon, A. J. and Whitfield, P. H.: Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models, *J. Hydrol.*, 259, 136–151, 2002.
- Carreau, J. and Vrac, M.: Stochastic downscaling of precipitation with neural network conditional mixture models, *Water Resour. Res.*, 47 W10502, doi:10.1029/2010WR010128, 2011.
- Cho, K. K., Bowman, K. P., and North, G. R.: A Comparison of Gamma and Lognormal Distributions for Characterizing Satellite Rain Rates from the Tropical Rainfall Measuring Mission, *J. Appl. Meteorol.*, 43, 1586–1597, 2004.
- Christensen, J., Boberg, F., Christensen, O., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophys. Res. Lett.*, 35, L20709, doi:10.1029/2008GL035694, 2008.
- Christensen, J., Kjellström, E., Giorgi, F., Lenderink, G., and Rummukainen, M.: Assigning relative weights to regional climate models: Exploring the concept, *Clim. Res.*, 44, 179–194, 2010.
- Colin, J., Déqué, M., Radu, R., and Somot, S.: Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique, *Tellus*, 62, 591–604, 2010.
- Darling, D. A.: The Kolmogorov-Smirnov, Cramer-Von Mises tests, *Ann. Math. Statist.*, 28, 823–838, 1957.
- Das, S. C.: The Fitting of a Truncated Log-Normal Curve to Daily Rainfall Data, *Australian J. Phys.*, 9, 151–155, 1956.
- Déqué, M.: 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, 2007.
- Déqué, M.: Regional climate simulation with a mosaic of RCMs, *Meteorol. Z.*, 19, 259–266, 2010.
- Déqué, M. and Somot, S.: Extreme precipitation and high resolution with Aladin, *Idjaras Quaterly Journal of the Hungarian Meteorological Service*, 112, 179–190, 2008.
- Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M., and van den Hurk, B.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, *Climatic Change*, 81, 53–70, doi:10.1007/s10584-006-9228-x, 2007.
- Drobinski, P., Carlotti, P., Newsom, R. K., Banta, R. M., Foster, R. C., and Redelsperger, J. L.: The Structure of the Near-Neutral Atmospheric Surface Layer, *J. Atmos. Sci.*, 61, 699–714, 2004.
- Drobinski, P., Bastin, S., Guénard, V., Caccia, J. L., Dabas, A. M., Delville, P., Protat, A., Reitebuch, O., and Werner, C.: Summer mistral at the exit of the Rhône valley, *Q. J. R. Meteorol. Soc.*, 131, 353–375, 2005.
- Ducrocq, V., Nuissier, O., Ricard, D., Lebeauapin, C., and Thouvenin, T.: A numerical study of three catastrophic precipitating events over Western Mediterranean region (Southern France). Part II: Mesoscale triggering and stationarity factors, *Q. J. R. Meteorol. Soc.*, 134, 111–130, 2008.
- Dudhia, J.: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, *J. Atmos. Sci.*, 46, 3077–3107, 1989.
- Durand, Y., Brun, E., Mériindol, L., Guyomarc'h, G., Lesaffre, B., and Martin, E.: A meteorological estimation of relevant parameters for snow models, *Ann. Glaciol.*, 18, 65–71, 1993.
- Durand, Y., Giraud, G., Brun, E., Mériindol, L., and Martin, E.: A computer-based system simulating snowpack structure as a tool for regional avalanche forecasting, *J. Glaciol.*, 45, 469–484, 1999.
- Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., and Mériindol, L.: Reanalysis of 44 years of climate in the French Alps (1958–2002): methodology, model validation, climatology and trends for air temperature and precipitation, *J. Appl. Meteorol. Climatol.*, 48, 429–449, 2009.
- Farda, A., Déqué, M., Somot, S., Horanyi, A., Spiridonov, V., and Toth, H.: Model Aladin as regional climate model for Central and Eastern Europe, *Studia Geophysica et Deodaetica*, 54, 313–332, 2010.
- Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Luthi, D., and Schär, C.: Soil moisture-atmosphere interactions during the 2003 European summer heat wave, *J. Climate*, 20, 5081–5099, 2007.
- Frei, C., Christensen, J. H., Déqué, M., Jacob, D., Jones, R. G., and Vidale, P. L.: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res.*, 108, 4124, doi:10.1029/2002JD002287, 2003.
- Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L.: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, *J. Geophys. Res.*, 111, D06105, doi:10.1029/2005JD005965, 2006.
- Frost, A. J., Charles, S. P., Timbal, B., Chiew, F. H. S., Mehrotra, R., Nguyen, K. C., Chandler, R. E., McGregor, J. L., Fu, G., Kirono, D. G. C., Fernandez, E., and Kent, D. M.: A comparison of multi-site daily rainfall downscaling techniques under Australian conditions, *J. Hydrol.*, 408, 1–18, doi:10.1016/j.jhydrol.2011.06.021, 2011.
- Geleyn, J.: Interpolation of wind, temperature and humidity values from model levels to the height of measurement, *Tellus*, 40A, 347–351, 1988.
- Ghosh, S. and Mujumdar, P. P.: Statistical downscaling of GCM simulations to streamflow using relevance vector machine, *Adv. Water Resour.*, 31, 132–146, 2008.
- Goubanova, K. and Li, L.: Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations, *Global Planet. Change*, 57, 27–42, 2007.
- Guénard, V., Drobinski, P., Caccia, J. L., Tedeschi, G., and Currier, P.: Dynamics of the MAP IOP-15 severe mistral event: observations and high-resolution numerical simulations, *Q. J. R. Meteorol. Soc.*, 131, 353–375, 2005.

- rol. Soc., 132, 757–778, 2006.
- Guldberg, A., Kaas, E., Déqué, M., Yang, S., and Vester Thorsen, S.: Reduction of systematic errors by empirical model correction: impact on seasonal prediction skill, *Tellus*, 57A, 575–588, 2005.
- Harpham, C. and Wilby, R. L.: Multi-site downscaling of heavy daily precipitation occurrence and amounts, *J. Hydrology*, 1–4, 235–255, doi:10.1016/j.jhydrol.2005.02.020, 2005.
- Haylock, M. R., Cawley, G. C., Harpham, C., Wilby, R. L., and Goodess, C.: Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and Statistical methods and their future scenarios, *Int. J. Climatol.* 26, 1397–1415, 2006.
- Herrmann, M., Somot, S., Calmanti, S., Dubois, C., and Sevault, F.: Representation of spatial and temporal variability of daily wind speed and of intense wind events over the Mediterranean Sea using dynamical downscaling: impact of the regional climate model configuration, *Nat. Hazards Earth Syst. Sci.*, 11, 1983–2001, doi:10.5194/nhess-11-1983-2011, 2011.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, *Clim. Dynam.*, 27, 787–813, 2006.
- Huth, R.: Disaggregating climatic trends by classification of circulation patterns, *Int. J. Climatol.*, 21, 135–153, 2001.
- Janjic, Z. I.: The step-mountain coordinate: Physical package, *Mon. Wea. Rev.*, 118, 1429–1443, 1990.
- Janjic, Z. I.: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP meso model, *Off. Note*, 437, 61 pp., Natl. Cent. For Environ. Predict., Camp Springs, Md., 2002.
- Kain, J. S. and Fritsch, J. M.: A one-dimensional entraining/detraining plume model and application in convective parameterization, *J. Atmos. Sci.*, 47, 2784–2802, 1990.
- Kain, J. S. and Fritsch, J. M.: Convective parameterization for mesoscale models: The Kain-Fritsch scheme, in the Representation of Cumulus Convection in Numerical Models, *Meteorol. Monogr.*, 24, 165–170, Am. Meteorol. Soc., Boston, Mass., 1993.
- Kallache, M., Vrac, M., Naveau, P., and Michelangeli, P.-A.: Non-stationary probabilistic downscaling of extreme precipitation, *J. Geophys. Res. Atmos.*, 116, D05113, doi:10.1029/2010JD014892, 2011.
- Laprise, R.: Regional climate modelling, *J. Comput. Phys.*, 227, 3641–3666, 2008.
- Lavaysse, C., Vrac, M., Drobinski, P., Lengaigne, M., and Vischel, T.: Statistical downscaling of the French Mediterranean climate: assessment for present and projection in an anthropogenic scenario, *Nat. Hazards Earth Syst. Sci.*, 12, 651–670, doi:10.5194/nhess-12-651-2012, 2012.
- Lebeaupin Brossier, C. and Drobinski, P.: Numerical high-resolution air-sea coupling over the Gulf of Lions during two tramontane/mistral events, *J. Geophys. Res.*, 114, D10110, doi:10.1029/2008JD011601, 2009.
- Lebeaupin, C., Ducrocq, V., and Giordani, H.: Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts, *J. Geophys. Res.*, 111, D12110, doi:10.1029/2005JD006541, 2006.
- Le Moigne, P.: Description de l'Analyse des Champs de Surface sur la France par le Système SAFRAN (Description of the Analysis of Near-Surface Atmospheric Fields over France with SAFRAN System), CNRM/GAME report, 77 [available from CNRM/GAME, Météo-France/CNRS, Toulouse, France], 2002 (in French).
- Li, Z. X.: Ensemble atmospheric GCM simulation of climate inter-annual variability from 1979 to 1994, *J. Climate*, 12, 986–1001, 1999.
- Louis, J.: A parametric model of vertical eddy fluxes in the atmosphere, *Bound.-Lay. Meteorol.*, 17, 187–202, 1979.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change. Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, doi:10.1029/2009RG000314, 2010.
- Mass, F. C., Ovens, D., Westrick, K., and Colle, B. A.: Does increasing horizontal resolution produce more skillful forecast?, *B. Am. Meteorol. Soc.*, 83, 407–430, 2002.
- Michelangeli, P. A., Vrac, M., and Loukos, H.: Probabilistic downscaling approaches: application to wind cumulative distribution functions, *Geophys. Res. Lett.*, 36, L11708, doi:10.1029/2009GL038401, 2009.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: A validated correlated k-model for the longwave, *J. Geophys. Res.*, 102, 16663–16682, 1997.
- Oetli, P., Sultan, B., Baron, C., and Vrac, M.: Are regional climate models relevant for crop yield prediction in West Africa?, *Environ. Res. Lett.*, 6, 014008, doi:10.1088/1748-9326/6/1/014008, 2011.
- Piani, C., Haerter, J. O., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, *Theor. Appl. Climatol.*, 99, 187–192, 2010.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchistéguy, L., and Morel, S.: Analysis of near surface atmospheric variables: validation of the SAFRAN analysis over France, *J. Appl. Meteorol. Climatol.*, 47, 92–107, 2008.
- Quintana Seguí, P., Ribes, A., Martin, E., Habets, F., and Boé, J.: Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins, *J. Hydrology*, 383, 111–124, doi:10.1016/j.jhydrol.2009.09.050, 2010.
- Radu, R., Déqué, M., and Somot, S.: Spectral nudging in a spectral regional climate model, *Tellus*, 60A, 885–897, 2008.
- Raje, D. and Mujumdar, P. P.: A comparison of three methods for downscaling daily precipitation in the Punjab region, *Hydrol. Process.*, 25, 3575–3589, doi:10.1002/hyp.8083, 2011.
- Rummukainen, M.: State-of-the-art with regional climate models, *WIRE (Wiley Interdisciplinary Reviews: Climate Change)*, *Adv. Rev.*, 1, 82–96, 2010.
- Ruti, P. M., Marullo, S., D'Ortenzio, F., and Tremant, M.: Comparison of analysed and measured wind speeds in the perspective of oceanic simulations over the Mediterranean basin: Analyses, QuikSCAT and buoy data, *J. Mar. Syst.*, 70, 33–48, doi:10.1016/j.jmarsys.2007.02.026, 2007.

- Salameh, T., Drobinski, P., Vrac, M., and Naveau, P.: Statistical downscaling of near-surface wind over complex terrain in Southern France, *Meteorol. Atmos. Phys.*, 103, 243–256, 2009.
- Salameh, T., Drobinski, P., and Dubos, T.: The effect of indiscriminate nudging time on large and small scales in regional climate modelling: Application to the Mediterranean basin, *Q. J. R. Meteorol. Soc.*, 136, 170–182, 2010.
- Sanchez-Gomez, E., Somot, S., Josey, S., Dubois, C., Elguindi, N., and Déqué, M.: Evaluation of the Mediterranean sea water and heat budgets as simulated by an ensemble of high resolution regional climate models, *Clim. Dyn.*, doi:10.1007/s00382-011-1012-6, 2011.
- Schär, C., Vidale, P. L., Luthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in the European summer heatwaves, *Nature*, 427, 332–336, 2004.
- Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, 26, 679–689, doi:10.1002/joc.1287, 2006.
- Schmidli, J., Goodess, C. M., Frei, C., Haylock, M. R., Hurrell, J. W., Ribalaygua, J., and Schmith, T.: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps, *J. Geophys. Res.*, 112, D04105, doi:10.1029/2005JD007026, 2007.
- Shabalova, M. V., Van Deursen, W. P. A., and Buishand, T. A.: Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model, *Clim. Res.*, 23, 223–246, 2003.
- Simmons, A. J. and Gibson, J. K.: The ERA-40 project plan, ERA-40 project Rep. Series 1, ECMWF, Shinfield Park, Reading, United Kingdom, 63 pp., 2000.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y. W., Wang Powers, J. G.: A description of the advanced research WRF Version 3, 125 pp., NCAR Technical Note NCAR/TN-475+STR, 2008.
- Sotillo, M., Ratsimandresy, A. W., Carretero, J., Bentamy, A., Valero, F., and Gonzalez-Rouco, F.: A high resolution 44-year atmospheric hindcast for the Mediterranean Basin: contribution to the regional improvement of global reanalysis, *Clim. Dynam.*, 25, 219–236, doi:10.1007/s00382-005-0030-7, 2005.
- Spak, S., Holloway, T., Lynn, B., and Goldberg, R.: A comparison of statistical and dynamical downscaling for surface temperature in North America, *J. Geophys. Res.*, 112, D08101, doi:10.1029/2005JD006712, 2007.
- Vautard, R., Yiou, P., D'Andrea, F., de Noblet, N., Viovy, N., Casou, C., Polcher, J., Ciais, P., Kageyama, M., and Fan, Y.: Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit, *Geophys. Res. Lett.*, 34, L07711, doi:10.1029/2006GL028001, 2006.
- Vidal, J. P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyrou, J. M.: A 50-year high-resolution atmospheric reanalysis over France with the SAFRAN system, *Int. J. Climatol.*, 30, 1627–1644, 2010.
- Vigaud, N., Vrac, M., and Caballero, Y.: Probabilistic downscaling of GCMs scenarios over southern India, *Int. J. Climatol.*, online first, doi:10.1002/joc.3509, 2012.
- Vrac, M. and Yiou, P.: Weather regimes designed for local precipitation modelling: Application to the Mediterranean basin, *J. Geophys. Res.-Atmos.*, 115, D12103, doi:10.1029/2009JD012871, 2010.
- Wilby, R. L., Dawson, C. W., and Barrow, E. M.: SDSM - a decision support tool for the assessment of regional climate change impacts, *Environ. Modell. Softw.*, 17, 145–157, 2002.
- Wilks, D. S. and Wilby, R. L.: The weather generation game: a review of stochastic weather models, *Prog. Phys. Geogr.*, 23, 329–357, 1999.
- Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern United States, *J. Geophys. Res.*, 107, 4429, doi:10.1029/2001JD000659, 2002.
- Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change*, 62, 189–216, 2004.
- Yang, C., Chandler, R. E., and Isham, V. S.: Spatial-temporal rainfall simulation using generalized linear models, *Water Resour. Res.*, 41, W11415, doi:10.1029/2004WR003739, 2005.