DESCRIPTION OF THE FICLIMA STATISTICAL DOWNSCALING METHODOLOGY

TABLE OF CONTENT

1. THE SIMULATION OF FUTURE CLIMATE

General Circulation Models (GCMs) are the most powerful tools for producing future climate projections (Huebener et al., 2007). However, GCMs currently operate at spatial resolutions of about 200 km and this low resolution is frequently unsuitable as input for impact models (von Storch, 1994; Mearns et al., 1997). These impact models are essential for designing adaptation policies that seek to minimize the negative impacts of climate change and to exploit the positive ones. For this reason, an important effort has been put into the development of strategies to infer high-resolution information from low-resolution variables, i.e., 'sensibly projecting the large-scale information on the regional scale' (von Storch et al., 1993). All these strategies fall into the overall denomination of *downscaling techniques*.

There are two main downscaling approaches (Murphy 1999; Fowler et al. 2007). In the so-called dynamical downscaling, high-resolution fields are obtained by nesting a Regional Climate Model (RCM) into the GCM (Giorgi et al., 2001; Christensen et al., 2007; Giorgi et al., 1994; Jones et al., 1997), or using a GCM with variable resolution (stretching technique) (Déqué and Piedelievre, 1995). In the statistical approach, high-resolution predictands are obtained by applying the relationships identified in the observed climate between these predictands and large-scale predictors to the GCM output (Wilby et al., 2004; Imbert and Benestad, 2005).

Dynamic and statistical approaches have advantages and disadvantages. Both need assumptions that cannot be verified in a climate change context (Giorgi et al., 2001) contributing to the uncertainty cascade leading to the climate simulations. Several criteria can be used to assist in the selection of the most suitable approach depending on the application (Wilby et al., 2004).

Climate scenario uncertainties have to be considered in a risk assessment framework and it needs to be done through probabilistic climate projections. In this regard, statistical methods seem to provide a good downscaling option, because they need relatively less GCM driving data and computational resources for processing the growing number of available GCM simulations. In addition, when high resolution (local) information is demanded, statistical methods can perform better than dynamical ones (Van der Linden and Mitchell 2009), due to the present coarse resolution of nested or stretched models (and to the fact that RCMs do not use local observations which capture local meteorology). Higher diagnostic capacity of statistical methods at the local scale is generally accepted in the meteorological operational forecasting framework, where statistical reinterpretation systems are the main tool for obtaining local information.

In the last decades, long-term statistics of climate have experienced relatively small changes compared to inter-annual variability. This variability offers an indirect way to assess the stability in a future climate context of statistical relationships used for downscaling. In this regard, two requirements have been identified for statistical downscaling performance: (i) performance should be good at different time scales (daily, seasonal, annual, decadal...) (Wilby and Wigley, 1997) and (ii) almost all of predictor situations in the GCM future climate must be within the applicability range of the statistical relationships determined for the calibration period of the method.

Considering drawbacks and advantages of both methods, FIC researchers decided to develop a statistical downscaling technique named *FICLIMA* (Ribalaygua et al, 2013). FICLIMA has been tested against other techniques in different scientific projects, both European[1] and National (Spain)[2] and applied in Europe and Latin America for different meteorological variables (temperature, precipitation, etc) with excellent results.

Statistical downscaling techniques establish empirical relationships between fields of low resolution, called *predictors*, and surface variables, called *predictands*. Such techniques can be classified into four groups: generators of time (simulated stochastically, i.e. intrinsically, series of daily values consistent with climatology), transfer functions (based on regression models, both linear and nonlinear, establishing long scale relationships between local predictands and predictors), selection of analogues (consisting in selecting from among a set of data the "n" atmospheric patterns more similar to the day problem) and types of time (based on a pre-qualification of a finite number of groups obtained as a synoptic similarity between the fields of low resolution).

The FICLIMA method combines two types of statistical techniques: the first step is a selection of analogs (or analog stratification), and in the second step some transfer functions are used on the days selected in the previous step.

The purpose of this paper is to present an overview of the FICLIMA methodology and its rationale, in any case, if a deeper explanation of the theoretical method is needed then it is possible to access to

http://www.ficlima.org/generacion-de-escenarios-locales/

where you can find more complete information on the methodology and some concrete examples of international projects in which it has been applied.

[1] Statistical and Regional dynamical Downscaling of Extremes for European regions (STARDEX, 2002-2005), Climate change and its impacts at seasonal, decadal and centennial timescales (ENSEMBLES, 2004-2009).

[2] First National Generation of Downscaled Future Climate Scenarios (2009), National Coordinated Programme for the generation of Downscaled Scenarios of Climate Change (ESTCENA, 2009-2012).

2. GENERAL AND THEORETICAL CONSIDERATIONS ON STATISTICAL DOWNSCALING

The development of a statistical downscaling methodology and the selection of predictors is based on theoretical considerations, in which four basic aspects need to be considered:

1. The stationarity problem: in a climate change scenario the relationships between predictors and predictands might change. Therefore, predictors should always be physically linked to predictands (because these linkages will not change) and consider the physical forcings of these predictands.
2. The characteristics and limitations of the GCMs: the methodologies to be developed will be finally applied to GCM outputs. Predictors selected should be well simulated by the GCMs and temporal and spatial resolution of the GCM should also be considered.
3. The statistical tool must reflect strong non-linear relationships that link predictors with most local surface weather predictands.
4. For climate change applications it is advisable not to use seasonal stratification in the selection of predictors: in climate change scenarios, climatic characteristics of calendar seasons may change. Thus the predictors / predictands relationships detected in a group of days with concrete climatic characteristics belonging to a specific season, would not be applicable for future days if climatic characteristics of the season have changed.

According to these aspects, the following conditions for selection of predictors have been identified:

1. Selection of predictors should be undertaken based on theoretical considerations, rather than using empirical analyses (which could result in non-physically based relationships that might not be applied in the future due to the stationarity problem). Predictors should be physical forcings of the predictands or, at least, physically linked to the predictands. Furthermore, identified links between predictors and predictands should be those that best reflect the physical links between them. If these requirements are fulfilled, a good diagnostic capability will be obtained at the daily scale.
2. Regarding GCMs, predictors should be:
    - Field variables rather than point values, because the former are more reliably simulated by GCMs.
    - Free-atmosphere rather than boundary layer variables because the former are more reliably simulated by GCMs.
    - Variables that are well simulated by GCMs. FICLIMA downscaling method have been adapted from a methodology used on a daily basis to produce operational meteorological forecasts. Many predictors are used in operational forecasts because they improve the forecasting skills, but some of them cannot be used in climate simulations because are too dependent on initial conditions to be well simulated by GCMs for the next decades.
    - Adapted to the scales in which GCMs provide information. Working with coarser temporal and / or spatial detail than those provided by GCMs means that some information is not used. Many of the physical forcings of the predictands can only be captured working at detailed temporal and spatial resolutions. This could be especially relevant for the simulation of some extreme precipitation events, which sometimes are produced by small and short-life atmospheric structures (convective structures). For

these reasons, downscaling methodologies should work at daily and synoptic resolutions, those provided by GCMs.

3. The statistical method should include strategies to take into account the non-linearity of the relationships between many of the predictors and predictands.
4. It is advisable not to make any seasonal stratification in the identification of the relationships between predictors/predictands. Sensitivity analyses performed with FICLIMA downscaling method shows that seasonal stratification does not improve forecast skills (because relationships reflect the physical links between predictors and predictands).

The development of the FICLIMA statistical downscaling methodology and the selection of the chosen predictors have taken into account the conceptual framework presented above.

3. THE *FICLIMA* DOWNSCALING METHODOLOGY

In general terms, the FICLIMA methodology estimates high-resolution surface meteorological fields for a day "x" (the problem day) in two steps: the first step is an analogue technique (Zorita et al. 1993); in the second step, high-resolution surface information is estimated in a different way for precipitation (using a probabilistic approach) and for temperature (using multiple linear regressions). A scheme of the methodology is shown in figure A.3.1.

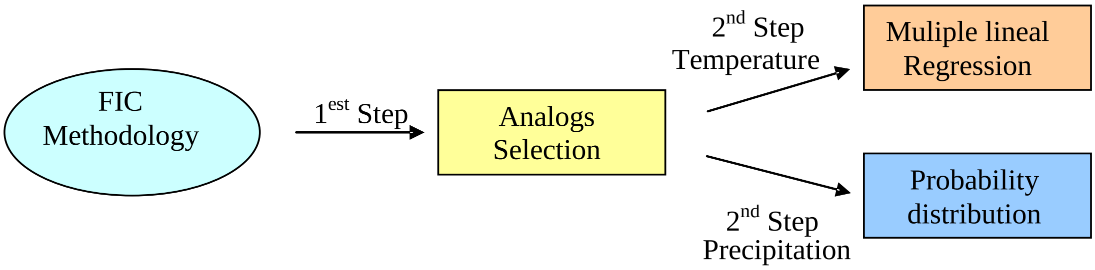

Figure A.3.1. Scheme of the FICLIMA methodology.

Similar two-step approaches have been applied in operational forecasting (Woodcock 1980; Balzer 1991). For climate change applications, Enke and Spekat (1997) adopted a similar technique, but where the first step of analogue stratification is replaced by stratification using a predefined clustering of atmospheric patterns. Analogue techniques can be considered as a special form of the clustering approach, where a specific type is determined for each problem day, containing the *n* most analogous days. This strategy greatly reduces the variability within a predefined cluster, which includes days with quite different atmospheric configurations. As a result, analogue techniques generally offer higher diagnostic capability regarding high resolution effects than do predefined clustering schemes. Figure A.3.2 describes a general representation of the FICLIMA methodology.

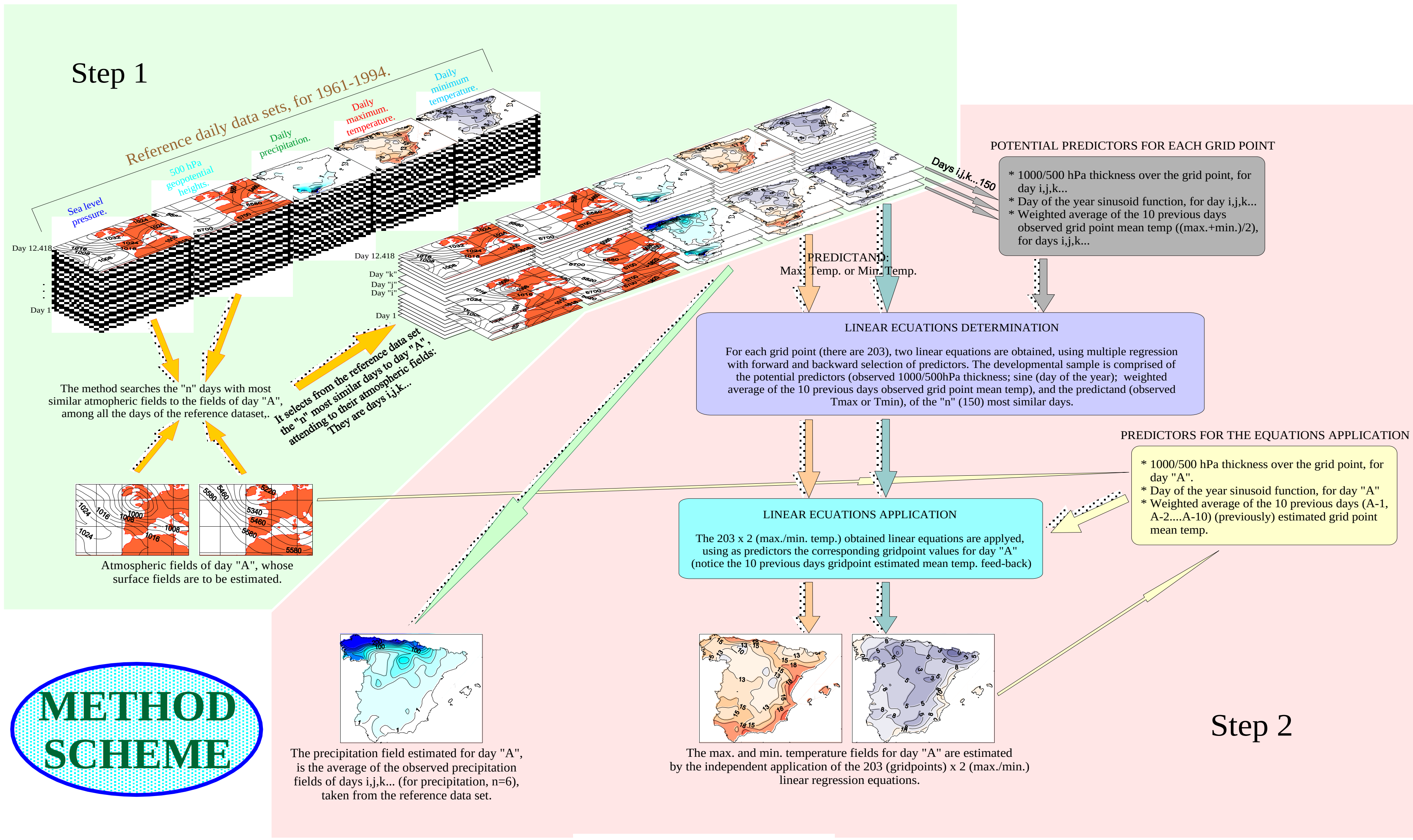

Figure A.3.2 General scheme of the FICLIMA methodology (represented for the Iberian Peninsula).

3.1 FIRST STEP: THE ANALOGUE TECHNIQUE

In the first step, the *n* most similar days to day “x”, identified on the basis of their low-resolution atmospheric fields, are selected from a reference dataset. The skill of the method depends on the spread and quality of the atmospheric and surface reference datasets and, in particular, on the measure used to determine the similarity between days (Matulla et al, 2008). Consequently, according to the ideas mentioned above, the similarity measure must contain diagnostic capability regarding high-resolution precipitation fields (i.e., low-resolution atmospheric fields considered to be similar according to the measure must be associated with similar high-resolution precipitation fields). Thus the similarity measure must assess the likeness of as many as possible precipitation physical forcings (see first aspect addressed in previous section) associated with the low resolution atmospheric configurations of the days being compared. In addition to diagnostic capability, the predictor variables of the measure must be reasonably well simulated by GCMs (see idea 2).

Many statistical methods entail strongly automated procedures to select the best predictors and to adjust the optimum predictors/predictand relationships. This is not, however, easy for analogue techniques for which calibration entails a laborious task of testing different combinations of predictors and similarity measures. Nevertheless, this allows the selection of predictors and similarity measures under theoretical considerations, with the aim of capturing physical forcings between predictors and predictands in order to guarantee the stationarity of the relationships (see first aspect addressed in previous section).

The similarity measure between two days must be a scalar magnitude (to allow ordering) and summarizes the resemblance of these two days with regard to their predictor fields.

Different algorithms which have traditionally been used to assess similarity between fields were tested in the calibration process: Pearson correlation coefficients and several Euclidean and pseudoeuclidean distances. Similarity measures were required to not only deal with the general pattern of the days being compared, but also with the values of the corresponding individual points of both fields. For the latter requirement, Pearson correlation coefficients perform worse than Euclidean distances and thus provide lower precipitation diagnostic capabilities. The good performance of Euclidean distances is supported by the analogue technique literature (Martin et al. 1997; Kruizinga and Murphy 1983).

The similarity between two days is calculated by determining (and standardizing) independently those days likeness with respect to each of the four final predictors fields. The unlikeness of days x_i and x_j regarding each predictor field P is calculated as a pseudoeuclidean distance with (eq. 1):

$$D_p(x_i, x_j) = \sqrt{\frac{\sum_{k=1}^{N}(P_{ik} - P_{jk})^2 \cdot W_k}{\sum_{k=1}^{N} W_k}} \qquad equation\ [1]$$

Where P_{ik} is the value of the predictor *P* of the day x_i at the grid point *k*; W_k is the weighting coefficient of the *k* grid point; and *N* is the number of the atmospheric grid points.

Once $D_p(x_i, x_j)$ has been calculated, it must be standardized. The standardization is carried out by substituting $D_p(x_i, x_j)$ by percentile *p*, which is the closest centile of the reference population of Euclidean distances among predictor fields *P* to the $D_p(x_i, x_j)$ value. The centile values are previously determined, independently for each *P* predictor field, over a reference population of more than 3,000,000 values of D_p. The reference population is calculated by applying equation 1, with the same W_k values, to randomly selected pairs of days. If the closest value to $D_p(x_i, x_j)$ is $cent_{i,j,P}$, it means that about the $cent_{i,j,P}$ % of the 3,000,000 of D_p values are lower than $D_p(x_i, x_j)$. The use of centile instead of the original distance D_p allows consideration of dimensionless and initially equally weighted variables for each predictor *P* in the measure.

After the $D_p(x_i, x_j)$ independent calculation and standardization (determination of the closest $cent_{i,j,P}$), the final similarity ($sim_{i,j}$) measure between days x_i and x_j is given by the inverse of a weighted average of the centile for the *P* predictors (eq. 2):

$$sim_{i,j} = \left(\sum_{P=1}^{4} w_p cent_{i,j,P}\right)^{-1} \qquad equation\ [2]$$

where w_p is the weighting coefficient of the predictor field *P*.

Once the theoretical methodology has been set, *P* predictors will be selected for the study area (step 2).

3.2 SECOND STEP

3.2.1 Temperature: Multiple Linear Regression Analysis

The procedure of estimation for temperatures requires, after the selection of the *n* analogous days described above (for temperature, *n* = 150), further diagnosis using multiple linear regression. Although predictor/predictand relationships determined in this second step are not linear, an important part of the non-linearity of the links between free atmosphere variables and surface temperatures is reduced with the first step (analogue) stratification, which selects the most similar days with respect to precipitation and cloudiness (two of the variables which introduce most non-linearity in the relationships). Linear regression performs quite well for the estimation of surface maximum and minimum temperatures due to the near-normal statistical distribution of these variables. It is important to remember that when using linear regression the predictand quantity is bound to have essentially the same statistical distribution as the predictor(s) variable(s) (Bürguer 1996). In this regard, potential predictors should possess close-to-normal distributions.

Multiple linear regression is performed independently for each surface point, and uses forward and backward stepwise selection of predictors. We use at least four potential predictors:

1. 1000/500 hPa thickness above the surface station.
2. 1000/850 hPa thickness above the surface station.
3. The solar irradiation of the day in the year associated to the studied day; it depends not only of the day in the year but also of the latitude of the studied station.
4. A weighted average of the daily mean of the temperatures of the ten previous days at the studied station.

Both thicknesses are used to include the strong relationship between lower troposphere and surface temperatures (a meteorological factor). The solar irradiation of the day of the year is used to consider the number of sunlight hours and its effect on the warming/cooling of the surface air (a seasonal factor). And the ten days weighted average of temperature is used to account for the soil thermal inertia influence (a soil memory factor).

The non-linear influence of other important meteorological factors, such as cloudiness, precipitation and low troposphere wind speed, is considered through the first-step of analogue stratification. The regression is performed for a population of *n* days which present very similar precipitation, and subsequently very similar cloudiness, conditions.

For each station (and each problem day) the regression is performed twice using as predictands maximum and minimum temperatures. Thus two diagnostic equations are calculated (using the predictand and predictor values of the *n* analogous days population) and applied to estimate both daily temperatures for each station and problem day.

3.2.2 Precipitation: Probabilistic Approach

Every problem day (x_i) has *n* analogues (a_{ij}), each one with a certain similarity ($sim(a_j, x_i)$) (*n*=30 for precipitation). Each analogue (a_{ij}) has an observed precipitation (ρ_{ij}) and an estimated probability (π_{ij}), according to equation 3.

$$\pi_{ij} = \frac{sim(a_j, x_i)}{\sum_{k=1}^{n} sim(a_k, x_i)} \qquad equation[3]$$

Thus each problem day (x_i) has n pairs of [ρ_{ij}, π_{ij}], and a preliminary estimate of precipitation (p_i) can be obtained by combining the n pairs according to equation 4.

$$p_i = \sum_{j=1}^{n} \rho_{ij}\pi_{ij} \qquad equation[4]$$

Since it is calculated as an average, this preliminary estimate greatly smoothes the extreme values of precipitation and underestimates the number of dry days.

In order to solve this problem the methodology designs an approach to obtain (e.g. a month) precipitation time-series for a certain period with a probability distribution similar to that obtained for the precipitation of all the analogues associated with the problem days of that period.

In this approach, for a problem month with m problem days, there are $n\times m$ pairs of precipitation and probability [ρ_{ij}, π_{ij}]. These pairs are sorted by ρ_{ij}, and ranked according to equation 5, until they form groups of k pairs whose sum of probabilities π_k is 1; in this way we obtain m new precipitation values (p'_h).

$$p'_h = \sum_{k|\sum \pi_k = 1} \rho_k \pi_k \qquad equation[5]$$

The m new precipitation values (p'_h) are assigned to the m days (x_i) of the month according to the preliminary precipitation estimates (p_i) obtained by equation 4, so that the highest p'_h is associated with the day (x_i) with the highest p_i; the second highest p'_h with the day with the second highest p_i; and so on.

Proceeding this way, the probability distribution of the m new precipitation values (p'_h) is similar to the probability distribution of $n\times m$ values of precipitation (ρ_k) - as desired. This method allows an empirical distribution of rain amounts for each day of the month to be constructed without assuming any *a priori* hypothesis about the probability distribution of each month (or assuming a particular associated analytical probability function such as the gamma function).

4. REFERENCES

- Balzer K (1991) Recent improvements in medium-range local weather forecasting in the Deutscher Wetterdienst. In: Lectures of the WMO training workshop on the interpretation of NWP products in terms of local weather phenomena and their verification, pp. 225-229
- Bürguer G (1996) Expanded Downscaling for Generating Local Weather Scenarios. Clim Res 7:118-28
- Christensen JH, Carter TR, Rummukainen M, et al. (2007) Evaluating the performance and utility of regional climate models: the PRUDENCE project. Clim Change 81:1-6
- Déqué M, Piedelievre JP (1995) High resolution climate simulations over Europe. Clim Dyn 11:321-339
- Enke W, Spekat A (1997) Downscaling climate model outputs into local and regional weather elements by classification and regression. Clim Res 8:195-207
- Fowler, H. J., Blenkinsop, S. and Tebaldi, C. (2007), Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol., 27: 1547–1578. doi: 10.1002/joc.1556
- Giorgi F, Shields Brodeur C, Bates GT (1994) Regional climate change scenarios over the United States produced with a nested regional climate model. J Climate 7:375-399
- Giorgi F. and Francisco R. (2001) Uncertainties in the prediction of regional climate change. Global Change and Protected Areas 9:127-139
- Huebener H, Cubasch U, Langematz U, Spangehl T, Niehörster F, Fast I and Kunze M (2007) Ensemble climate simulations using a fully coupled ocean–troposphere–stratosphere general circulation model. Phil Trans Roy Soc A 365: 2089–210. doi:10.1098/rsta.2007.2078.
- Imbert A. and Benestad R. (2005) An improvement of analog model strategy for more reliable local climate change scenarios. Theor Appl Climatol 82:245–255. doi:10.1007/s00704-005-0133-4
- Jones RG., Murphy JM., Noguer M., and Keen B. (1997) Simulation of climate change over Europe using a nested regional-climate model. II: Comparison of driving and regional model responses to a doubling of carbon dioxide. Q J R Meteorol Soc 123:265-292
- Kruizinga S. and Murphy AH. (1983) Use of an Analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands. Mon Wea Rev 111:2245-2254
- Martin E., Timbal B., and Brun E. (1997) Downscaling of general circulation model outputs simulation of the snow climatology of the French Alps and sensitivity to climate change. Clim Dyn 13:45-56.
- Matulla C., Zhang X., Wang XL., Wang J., Zorita JE., Wagner S. and von Storch H. (2008) Influence of similarity measures on the performance of the analog method for downscaling daily precipitation, Climate Dynamics, 30,133-144
- Mearns, LO. (1997) On the statistical evaluation of climate model experiments - Comment. Clim Change 37:443-448
- Murphy J. (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. J Climate 12:2256-2284

- Ribalaygua, J., Torres, L., Pórtoles, J., Monjo, R., Gaitán, E. and Pino, M.R. (2013) Description and validation of a two-step analog/regression downscaling method. Theoretical and Applied Climatology. Doi: 10.1007/s00704-013-0836-x. URL: http://link.springer.com/article/10.1007/s00704-013-0836-x [2]
- Van der Linden P. and Mitchell JFB. (eds.) 2009: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, UK. 160pp. Available from http://ensembles-eu.metoffice.com/docs/Ensembles_final_report_Nov09.pdf (see. p 68)
- Von Storch H (1994) Inconsistencies at the interface of climate impact studies and global climate research. Max Planck Institute for Meteorology Technical Report 122.
- Von Storch, H., Zorita, E., and Cubasch U. (1993) Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime. Journal of Climate, Volume 6, pp. 1161-1171
- Wilby, RL. and Wigley, TML. (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr 21:530-548
- Wilby, RL., Charles SP., Zorita E., Timbal B., Whetton P., and Mearns LO. (2004) Guidelines for use of Climate Scenarios Developed from Statistical Downscaling Methods. Itergovernmental Panel on Climate Change (IPCC) Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA), available from http://ipcc-ddc.cru.uea.ac.uk/guidelines/StatDown_Guide.pdf.
- Woodcock, F. (1980) On the use of analogues to improve regression forecasts. Mon Wea Rev 108:292-297
- Zorita, E., Hughes, J., Lettenmaier, D. and Storch, Hv. (1993) Stochastic downscaling of regional circulation patterns for climate model diagnosis and estimation of local precipitation. Max Planck Institute for Meteorology Technical Report 109.