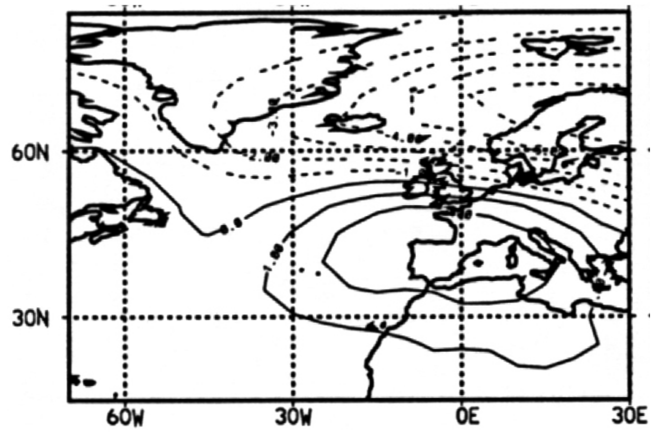
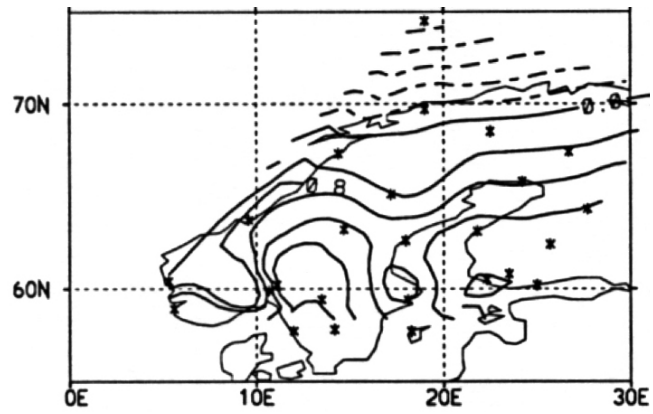
GKSS 97/E/20

A survey of statistical downscaling techniques

Pair 1





Authors:
E. Zorita
H. von Storch
(Institute of Hydrophysics)

GKSS-Forschungszentrum Geesthacht GmbH • Geesthacht • 1997

GKSS 97/E/20

A survey of statistical downscaling techniques

E. Zorita, H. von Storch

44 pages with 10 figures

Abstract

The derivation of regional information from integrations of coarse-resolution General Circulation Models (GCM) is generally referred to as downscaling. The most relevant statistical downscaling techniques are described here and some particular examples are worked out in detail. They are classified into three main groups: linear methods, classification methods and deterministic non-linear methods. Their performance in a particular example, winter rainfall in the Iberian peninsula, is compared to a simple downscaling analog method. It is found that the analog method performs equally well than the more complicated methods. Downscaling analysis can be also used as a tool to validate regional performance of global climate models by analyzing the covariability of the simulated large-scale climate and the regional climates.

Ein Überblick über statistische Regionalisierungsverfahren

Zusammenfassung

Die Ableitung regionaler Information aus Integrationen grob aufgelöster Klimamodelle wird als „Regionalisierung" bezeichnet. Dieser Beitrag beschreibt die wichtigsten statistischen Regionalisierungsverfahren und gibt darüber hinaus einige detaillierte Beispiele.
Regionalisierungsverfahren lassen sich in drei Hauptgruppen klassifizieren: lineare Verfahren, Klassifikationsverfahren und nicht-lineare deterministische Verfahren. Die Methoden werden auf den Niederschlag auf der iberischen Halbinsel angewandt und mit den Ergebnissen eines einfachen Analog-Modells verglichen. Es wird festgestellt, dass die Ergebnisse der komplizierteren Verfahren im wesentlichen auch mit der Analog-Methode erzielt werden können.
Eine weitere Anwendung der Regionalisierungsmethoden besteht in der Validierung globaler Klimamodelle, indem die simulierte und die Kovariabilität zwischen dem großskaligen und dem regionalen Klima miteinander verglichen wird.

Manuscript received / Manuskripteingang in der Redaktion: 23. Mai 1997

1 Introduction.

One of the most important tools in the study of climate variability and climate change are General Circulation Models (GCMs). These models are state-of-the art numerical coupled models that represent several subsystems of the earth's climate (atmosphere, oceans, sea-ice, land surface processes) that are thought to be capable to simulate the large scale state of the climate. At planetary scales, GCMs are able to simulate reliably the most important mean features of the global climate, for instance the Intertropical Convergence Zones, the three dimensional atmospheric circulation cells, the jet streams, etc. With some limitations, they also simulate reasonably well essential features of the ocean circulation like the western boundary ocean currents and the conveyor belt driven by the thermohaline circulation. Some of the latest GCMs also produce ENSO- type atmosphere-ocean coupled variability in the Pacific basin. With respect to the interannual variability, it has been found that some GCMs also reproduce satisfactorily the most important patterns of variability of the atmospheric flow and of the sea surface temperatures (SST) at mid-Iatitudes. However, at finer spatial resolutions, with scales of a few grid distances, climate models have much smaller skill (Grotch and MacCracken, 1991). Many examples of the deficiencies of the global GCMs in simulating basic local climatic variables like surface-air temperature and precipitation have been presented, two of which will be mentioned here.

- A detailed comparison among the regional performance of several low-resolution GCMs in the Mediterranean basin can be found in Cubasch et al. (1996). Therein, it was concluded that the skill of the models in simulating the observed climate is much higher for near-surface air temperature than for precipitation, but that even for the former variable clear discrepancies are detected. With respect to climate change the responses simulated by the models to a doubling of atmospheric CO_2 concentration are not univocal. In some cases two versions of the same atmospheric model coupled to a different ocean model produce temperature change patterns with the reversed sign. Concerning the changes in simulated precipitation each model actually predict patterns that are quite different to one another.

- Another example is provided by Risbey and Stone (1996), who analysed the performance of the Climate Community Model model with T42 and T106 resolutions in the Sacramento

River basin in California. They found that although the model reproduces the right mean annual rainfall, its probability distribution differs markedly from the observations: whereas the simulated rainfall occurs mainly in form of drizzle distributed of many rainy days, the observed rainfall is measured in much stronger precipitation events distributed over much fewer days.

The fact that the models do a credible job on the global scale and fail on the regional scale seems to be a contradiction. However, the global climate is to a great extent the response to the differential solar forcing, earth rotation and the large-scale structure of the earth's surface (land-sea distribution, topography). The regional climates, on the other hand, are the response of the global climate to regional details. Therefore, it seems reasonable to simulate the global climate adequately even though none of the regional climates is simulated skillfully.

There are at least three reasons for the failure of the models on this regional or local scale:

- The spatial resolution provides an inadequate description of the structure of the earth's surface. The land-sea distribution is heavily smeared out: the mountains appear as broad flat hills. For spectral models the truncated representation of the topography is also a source of additional difficulties which may be severe at the local scale (Lindberg and Broccoli, 1996). A clear example is provided by the real annual cycle of precipitation in the Alps: in the northern side a summer rainfall maximum is observed, whereas some hundreds of kilometers southward a winter maximum is apparent (Fliri, 1974). It is reasonable to think that it will be quite difficult for the GCMs to simulate properly those small-scale features of the actual climate and therefore the climate change assessment at those scales will have to be considered with care.

- The hydrodynamics of the atmosphere are non-linear and the energy, which is fed into the system at the cyclonic scale, is cascaded through nonlinear interactions to the smallest scales. Because of the numerical truncation this cascade is interrupted and the flow to smallest scales is parameterized. This parametrizations affect the smallest resolved scales most strongly.

- Sub-grid scale processes in the models, such as cloud formation, rainfall, infiltration, evaporation, runoff, etc., are all parametrized. These processes are calculated by means of bulk formulae, the parameters of which may not have been fitted for the region of interest. These parametrizations may include additional errors in the GCM simulation. There are indications (Risbey and Stone, 1996; Machenhauer et al., 1996) that this may be the most important source of error of the GCMs, perhaps even more than its inadequate resolution.

However, these subgrid processes are actually those with the greatest ecological or societal impact, since they strongly affect the local climate at the scales of the human and ecological environment. Therefore, there is a broad consensus about the need to simulate the subgrid-process and the local climates properly, perhaps beyond the capabilities of the current GCMs.

1.1 Strategies to bride the scale gap.

The efforts to improve GCM simulations have been directed in two directions. In the last years, with increasing computer power, there has been a clear tendency to finer and finer GCM horizontal resolutions. For instance, while in 1990 a T21 resolution (about 5.6° x 5.6°) for the atmospheric submodel was considered as state-of-the-art, some of the last integrations with atmospheric models have been carried out with a resolution of T106 (about 1.125°). This resolution is however quite costly and for climate change estimation the applications so far have been restricted to the “time slice modus”. In this modus the atmospheric high-resolution model is forced by the mean boundary conditions simulated in a low-resolution atmosphere-ocean coupled model. On the other hand the use the so called limited area models (LAMs) (Giorgi and Mearns, 1991), is becoming more frequent and is being also applied to the ocean component of the climate model (Kauker, personal communication). These LAMs are sophisticated atmospheric (or oceanic) models of a limited geographical area (of the order of 10^7 km^2) with a resolution of the order of 20-50 km, which use the large-scale fields simulated by the GCMs as boundary conditions, but that take the regional characteristics, such as topography, into account. Although an increased resolution in the region of interest usually improves the simulation of orographycally induced precipitation and the cyclonic activity at mid-latitudes is better reproduced (Machenauer et al., 1996), it may not go hand in hand with an improved regional climate simulation. For instance, the LAMs developed at the UK Meteorological Office (UKMO) and in Meteo-France (based on the UKMO GCM and the ARPAGE GCM, respectively) show systematic errors that are not solved by increasing the resolution. These are probably associated with the parametrizations of sub-grid processes, which are taken over from the parent GCMs and with the large-scale errors of the coarse-resolution GCMs themselves (Machenhauer et al., 1996). Therefore, there seems to exist also a need not only for finer resolutions, but also for better sub-grid parametrizations (Risbey and Stone, 1996).

The alternative approach to overcome the scale mismatch between the skill of climate simulations and the needs of ecosystem and sector models is statistical downscaling. This technique is becoming quite popular, due to relative simplicity and lower costs compared to the use of LAMs. This paper is focused on the description of some types of statistical downscaling and a discussion of the their most important merits and caveats.

1.2 Statistical downscaling.

Essentially the idea of the statistical downscaling consists in using the observed relationships between the large-scale circulation and the local climates to set up statistical models that could translate anomalies of the large-scale flow into anomalies of some local climate variable (Storch 1993). There exist quite different statistical models, depending on the nature of the local variable.

An important assumption that underlies the statistical approach to climate impact assess· ment is that the link between the large-scale circulation and the local climate remains unchanged in an altered climate, which is by no means guaranteed. However, if the time series used to tune the statistical model are long enough it is reasonable to assume that they contain many different situations, including those that will be stronger or more probable in an altered climate. If these situations are important for the local climate, the statistical model should be able to identify them in the historical observations and estimate with some skill the probable impact on the local climate. This assertion is of course only valid if the expected shift in the mean state lies roughly within the natural variability of the present climate, which is the information used by the statistical model. If changes in the mean climate are larger than the observed natural variability the estimation via statistical downscaling may still be useful but it should be considered with care. This draw-back is in some sense also present in climate change estimation with GCM experiments, since these models contain many parametrizations which in principle are only valid for the present climate. However, the functional form of these parametrization schemes are in many cases based on sensible physical reasoning, so that they hopefully will remain more or less valid in an altered climate.

To avoid to some degree the uncertainties of GCMs and LAMs simulations, the ability of these models to simulate past climates is, or should be, tested. The counterpart requirement for statistical downscaling models is that they should be able to reproduce the historical evolution of the local variables when they are driven by the observed large-scale circulation in the past. However, the statistical models are seldom so good that they can replicate with accuracy the recorded historical evolution at the local scale. Often, other factors that are not present in the fitting period, like changes in the station location 01' in measurement procedures, hinder the validation of the model. These should ideally be identified and corrected when possible. More interesting in the context of climate change, is the replication of the low- frequency (decadal and longer) natural variability. The low-frequency natural variations in the form of trends or oscillations can be considered as natural climate changes and a good statistical model should be able to reproduce them.

A recent example of the validity of the statistical downscaling approach has been presented by Busuioc and Storch (1996). These authors fitted a statistical model linking the atmospheric circulation and rainfall in Romania with data from a control run of the ECHAM3 climate model with an atmospheric resolution of T42 (about 2.5°). Then they used this statistical model to estimate changes in precipitation in a CO_2 climate based on the atmospheric circulation changes simulated by a CO_2 experiment. The estimated precipitation changes were consistent with those changes directly simulated at the grid-point level by the ECHAM 3 model at a T106 resolution (about 1.125°). Therefore, it can be concluded that their statistical downscaling model could bridge the scale gap in the world simulated by the ECHAM 3 model.

1.3 By-products of statistical downscaling.

The application of statistical downscaling is not restricted to the context of climate impact studies, but it can also be used for the validation of GCMs and LAMs. A statistical downscaling model describing the relationship between the local variable and the large-scale circulation in the observed climate can give useful insights about the physical causes of the natural variability of the local climate variable. In the ideal case this relationship should be also found in the

GCM/LAM simulations. The identification of a *wrong* statistical relationship in the simulation the absence of any, may help find the origin of the model deficiencies in particular regions and suggest improvements for the parametrizations of local processes. This type of validation is seldom carried out and we would like to underline its importance with the evaluation of the skill with which two climate models reproduce the well known connection between the North Atlantic Oscillation (NAO, van Loon and Rogers, 1978), the seasaw of the sea level pressure between Azores and Iceland in the North Atlantic in wintertime, and near-surface air temperatures in Scandinavia (see section 3)

Another useful by-product of a downscaling application is the verification of the historical data quality and the reconstruction of past local conditions. If the statistical downscaling model can be physically interpreted and it is successful in the replication of past climate features, such as trends or oscillations, their real existence will gain support. For instance, von Storch et al., (1993) found that the long-term upward trend of winter Iberian rainfall could be explained by a counterpart trend in the North-Atlantic SLP field. They reasonably concluded that both trends were real and not an instrumental artifact (see section 7).

Also, downscaling models can be a useful tool to fill data gaps of the local time series or even represent a first estimation of lacking information in relatively prolonged periods. It happens quite often that data observational networks present inhomogeneities due to changes of location or measuring techniques. The use of a downscaling model would yield in principle an estimation of the missing data which would be consistent with the large-scale forcing.

The paper is divided into 10 sections. Through the examples presented in this paper several data sources and climate model integrations have been used. These are described in section 2. In section 3 an example of the application of downscaling to the validation of climate models is presented. Section 4 deals with a general formalism of statistical downscaling models. In the following sections several families of statistical models are presented. To some degree it is a matter of taste to classify statistical models into families, but it seems that in downscaling literature three types of models have been used so far. These are the analog method (section 5), classification methods (section 6), linear methods (section 7) and neural networks (section 8). The problem of the level of simulated variability is discussed in section 9. A section with final considerations and an outlook closes the paper (section 10).

2 Data description

The sea-level pressure data that represent the observations originate in the National Center for Environmental Research (NCEP, previously NMC, Washington DC) analysis provided by National Center for Atmospheric Research (NCAR, Boulder). They have been used with a resolution of 5°x5° degrees. The monthly means of air-surface temperature in Scandinavia belong to the World Station Climatology data set provided by NCAR. The daily rainfall data in Spain were kindly supplied by the Instituto Nacional de Meteorologia (Madrid), and the monthly rainfall records by the Universidad Complutense (Madrid, Spain). The control integrations with the coupled atmosphere-ocean ECHAM1-LSG and the atmospheric model ECHAM3 were carried out at the Deutsches Klimarechenzentrum (Hamburg). The ECHAM1 model was run with a T21 resolution (about 5.6°x5.6°) and coupled to the LSG ocean model. A detailed description of this control run can be found in Cubasch et al. (1992). The integration with the atmospheric model ECHAM3 was run at a resolution T42 (about 2.8°x2.8° in a time slice modus, where the boundary conditions for sea-surface temperature and sea-ice distribution were taken from a ten-year me an of the ECHAM1-LSG control run (Cubasch et al., 1996).

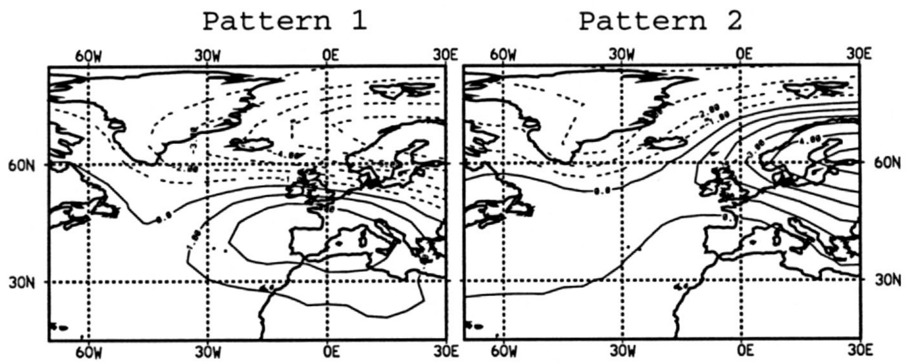
3 Statistical downscaling for the validation of climate model integrations.

In this section the skill of two climate models to reproduce the well-known relationship between the North-Atlantic SLP field and Scandinavian air temperature in winter (DJF) is investigated by means of Canonical Correlation Analysis (CCA). This CCA identifies patterns of two fields, in this case pairs of SLP and Scandinavian temperature patterns, that tend to appear simultaneously (see section 7 for a more formal description). Two pairs of canonical patterns of observed SLP and Scandinavian air temperature are shown in Figure l.

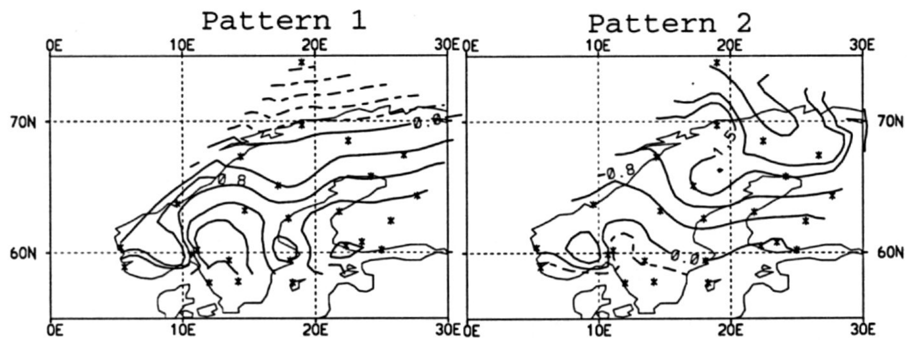
Both canonical pairs of patterns may be physically explained by temperature advection by the large-scale circulation. For instance, the first SLP pattern, describes the variations of the strength of the zonal flow over the North Atlantic that brings milder or colder winter temperatures to the whole Scandinavian peninsula, and is related to the NAO pattern. The CCA also identifies a weaker, more local, relationship between both fields, represented by the

Figure 1: *The first two canonical correlation patterns of observed winter (DJF) Scandinavian air temperature (bottom) and simultaneous SLP in the North Atlantic area (top). The correlation between the time coefficients is .73 and .63 respectively.*

[North Atlantic winter SLP (hPa)]



[Scandinavian air temperature (K). The temperature patterns explain 56% and 28% 01 the variance, respectively. The stars indicate the station positions.]



second SLP canonical pattern. This second SLP pattern describes the strength of advection anomalies from higher latitudes to the Northern part of the peninsula and of continental origin to the southern part.

One of the main advantages of the method based on CCA is that it delivers spatial patterns that normally lend themselves to a clear physical interpretation. In this sense it is also simple to investigate if the GCM is able to reproduce the observed linear relations hip between both fields and therefore to infer a subjective level of confidence of the performance of the model simulations in a sensitivity experiment. In Fig 2 the same CCA analysis between DJF North Atlantic SLP and simultaneous Scandinavian near surface temperature is shown for a GCM simulation of the present climate with the low-resolution (T21) ECHAM1-LSG coupled climate model. It can be seen first that the correlations between the time coefficients are much weaker than in the observations. Concerning the SLP patterns, they look qualitatively different: none of them can be associated to the strength of the zonal flow in the North Atlantic and their physical interpretation is not straightforward. The first pair seems to be associated with blocking activity over Scandinavia; the second pair may however be explained by air-temperature advection. Therefore, it is to be expected that this model will not properly simulate changes in Scandinavian temperature in a future climate, although surface air temperature is perhaps the variable best simulated by climate models.

A more recent version of the ECHAM atmospheric model is the ECHAM3. In this version several changes were introduced, in particular a different gravity-wave drag parametrization that considerably improved the simulation of the atmospheric circulation over the North Atlantic. The same CCA analysis based on a control integration of this model with a resolution T42 (about 2.8°x2.8°) run in a “time slice” modus are shown in fig 3. The canonical patterns based on the T42 simulations are much more realistic, and both pairs of canonical patterns are satisfactorily reproduced. It is therefore to be expected that the regional climate change estimation over Scandinavia with the ECHAM3 model will be much more credible than with the ECHAM1 model.

4 Statistical downscaling procedures.

The statistical downscaling strategy can be formulated in a general way as follows:

1. MODEL DESIGN

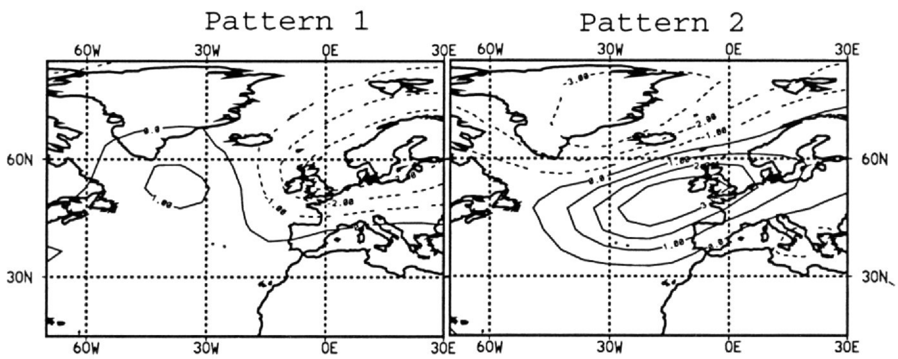
   (a) Identify regional climate parameter of interest, *R*.

   (b) Find climate parameter *L* which

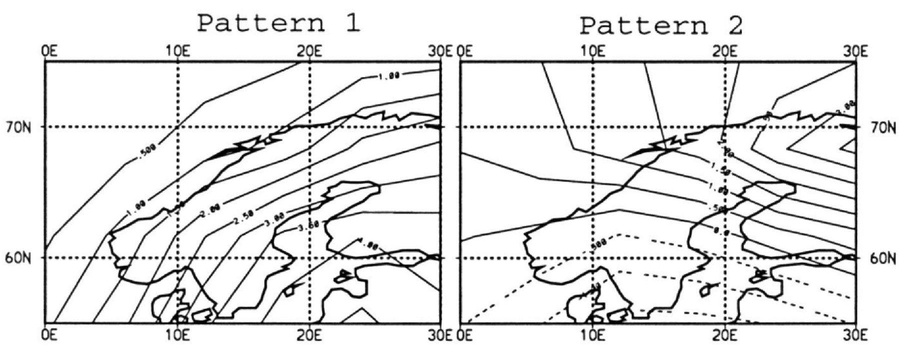
      i. controls *R by* $R = F(L, \alpha) + \varepsilon$ with a vector of unknown stochastic parameters $(\alpha_1 \ldots \alpha_m)$. The ε represents the part of *R* not described by F (see section 9).

Figure 2*: As in Figure 1, based on control simulations by the ECHAMI-LSG climate model (resolution T21, about 5.6°x 5.6°). The correlations between the time coefficients is .44 and .19.*

[North Atlantic winter SLP (mb)]



[Scandinavian air temperature (K). The temperature patterns explain 51 % and 21 % 01 the variance, respectively]



ii. is reliably simulated in a climate model.

(c) use paired samples (R, L) from historical records to fit α such that

$$\| \varepsilon \| R - F(L, \alpha \| = \min$$

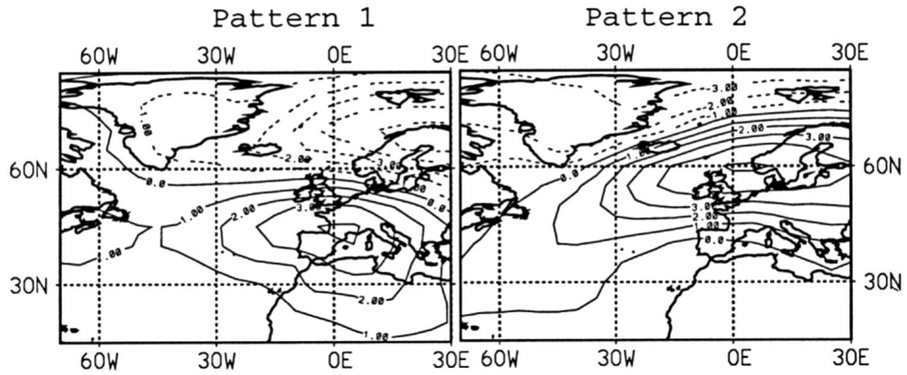
(d) Verify the fitted model $R = F(L, \alpha)$ by means of independent historical data.

2. MODEL APPLICATION

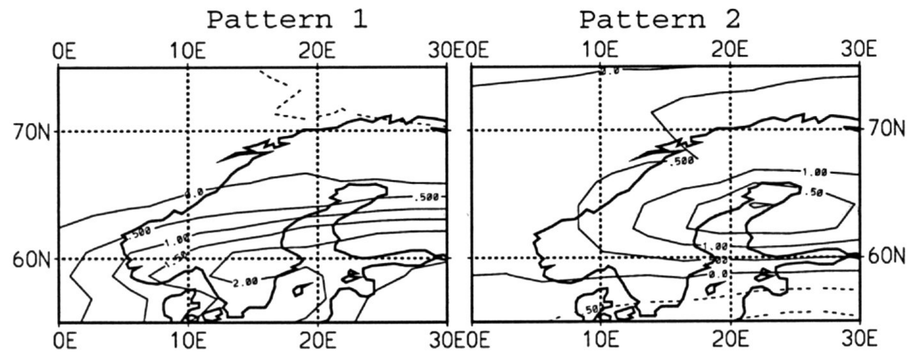
(a) Get the climate parameter L from the output of a climate model.

Figure 3*:* As in figure 1 but based on the ECHAM3 atmospheric model (T42, about 2.8° x 2.8°). The correlations between the time coefficients is .86 and .68.

[North Atlantic winter SLP (mb)]



[Scandinavian air temperature (K). The temperature patterns explain 42% and 17% of the total variance.]



(b) Check if the climate model reproduces the link between R and L

(c) Calculate R by $R = F(L, \alpha)$.

(d) Use R as forcing function for an impact model.

The most difficult step is the design of the statistical model. The type of model depends on the desired output needed for the climate scenario. Obviously, a model required to yield daily time series of temperature, rainfall, cloudiness, etc (sometimes called a weather generator) will be more complicated than a model set up to estimate changes in total annual precipitation.

The statistical models found in the literature can be roughly classified into three groups: linear methods, classifications methods and the quite recent applications of neural-networks. But perhaps the most simple technique is the analog method. The analog method is also a extreme form of the classification methods, hut due to its simplicity it can be taken as a kind of benchmark for the other methods. In the following sections the results of the analog method for a particular example (winter rainfall in Spain) are presented and compared with the results obtained with more complicated techniques.

5 The analog method.

The most simple downscaling scheme is the analog method. This method has been essentially applied in the field of weather forecasting (Lorenz (1969; Kruizinga and Murphy, 1983), and in short-term climate prediction (Barnett and Preisendorfer, 1978; van den Dool et al., 1994). For downscaling purposes the idea of the analog method is simple. The large-scale atmospheric circulation simulated by a GCM is compared to each one of the historical observations and the most similar, in a sense that has to be still defined, is chosen as its analog. The simultaneously observed local climate is then associated to the simulated large-scale pattern.

The most relevant problem associated with this method is the need for sufficiently long observations, so that a reasonable analog of the large-scale circulation can be always found. Due to the amount of degrees of freedom of the large-scale atmospheric circulation, it has been pointed out (Van den Dool et al., 1994) that on a global basis and for prediction purposes several thousand years would be needed to guarantee that an analog can always be found. However, many of these degrees of freedom represent just background noise that can be previously filtered out, for instance by a standard Empirical Orthogonal Function Analysis (EOF) and the area of interest is not global but normally covers a continent or an ocean basin. Furthermore for downscaling purposes the analog method is not used in a prediction modus, but rather for hindcast. In this case, analogs are indeed found for most of the downscaling applications that we have explored. In this slightly modified form, the anomalies of the atmospheric circulation, for instance represented by the anomalies of the SLP field f, are described by the few leading

EOF patterns:

$$f(i,t) = \sum_k x_k(t) g_k(i) \tag{1}$$

where i is a grid point index, t is the time step, g_k is the kth EOF pattern and x_k is the amplitude of this pattern at time t. The analogs are searched only within the space spanned by these EOF patterns.

Consider an atmospheric anomaly pattern Δf(r) (for instance the difference between a CO_2 and control simulation with a GCM or the anomaly between an observed circulation and the long-term mean). This pattern may have coordinates z_i in this EOF space. Its analog is defined as the month t that minimizes the distance in EOF space:

$$\sum_{k=1}^{n} \left\{ z_k - x_k(t) \right\}^2 \tag{2}$$

where $x_k(t)$) are the projection of the SLP anomalies onto the kth EOF pattern, and n is the number of the retained EOFs. The method can be generalised by introducing different weights d_k on the EOF coordinates:

$$\sum_{k=1}^{n} \left\{ d_k (x_k(t_0) - x_k(t))^2 \right\} \tag{3}$$

The weights d_k may be optimised, so that the normalized squared deviation E between the simulated and observed local rainfall is as small as possible:

$$E = \sum_{j=1}^{N} \sum_{t=1}^{T} \left\{ f_j^o(t) - f_j^s(t) \right\}^2 / \sigma_j^2 \tag{4}$$

where f_j^o *and* f_j^s are the observed and simulated rainfall at station j, respectively, and σ_j^2 is the observed variance at station j.

The optimization procedure can be considerably difficult, since the function E has in general a complicated topography. We will restrict here to the simpler method contained in (2), with n=5.

This method is illustrated in detailed by the following example. We are interested in the win-ter (DJF) precipitation over the Iberian Peninsula, given at a number of irregularly distributed meterological stations. It is assumed that this regional variable is controlled to a great extent by the atmospheric variability in the European-North Atlantic sector. The large-scale variable will be the SLP field, which for this purpose offers some advantages compared to geopotential height data. First, there exist long homogeneous time series of this variable that allow to set up the statistical model in some period and check it in an independent data set. Second, in climate change GCM experiments, geopotential heights tend to be much more affected by the global warming, but these changes may be related to changes in the mean atmospheric density and not necessarily to changes of the atmospheric circulation.

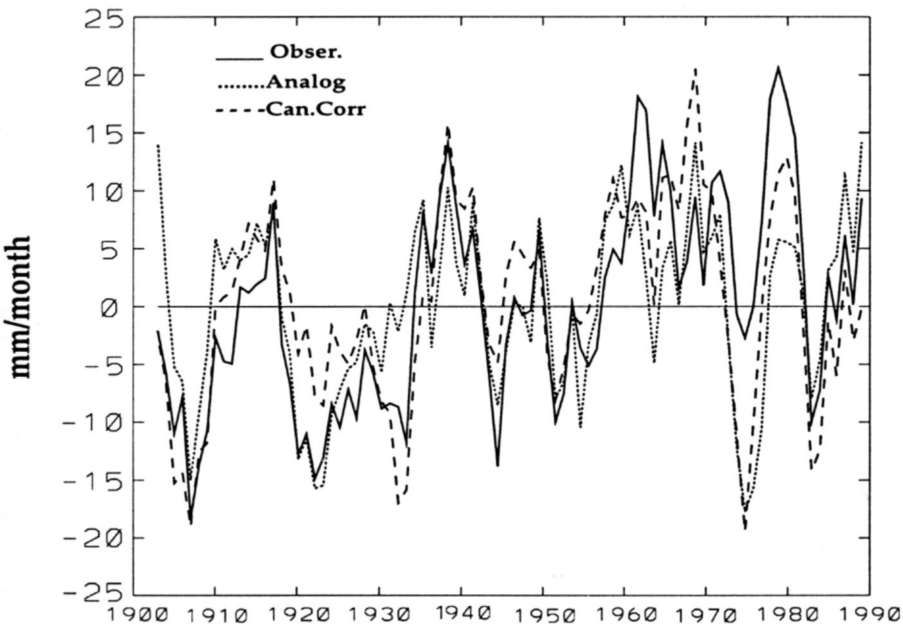
To check the quality of the analog method we try to reconstruct the time series of the winter (DJF) average precipitation in the period 1901-1989 by looking for analogs of the atmospheric circulation in the period 1951-1989. In the overlapping period the analog for a particular month is searched in the whole data set available, but always in a season different to the target pattern. Since the interannual autocorrelation of the atmosphere is negligible this procedure should amount to searching the analogs in an independent data set.

To reduce the number of atmospheric degrees of freedom the EOFs of the winter SLP are calculated. Then, the winter SLP anomalies are projected onto the first few EOFs. The results obtained by the analog method are subsequently presented in the following sections, compared directly with those achieved by the more complicated techniques.

6 Classification methods.

The general principle underlying the classification methods is also simple, although the practical implementation can become quite complicated. A classification scheme of the atmospheric circulation in the area of interest is developed and a pool of historical observations is distributed into the defined classes. The classification criteria are than applied to atmospheric circulations simulated by a GCM, so that each circulation can be classified as belonging to one of the classes. To each observed circulation there exist a simultaneous observation of the local variable. The value of the local variable to be associated with the simulated large-scale circulation can be chosen as either the average of all regional observations simultaneous to the elements of that class, or only the regional observations simultaneous to *one* element of the class, selected at random. Which of both strategies is best suited depends on the particular problem. For instance, if one is interested in simulating local daily rainfall, averaging over all the elements of a class will

Figure 4: Five-year-running mean time series of area-averaged winter (DJF) rainfall anomalies (mm/month) as derived from station data in the Iberian Peninsula (solid line); derived from the North Atlantic SLP pressure field: analog method (dotted line) and Canonical Correlation Analysis (dashed line)



in general lead to an underestimation of the local rainfall variance, of extreme events, and in general, to a local rainfall probability distribution different to the observed. On the other hand, averaging over several elements of the class will filter out more effectively measurements errors at that station.

Another possibility is to apply a fuzzy-logic approach (Ozelkan, 1996). Each circulation pattern belongs to *several* classes simultaneously, with an intensity that can be calculated by previously defined fuzzy-rules. The local variable is estimated as a weighted average of the mean local climates corresponding to each class.

The practical problem remains to define a method to classify the large-scale patterns. There are many classification methods. However, it should be pointed out that all classification schemes are to some degree subjective, although so me of them, once defined, allow for an automatic clas-ification of circulation patterns. In the most objective schemes only the number of independent classes has to be subjectively fixed at some stage of the model design.

The typical example of subjective classification schemes is the traditional Grosswetterlagen classification of the German Weather Service for Western Europe-North Atlantic sector. This classification has been used for downscaling purposes (Bardossy and Plate, 1992). Weather typing procedures have been developed by many national weather services for their particular regions based on the local expertise. Other well known subjective schemes are the Schuepp's scheme for Switzerland and the Lamb classification for the British Isles daily weather patterns. There exists an automatic procedure (Jones et al., 1993) that takes into account the empirical rules first proposed by Lamb.

More complicated technically is the design of an objective classification scheme. In this respect several types of schemes can be distinguished: classifications that only depend on the large-scale circulation data, classifications that depend on the local variable, and schemes that use information from both data. A typical example of the first and second group is traditional cluster analysis of atmospheric circulation patterns (Cheng and Wallace, 1993) or of the local variable. In the first case the classes of the large-scale circulation are given directly by the analysis. In the second case the elements of the large-scale circulation class are defined as the simultaneous circulation to each element of the local climate class. This second method has the advantage that the resulting large-scale classes should really correspond to different local situations, which is not necessarily the case in a classification based only on the large-scale patterns.

An example of a quite complicated classification based simultaneously on the large-scale circulation and the local variable is represented by CART (Classification and Regression Trees, Breimann et al., 1990) analysis. This method has been mainly applied to the simulation of local daily rainfall (Hughes et al., 1993, Schnur and Lettenmaier, 1997). However, due to its extensive needs of computer time, it has been applied to a limited number of stations and it has been assumed that rainfall is just a two-outcome process, wet or dry. The CART analysis searches recursively for a binary decision tree, whose decision nodes are based on the values of the atmospheric variables at some key locations, or the values of key atmospheric indices. Each terminal node of the tree represents a weather state. A weather state resulting from the CART analysis is such that the joint probability distribution (including all stations) based on the days belonging to that weather state is, in some sense, maximally different from the probability of the other weather states.

The weather states defined by any classification scheme can be also used to validate the performance of a GCM and eventually investigate the reason why a particular GCM may not be simulating properly the local climate in a certain region. Notwithstanding the fact that the weather classes are the result of a more or less subjective definition, and if these classes are not defined in a very restrictive manner, the GCMs in general are able to simulate reasonably their probabilities of occurrence (Hulme et al., 1993). Much more problematic is the simulation of the local climate associated with the individual large-scale classes, since the local climate is for instance quite sensitive to the exact location of anticyclones or deviations of storm tracks from the long-term mean. The best performance is usually found for surface temperature in the winter season, since this local variable is to a great extent determined by large-scale advection. The associated summer temperature and rainfall in both seasons, however, are normally not so satisfactorily reproduced.

If the large-scale circulation is classified on a daily basis, an important aspect of the validation is the dynamical behavior of these classes, for instance, their mean life times and transition probabilities. Since the weather typing is not usually defined from a dynamical point of view, this aspect of the validation can give more objective information on the ability of the models to simulate the regional weather. The GCMs do show some skill in reproducing the transition probabilities of the weather types (Zorita et al., 1995), but their performance has still to be considerably improved if realistic local weather time series are to be directly used in climate impact studies. Therefore, some of the deficiencies in the simulation of local weather scenarios lie clearly in the deficiencies of the GCMs to simulate the evolution of the large-scale atmospheric patterns. However, for daily rainfall there exists a more serious problem, which remains still unsolved. Downscaling procedures based on classification schemes and using the *observed* large-scale circulation produce daily rainfall time series with clearly less persistence than in the observation, i.e. the observed time-clustering of precipitation is not replicated by the downscaling techniques (Hughes et al., 1993). This problem can be partially, but not completely, reduced by simulating daily rainfall not only conditional on the daily weather-class, but also on the evolution in the previous few days (Zorita et al., 1995). Therefore, there seems to exist in the real rainfall process some kind of local persistence that presumably cannot be taken into account only by large-scale processes alone (Hughes et al., 1993). One possibility that should be worth exploring could be to assume that the rainfall events resulting from the passage of mid-latitude depressions follow some kind of spatial-temporal stochastic model, such as the Newmann-Scott or Barlett-Lewis models (Marroquín et al., 1996), and try to find a link between the adjustable parameters of these models and the monthly or seasonal circulation statistics, for instance, mean circulation and intramonthly variability. The daily rainfall amounts would be generated in a second step, once the stochastic model parameters have been estimated conditioned upon the large-scale state.

The classification methods can also be combined with a dynamical LAM for regional climate change estimation (Frey-Bunness et al., 1995). Once the large-scale circulation patterns have been classified, several integrations with a LAM model are carried out, in which the dynamical model is driven by boundary conditions that represent the large-scale circulation of each class. The local climates corresponding to each large-scale class simulated by the LAM are averaged, weighted by the probabilities with which each class occurs, either in the observations, in a control integration of a GCM or in a climate change experiment with a GCM. The advantage of this method is that the computation costs only depend on the number of classes. If one is interested in the broad features of the local climate, integrations with the boundary conditions of only the most frequent classes may be sufficient. However, for a more reliable representation of the local climate, for instance of the extreme events statistics, one is forced to carry out simulations also with the classes that appear only with low probability.

7 Linear methods.

Linear models are perhaps the most popular in the downscaling context. They apply the huge battery of already existing linear methods, for instance from the simple linear regression up to multivariate singular value decomposition, to the concept of *teleconnection*. One of the first physical teleconnections identified in climate research was the link between the North Atlantic Oscillation (NAO) and the surface-air temperature in Scandinavia (see introduction and Figure 1). This link can be applied to downscaling by setting up a linear regression between the anomalies of the NAO index (the SLP difference between Azores and Iceland) and the anomalies of the temperature in a Scandinavian station. The changes in the strength of the N AO in a future climate can be then translated to changes in local temperature by means of linear regression, apart from the global temperature increase.

The general idea of the linear methods is the same as in the above example, namely to link anomalies of the large-scale circulation to anomalies of the local climate. However, the technical complexity of the method can be considerably increased, as shown in this section in a more sophisticated treatment of the relationship between the North-Atlantic SLP and winter Iberian rainfall (von Storch et al., 1993).

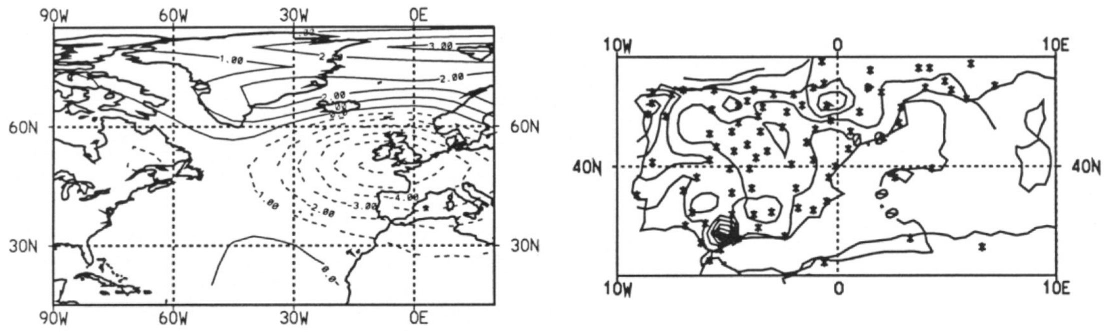
For this purpose we have chosen Canonical Correlation Analysis (CCA). Given two random vector time series x and y, in this case the winter rainfall at the Iberian stations and the SLP field defined on a grid over the North Atlantic, CCA identifies pairs of patterns whose time evolution is optimally correlated (Barnett and Preisendorfer, 1987; Bretherton et al., 1992). These spatial patterns are the eigenvectors of the matrices:

$$\begin{aligned} M_X &= C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx} \\ M_Y &= C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy} \end{aligned} \tag{5}$$

where C's are the respective cross-covariance matrices involving the variables of the SLP (x) and local temperature (y) time series. It can be shown that both matrices M_X and M_Y have the same non-zero eigenvalues c_k, which are the squared correlation between the time series associated with the kth eigenvector of M_X and M_Y. Otherwise the time series are pairwise uncorrelated. This does not necessarily mean that the processes represented by the different CCA patterns are physically independent but they normally represent, at least in a first approximation, different aspects of the variability. For more technical details about CCA and other similar techniques the reader is referred to Bretherton et al. (1992). Prior to the CCA analysis both fields, North-Atlantic SLP and Iberian rainfall, are filtered by standard EOF analysis to reduce the number of degrees of freedom of both fields. We will compare the results of the CCA with the analog method in the previous section, and therefore the same number of EOFs for the SLP field are retained, so that the same large-scale information enters both methods. For Iberian rainfall the first two EOFs, that describe about 80% of the total variance, are retained for the CCA.

The CCA identifies one dominant pair of patterns (Fig 5). The rainfall pattern has the same sign at all stations, with highest values near the Atlantic coast and decreasing values towards

Figure 5*: The first pair of canonical correlation pattern of observed winter (DJF) Iberian rainfall (isolines 25 mm/month) and simultaneous SLP (hPa) in the North Atlantic sector. The correlation between the time coefficients is .86. The rainfall pattern explains 50% of the variance.*



the Mediterranean. The associated SLP pattern consists of a low-pressure cell located over the British Isles. The canonical pair of patterns may be physically explained by advection by the large-scale circulation. For instance, the SLP pattern describes variations of the advection of air masses of Atlantic origin to the Iberian Peninsula.

The result of the CCA provides a method for estimating a regional rainfall anomaly $R_t(y)$ at a set of stations from a given large-scale pressure anomaly field $f_t\,(x)$ in a consistent way. Mathematically, this is accomplished in three steps: If the pairs of CCA patterns for the large-scale and local variable are denoted by p_k and q_k, respectively the first step is to calculate the amplitude a_k with which the k^{th} large-scale CCA pattern p_k appears in the new SLP data $f_t\,(x)$. This is achieved by minimizing the sum of squares:

$$E = \sum_{t,i}\left\{f_t(x_i) - \sum_k a_k(t)p_k(x_i)\right\}^2 \tag{6}$$

Equating to zero the derivatives of E with respect to a_k leads to a linear system of equation that can be solved by standard methods. The estimated rainfall anomalies at station y_j associated with f (t) is just the sum of the estimated amplitudes of the local CCA patterns.

$$R\left(y_j\right) = \sum_k c_k a_k(t) q_k\left(y_j\right) \tag{7}$$

where c_k are the canonical correlations. This procedure is, of course, capable of describing only the part of rainfall variance that can be traced to the atmospheric circulation. The implicit assumption is that the observed intermonthly SLP-rainfall relationship can be extrapolated to longer time scales, which is reasonable as long as the climate variations are considered small. We will later see that this condition is fulfilled.

The reliability of the suggested statistical relationship is tested by reconstructing the patterns of Iberian rainfall from the beginning of this century. The North Atlantic pressure field has undergone significant changes during this century (Hense et al., 1990; Schabbar et al., 1990) and this variability should have had an effect on rainfall. Note that the data prior to 1950 have not been used to perform the CCA and, thus, represent independent samples.

The area-averaged rainfall, obtained indirectly from the pressure fields and from the in-situ meteorological observations are displayed in Fig. 4. The two time series, both smoothed with a 5-year running-mean filter, vary coherently on all resolved time scales. Interestingly, the method is able to reproduce the low-frequency oscillations with a time-scale of about 20 years and the positive rainfall trend. This confirms that, to large extent, winter rainfall in this region is controlled by the large-scale circulation, and that the trend, as well as the inter-decadal variations, are real. This follows since the two data sets, rainfall and pressure, are from independent sources.

The linear methods can also been combined with other high-resolution empirical information to produce the desired downscaling. The model proposed by Stamm and Gettelmann (1995) lies in some sense between the purely empirical model described in the previous example and nested LAMs. These authors also focus on rainfall and take the large-scale variables that may have some influence, namely observed sea surface temperature and winds. They set up a numerical advection scheme that takes also into account high-resolution topographic data and that tries to estimate the precipitable water budget of an air parcel in its way towards the region or station of interest. Then, a regression equation is used to link this value with the observed rainfall. For climate change assessment the winds (and SSTs) simulated by a GCM are used to estimate the amount of precipitable water and then the regression equation is used to estimate local rainfall. This type of model, however, has not been yet profusely used.

An important difference between the linear method and classification methods is that the linear regression methods produce time series with smaller variance than the observations, which can be a quite serious problems for some ecological applications. This occurs because the linear models filter out the noise inherent in the observations that is not related to the atmospheric circulation. This problem is discussed in more detail in section 10.

A caveat of the linear methods is that they cannot be used directly when the local variables are not normally distributed. Since the variability of the large-scale atmospheric circulation is usually normally distributed (at least it is quite difficult to detect deviations from normality) the output of any linear method is bound to be also normally distributed. There are, however, local variables that strongly deviate from normality, the most important example being perhaps *daily* precipitation. The standard solution to this problem has been to transform the local variable so that the distribution of the transformed variable is approximately gaussian. This is always possible for an unbound random variable, but not for rainfall which is always bounded by zero. This may be not such a serious problem if the bound is sufficiently separated from the median in terms of the standard deviation, but this type of transformation would always induce additional errors. Perhaps more serious is the fact that such a variable transformation would probably shrink the long tails of a non- normal distribution (for instance the usual ($1+\log x$) transformation in the case of rainfall), but then the inaccuracy in the estimation of changes in the mean have to be expressed in physical units, and therefore has to be back-transformed, widening the confidence intervals.

In the next section we present another possibility to apply linear methods even if the local variables are non-normally distributed.

8 Linear methods applied to non-normally distributed local variables.

Some climate impact models (e.g., agricultural models) do not need time series of local forcing functions that are consistent with the simultaneous large-scale fields, but just a few statistical properties of those time series suffice: it is the mean rainfall, distribution of rainfall amounts, the distribution of the length of dry periods, etc., and their changes in a new climate, that are important. The idea is then to establish relationships between the large-scale fields and, for instance, the probability distribution or the distribution of storm interarrival times, etc. The interannual variability of these functions is expected to be more normally distributed, since they are calculated by suitable time-averages. In this case a linear technique may have chance to be successful.

In this example we focus on the probability distribution of daily rainfall and of the storm interarrival times of the station Cáceres in the Iberian Peninsula (30.5°N, 6.3°W) in winter (DJF), The probability of storm interarrival time $p_{sit}(\tau)$ is defined as the probability that the length of periods with no rain equals or exceeds τ days.

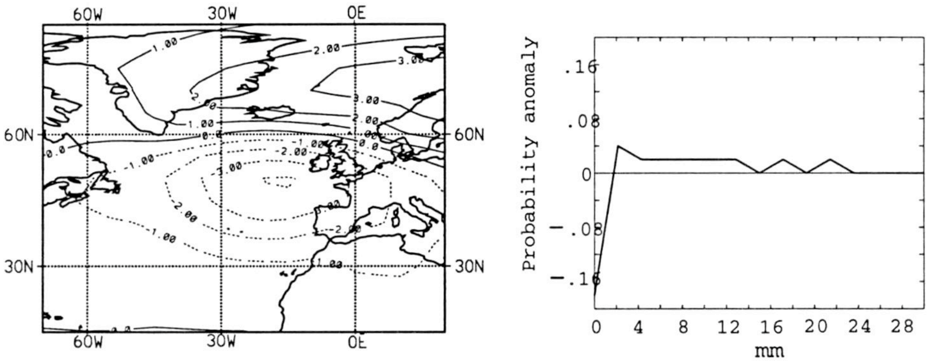
We choose in this case the *seasonally averaged* winter SLP fields as large-scale variables. Two CCA calculations have been performed: one in which the second field for the CCA is the winter daily rainfall distributions at this station; and another one in which the local field is the frequency of storm interarrival times. These frequency distributions are determined for each winter season from the observations. It is clear that the long-tails may not be reliably estimated. For instance, since each winter contains 90 days, the 98% quantile of the rainfall distribution depends on a single value, namely the highest amount in that winter. The 90% quantile depends on 9 values.

The CCA is then performed as in the previous section in the period 1965-85. The results of both calculations are shown in figure 6. The SLP canonical pattern looks similar in both cases to the result of the CCA analysis between SLP and monthly rainfall. This supports the validity of this approach. The canonical patterns of rainfall distribution and storm interarrival times indicate that the increase of mean rainfall that occurs when the SLP in the North Atlantic is lower than normal (Fig 5) and is mainly caused by an increase of days with weak rainfall at the cost of dry days, whereas the number of days with heavy rainfall (> 3 mm/day) remains almost unchanged. With respect to the storm interarrival times the canonical patterns indicate that the low pressure cell over the North Atlantic is associated with a reduction of the probability of moderate dry spells (< 5days), where the changes in the probability of longer dry spells is smaller.

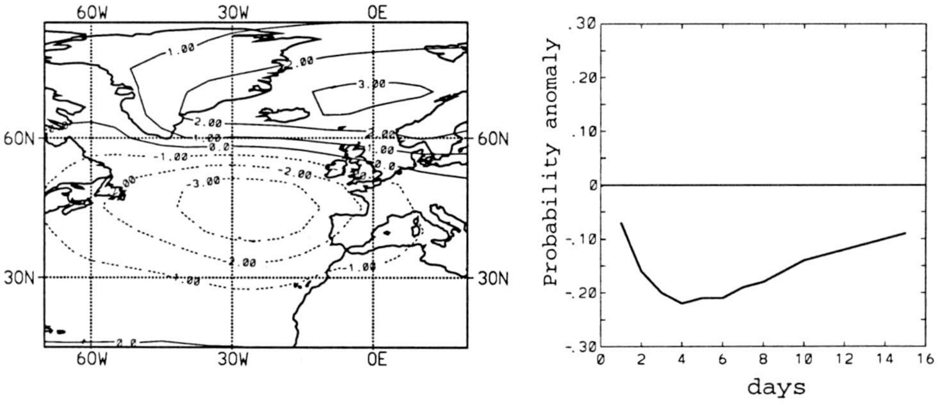
To validate the statistical model the distribution of interarrival times for the period 1942-1989 may be estimated based on the linear model and the SLP data in this period. We show in Figure 7 the evolution of the 5-day storm-interarrival time (i.e. the probability that a dry period would last a least 5 days), both from the observations and from the estimation with the CCA method. The analog method has been also applied to the same data set. In this case the analogs for the whole period are searched *on a daily basis* in the winter months between 1965-1985. For

Figure 6: Results of the two Canonical correlation analysis of seasonal North Atlantic SLP in winter time and daily rainfall amounts in Cáceres, Spain: SLP and daily rainfall probability distribution; and SLP and the probability of storm interarrival times.

[SLP (mb) and probability of daily rainfall amounts in mm/day. The correlation between the time coefficients is .87, the probability pattern explains 30% of the variance.]

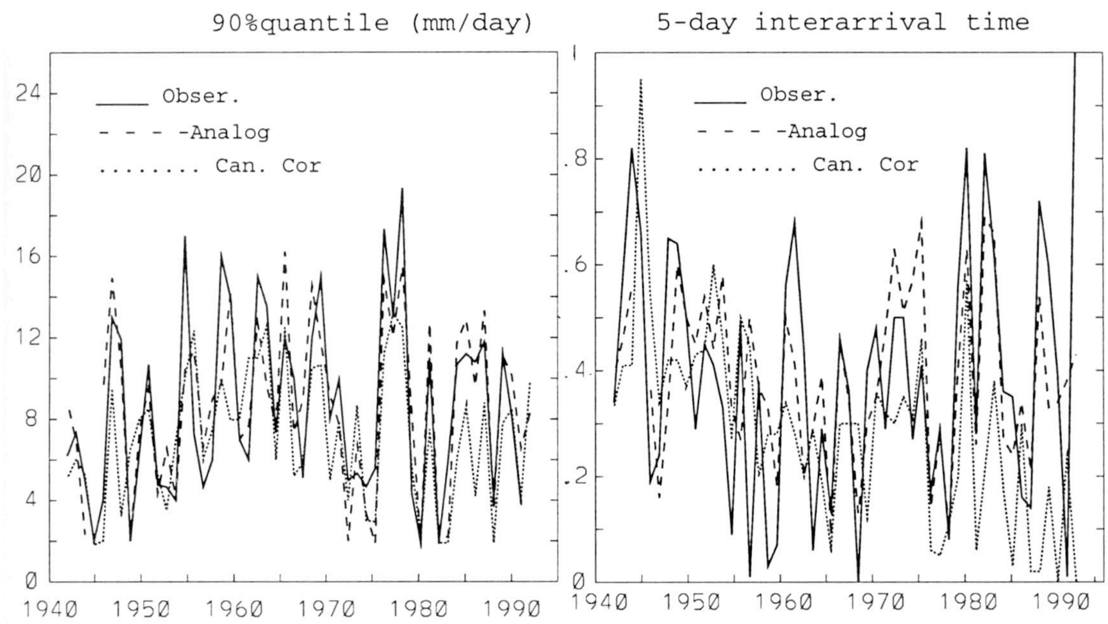


[SLP (mb) and storm interarrival times in days. The correlation between the time coefficients is .62, the probability pattern explains 38% 0/ the variance.]



the training period the analogs were again searched in a season different from the target day. The agreement between simulations and observations is clear, especially at low frequencies. The 90% quantile time series seems to be negatively correlated with the 5-day interarrival time time series also at low frequencies, indicating that periods with higher (lower) rainfall are associated with shorter (longer) dry spans. This relationship is also captured by both statistical methods.

Figure 7: Time series 0/ probability of 5-day storm interarrival time and quantile of daily rain/all amount in Gáceres (Spain) in wintertime (DJF) , as observed and reconstructed from the winter North Atlantic SLP field, by the analog method and the CCA method.



9 Neural networks.

Neural networks have found in the last years a wide range of applications. A quite complete review about this subject can be found in (Lau and Widow, 1990). For applications in meteorology the reader is referred to Elsner and Tsonis (1992) and the references therein. In climatology recent applications comprise the EI Niño-Southern Oscillation phenomenon (Grieger and Latif, 1994; Tangang et al., 1997) and Indian monsoon rainfall (Navone and Ceccatto, 1994). Neural networks have a great potentiality in many contexts, but they have been applied to the downscaling problem only in a few cases (Hewitson and Crane, 1992 and 1996). Although promising, it remains to be demonstrated that they are generally a useful downscaling method.

Only the basic concepts necessary to follow this section will be given here. Very briefly, a neural network is an algorithm that transforms an input vector $\mathbf{x}^{in}$ into an output vector $\mathbf{x}^{out}$ by a stepwise non-linear transformations, as illustrated in figure 8. Each transformations is carried out in two steps . In a first step each component of the input vector x_i^{in} is separately transformed by a nonlinear function f:

$$x_i^* = f\left(x_i^{in}\right) \tag{8}$$

In a second step a linear transformation is applied to **x***:

$$x_j^1 = \sum_i w_{ij}^1 x_i^* \tag{9}$$

The resulting vector $\mathbf{x}^1$ is in turn the input for the next non-linear transformations. It is useful to think as if these two step process is performed by one layer of "neurons" (fig 8), the whole neural network containing several layers. Finally, the output vector $\mathbf{x}^{\underline{out}}$ is the result of the operation of the last layer.

A complete model can be build when the parameters $\tilde{w}$ are known. This can be achieved by fitting them with a set of known inputs $x_i^{in}(t)$ and outputs $y_i(t)$, by minimizing the deviations:

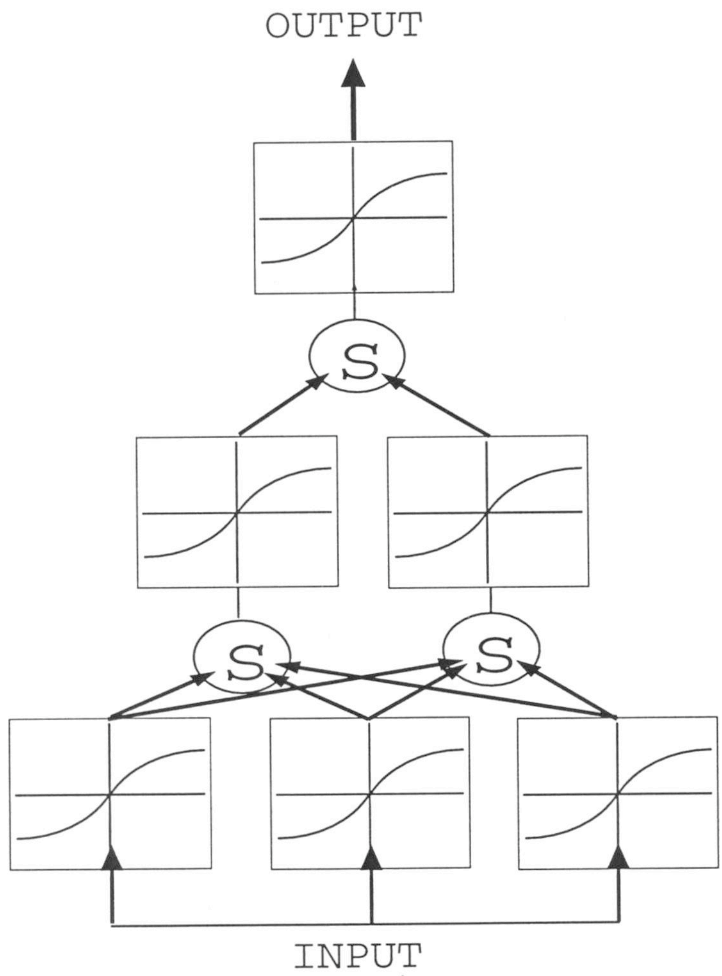
$$\sum_{i=1}^{N} \sum_{t=1}^{T} \left\{x_i^{out}(t) - y_i(t)\right\}^2 \tag{10}$$

where N is the number of stations and T the length of the time series.

We have used in this example a neural network of three layers to construct a nonlinear model that links the daily SLP (as predictor) and daily rainfall amounts (as predictand). The input vector is composed of the principal components associated with the leading EOFs of the daily SLP field.

The number of elements in the intermediate layer (sometimes called the hidden layer) is somewhat arbitrary but constrained by the following considerations. First, as in any statistical model the number of parameters in w $\tilde{w}$ should be kept to a minimum to avoid overfitting of the noise in the training period. Otherwise the skill of the network falls abruptly when it is applied to a set of predictors in an independent data set. To understand the second consideration more easily, consider for the moment a linear network (with a linear filter function f) with just a single element in the hidden layer and assume that the desired output time series are normalized ($\mu = 0$, $\sigma = 1$). Then this linear model is equivalent to a CCA model with the first pair of canonical loadings given by w_{ij}^1 and w_{ij}^2. If we think of a canonical pair of patterns as representing a physical process, as we did in the Scandinavian temperature example, then we should include so many neurons in the hidden layer as many physical situations giving rise to rainfall. As a rule of thumb, this number should be *of the order* of the rainy *Grosswetterlagen*

Figure 8: Schematic structure of an algorithmic neural network.



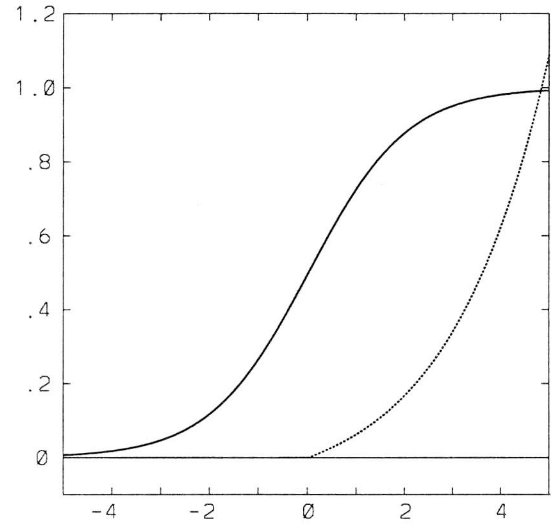
for this region or the number of significant rainfall EOFs. In this example we have included five elements in the hidden layer.

The last question that needs to be solved is the form of the filter function f. For many applications, sigmoidal-type functions of the form:

$$f(x) = \frac{1}{1 + e^{x - x_c}} \tag{11}$$

have been used (fig) 9, but they are not suitable for our downscaling purposes: the net could only generate rainfall in the interval (0,1), or by re-scaling in some apriori finite interval; furthermore

Figure 9: Two possible nonlinear filter functions that relates the input and output of a neuron in a neural net. Bold line is a classical sigmoidal type; the dashed line represent Eq. 12, more suitable for the simulation of daily rainfall. In this plot $x_c = 0$



it could not generate truly dry days. We have found that the function:

$$f(x) = \begin{cases} 0 & : x \le x_c \\ e^{r(x-x_c)} - 1 & : x \ge x_c \end{cases} \tag{12}$$

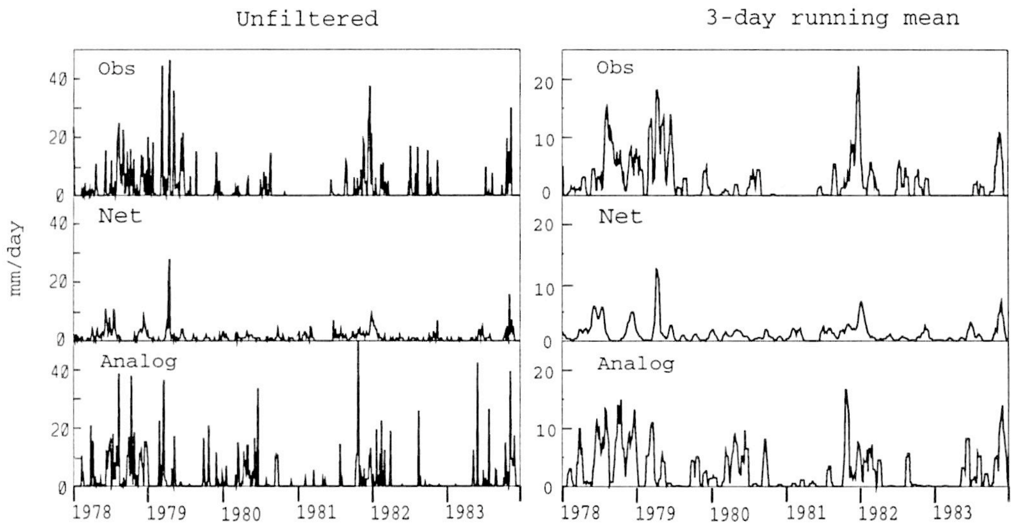
gives reasonable results (fig 9, where the cut-off x_c is the value of the input for which the neuron becomes *active*. The fact that it is not strictly differentiable at $x = x_c$ is not a too big practical problem. This choice is dictated by the nature of rainfall and is not universally applicable. For other applications other forms for the non-linear transformation f could be explored. Of course the form of f for the last layer of neurons strongly influences the probability distribution of the simulated local variable, so that a good choice for f has to take the real distribution into account. Note that in principle one could use different filter functions for the different neurons, so that considerable flexibility is possible.

Following this considerations a neural network has been design to describe the daily rainfall in wintertime in Cálceres (Spain) in the period 1978-1983. These winters include relatively well defined wet and dry periods, so that the skill of the net can be better illustrated. The input variables are the coefficients of the leading 5 SLP EOFs, calculated on a daily basis. The coefficients for day t, t-l, t-2 are used to estimate the rainfall at day t. Therefore the input layer has 15 neurons. An intermediate layer of 5 neurons has been intercalated. The output layer consists of a single neuron. The form of filter functions is the same for all neurons (see figure 9, dashed line), but for the output neuron the exponential parameter is about 5 times larger. The weights wand the values of the cut-offs x_C (different for each neuron) of the network have been fitted with daily data in the winters between 1970-1975 by the *back propagation error algorithm* (Rummelhart et al., 1986). Then the winter daily SLP data in the period 1978-83 are used to simulate rainfall at Cáceres. The results with the neural net are also compared with those obtained with the analog method (see section 4), by looking for three-day evolution analogs in the whole data set available except in the winter months in 1978-1983. Fig 10 (left) shows the observed rainfall and the rainfall estimated by these two methods. Figure 10 (right) shows the same time series but smoothed with a 5-day running mean filter. It can be seen that the neural net shares a common drawback with other "deterministic models": they produce time series with less variance than in the observations and therefore they tend to underestimate the frequency or intensity of heavy rainfall and the frequency of dry days, although the general distribution of dry and wet periods is reasonable. The agreement with the observations is clear in the smoothed data time series, indicating that there exists high-frequency variability that is not captured by the neural net. Note that only information from the SLP has been used and that incorporation of geopotential heights or temperature in upper atmospheric layers (also large-scale fields) is likely to improve the results. It is also suspected (Hughes et al., 1993, Zorita et al., 1995) that the daily rainfall process is not only conditioned by the large-scale meteorological fields but also by the local rainfall in previous days, and this information is only indirectly and not completely available to the net through the SLP field in the previous days. But perhaps the biggest drawback is the difficulty to assign a physical interpretation to the weights.

10 The problem of the level of simulated variability.

The statistical models presented here describe in general a partial relationship between independent variables representing the large-scale climate variability, and dependent local variables. The part of the local variables that remains undescribed by the independent variables is normally referred to as noise. From this point of view the observed local variable at time t *R(t)*

Figure 10: Daily rainfall (mm/day) time series at Gáceres (Spain) in the winter months (DJF) in the years 1978-83 (first point is 1st January 1978, last 31st December 1983, simulated by the neural network and the analog method. The input for both models are the coefficients of the five leading SLP EOFs in the current and two previous days.



is the outcome of a stochastic variable $\tilde{R}(t)$, with a probability distribution P that depends on the simultaneous large scale forcing F(t):

$$\tilde{R}(t) \quad P(F(t)) \tag{13}$$

The parameters of the downscaling models presented in this paper (with the exception of the analog method in its simple version where all EOFs are treated equally) are fitted by minimizing a sum of squared deviations between the simulated and observed values, what means that the fitted model yields the best (in the sense of minimum uncertainty) estimation of the mean of $E\{P(F(t)\} = \overline{P(F)}$. We denote this optimal fitted model by M and the best estimation of $\overline{P(F)}$.by $\hat{R}(t)$.With this notation:

$$\hat{R}(t) \quad P(F(t)) \tag{14}$$

However, the fitted model M is not optimized with respect to the the variance of $\tilde{R}(t)$. There are two contributions to the variance of $\tilde{R}(t)$: a local one:

$$Var_{local} = E\{P(F - \overline{P(F)}).2\} \tag{15}$$

caused generally by local processes and measurements errors, and the variance forced externally by F:

$$Var_{external} \alpha E\{(F - \overline{F})^2\} \tag{16}$$

where we assumed that the variability of the external forcing is independent of the internal variability.

When the statistical model is applied to an external forcing simulated by a GCM or taken from the observations in a verification period, the variance of the simulated output is less than the observed variance. This occurs because the variance of the simulated local variable is caused only by the variance of the external forcing $Va_{external}$, and does not contain the *internal* contribution to the variance of $\tilde{R}(t)$. This is not important if the aim of the model is just the estimation of changes in the mean local climate, but it is really important if the output of the statistical model is used to drive an ecosystem or sector model. In this case the level and structure of this noise may need to be addressed.

Some authors (Karl et al., 1990) have used inflated regression coefficients in linear models to increase the variance of the simulated output. However, in doing so one is artificially enlarging the signal-to-noise ratio in the simulated time series and also modifying the cross-covariance structure of the regional climate. This can be important for ecological modelling, for instance if temperature and rainfall are downscaled simultaneously. Another approach to this problem has been recently proposed by Bürger (1996). According to this author the second step in the design of the downscaling model (section 4), namely the estimation of the models parameters by minimizing the differences between the model response and the observations, is replaced by a constrained minimization procedure. The simulated local variables are forced to have the same covariance structure, and therefore the same individual variances, than the local observations. The price that has to be paid is that the fitting between simulations and observations in the training period is not as good as with an unconstrained minimization. Therefore, the statistical model produces a simulated output with the right level of local variability, but it is less consistent with the large-scale forcing. One has to find a compromise between both requirements that surely will depend on the particular application.

A more consistent way would be to acknowledge our ignorance about the origin of this unexplained part of the local variability and try to take it into account as an additional and independent random component. Therefore, for the purposes of obtaining estimates of the local variable, for instance to be subsequently used as forcing for an ecosystem model, the estimation of (14) could be replaced by a more useful estimation of the local variable conditional on the external forcing:

$$\hat{R}(t) = M(F(t)) + \varepsilon \tag{17}$$

where ε is a normally distributed random variable with mean zero and variance given by:

$$Var(\varepsilon) = Var(R) - Var(M(F)) \tag{18}$$

In other words, the variance of ε is the difference between the variance of the observed local variable and the variance of the simulated response.

11 Concluding remarks.

A variety of statistical downscaling examples has been presented and discussed. It is clear that there is no universal method valid for all variables and all regions and that one is bound to design statistical models on a case-by-case basis. This should not be a too large disadvantage for the investigator interested in a single region but it is certainly impracticable for an assessment of climate change on a global basis. In this respect Limited area models (LAMs) are more suitable. On the other hand, statistical models should be in most cases easy to develop and *test*. If they are able to reproduce the observed low-frequency variability of the regional climate, they will likely estimate correctly regional climate changes (provided that the GCMs correctly simulate the large-scale climate changes). LAMs are much more difficult to test. They require high quality large-scale forcing fields, that are normally available only for a couple of decades. This means that one cannot be sure if they can simulate regional climates other than the present one. Statistical methods can in this respect be of some help: LAMs should also be able to represent the statistical relationship between the large-scale fields and the regional climate. A study of these relationships as simulated by the LAM can be helpful in improving the dynamical model itself. The number of applications of statistical downscaling is rapidly growing, but unfortunately several problems have been only marginally considered: One aspect is the error in the estimates of local climate change due to the statistical analysis (Nonhebel, 1994; Gyalistras, 1996). The final output of a climate impact study is the result of a regional climate change estimation and the application of this estimation to a climate impact model. The estimation of climate change at regional scale contains a certain degree of uncertainty, which is the result of the intrinsic uncertainty of the global climate simulations superimposed on the uncertainty introduced by the statistical or the dynamical downscaling technique.

Other aspect is related to the sensitivity of climate impact models. Much effort is being put to reduce the uncertainties of the regional climate change estimations, but these efforts have to be matched from the side of climate impacts studies by an assessment of the sensitivity of the climate impact models. For a meaningful climate impact study the error bars of the estimations of the local climate change have to be smaller than the sensitivity of the climate impact models, defined roughly as the minimum increment in the forcing fields necessary to produce a significantly different output. Some studies in this direction have been recently published (e.g. Bugmann, 1997).

Therefore, the solution of the spatial scale gap between climate research and climate impact studies has to be bridged not only by *downscaling* on the side of the climate research but also by *upscaling* on the side of the climate impact research. Upscaling means to design the impact models in such a manner that they can be run with forcing fields with the considerable uncertainty that is to be expected from quasi-realistic climate models. This can mean in practice that internal model parameters cannot be derived from detailed case studies. A more reasonable strategy for the application to climate impact studies would be to fit those internal parameters, so that the model can reproduce the low-frequency natural variability of the ecosystem.

Acknowledgments. We thank Maria Casado from the Instituto Nacional de Meteorologia (Madrid) who kindly supplied the Spanish daily rainfall data, and Fidel Gonzalez from the Universidad Complutense (Madrid, Spain) who provided the monthly rainfall records. Marisa Montoya from the University of Hamburg and Stefan Güss, Hauke Heyen and Ralf Weisse from the GKSS-Research Center helped us with valuable comments to correct and improve earlier versions of this manuscript.

REFERENCES

Bardossy, A., and J. Plate, 1992: Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resour. Res.* **28**, 1247-1259.

Barnett, T., and R Preisendorfer, 1978: Multifield analog prediction of short-term climate fluctuations using a climate state vector. *J. Atmos. Sci.* **35**, 1771-1787.

______________, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface-air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.* **115**, 1825-1850.

Breimann, L., J.H. Friedmann, R.A., Olsen, and J.C. Stone, 1984: *Classifications and Regression Trees.* Wadsworth, 391 pp.

Bretherton, C., C. Smith and J. Wallace, 1992: An intercomparison of methods to find coupled patterns in climate data. *J. of Climate* **5**, 541,560.

Bugmann, H., 1997: Sensitivitu of forests in the European Alps to future climate change. *Climate Research* **8**, 35-44.

Bürger, G., 1996: Expanded downscaling for generating local weather scenarios. *Clim. Res.* **7**,111-128.

Busuioc, A., and H. von Storch, 1997: Verification of GCM generated regional precipitation and of statistical downscaling estimates. *Accepted in Climate Research*

Cheng, X., and J.M. Wallace, 1993: Cluster analysis of the Northern Hemisphere wintertime 500 hPa height fields: Spatial patterns. *J. Atmos. Sci.* **50**, 2647-2696.

Cubasch, U. K. Hasselmann, H. Hoeck, E. Maier-Reimer, U. Mickolajewicz, B.D. Santer, and R. Sausen, 1992: Time-dependent greenhouse warming computations with a coupled ocean-atmosphere model. *Climate Dynamics* **8**, 55-59.

Cubasch U., H. von Storch, J. Waszkewitz and E. Zorita, 1996: Estimates of climate changes in southern Europe using different downscaling techniques. *Climate Research* **7**, 129-149.

Elsner, J.B, and A.A. Tsonis, 1992; Nonlinear predictions, chaos and noise. *Bull. Am. Meteor. Soc*. **78**, 49-60.

Fliri, F, 1974: Niederschlag und Lufttemperatur im Alpenraum. *Wissenschaftl. Alpenverein-hefte* **24**, 111-125.

Frey-Buness A., M. Heimann, and R Sausen, 1995: A statistical dynamical downscaling procedure for global climate model simulations. *Theor. App. Climatol*. **50**, 117-131

Giorgi F., and L. O. Mearns, 1991: Approaches to the simulations of regional climate change: a review. *Rev. Geophys*. **29**, 191-216.

Grieger, B., and M. Latif, 1994: Reconstructing the EI Niño attractor with neural networks. *Climate Dynamics* **10**, 267-276.

Grotch, S., and M. MacCracken, 1991: The use of general circulation models to predict regional climate change. *J. Climate* **4**, 286-303

Gyalistras, D., A. Fischlin, and M. Riedo 1996: Herleitung stündlicher Wetterszenarien unter zukünftiger Klimabedingungen. In *Klimaveränderungen und Grünland-Eine Modellstudie über die Auswirkungen zukünftiger Klimaveränderungen auf das Dauergrünland in der Schweiz,* J. Fuhrer(ed), Verlag der Fachvereine, Zürich.

Hense A., R Glowienka-Hense, H. von Storch, and U. Stahler, 1990: Northern Hemisphere atmospheric response to changes of Atlantic Ocean SST on decadal time scales: A GCM experiment. *Clim. Dyn*. **4**, 157-174.

Hewitson, B., and R.G. Crane, 1992: Large-scale atmospheric controls on local precipitation in tropical Mexico. *Geophys. Res. Lett*. **19**, 1835-1838

______, 1996: Climate Downscaling: techniques and application. *Climate Research* **7**, 85-95.

Hughes, J., D. Lettenmaier, and P. Guttorp, 1993: A stochastic approach for assessing the effect of changes in regional circulation patterns on local precipitation. *Water Res. Research* **29**, 3303-3315.

Hulme, M., K. Briffa, P. Jones, and C. Senior, 1993: Validation of GCM control simulations using indices of daily airflow types over the British Isles. *Climate Dynamics* **9**, 95-105.

Jones, P., Hulme M., and Briffa K., 1993: A comparison of the Lamb circulation types with an objective classification derived from grid-point mean-sea-level pressure data. *Int. J. of Climatol.* **13**, 655-665.

Karl, T.M., W-C Wang, M.E. Schlesinger, D.E. Knight, and D. Portman, 1990: A method of relating General Circulation Model simulated climate to the observed local climate. Part I: seasonal statistics. *J. Climate* **3**, 1053-1079.

Kruizinga, S., and A. Murphy, 1983: Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands. *Mon. Wea. Rev.* **111**, 2244-2254.

Lamb, P. and R. PeppIer, 1987: The North Atlantic Oscillation: concept and an application. *Bull. Amer. Meteor. Soc.* **68**, 1218-1225

Lau C., and B. Widow (eds), 1990: Special Issue on neural networks. *Proceedings of the IEEE* **78**

Lindberg C., and A. Broccoli, 1996: Representation of topography in spectral climate models and its effect on simulated precipitation. *Journal of Climate* **9**, 2641-2659.

Lorenz,E.N., 1969: Atmospheric predictability as revealed by naturally occurring analogs. *J.Atmos. Sci* **26**, 639-646.

Machenhauer, B., M. Windelbrand, M. Bozet, and M. Deque, 1996: Validation of present-day regional climate simulations over Europe: nested LAM and variable resolution global model simulations with observed or mixed ocean boundary conditions. Max-Planck-Institut Re- port No 191, 60 pp. Max-Planck-Institut für Meteorologie, Hamburg, Germany.

Marroquín, A., J. García, J. Garrido, and V. Mateos, 1996: Newmann-Scott cluster models for daily rainfall process in lower Extremadura (Spain): Rainfall generating mechanisms. *Theor. App. Climatol.* **52**, 183-19.

Navone, D., and H.A. Ceccatto, 1994: Predicting Indian monsoon rainfall - a neural network approach. *Climate Dynamics* **10**, 305-312.

Nonhebel, S., 1994: Inaccuracies in weather data and their effects on crop growth simulations results-I. Potential production. *Climate Research* **4**, 61-74.

Ozelkan, A, 1996, 1996: Relationships between monthly atmospheric circulation patterns and precipitation: fuzzy-logic and regression approach. *Water Resour. Res*. **32**, 2097-2103.

Risbey, J, and P. Stone, 1996: A case study of the adequacy of GCM simulations for input to regional climate change. *J. Climate* **9**, 1441-146.

Rummelhart, D.E., G.E. Hinton, and R.J. Williams, 1986: Learning representations by back-propagating errors. *Nature* **323**,533-536.

Shabbar A., K. Kiguchi, and K. Knox, 1990: Regional analysis of the Northern Hemisphere 50 kPa geopotential height from 1946 to 1985. *J. Climate* **3**, 543-557.

Schnur, R, and D.P. Lettenmaier, 1997: A case study of Statistical Downscaling in Australia using Weather Classification by recursive Partitioning. *J. of Hydrology*, to be published.

Stamm, J. and A. Gettelmann, 1995: Simulation of the effect of doubled atmospheric CO_2 on the climate of Northern and Central California. *Climate Change* **30**, 295-325

Tangang, F.T, W.W Hsieh, and B. Tang, 1997: Forecasting the equatorial Pacific sea surface temperatures by neural networks models. *Climate Dynamics* **13**, 135-147.

van den Dool, 1994: Searching for analogs, how long must we wait? *Tellus* **46A**, 314-324.

van Loon H. and Rogers, 1978: The seasaw in winter temperature between Greenland and Northern Europe. *Mon. Wea.Rev*. **106**, 296-310.

von Storch, H, 1995: Inconsistencies at the interface between climate research and climate impact studies. *Metereologische Zeitschrift* **4**, 72-80.

___, H., E. Zorita and U. Cubasch, 1993: Downscaling of climate change estimates to regional scales: an application to Iberian winter rainfall in winter time. *J. Climate* **6**, 1161-1171.

Walker, G. 1924: Correlations in seasonal variations of weather. IX. Mem. Indian Meteor. Dept. **24**, 275-332.

Widmann, M., and C. Schär, 1997: A principal component and long-term trend analysis of daily precipitation in Switzerland. *Theor. and Appl. Climatol*., submitted.

Zorita E., J. Hughes, D. Lettenmaier, and H.von Storch, 1995: Stochastic downscaling of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *J. Climate* **8**, 1023-1042.