

An artificial neural network technique for downscaling GCM outputs to RCM spatial scale

R. Chadwick^{1,*}, E. Coppola², and F. Giorgi²

¹Department of Meteorology, University of Reading, UK

²The Abdus Salaam International Centre for Theoretical Physics, Trieste, Italy

*now at: Met Office Hadley Centre, Fitzroy rd, Exeter, UK

Received: 23 September 2011 – Revised: 7 December 2011 – Accepted: 12 December 2011 – Published: 22 December 2011

Abstract. An Artificial Neural Network (ANN) approach is used to downscale ECHAM5 GCM temperature (T) and rainfall (R) fields to RegCM3 regional model scale over Europe. The main inputs to the neural network were the ECHAM5 fields and topography, and RegCM3 topography. An ANN trained for the period 1960–1980 was able to recreate the RegCM3 1981–2000 mean T and R fields with reasonable accuracy. The ANN showed an improvement over a simple lapse-rate correction method for T , although the ANN R field did not capture all the fine-scale detail of the RCM field. An ANN trained over a smaller area of Southern Europe was able to capture this detail with more precision. The ANN was unable to accurately recreate the RCM climate change (CC) signal between 1981–2000 and 2081–2100, and it is suggested that this is because the relationship between the GCM fields, RCM fields and topography is not constant with time and changing climate. An ANN trained with three ten-year “time-slices” was able to better reproduce the RCM CC signal, particularly for the full European domain. This approach shows encouraging results but will need further refinement before becoming a viable supplement to dynamical regional climate modelling of temperature and rainfall.

1 Introduction

Studies of climate impacts usually require estimates of future climate parameters at a higher spatial scale than can be provided by the current generation of General Circulation Models (GCMs). One method of producing higher resolution estimates is to run a Regional Climate Model (RCM) for a particular sub-region of the globe, forced with lateral and surface boundary conditions from a GCM (Giorgi and Mearns,

1999). However, current RCMs, which have a resolution of around 10–50 km, are still computationally expensive to run, even for only a fraction of the globe.

Uncertainties in future climate projections derive from various sources, such as future emission scenarios, internal variability of the climate system and the accuracy and configuration of the climate models themselves, both global and regional (e.g. Giorgi and Francisco, 2000; Deque et al., 2005). These uncertainties can be quantified to some extent by examining the climate projections for a range of different scenarios, global climate models, climate realizations, and regional downscaling tools (Deque et al., 2005). This process might require the completion of a very large number of simulations (Giorgi et al., 2008), which may not currently be feasible due to the computational expense of each GCM and RCM model run. Therefore there is interest in possible alternative, less expensive approaches to fill the GCM-RCM climate change simulation matrix.

Statistical and statistical/dynamical GCM downscaling techniques require much less computational time than a pure dynamical approach. Statistical downscaling methods use empirical relationships established between large-scale and fine-scale variables using historical data, and a number of these techniques are described in Giorgi et al. (2001) and Maraun et al. (2010). These methods rely on these statistical relationships remaining constant with future climate, an assumption that is open to question (Wilby et al., 1998).

Statistical/dynamical approaches construct a statistical relationship between RCM and GCM outputs over a limited RCM run, then extrapolate this relationship to produce fine-scale estimates for other periods of the GCM run (e.g. Busch and Heimann, 2001). This type of approach is computationally less expensive than running the RCM for the whole time period of interest, but relies on the relationship between GCM and RCM output parameters remaining constant with time and changing climate.



Correspondence to: R. Chadwick
(robin.chadwick@metoffice.gov.uk)

In this paper we explore an approach similar to statistical/dynamical downscaling to provide a means to populate a downscaled climate projection matrix without having to run RCMs for the full range of GCM simulations. Instead of a statistical method, an Artificial Neural Network (ANN) is used to construct a relationship between a GCM and corresponding nested RCM fields. If such a relationship, which likely depends on the given GCM-RCM pair considered, can be shown to be accurate, it can be used to “downscale” the information of a given GCM to the grid of a given RCM without having to re-run the RCM for each GCM experiment. For example this could be used to downscale a given GCM onto a given RCM grid for different greenhouse gas emission scenarios or for different realisations of the same scenario. This approach of course relies on the ability of the derived ANN relationship to be valid in future climate conditions as well as the present. The concept can be seen as a type of emulator (Bastos and O’Hagan, 2008; Conti and O’Hagan, 2010), where an ANN is used to emulate RCM output for future scenarios without needing to run the RCM.

We develop the ANN model and test its accuracy for present and future climate conditions using a transient climate projection (1960–2100) over the European region completed with the Abdus Salam International Centre for Theoretical Physics (ICTP) RCM, RegCM3 (Pal et al., 2007), nested within the Max Plank Institute global model ECHAM5 (Roeckner et al., 2003). We test the accuracy of the ANN model for different historical periods and future time slices in the simulation, and suggest optimal approaches to maximize the model reliability. The paper begins in the next section with a description of the climate and ANN models and of the experiment design.

2 Model and simulation description

The regional climate model used here is RegCM3, described by Pal et al. (2007) and references therein. RegCM3 is based on the hydrostatic primitive equation dynamical core of the NCAR mesoscale 95 meteorological model MM5 (Grell et al., 1994). Resolvable precipitation processes are treated with a sub-grid explicit moisture scheme (Pal et al., 2000), and the convection scheme used here is that of Grell et al. (1994), with the Fritsch and Chappell (1980) closure assumption. The Kiehl et al. (1996) CCM3 radiation scheme, biosphere-atmosphere transfer scheme (Dickinson et al., 1993) and Zeng et al. (1998) ocean flux scheme are also used.

Initial and 6-hourly boundary conditions and SSTs are obtained from the ECHAM5 AOGCM (Roeckner et al., 2003). This parallel ECHAM5 run also provides the GCM data used to develop and test the artificial neural network in this study.

The RCM domain is that used for the EU ENSEMBLES project (Van Der Linden and Mitchell, 2009), covering the entire European region. The domain covers the entire

Table 1. ANN inputs.

Input Number	Input
1	NE GCM temperature/rainfall
2	NW GCM temperature/rainfall
3	SW GCM temperature/rainfall
4	SE GCM temperature/rainfall
5	NE GCM elevation
6	NW GCM elevation
7	SW GCM elevation
8	SE GCM elevation
9	NE GCM gridpoint to RCM gridpoint distance
10	NW GCM gridpoint to RCM gridpoint distance
11	SW GCM gridpoint to RCM gridpoint distance
12	SE GCM gridpoint to RCM gridpoint distance
13	RCM elevation
14	RCM landuse
15	Month
16	Year

European region at a grid spacing of 25 km. RegCM3 was integrated for the period 1960–2100. ECHAM5 was integrated for the same period at a resolution of 1.89 degrees, using the A1B scenario of greenhouse gas forcing (Solomon et al., 2007). The lateral buffer zone is 15 grid points in width, and uses an exponential relaxation term (Giorgi et al., 1993). The size of the buffer zone is small compared to the RCM domain, and RegCM3 is therefore free to produce its own internal circulations and climatology.

3 ANN methodology

Artificial neural networks (ANNs) are a pattern recognition tool capable of reproducing empirical, possibly non-linear relationships between a set of “input” variables and some corresponding “output” variables. ANNs are based on the physiology of the brain, with a series of “nodes” able to pass information between one another in a similar way to cells in the brain. For a full description of ANN theory, see Bishop (2000) or Picton (2000). ANNs have been applied on several previous occasions to problems within climate science (see Hsieh, 2009; Haupt et al., 2009 for recent reviews of these applications), including to the field of global and regional attribution of climate change (Pasini et al., 2006; Pasini and Langone, 2010).

In this case, the interest was in establishing the relationship between GCM and RCM output parameter fields, specifically temperature and rainfall (T and R). This was achieved by presenting properties of the GCM and RCM fields to an ANN, and “training” it to be able to replicate the relationship between the two. The method used is described in detail below.

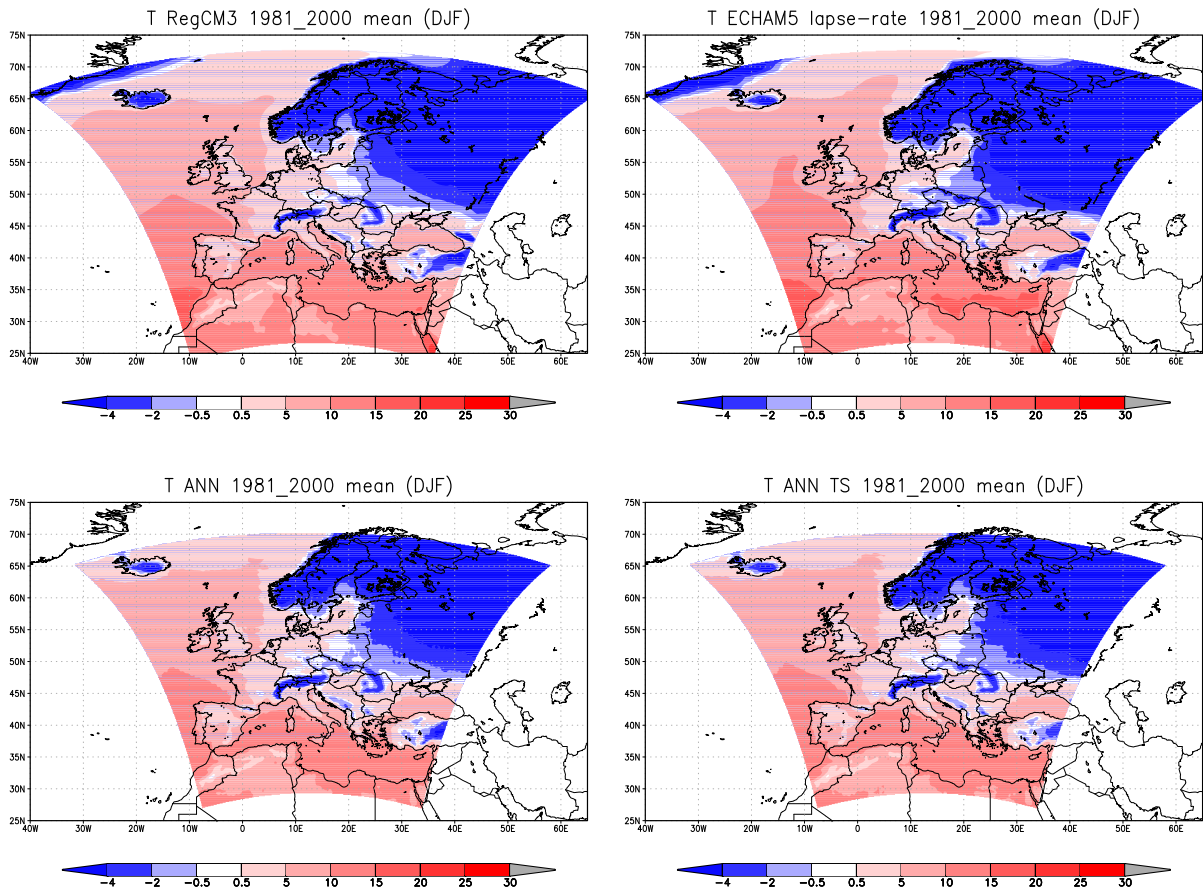


Fig. 1. DJF climatological mean T ($^{\circ}\text{C}$) for Europe, 1981–2000. From RCM, GCM with lapse-rate correction, EurANN (ANN trained over the full European domain), and EurANNTS (Time-slice ANN trained over the full European domain).

The particular type of ANN used here is known as a feed-forward multi-layer perceptron (MLP). MLPs consist of a number of layers, each containing several nodes. Each node in a particular layer is linked to every node in the previous and subsequent layers. For a feed-forward network such as the one used here, nodes can only pass information one way, in the direction leading from the first, input layer to the final, output layer. The number of layers, combined with the number of nodes in each layer, is collectively known as the “ANN architecture”.

Each connection between nodes has an associated “weight” W , the value of which can be modified to strengthen or weaken the connection between nodes. It is the value of these weights, together with the ANN architecture, that largely determine the properties of an MLP. The input to a particular node consists of the weighted sum of the outputs from all nodes in the previous layer. So the input to a particular node y in the j -th layer of an ANN is given by:

$$\text{Input}_{jy} = \sum_{x=1}^{N_{j-1}} W_{(j-1)xy} \text{Output}_{(j-1)x}, \quad (1)$$

where $\text{Output}_{(j-1)x}$ is the Output of node x in the $(j-1)$ -th layer, $W_{(j-1)xy}$ is the weight connecting node x in the $(j-1)$ -th layer and node y in the j -th layer, and N_{j-1} is the total number of nodes in the $(j-1)$ -th layer.

This input is mapped to an output for that node by use of a sigmoid function, where:

$$\text{Output}_{jy} = \frac{1}{1 + \exp(-\text{Input}_{jy})}. \quad (2)$$

This sigmoid function is similar to a simple threshold function but is also continuous and differentiable, which is necessary for the use of a back-propagation algorithm. The output Output_{jy} is then passed on to each node in the $j+1$ -th layer, and so on until the final layer of output nodes is reached.

A network is prepared for a certain purpose by means of a calibration, or “training” process. A large number of training patterns are used, each pattern consisting of a set of input variables and the corresponding “target” output variable. In

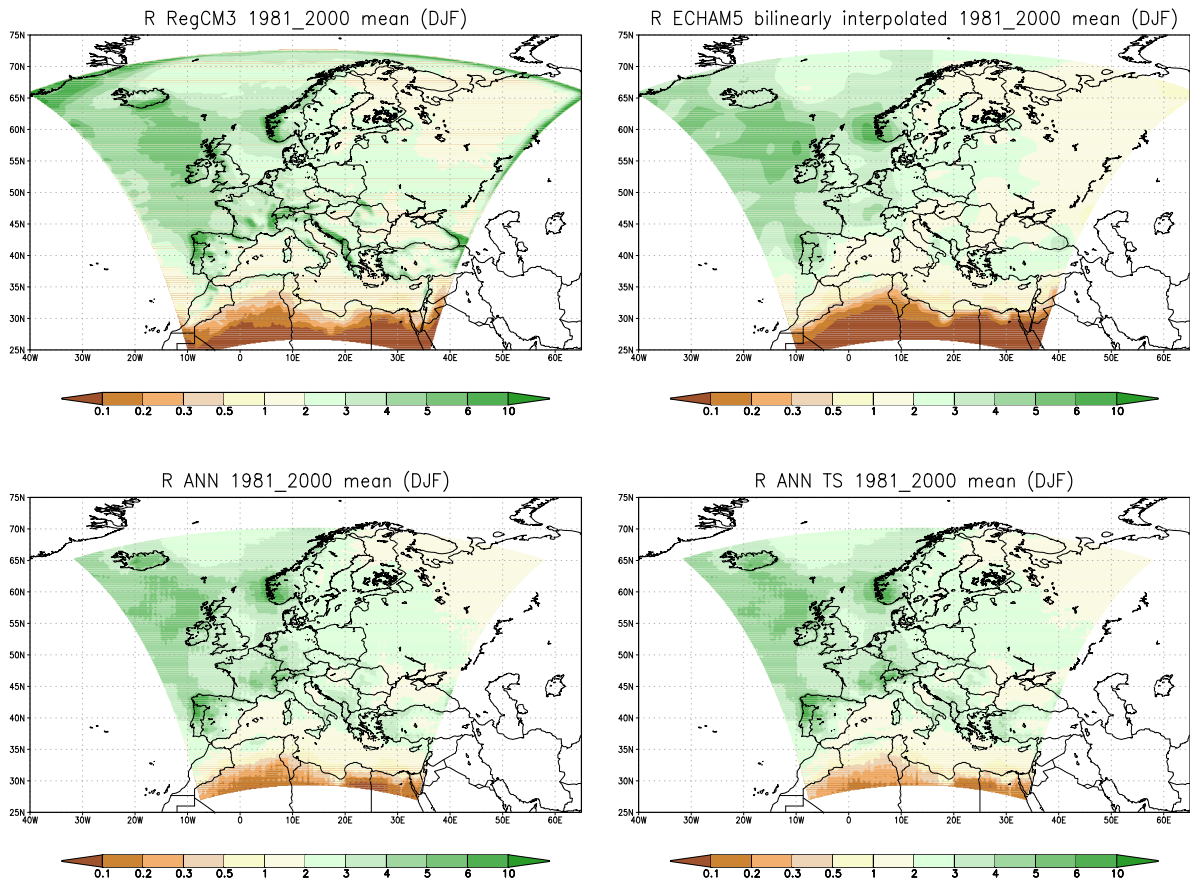


Fig. 2. DJF climatological mean R (mm day^{-1}) for Europe, 1981–2000. From RCM, GCM bilinearly interpolated to RCM scale, EurANN (ANN trained over the full European domain), and EurANNTS (Time-slice ANN trained over the full European domain).

this case it was desired to produce an output of T or R for each point in the RCM grid. ANNs were trained separately for T and R , so each network produced only one output variable, corresponding to a single node in the output layer of the ANN.

The ANN weights are randomly initialised, and the inputs for the first pattern are fed into the ANN, producing a corresponding output. This output is compared with the corresponding target output for the pattern, resulting in an “error” value for that pattern. This process is repeated for all remaining patterns, and the total “error”, E is taken to be the sum of the squared errors from all patterns. One cycle of this process involving all training patterns is known as an “epoch”.

The goal of ANN training is to minimise E by adjusting the ANN weights. This is achieved by use of a “back-propagation algorithm” (Rumelhart and McClelland, 1986). A relationship $\delta E/\delta W$ can be analytically established for every weight W in the ANN. The value of each weight is then modified by an amount proportional to $\delta E/\delta W$. The process is then repeated, with all training patterns processed through the modified network, E calculated, and weights modified, until the value of E is regarded to be at a minimum. In

practise it can take several thousand epochs before a true minimum is reached. A “momentum” term is also included in the training process. This reduces oscillations of the solution around the absolute minimum of E and reduces the number of training epochs needed.

A crucial property of ANNs is that they should be able to generalise, meaning that after training they can process a set of inputs not used in the training process (and hence not “seen” before by the network) into a reasonable output value. In some cases when too many epochs of training are used, an ANN can become “overtrained” to its training dataset and loses the ability to generalise. In order to avoid this, the available input/output patterns were separated into a training dataset and a smaller “test” dataset. After each epoch, E was calculated for both the training and test datasets. If E continued to fall or stabilised for the training dataset but began to rise for the test dataset, then overtraining was diagnosed and the training process was stopped.

The ANN code used here was based on a modified version of that developed by Lönnblad et al. (1991) for pattern recognition problems in particle physics.

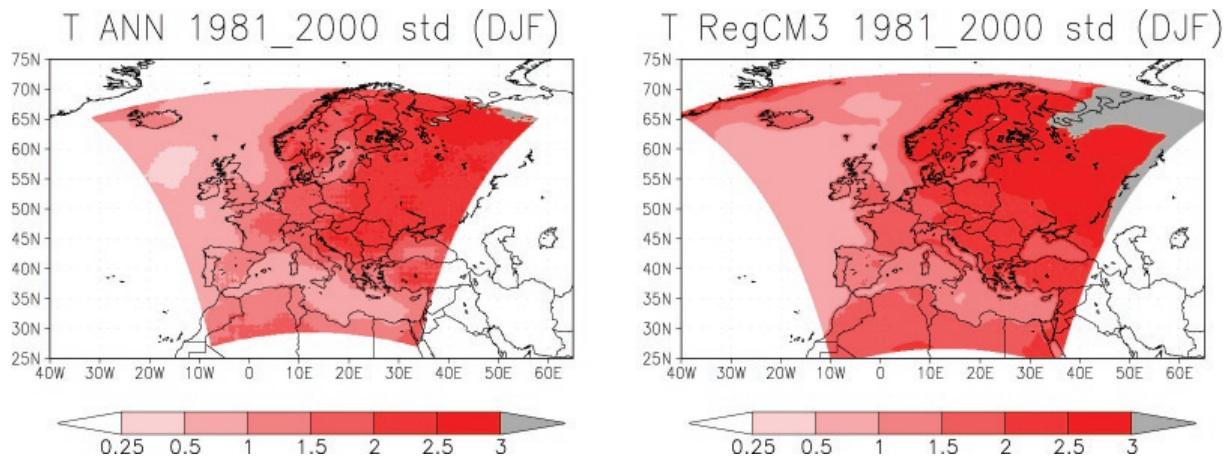


Fig. 3. DJF standard deviation of seasonal mean T ($^{\circ}\text{C}$) for Europe, 1981–2000. From EurANN (ANN trained over the full European domain) and RCM.

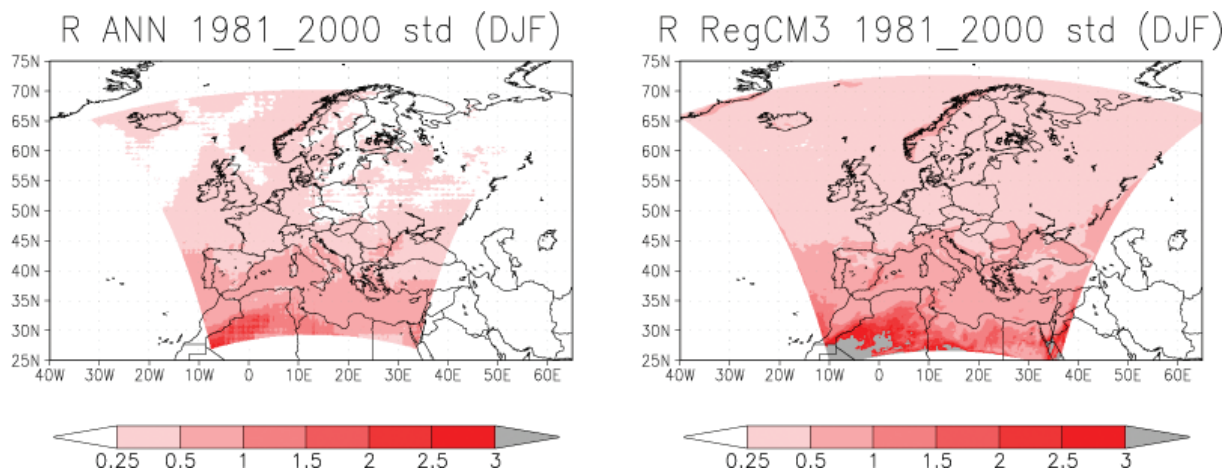


Fig. 4. DJF standard deviation of seasonal mean R (mm day^{-1}) for Europe, 1981–2000. From EurANN (ANN trained over the full European domain) and RCM.

An MLP with at least 3 layers should be able to reproduce any mapping between input and output variables (Hornik et al., 1989), but 4 layers are used here as it has been found previously that an extra layer leads to more efficient ANN training for problems similar to this one (Tomassetti et al., 2009; Chadwick and Grimes, 2011; Coppola, 2006).

As inputs it was possible to use any information from the GCM output field, and non-time-variant properties of the RCM field such as gridpoint elevation and land-use. As GCM and RCM gridpoints are not in general spatially coincident, and the RCM field has a higher spatial resolution, a decision had to be taken on which GCM gridpoints to use as input data for each RCM gridpoint.

Tomassetti et al. (2009) used an ANN approach to downscale precipitation fields from 27 km to 3 km, and their method of matching the two fields of differing spatial resolution has been adopted here. For each RCM gridpoint, the

surrounding 4 GCM gridpoints are used as inputs. So one GCM gridpoint will be to the NE of the RCM gridpoint, one to the NW, one to the SW and one to the SE. The inputs used are shown in Table 1. For RCM-scale T as output, GCM T input values were used, while GCM R values were used for RCM-scale R output.

To improve the efficiency of ANN training, inputs are standardised to values in the range [0 1] by the function:

$$x_{\text{final}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}, \quad (3)$$

where x_{max} and x_{min} are the min/max values for a particular input and x_{final} is the final standardised input for use in the ANN.

GCM and RCM T/R fields were used at monthly time-resolution. The available coincident RCM and GCM dataset was for the period 1960–2100.

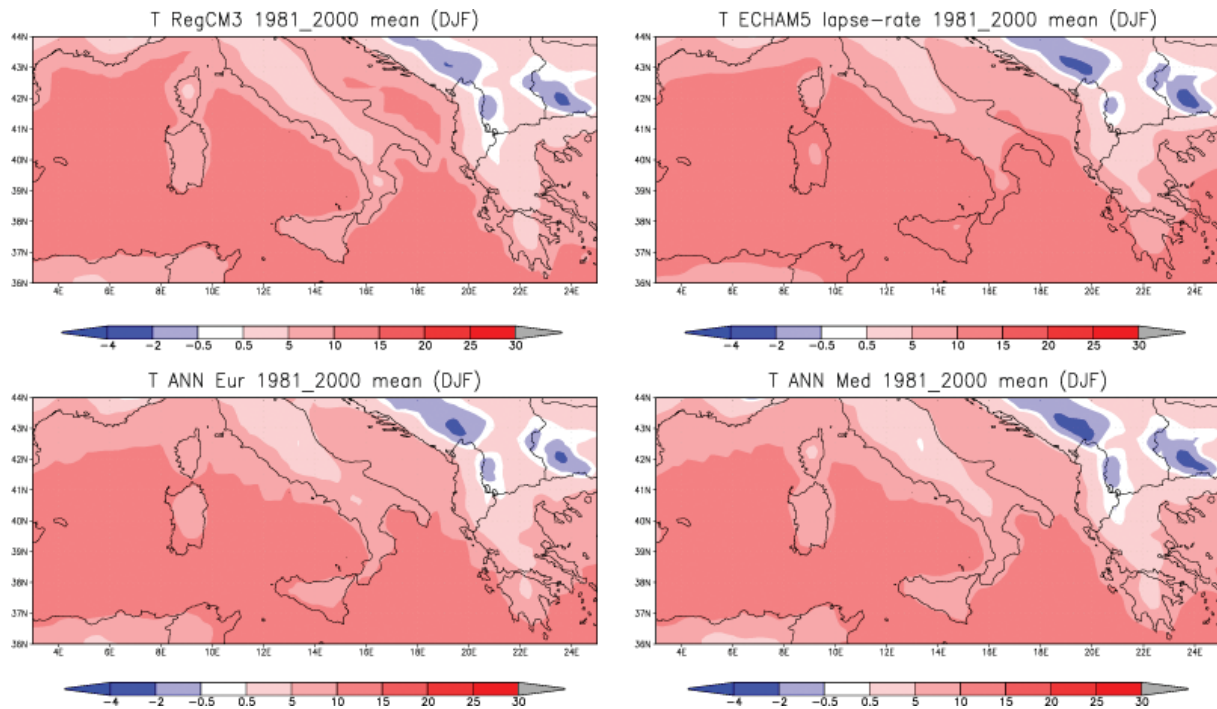


Fig. 5. DJF mean T ($^{\circ}\text{C}$) for the Mediterranean region, 1981–2000. From RCM, GCM with lapse-rate correction, EurANN, and MedANN.

4 Results and discussion

The ANN was trained initially with RCM data from 1960–1980 (separate randomly selected subsets of this dataset were used for ANN training and testing) and validated with RCM data from 1981–2000. Due to the large number of RCM grid-points available in the European domain over the 21 yr training period, and in order to reduce processing time, only a randomly selected selection of 20% of the total available grid-points for each month in the training period were used for training. The particular gridpoints used were selected separately for each month of the training period, in order to fully sample the RCM domain during training. Sensitivity tests established that increasing this proportion above 20% made no noticeable difference to ANN performance, so 20% of the data was considered to be a sufficient amount to properly train the ANN. ANN training proceeded with the training dataset until E was minimised, with the subset of test data used to prevent overtraining. Once trained, the ANN was applied to the validation dataset and ANN estimates produced for each validation month and RCM gridpoint.

The time-mean 1981–2000 ANN DJF output using the entire European RCM domain (excluding buffer zones for both training and validation) are shown in Figs. 1 and 2. The RCM fields are also shown.

The ANN downscaling method described here was also compared against a simple lapse-rate downscaling method for temperature fields similar to the one described in Engen-Skaugen (2007). The GCM T and elevation fields are first

bilinearly interpolated on to the RCM grid. The difference Δh in km between the interpolated GCM elevation and the RCM elevation is then calculated at each RCM grid-point. Finally a standard atmospheric lapse-rate correction of 6.5 K km^{-1} is applied to each point in the interpolated GCM T field:

$$T_{\text{lapse-rate}} = T_{\text{GCMinterp}} - 6.5 \times \Delta h \quad (4)$$

This lapse-rate corrected field is shown in Fig. 1. As no similar lapse-rate correction is appropriate for rainfall, Fig. 2 instead shows the GCM R field bilinearly interpolated on to the RCM grid.

As the purpose here was to reproduce RCM T and R fields, ANN performance was measured by how well it could reproduce the large and small-scale features of these fields. Although RCM fields may not necessarily reflect reality, if the ANN reproduced these fields accurately then it was considered to be performing well. As a quantitative measure of model performance we use the area averaged bias, root mean square error (RMSE) and spatial coefficient of correlation between the ANN or GCM models against the RCM results. These are shown in Tables 2 and 3.

It was found that the ANN was able to successfully reproduce the large scale features of the climatological mean RCM T and R fields (see Figs. 1 and 2). The representation of detailed features in the mean T field is good in the ANN, but there is no obvious qualitative improvement over the lapse-rate down-scaling method here. The ANN R field does represent several features of the RCM field better than

Table 2. Area averaged bias, RMSE and spatial correlation coefficient of temperature, for EurANN (ANN trained over the full European domain), EurANNTS (Time-slice ANN trained over the full European domain), MedANN (ANN trained over the Mediterranean region), MedANNTS (Time-slice ANN trained over the Mediterranean region), lapse-rate corrected ECHAM5 and ECHAM5, taking RegCM3 fields as “truth”, for the period 1981–2000 over the Mediterranean region.

	EurANN	EurANNTS	MedANN	MedANNTS	Lapse-rate	GCM
<i>T</i> Bias (°C)						
DJF	0.041	0.234	−0.186	0.045	1.244	1.141
MAM	0.263	0.116	0.012	0.107	1.364	1.262
JJA	−0.198	−0.091	−0.064	−0.073	1.042	0.941
SON	0.359	0.377	0.027	0.088	1.960	1.857
<i>T</i> RMSE (°C)						
DJF	0.730	0.722	0.427	0.411	1.467	1.875
MAM	0.362	0.319	0.205	0.219	1.364	1.448
JJA	0.746	0.652	0.461	0.475	1.343	1.296
SON	0.627	0.562	0.355	0.303	1.974	2.114
<i>T</i> corr.						
DJF	0.970	0.973	0.991	0.990	0.953	0.852
MAM	0.990	0.989	0.995	0.995	0.982	0.760
JJA	0.872	0.896	0.946	0.945	0.796	0.478
SON	0.979	0.986	0.991	0.994	0.970	0.834

Table 3. Area averaged bias, RMSE and spatial correlation coefficient of rainfall, for EurANN (ANN trained over the full European domain), EurANNTS (Time-slice ANN trained over the full European domain), MedANN (ANN trained over the Mediterranean region), MedANNTS (Time-slice ANN trained over the Mediterranean region) and ECHAM5, taking RegCM3 fields as “truth”, for the period 1981–2000 over the Mediterranean region.

	EurANN	EurANNTS	MedANN	MedANNTS	GCM
<i>R</i> Bias (mm day ^{−1})					
DJF	−0.158	−0.117	0.078	0.095	−0.615
MAM	−0.262	−0.153	−0.021	−0.074	−0.742
JJA	0.049	0.082	0.154	0.163	−0.409
SON	−0.160	−0.074	−0.003	−0.050	−0.743
<i>R</i> RMSE (mm day ^{−1})					
DJF	0.634	0.630	0.472	0.526	0.801
MAM	0.387	0.338	0.282	0.297	0.763
JJA	0.195	0.203	0.226	0.214	0.411
SON	0.445	0.433	0.348	0.373	0.777
<i>R</i> Corr					
DJF	0.591	0.601	0.853	0.831	0.327
MAM	0.814	0.823	0.906	0.902	0.595
JJA	0.877	0.883	0.848	0.869	0.818
SON	0.648	0.656	0.836	0.806	0.461

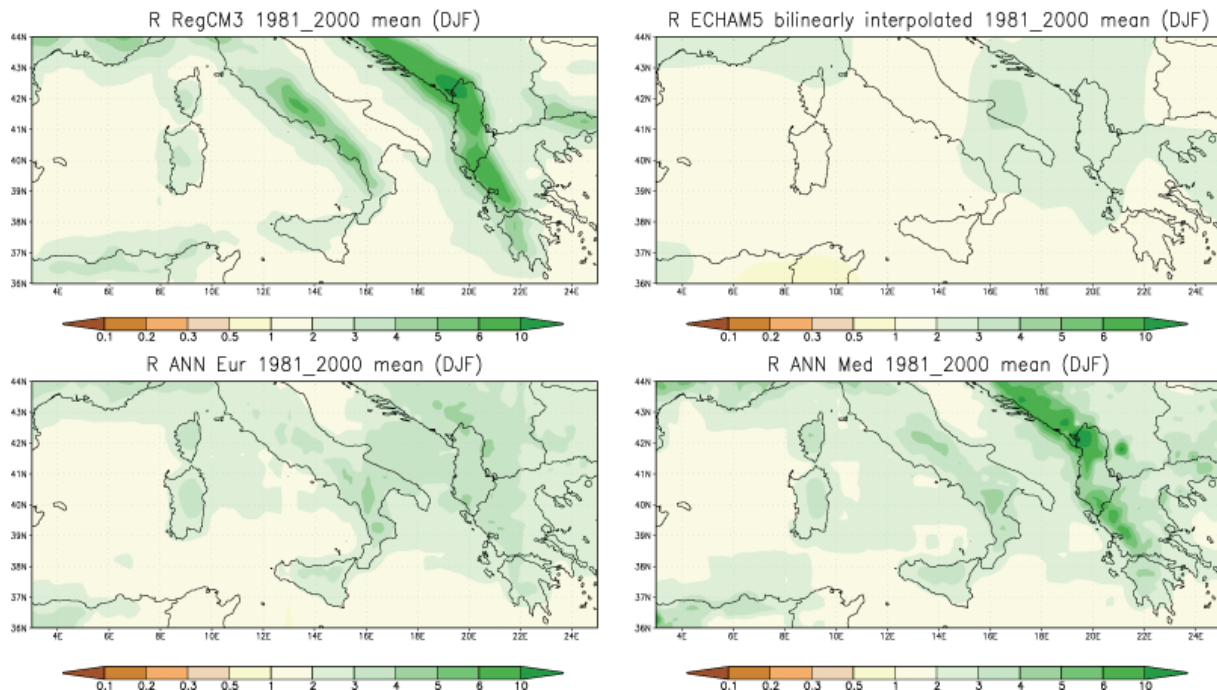


Fig. 6. DJF mean R (mm day^{-1}) for the Mediterranean region, 1981–2000. From RCM, GCM bilinearly interpolated to RCM scale, EurANN, and MedANN.

the GCM, such as the high precipitation over North-West Spain and Portugal (see Fig. 2). However some of the more detailed features such as enhanced rainfall over the mountainous coast of Croatia, and the extent of enhanced rainfall over the Alps, are not captured by the ANN. Quantitatively, it can be seen from the bias, RMSE and spatial correlation shown in period 1 (1981–2000) of Fig. 12 that the ANN shows an improvement over the both the GCM and the lapse-rate correction, for both T and R .

Figures 3 and 4 show the standard deviation of the RCM and ANN fields over the period DJF 1981–2000. In general the two fields have the same pattern, with the ANN slightly underestimating the magnitude of the RCM standard deviation, particularly for rainfall. This is consistent with previous work showing that ANNs often fail to capture the full range of values that they are attempting to simulate (Chadwick and Grimes, 2011).

In order to try and improve ANN performance for smaller regional rainfall features, a European sub-region covering the Mediterranean was focused on (defined as $36\text{--}44^\circ\text{E}$, $3\text{--}25^\circ\text{N}$). The ANN was retrained specifically for this reduced area in the same way as before, but using only data from within the region. T and R output from this more localised ANN (referred to as MedANN) are shown in Figs. 5 and 6.

It can be seen that MedANN captures the high rainfall along the Croatian coast far more effectively than either the original European ANN (EurANN) or the GCM, with some enhanced rainfall over mountainous areas of Italy also

apparent. The MedANN temperature field also shows some improvement over the lapse-rate method when observed at this level of detail, with better agreement with the RCM in Southern Italy, Sicily and Sardinia. Area-averaged bias, RMSE and spatial correlation statistics of T and R , for EurANN, MedANN, ECHAM5 lapse-rate corrected ECHAM5 (for T only) and ECHAM5 with respect to RegCM3 fields, for the period 1981–2000 over the Mediterranean region are shown in Tables 2 and 3. It can be seen that in general MedANN performs better than EurANN over the Mediterranean region (with the exception of JJA rainfall).

These results indicate that the performance of the ANN is sensitive to the choice of region of calibration. The more localised the region of ANN calibration, the more it is able to reproduce the specific relationship between orography, GCM variables and RCM variables found in that region. So the MedANN is more appropriate for estimating rainfall and temperature over the Mediterranean than the EurANN, but equally the MedANN would be expected to perform poorly if applied over the European region as a whole.

Our results concerning the performance of the ANN for the historical period show encouraging indication towards the use of the ANN to downscale GCM information at the RCM level. It then remains to be seen whether the ANN can correctly reproduce the climate-change (CC) signal shown by the RCM over Europe. Towards this goal, estimates were produced with the same ANN as above for the future time period 2081–2100, and the differences of the mean T and

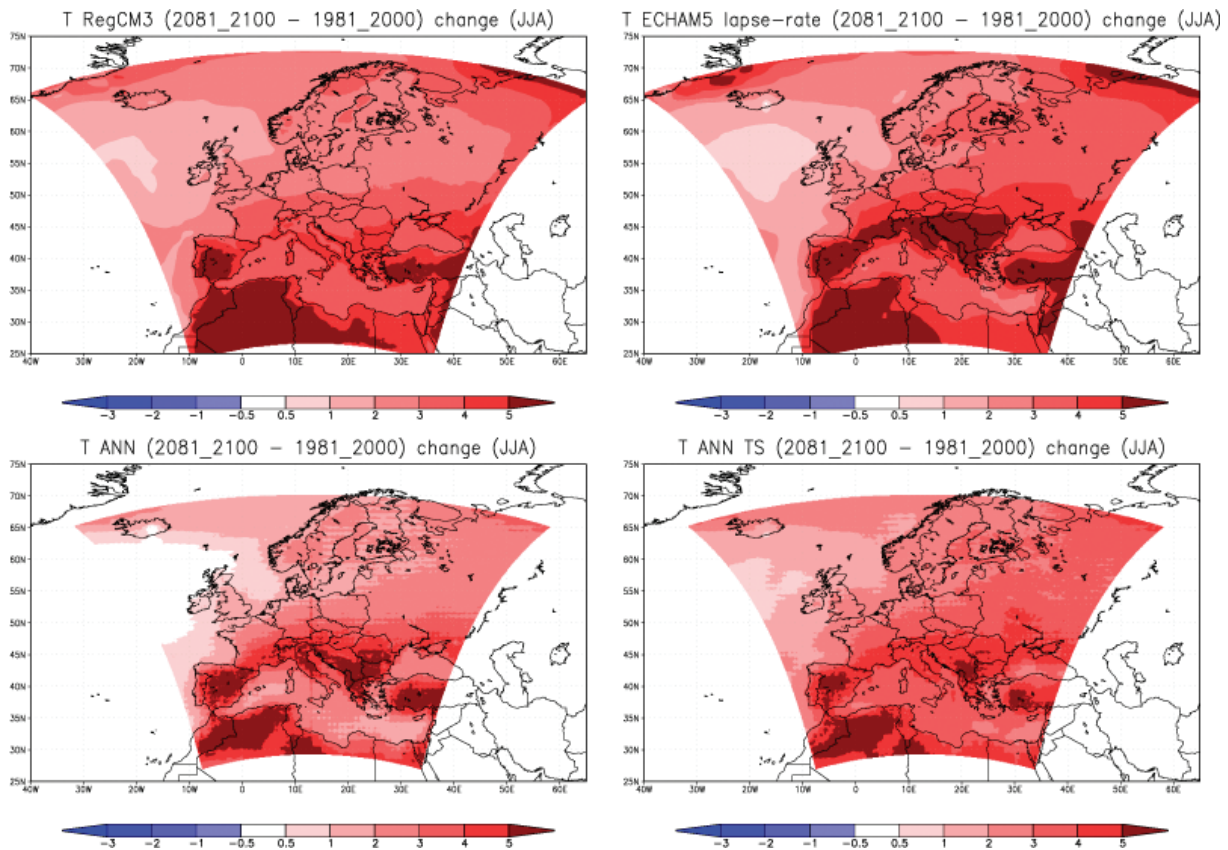


Fig. 7. JJA T ($^{\circ}\text{C}$) climate change signal, mean 2081–2100 – mean 1981–2000 for Europe. From RCM, GCM with lapse-rate correction, EurANN, and EurANNTS.

R fields between 2081–2100 and 1981–2000 were calculated for EurANN and MedANN, as shown in Figs. 7–9. Bias, RMSE and spatial correlation statistics for these climate change estimates are shown in Tables 4 and 5.

Figure 7 shows that EurANN reproduces the large-scale RCM CC signal for temperature relatively well, and is much more similar to the RCM field than the GCM is. However smaller scale T CC patterns are not so well represented. The ability of the ANN to reproduce the fine-scale RCM signal is variable between seasons, with Fig. 8 showing a reasonably accurate ANN estimate and Fig. 9 a less accurate example.

Figure 10 and Table 5 shows that the EurANN CC signal does not reproduce the RCM rainfall field well. In particular, the ANN shows a general drying over Southern Europe which is present in the GCM but not the RCM. The MedANN also performs relatively poorly, as can be seen in Fig. 11 and Table 5.

As the ANN is capable of reproducing the RCM 1981–2000 mean fields reasonably accurately and consistently, this failure to consistently reproduce the CC signal must be due to an inability to correctly reproduce the 2081–2100 fields. The ANN was trained with data from 1960–1980, and the

successful application of the ANN to other time periods relies on the relationship between input and output data remaining constant with time and changing climate. This assumption appears to hold for the validation period of 1981–2000, which is close to the training period, but may break down for time periods further in the future such as 2081–2100.

In order to try and improve the CC signal of the ANN, retraining was performed using a “time-slice” approach. In this case, training data was taken from 3 separate 10 yr time-periods (1961–1970, 2011–2020 and 2061–2070), and this time-slice ANN (ANNTS) was then used to produce T and R estimates for 1981–2000 and 2081–2100. The ANNTS was trained both for the whole European domain (EurANNTS) and for the Mediterranean (MedANNTS) in a similar way to the original ANN.

The ANNTS output for 1981–2000 is shown in Figs. 1–6, and it can be seen that this new approach produces results similar to the ones obtained from the original ANN. The ANNTS seems able to replicate the 1981–2000 RCM output as well as the original non-timeslice ANN, and this can also be seen in the statistics in Tables 2 and 3 and Figs. 12 and 13.

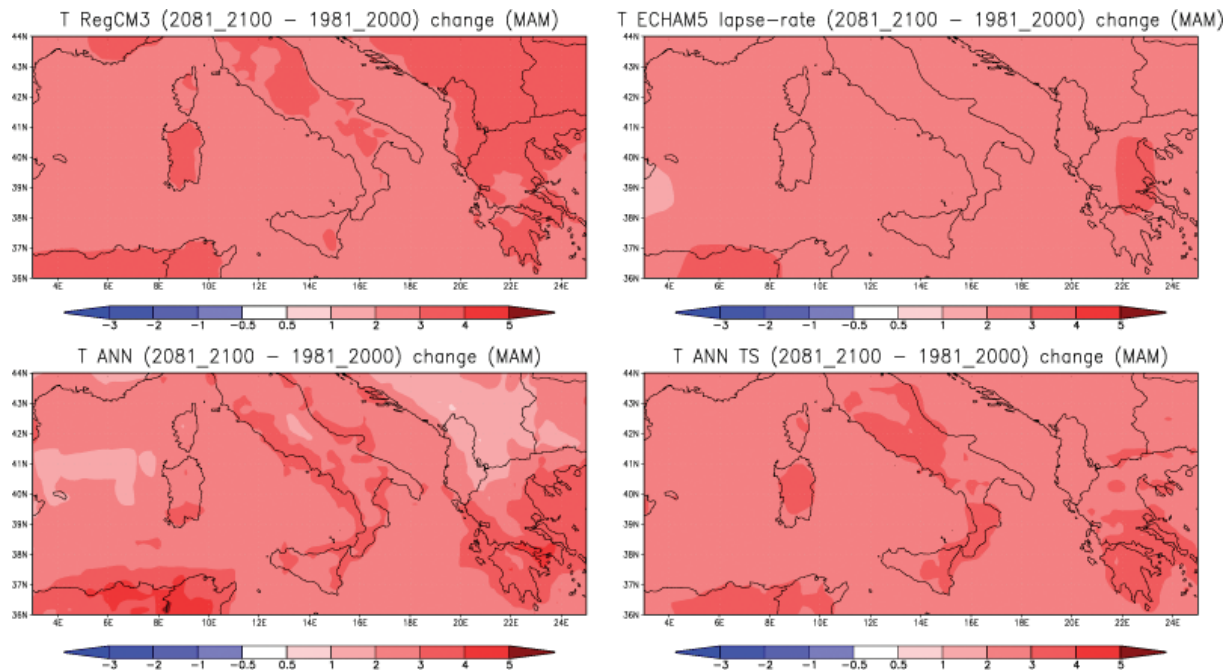


Fig. 8. MAM T ($^{\circ}\text{C}$) climate change signal, mean 2081–2100 – mean 1981–2000 for the Mediterranean region. From RCM, GCM with lapse-rate correction, MedANN, and MedANNTS.

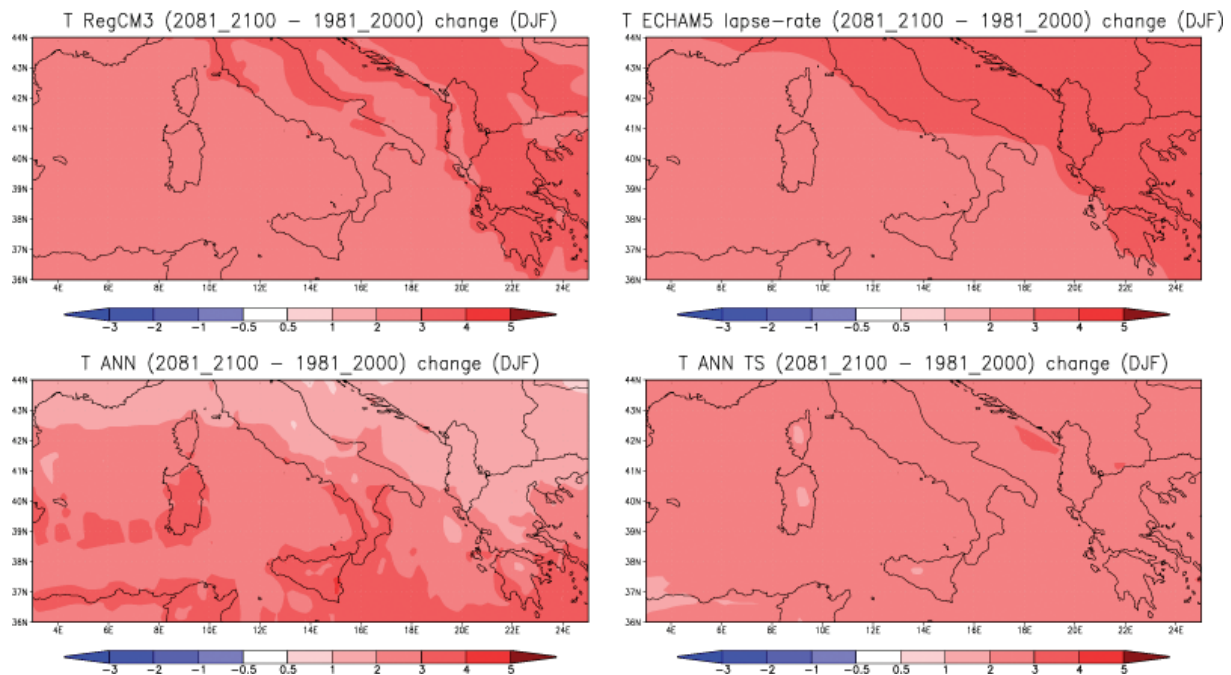


Fig. 9. DJF T ($^{\circ}\text{C}$) climate change signal, mean 2081–2100 – mean 1981–2000 for the Mediterranean region. From RCM, GCM with lapse-rate correction, MedANN, and MedANNTS.

The ANNTS CC signal is shown in Figs. 7–11, and here there is some improvement over the original ANN (see Tables 4 and 5 for a comparison of skill statistics). The EurANNTS CC signal shows that on the large scale the RCM

CC signal is well reproduced for both T and R (see Figs. 7, 10 and Tables 4 and 5), with the rainfall field in particular much improved over the original ANN and a slight improvement over the GCM. For the MedANNTS, the detail of the

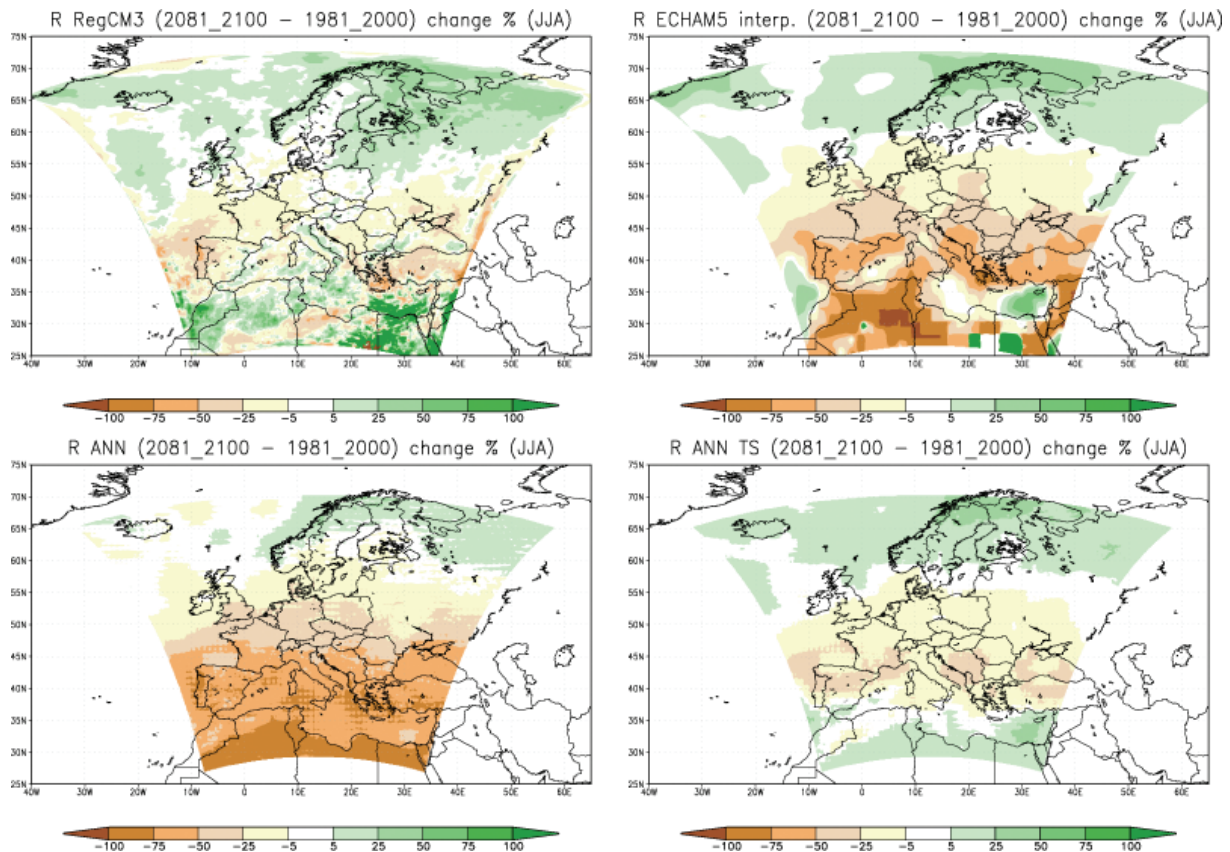


Fig. 10. JJA R (%) climate change signal, mean 2081–2100 – mean 1981–2000 for Europe. From RCM, GCM with lapse-rate correction, EurANN, and EurANNTS.

RCM CC signal is often reproduced well (and slightly better than the GCM), but in some cases is not an improvement over the GCM CC signal (e.g. DJF T and P). In most cases, the MedANNTS is an improvement over the original non-timeslice ANN. Although the ANNTS is able to correctly reproduce the RCM CC signal in many cases, this is not always true.

To examine why a timeslice training approach produces a more accurate CC signal than training with a single period, ANN and ANNTS estimates were produced for a series of six different 20 yr periods (shown in Fig. 12). Area-averaged Bias, RMSE and spatial correlation values of EurANN, EurANNTS, MedANN and MedANNTS (taking RegCM3 as “truth”) for the time-mean output of each of these periods are shown in Figs. 12 and 13. Figure 12 shows that while the GCM bias and RMSE remain approximately constant with time, the EurANN estimates degrade as the time period gets further from the ANN calibration period. In contrast the EurANNTS bias and RMSE do not drift with time. For MedANN and MedANNTS similar behaviour can be seen for the RMSE in Fig. 13, though there is no drift in bias of the MedANN.

This drift in accuracy explains the poor replication of the CC trend by the ANN, despite the accuracy of the ANN for any single time period remaining better than that of the GCM. The stationarity in bias and RMSE of ANNTS should make the CC trend obtained by this method more reliable than that of the original ANN.

5 Conclusions

In this paper we present an ANN approach for the downscaling of GCM temperature and precipitation fields to reproduce corresponding fields obtained with a nested RCM, and apply this approach to a nested RCM transient climate change simulation over Europe. The ANN appears able to realistically recreate RCM output fields from GCM input data for both T and R . An ANN trained with the full European domain represents large scale patterns well, but misses smaller rainfall features such as high rainfall over the mountains of Croatia. The EurANN appears to represent the RCM rainfall field more closely than the GCM rainfall field does (see Fig. 2). For temperature the lapse-rate corrected GCM, the EurANN and the RCM fields are all qualitatively similar (Fig. 1), but

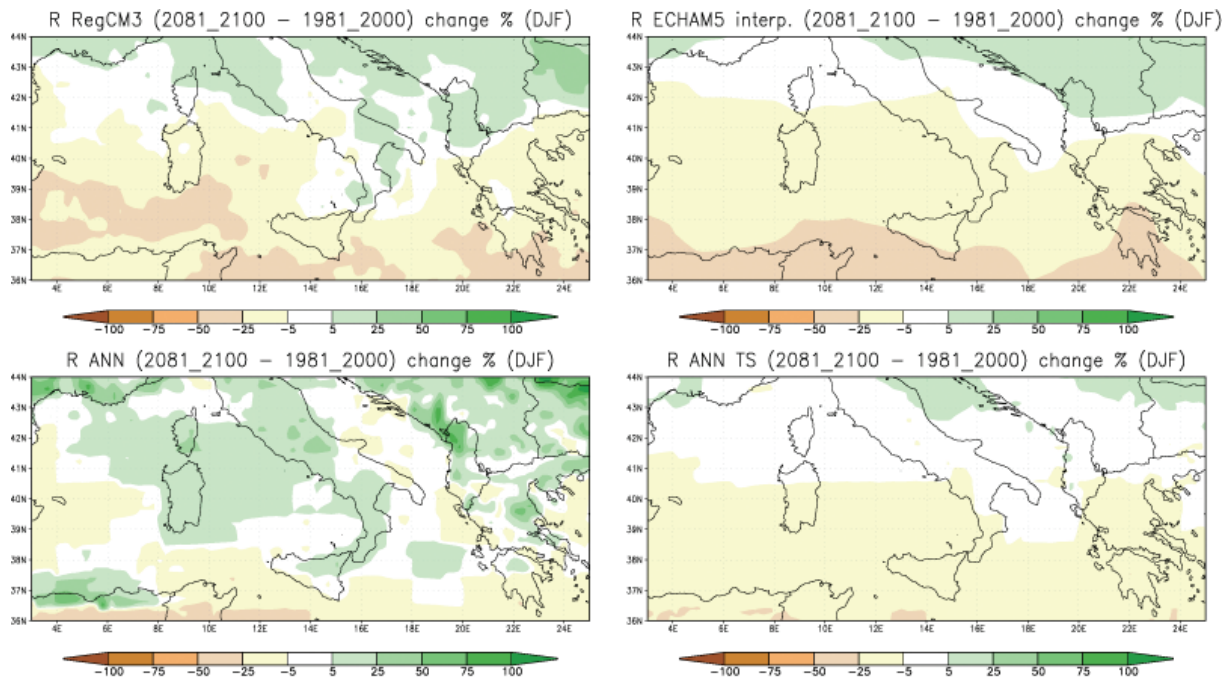


Fig. 11. DJF R (%) climate change signal, mean 2081–2100 – mean 1981–2000 for the Mediterranean region. From RCM, GCM bilinearly interpolated to RCM scale, MedANN, and MedANNTS.

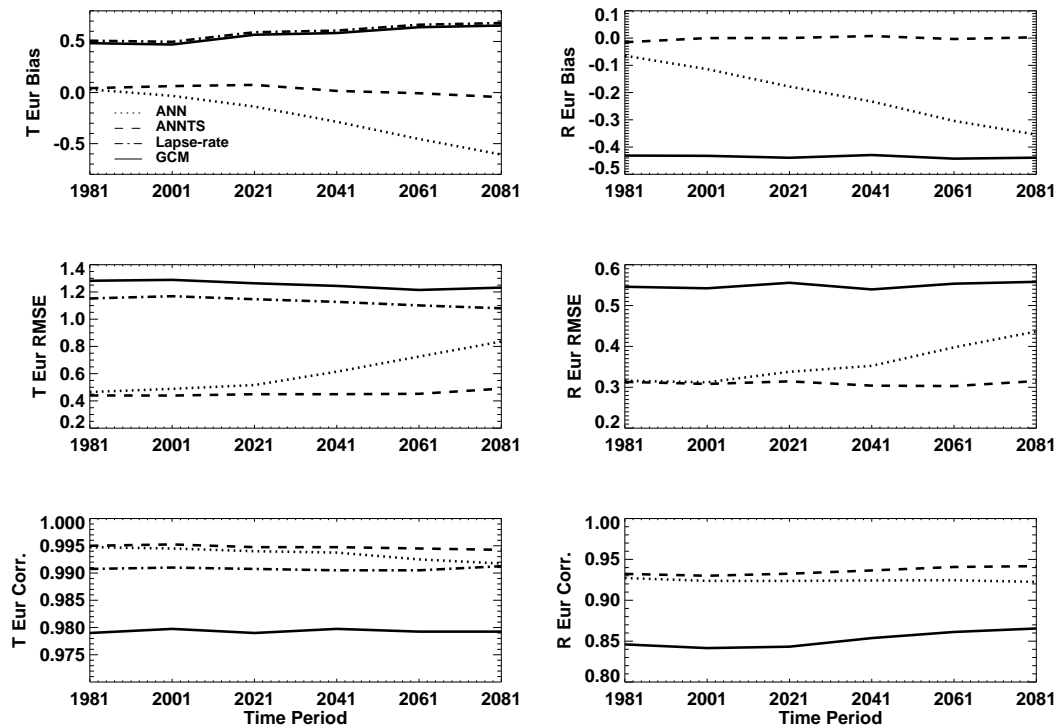


Fig. 12. Area averaged bias (top), RMSE (middle) and spatial coefficient of correlation (bottom) of EurANN (ANN trained over the full European domain), EurANNTS (Time-slice ANN trained over the full European domain), ECHAM5 lapse-rate corrected ECHAM5 (for T only) and ECHAM5 with respect to RegCM3 fields for 6 different time periods over Europe. Statistics are averaged over all four seasons. Left-hand plots show temperature statistics, right-hand plots show rainfall. Time periods are 20 yr means starting from the year shown on the x axis.

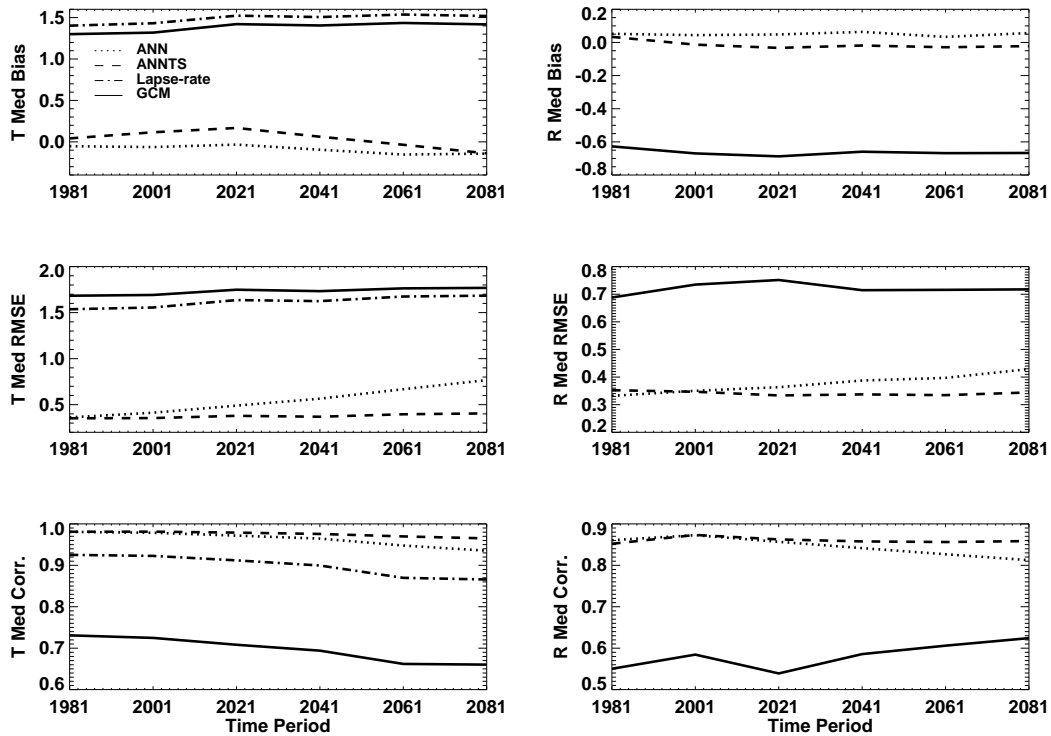


Fig. 13. Area averaged bias (top), RMSE (middle) and spatial coefficient of correlation (bottom) of MedANN, MedANNTS, ECHAM5 lapse-rate corrected ECHAM5 (for T only) and ECHAM5 with respect to RegCM3 fields for time-means of 6 different 20-yr periods over the Mediterranean region. Statistics are averaged over all four seasons. Left-hand plots show temperature statistics, right-hand plots show rainfall. Time periods are 20 yr means starting from the year shown on the x axis.

the EurANN skill scores are an improvement on the lapse-rate method (see period 1 (1981–2000) of Fig. 12).

Training the ANN over a smaller regional domain improves the level of detail that it can capture. In this case, the MedANN shows an improvement over the lapse-rate method for recreating RCM temperature fields (Fig. 5), and the MedANN rainfall fields represent RCM rainfall much more closely than the GCM rainfall does (Fig. 6). MedANN also shows an improvement over EurANN for this region.

Although the ability of ANNs to reproduce RCM mean fields for a certain time period is interesting, the real property of note here is whether an ANN approach can correctly replicate the climate change signal of an RCM. The original ANN, trained only with data from 1960–1980, does not reproduce the 2081–2100 – 1981–2000 CC signal well for either R or T , and for either the full European domain or the Mediterranean sub-region.

Figures 12 and 13 show that this is due to a drift with time in the accuracy of the ANN. As the validation time period gets further from the ANN training period the accuracy of the ANN decreases. This may be because the relationship between input and output data changes with time and changing climate, meaning that the patterns established during ANN training for 1960–1980 are not suitable for 2081–2100. One possibility is that regional circulation changes and/or

land-atmosphere feedbacks (e.g. soil moisture-precipitation) occur within the RCM during climate change that are not also present in the GCM. This could alter the relationship between RCM and GCM T and R values at coincident grid-points, and render the 1960–1980 calibration unreliable for later time periods.

We attempted to mitigate this problem by training the ANN model using different time slices in the future climate simulation (ANNTS model). With this approach the error remains stationary throughout the range of time periods. The improvement could be either because the ANNTS has access to a wider range of T and R data during training, because “year” is included in the ANN inputs and the ANN is able to apply a time-variable relationship between GCM input, elevation and output, or a combination of the two. In this way the ANNTS might be able to capture time-varying non-linear dynamical processes in the RCM.

The non-stationarity of ANN error shown here has implications for statistical-dynamical (and possibly statistical) GCM downscaling methods. In particular, any method which relies on the relationship between GCM variables, orography and RCM variables remaining constant with changing climate might also be expected to be subject to the same drift in accuracy as seen here. Training using future time slices improves the performance and applicability of the ANN

Table 4. Area averaged bias, RMSE and spatial correlation coefficient of temperature climate change signal between 1981–2000 and 2081–2100, for ANN, ANNTS, and lapse-rate corrected ECHAM5, taking RegCM3 fields as “truth”. First three columns show results for EurANN and EurANNTS validated over the European region, and the remaining columns show results for MedANN and MedANNTS validated over the Mediterranean region.

	Europe			Med		
	EurANN	EurANNTS	Lapse-rate/GCM	MedANN	MedANNTS	Lapse-rate
<i>T</i> Bias (°C)						
DJF	−0.692	−0.112	0.309	−0.359	−0.268	0.129
MAM	−0.524	−0.152	−0.002	−0.122	−0.093	−0.228
JJA	−0.673	−0.057	0.151	−0.176	−0.224	0.320
SON	−0.672	−0.033	0.235	0.299	−0.137	0.247
<i>T</i> RMSE (°C)						
DJF	0.697	0.206	0.413	0.695	0.274	0.220
MAM	0.562	0.253	0.325	0.410	0.160	0.239
JJA	0.887	0.365	0.457	0.719	0.346	0.613
SON	0.678	0.213	0.377	0.507	0.241	0.315
<i>T</i> corr.						
DJF	0.936	0.955	0.956	−0.359	0.654	0.847
MAM	0.932	0.944	0.947	0.209	0.718	0.788
JJA	0.894	0.917	0.908	0.722	0.580	0.562
SON	0.952	0.946	0.917	0.546	0.782	0.797

Table 5. Area averaged bias, RMSE and spatial correlation coefficient of temperature and rainfall climate change signal between 1981–2000 and 2081–2100, for ANN, ANNTS, and ECHAM5 interpolated to RegCM3 spatial scale, taking RegCM3 fields as “truth”. First three columns show results for EurANN and EurANNTS validated over the European region, and the remaining columns show results for MedANN and MedANNTS validated over the Mediterranean region.

	Europe			Med		
	EurANN	EurANNTS	Lapse-rate/GCM	MedANN	MedANNTS	GCM
<i>R</i> Bias (mm day ^{−1})						
DJF	−0.247	0.068	0.023	0.316	−0.023	−0.037
MAM	−0.255	0.065	0.035	0.071	0.087	0.131
JJA	−0.358	−0.055	−0.073	−0.319	−0.105	−0.053
SON	−0.300	−0.005	−0.016	0.278	0.599	0.517
<i>R</i> RMSE (mm day ^{−1})						
DJF	0.292	0.161	0.169	0.417	0.187	0.152
MAM	0.286	0.162	0.163	0.245	0.173	0.207
JJA	0.367	0.149	0.169	0.320	0.124	0.091
SON	0.328	0.147	0.173	0.287	0.231	0.247
<i>R</i> Corr						
DJF	0.750	0.887	0.803	0.395	0.696	0.831
MAM	0.793	0.889	0.850	−0.072	0.562	0.207
JJA	0.728	0.764	0.654	0.561	0.636	0.567
SON	0.556	0.756	0.622	0.278	0.599	0.517

approach. However our results do point to the importance of the issue of stationarity in statistical downscaling models, which needs to be addressed in further research.

The use of local inputs could in principle reduce the ability of the ANN to reproduce precipitation and temperature changes associated with changes in synoptic scale circulation. As the ANN performs well for the validation period 1981–2000, the use of local inputs does not appear to be a problem when the climate change signal between training and validation period is small. However, as noted above, one possible cause of the poor performance of the original non-timeslice ANN in the 2081–2100 period is that regional circulation changes occur in the RCM that are not present in the GCM. It is possible that the introduction of non-local inputs into the time-slice ANN could provide information to the ANN about any divergence between the RCM and GCM circulations with time, and this is something that merits further investigation.

The time-slice ANN was in general able to reproduce the RCM climate-change signal more accurately than the GCM, but not in all cases. We plan to further refine the ANNTS approach in order to use it as a tool to fill the GCM-RCM simulation matrix necessary to assess uncertainties in regional climate change projections.

Acknowledgements. This study was performed while RC was a visiting scientist in the Earth Systems Physics group of the ICTP, and was made possible by the kind invitation for him to spend two months working there. Helpful comments from two anonymous reviewers also helped to improve the manuscript.

Edited by: S. Vannitsem

Reviewed by: two anonymous referees

References

- Bastos, L. and O'Hagan, A.: Diagnostics for Gaussian process emulators, *Technometrics*, 51, 425–438, 2008.
- Bishop, C.: *Neural Networks for Pattern Recognition*, Clarendon Press, UK, 2000.
- Busch, U. and Heimann, D.: Statistical-dynamical extrapolation of a nested regional climate simulation, *Clim. Res.*, 19, 1–13, 2001.
- Chadwick, R. and Grimes, D.: An artificial neural network approach to multi-spectral rainfall estimation over Africa., *J. Hydrometeorol.*, submitted, 2011.
- Conti, S. and O'Hagan, A.: Bayesian emulation of complex multi-output and dynamic computer models, *J. Stat. Plan. Inf.*, 140, 640–651, 2010.
- Coppola, E.: Validation of improved TAMANN neural network for operational satellite rainfall estimation in Africa, *J. Appl. Meteorol.*, 45, 1557–1572, 2006.
- Deque, M., Jones, R., Wild, M., Giorgi, F., Christensen, J., Hassell, D., Vidale, P., Rockel, B., Jacob, D., Kjellstrom, E., de Castro, M., Kucharski, F., and van den Hurk, B.: Global high resolution versus Limited Area Model climate change projections over Europe: quantifying confidence level from PRUDENCE results, *Clim. Dynam.*, 25, 653–670, 2005.
- Dickinson, R., Henderson-Sellers, A., and Kennedy, P.: Biosphere-Atmosphere Transfer Scheme, BATS: version 1E as coupled to the NCAR Community Climate Model, Tech. Rep. NCAR/TN-387+STR, NCAR, 1993.
- Engen-Skaugen, T.: Refinement of dynamically downscaled precipitation and temperature scenarios, *Climatic Change*, 84, 365–382, 2007.
- Fritsch, J. and Chappell, C.: Numerical prediction of convectively driven mesoscale pressure systems, *J. Atmos. Sci.*, 37, 1722–1733, 1980.
- Giorgi, F. and Francisco, R.: Evaluating uncertainties in the prediction of regional climate change, *Geophys. Res. Lett.*, 27, 1295–1298, 2000.
- Giorgi, F. and Mearns, L.: Introduction to special section: Regional climate modeling revisited, *J. Geophys. Res.-Atmos.*, 104, 6335–6352, 1999.
- Giorgi, F., Marinucci, M., Bates, G., and DeCanio, G.: Development of a second generation regional climate model (RegCM2) 2: Convective processes and assimilation of lateral boundary conditions, *Mon. Weather Rev.*, 121, 2814–2832, 1993.
- Giorgi, F., Hewitson, B., Christenson, J., Fu, C., Jones, R., Hulme, M., Mearns, L., Storch, H. V., and Whetton, P.: Regional climate information – evaluation and projections., In *IPCC WG1 TAR*, 2001.
- Giorgi, F., Diffenbaugh, N., Gao, X., Coppola, E., Dash, S., Frumento, O., Seidou Sanda, I., Rauscher, S., Remedio, A., Steiner, A., Sylla, B., and Zakey, A.: Exploring uncertainties in regional climate change: The Regional Climate Change Hyper-Matrix Framework., *Eos*, 89, 445–446, 2008.
- Grell, G., Dudhia, J. and Stauffer, D.: Description of the fifth generation Penn State/NCAR Mesoscale Model, MM5, Tech. Rep. NCAR/TN-398+STR, NCAR, 1994.
- Haupt, S., Pasini, A., and Marzban, C. (Eds.): *Artificial Intelligence Methods in the Environmental Sciences*, Springer, 2009.
- Hornick, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359–366, 1989.
- Hsieh, W.: *Machine learning methods in the environmental sciences*, Cambridge University Press, Cambridge, UK, 2009.
- Kiehl, J., Hack, J., Bonan, G., Boville, B., Briegleb, B., Williamson, D., and Rasch, P.: Description of the NCAR Community Climate Model, CCM3, Tech. Rep. NCAR/TN-420+STR, NCAR, 1996.
- Lönnblad, L., Peterson, C., and Rögngvaldsson, T.: Using neural networks to identify jets, *Nuclear Phys. B*, 349, 675–702, 1991.
- Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M., Venema, V., Chun, K., Goodess, C., Jones, R., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, doi:10.1029/2009RG000314, 2010.
- Pal, J., Small, E., and Eltahir, E.: Simulation of regional-scale water and energy budgets: Representation of subgrid cloud and precipitation processes within RegCM, *J. Geophys. Res.*, 105, 29579–29594, 2000.
- Pal, J., Giorgi, F., Bi, X., Elguindi, F., Solomon, F., Gao, X., Francisco, R., Zakey, A., Winter, J., Ashfaq, M., Syed, F., Bell, J., Diffenbaugh, N., Kamacharya, J., Konare, A., Martinez, D., da Rocha, R., Sloan, L., and Steiner, A.: The ICTP RegCM3

- and RegCNET: Regional climate modelling for the developing world, *Bull. Amer. Meteorol. Soc.*, 88, 1395–1409, 2007.
- Pasini, A. and Langone, R.: Attribution of precipitation changes on a regional scale by neural network modelling: A case study, *Water*, 2, 321–332, 2010.
- Pasini, A., Loré, M., and Ameli, F.: Neural network modelling for the analysis of forcings/temperatures relationships at different scales in the climate system, *Ecol. Model.*, 191, 58–67, 2006.
- Picton, P.: *Neural Networks*, 2nd Edn., Palgrave, Basingstoke, UK, 2000.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornbluh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmosphere general circulation model ECHAM5. Part 1: Model description, Tech. Rep. Rep. 349, Max Planck Institute for Meteorology, 2003.
- Rumelhart, D. and McClelland, J.: *Parallel Distributed Processing*, MIT press, Cambridge, MA, 1986.
- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H. (Eds.): *IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.
- Tomassetti, B., Verdecchia, M., and Giorgi, F.: NN5: A neural network based approach for the downscaling of precipitation fields – Model description and preliminary results, *J. Hydrol.*, 367, 14–26, 2009.
- Van Der Linden, P. and Mitchell, J.: *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project*, Tech. rep., Met Office Hadley Centre, Fitzroy rd, Exeter, EX1 3PB, UK, 2009.
- Wilby, R., Wigley, T., Conway, D., Jones, P., Hewitson, B., Main, J., and Wilks, D.: Statistical downscaling of general circulation model output: A comparison of methods, *Water Resour. Res.*, 34, 2995–3008, 1998.
- Zeng, X., Zhao, M., and Dickinson, R.: Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data, *J. Climate*, 11, 2628–2644, 1998.