Theor. Appl. Climatol. 76, 1–12 (2003)
DOI 10.1007/s00704-003-0008-5

Theoretical and Applied Climatology
Printed in Austria

[1] Johann Wolfgang Goethe-University, Institute for Meteorology and Geophysics, Frankfurt a.M., Germany
[2] Deutscher Wetterdienst, Offenbach a.M., Germany

Nonlinear statistical attribution and detection of anthropogenic climate change using a simulated annealing algorithm

A. Walter[1,2] and **C. D. Schönwiese**[1]

With 9 Figures

Received August 9, 2002; revised May 12, 2003; accepted June 1, 2003
Published online October 16, 2003

Summary

The climate system can be regarded as a dynamic nonlinear system. Thus, traditional linear statistical methods fail to model the nonlinearities of such a system. These nonlinearities render it necessary to find alternative statistical techniques. Since artificial neural network models (NNM) represent such a nonlinear statistical method their use in analyzing the climate system has been studied for a couple of years now. Most authors use the standard Backpropagation Network (BPN) for their investigations, although this specific model architecture carries a certain risk of over-/underfitting. Here we use the so called Cauchy Machine (CM) with an implemented Fast Simulated Annealing schedule (FSA) (Szu, 1986) for the purpose of attributing and detecting anthropogenic climate change instead. Under certain conditions the CM-FSA guarantees to find the global minimum of a yet undefined cost function (Geman and Geman, 1986).

In addition to potential anthropogenic influences on climate (greenhouse gases (GHG), sulphur dioxide (SO_2)) natural influences on near surface air temperature (variations of solar activity, explosive volcanism and the El Niño/Southern Oscillation phenomenon) serve as model inputs. The simulations are carried out on different spatial scales: global and area weighted averages. In addition, a multiple linear regression analysis serves as a linear reference.

It is shown that the adaptive nonlinear CM-FSA algorithm captures the dynamics of the climate system to a great extent. However, free parameters of this specific network architecture have to be optimized subjectively. The quality of the simulations obtained by the CM-FSA algorithm exceeds the results of a multiple linear regression model; the simulation quality on the global scale amounts up to 81% explained variance. Furthermore the combined anthropogenic effect corresponds to the observed increase in temperature Jones et al. (1994), updated by Jones (1999a), for the examined period 1856–1998 on all investigated scales. In accordance to recent findings of physical climate models, the CM-FSA succeeds with the detection of anthropogenic induced climate change on a high significance level. Thus, the CM-FSA algorithm can be regarded as a suitable nonlinear statistical tool for modeling and diagnosing the climate system.

1. Introduction

Modern climatology is facing the question whether an anthropogenic induced climate change is already observable in climatic variables, e.g. near-surface air temperature. Because the climate system can be regarded as a nonlinear system (Houghton et al., 2001), traditional linear statistical models are not capable of describing the climate system in its full complexitiy and thus fail to answer this question.

Since nonlinear Neural Network Models (NNM) provide a statistical solution to this problem their application towards analyzing the climate system has been studied for a couple of years, see for example Grieger and Latif (1993), Hsu et al. (1997), Hsieh and Tang (1998) and Walter et al. (1998). For their investigations most

authors use the standard NNM, the Backpropagation Network (BPN) (Rumelhart et al., 1986), although the BPN algorithm has one big disadvantage: it cannot guarantee to reach the global minimum of a yet undefined cost function and thus carries a certain risk of over-/underfitting.

In an earlier work (Schönwiese et al., 1997), where the BPN architecture was used, we obtained a GHG-signal of 0.8°–1.3 °C and a combined GHG + SU signal of 0.5–0.8 °C for the period 1866 to 1994. Here we will present an updated simulation (1856–1998) and furthermore use a more sophisticated simulation technique, the so-called Cauchy Machine (CM), for the purpose of simulating, attributing and detecting anthropogenic climate change signals in observed variations of near-surface air temperature. In contrast to the BPN, the CM and its implemented Fast Simulated Annealing (FSA) learning algorithm (Metropolis et al., 1953), (Szu, 1986), see Section 3.1, guarantee under certain conditions to reach the global minimum of any given cost function (Geman and Geman, 1986). Thus, in this work the CM-FSA architecture is applied for the attribution and detection of anthropogenic climate change.

The data used for this approach is described in Section 2, whereas Section 3 explains the basic concepts of neural network models, with an emphasis on the CM-FSA in Section 3.1. Section 4 deals with the crucial issue of statistical isolation of climatological cause-effect relations. The preprocessing of the data is briefly outlined in Section 5. The results of our approach are presented in Section 6 and the paper ends with some conclusions drawn from our results in Section 7.

2. Climatic fundamentals and data

A change in average net radiation at the top of the atmosphere, because of a change in either solar or infrared radiation, is called a *radiative forcing*.

Such a radiative forcing perturbes the balance between incoming and outgoing radiation. Over time climate responds to this pertubation to reestablish a radiative balance. Thus, a positive radiative forcing tends to warm the surface and vice versa. For example an increase in atmospheric CO_2 concentration leads to a reduction in outgoing infrared radiation and a positive radiative forcing. Therefore the global mean surface temperature change due to a change in a specific radiative forcing can be written as

$$\Delta T_s = \Delta F \cdot \lambda, \tag{1}$$

where ΔF is the change in the forcing and λ is the nearly invariant[1] climate sensitivity parameter. The invariance of λ has made the radiative forcing concept a convenient measure to estimate the global annual mean temperature response (ΔT_s) to a certain forcing mechanism.

The radiative effects of the major GHGs beside CO_2 (e.g. Methane (CH_4), Nitrous Oxide (N_2O), Halocarbons (mainly CFC-11) and Ozone (O_3)) are often represented by an equivalent CO_2 concentration which is the CO_2 concentration that gives a radiative forcing equal to the sum of the forcings for the individual GHG. We used a representation of anthropogenic GHG forcing in terms of CO_2 equivalents also used by Houghton et al. (2001).

Anthropogenic aerosols scatter and absorb short-wave and long-wave radiation thereby perturbing the energy budget of the atmosphere and exerting a direct radiative forcing (direct effect). In addition, aerosols serve as cloud condensation and ice nuclei thereby modifing the radiative properties and lifetime of clouds (difficult to estimate indirect effect). Because anthropogenic sulfate aerosols have only a atmospheric lifetime of a few days this forcing may be directly proportional to the corresponding SO_2 emissions and is therefore believed to be strongest over industrialized regions of the northern hemisphere. We used the updated SO_2 emission data from Charlson et al. (1992), the obtained signals are referred to as SU. Other than sulphate aerosols, e.g. black carbon aerosol, organic carbon aerosol, have not been considered in this work.

Radiative forcing may lead to climate variations but climate variations can also be initialized from internal interactions between components of the climate system. Therefore a distinction between externally and internally induced natural climate variability has to be made. Thus, a certain knowledge about natural climate variability is necessary for the isolation of anthropogenic cause-effect relations, see Section 4 for statistical details.

[1] Typically about $0.5\,\mathrm{K/Wm^{-2}}$ (Ramanathan et al., 1985).

Variations in the solar output are a source for externally induced natural climate variability. However, only since the late 1970s variations of solar irradiance have been measured directly and therefore it is necessary to use other proxy data, e.g. sunspot numbers (Stevens and North, 1996), of the solar activity to deduce variations at earlier dates. In the simplest type of reconstruction a proxy measure is calibrated against recent measurements and extrapolated backwards using a linear relationship. The time series used here to describe solar forcing is from Lean et al. (1995) and Lean and Rind (1999) respectively and shows clearly the well known 11-year solar cycle imposed by a longer modulation.

Episodic, explosive volcanic eruptions lead to a significant enhancement of the aerosol concentration in the stratosphere. The most dramatic recent volcanic event was the eruption of Mt. Pinatubo in 1991 which reached a peak forcing of about $-3\,\mathrm{W/m^{-2}}$ in late 1991 (Hansen et al., 1998), thus tending to cool the earth's surface. Stratospheric aerosol levels have meanwhile fallen well below the peak values of 1991 to 1993 and are comparable to the low levels seen in 1979 (Houghton et al., 2001). Explosive volcanism whose ejecta reach the stratosphere and form climate relevant sulfate aerosols is considered here in terms of heating anomalies as provided by Grieser and Schönwiese (1998).

Furthermore we used a reconstructed time series of the El Niño/Southern Oscillation (ENSO) phenomenon provided by Staeger (1998) based on Jones (1999b). The linear correlation between these two series for the time period 1866–1998 amounts to 0.96. ENSO is the primary natural climate variability factor in the 2–7 year domain. El Niño is defined by anomalies of sea surface temperatures (SST) in the eastern tropical Pacific, while the Southern Oscillation Index (SOI) is a measure of the atmospheric circulation response in the Pacific-Indian-Ocean region.

ENSO is generated by ocean-atmosphere interactions in the tropical Pacific but affects climate globally. Beside having fundamental consequences on local climate ENSO seems to have a global influence: during and following El Niño, the global mean surface temperature increases as the ocean transfers heat to the atmosphere (Sun and Trenberth, 1998). The shift of rainfall locations in the tropics due to an ENSO event alters the heating patterns of the atmosphere which forces large scale waves in the atmosphere. These establish meridional teleconnections, that extend to mid-latitudes altering the winds and changing the jet stream and storm tracks (Trenberth et al., 1998) which may lead to modified weather patterns in mid-latitudes as well.

Annual time series of observed surface temperature variations were used in our simulations

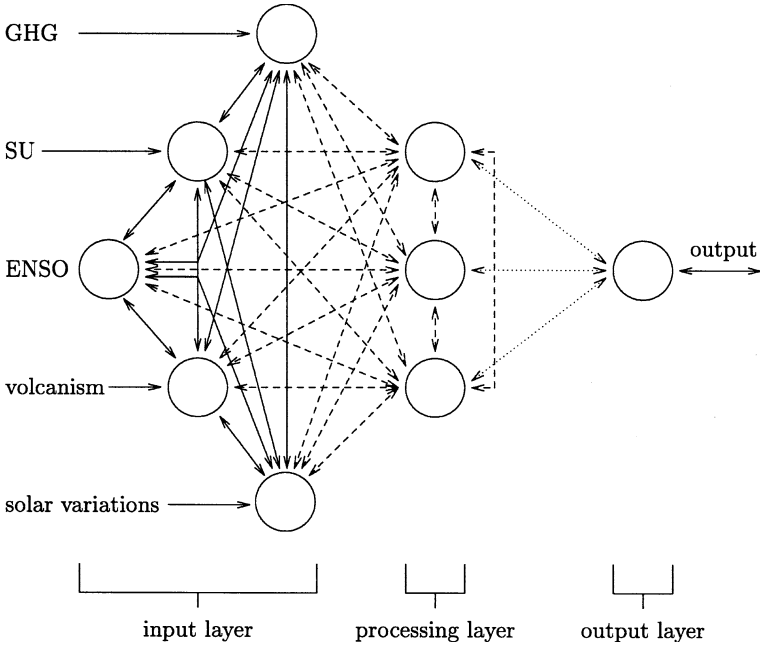

Fig. 1. Schematic illustration of the model configuration used. Weights within the input layer drawn solid, within the processing layer dashed and weights from processing to output layer drawn dotted. The number of processing units shown does not reflect the number of processing units for our simulations

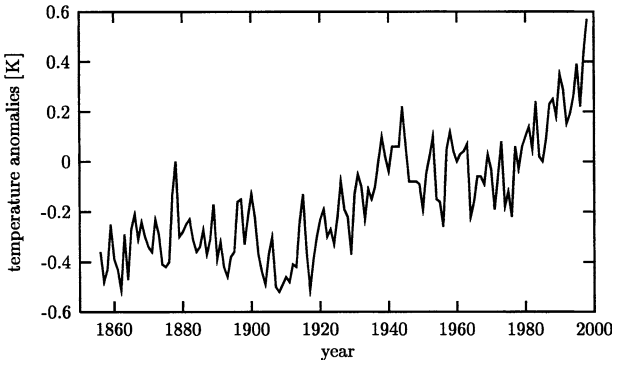

Fig. 2. Global temperature anomalies 1856–1998 (Jones et al., 1994), updated by (Jones, 1999a)

for the period 1856–1998, 1892–1995 for the area weighted averages respectively. A high quality data set of mean global and mean hemispheric surface air temperature provided by Jones et al. (1994) and updated regulary by Jones (1999a) served as the target function in our modeling approach. The area weighted time series used were derived from Jones et al. (1994) and Jones (1999a) respectively, the area design is according to Hansen and Lebedeff (1987).

A schematic model configuration is shown in Fig. 1, whereas global surface temperatures anomalies for the time period 1856–1998 (target function) are shown in Fig. 2. The forcing time series considered above (model inputs) are shown in Fig. 3.

3. Neural network models

The spirit of neural network modeling is to use fully nonlinear functions and use a large number of terms so that model mismatch errors are not a concern. Instead of matching the architecture of the model to a problem, a model is used that can describe almost anything, and careful training of the model is used to constrain it to describe the data.

NNM have their biological foundations in studying the learning mechanisms of the brain (Adrian, 1926), (Rosenblatt, 1958) and (Grossberg, 1982) and attempt to transfer these learning capabilities into the language of *Neurocomputing* (Anderson and Rosenfeld, 1986). NNM *learn* inherent data features using a data subset as training data and test these learned features using a unknown verification subset. This technique is called cross-validation (Stone, 1974).

The standard NNM is still the Backpropagation Network (BPN) (Rumelhart et al., 1986).

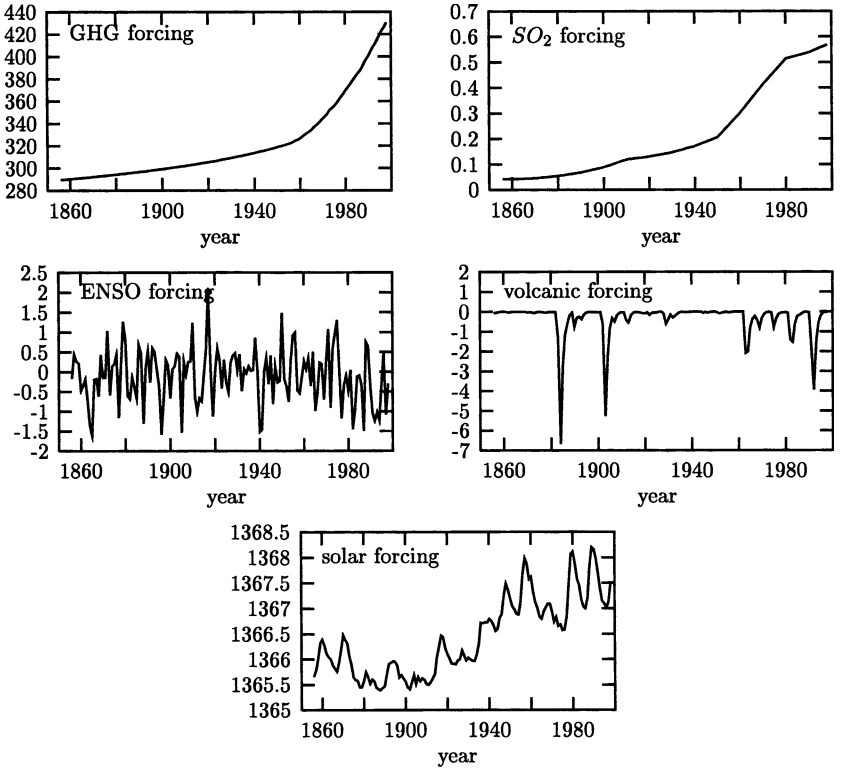

Fig. 3. Forcing mechanisms used in our simulations. Units are: GHG forcing, CO_2 equivalent concentrations [ppm] = parts per million by volume (Houghton et al., 2001); SO_2 forcing data from (Charlson et al., 1992) [mg/m^2]; ENSO forcing, normalized pressure anomalies [hPa] data from (Staeger, 1998) based on (Jones, 1999b); heating rate anomalies [W/m^2] due to explosive volcanism provided by (Grieser and Schönwiese, 1998) and solar forcing [W/m^2] (Lean et al., 1995), (Lean and Rind, 1999)

The BPN is based on a supervised learning algorithm to find the global minimum of a yet undefined cost function. There exists no objective criteria to adjust free parameters (e.g. learning constant, number of processing units) of the model properly (Freeman and Skapura, 1992), and thus the BPN carries a certain risk of being stuck in local minima of the cost function which might lead to false simulation results.

The CM, see Section 3.1 for details, uses a stochastic learning law instead, which, under certain conditions, guarantees to reach the global minimum of a cost function (Geman and Geman, 1986) and thus reduces the risk of overfitting.

The input to a typical NNM is a vector of elements (x_k), here the choosen climate forcings, see Section 2, therefore the NNM used in this application consists of five input neurons (potential forcings) and one output (surface temperature) neuron, see again Fig. 1 for a schematic illustration of the model configuration used.

One big disadvantage of non-linear statistical models is that their behavior for non-stationary processes is not well understood. The inputs in our model (GHG, SU, solar activity, volcanism and ENSO, see Fig. 3) as well as our target function (mean global surface air temperature, see Fig. 2) all reveal some characteristic time structures. For example, the GHG forcing shows a progressive trend, whereas the SU forcing shows a more unsteady behaviour. Even in mean global surface air temperature series a linear trend of roughly 0.6 °C for the period 1856–1998 is obvious. Trendy series are non-stationary. To ensure that our model captured all characteristics of the data anyhow, we had to make sure that the whole range of amplitudes of the time series considered is covered during the training process of the CM-FSA. Otherwise the model would fail to simulate a reasonable cause-effect relation for e.g. a high *GHG* forcing value if such a high value never occured during training. This can be done by *selecting* the values for training and validation so that the model is given the ability to *learn* the cause-effect relations associated with extreme values. We used 75% of the data for training and the remainder for validation. Using this approach the problems of modeling non-stationary time series with a non-linear model can be avoided for the better part.

3.1 Simulated Annealing and the Cauchy Machine

The BPN relies on the minimization of the mean square error χ^2

$$\chi^2 = \frac{1}{2}\sum_n [y(x_k) - Y(x_k)] \tag{2}$$

between the networks output (Y) and the given test data (y), which in the case considered here is observed surface temperature variations. x_k are the $k = 5$ dimensional input forcings introduced in Section 2 and n is the length of the record. The BPN then tries to reduce χ^2 by means of gradient descent down an error surface with a topology that is not well understood (Freeman and Skapura, 1992). This carries the risk of being caught in local minima of the χ^2-hyperplane because only downward steps were allowed.

If caught in a local minimum the effect is that the network appears to stop learning and the error does not decrease any further with additional training.

In this section a method for reducing the possibility of falling into local minima is presented. This method is called Simulated Annealing (SA) because of its strong analogy to the physical annealing process done to metals and other substances. A statistical method analogue to the one used in the physical annealing process was introduced by Metropolis et al. (1953).

With this now so-called Metropolis algorithm the first analogy between a physical thermodynamical system and mathematical function minimization has been introduced into statistics. Prior to the introduction of the Metropolis algorithm all other algorithms converged to the nearby solution as quick as possible. The Metropolis algorithm on the other hand is able to get out of local minima of the function to be minimized. Thus, with this algorithm the famous traveling salesman problem of finding the shortest cyclical itenirary for a traveling salesman who must visit each of N cities in turn has been effectively 'solved'.

To perform a simulated annealing process with a neural network we have to postulate that it is possible to extend the analogy between information theory and statistical mechanics to allow us to place our NNM in contact with some heat reservoir at some yet undefined, temperature.

If so, during the simulated annealing process we can gradually lower the system temperature while processing takes place in the network in the hopes of avoiding local minima on the energy landscape, i.e. the χ^2-hyperplane.

To perform this process, we have to simulate the effects of temperature on our model. In a physical system, molecules have an average kinetic energy proportional to the temperature of the system. Thereby individual molecules may have more or less kinetic energy than the average and random collisions may cause a molecule to gain or lose energy. This behaviour can be simulated in a NNM by adding a stochastic element to the processing. Instead of a deterministic procedure the system is heated to a certain temperature T and the output of each neuron is determined stochastically according to the Boltzmann distribution

$$\frac{P_\alpha}{P_\beta} = \exp^{-(E_\alpha - E_\beta)/T}, \tag{3}$$

where P_α is the probability of being in the αth global state and E_α is the energy of this state.

If only binary outputs are allowed to describe the state of the network, than for a single neuron, y_k, with the network energy E_a for $y_k = 1$ and E_b when $y_k = 0$, regardless of the previous state of y_k, we can set $y_k = 1$ with a probability of

$$p_k(y_k \equiv 1) = \frac{1}{1 + \exp^{-\Delta E_k/T}}, \tag{4}$$

where $\Delta E_k = E_b - E_a$. Equation (4) ensures that, every so often, a neuron will update so as to increase the energy of the system, thus helping the system to get out of local minima by moving upward on the χ^2-hyperplane. Because of the fact that a change of a single units output will change the state of the whole model this algorithm can be regarded as a local decision rule.

As processing continues, the control parameter T is reduced gradually. In the end, there will be a high probability that the system is in a global energy minimum which is corresponding to a global minimum of χ^2.

The system energy of such a network can be computed from

$$E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1_{j\neq i}}^{n} w_{ij} y_i y_j, \tag{5}$$

where w_{ij} is the weight between neuron i and neuron j, y_i and y_j are the outputs of neuron i and j, respectively, and n is the total number of processing neurons in the network.

The function to be minimized when using a CM is not the least square error Eq. (2), but the information theoretic quantity G, known as information gain or relative entropy (Ackley et al., 1986)

$$\begin{aligned} G &= \sum_{i=1}^{q} P_{1i} \log_2 \frac{P_{1i}}{P_{2i}} \\ &= \sum_{i=1}^{q} P_{1i} \log_2 P_{1i} - \sum_{i=1}^{q} P_{1i} \log_2 P_{2i}. \end{aligned} \tag{6}$$

Here P_{1i} and P_{2i} are two symbol probabilities of two sources S_1 and S_2 each containing q symbols.

The second term of the right side of Eq. (6) is not the entropy of a source. The $\log_2 P_{2i}$ terms are weighted by the S_1 probabilities, P_{1i}, rather than by the S_2 probabilities P_{2i}. Thus, G can be thought of as a measure of the distance, in bits, from source S_2 to source S_1. The term P_{2i} in Eq. (6) is dependant on the networks weights w_{ij}, so that G can be altered by altering these weights. The learning law of a CM can thus be written as

$$\frac{\partial G}{\partial w_{ij}} = -\frac{1}{T}(p_{ij}^- - p_{ij}^+), \tag{7}$$

where p_{ij}^- and p_{ij}^+ are the so called co-occurrence probabilities which compute the frequency that neurons i and j both are active, i.e. an output value of 1, if averaged over all possible combinations of patterns. The weight updates are then calculated according to

$$\Delta w_{ij} = \varepsilon(p_{ij}^+ - p_{ij}^-), \tag{8}$$

where ε is a learning constant which has to be carefully chosen ($[0 < \varepsilon \ll 1]$). From Eq. (8) it is obvious that the weights will continue to change as long as the two co-occurrence probabilities differ. For a more complete derivation of Eq. (7) the reader is referred to Rumelhart and Mc Clelland (1986).

As a suitable annealing schedule we used

$$T(t_n) = \frac{T_0}{1 + t_n} \tag{9}$$

given in (Szu, 1986), where T_0 is the initial temperature of the system and t_n is a discrete time variable corresponding to the n-th training step. In contrast to the annealing schedule given in (Geman and Geman, 1986) Eq. (9) is called fast simulated annealing (FSA) (Szu, 1986).

The training algorithm of a CM can thus be described as follows:

1. one training vector is clamped to the visible units of the network.
2. annealing of the network according to the annealing schedule Eq. (9) until equilibrium is reached at a desired maximum temperature.
3. The network is run for several more processing cycles, after each cycle the pairs of neurons with $y_k = 1$ (on) are simultaneously determined.
4. The co-occurrence results from step 3 are averaged.
5. Steps 1 to 4 are repeated for all training vectors. To get an estimate of p_{ij}^{+} the co-occurrence results for each pair of connected units are averaged.
6. The visible units are unclamped and the network is annealed until equilibrium is reached at a desired minimum temperature.
7. The network is run for several more processing cycles. After each cycle the pairs of connected units with simultaneous values $y_k = 1$ (on) are determined.
8. The co-occurrence results from step 7 are averaged.
9. Steps 6 through 8 are repeated as often as in step 5. The co-occurrence results are again averaged to get an estimate of p_{ij}^{-} for each connected pair of units.
10. The appropriate weight change is computed using Eq. (7)
11. Steps 1 through 10 are repeated until p_{ij}^{+} and p_{ij}^{-} are sufficiently small.

Beside the analysis of near-surface air temperature variations this work is aimed at the attribution and detection of anthropogenic climate change. Therefore, in a first step, the isolation of cause-effect relations is performed using the CM-FSA. After the successful statistical isolation of these relations detection studies based on a test of one-sided Gaussian distribution are performed.

4. Statistical isolation of cause-effect relations and detection strategy

This study is aimed at the statistical assessment of climatic cause-effect relations, especially the estimation of the anthropogenic influence on surface temperature. Thus, the results will reveal a characteristic time-structure and magnitude (in Kelvin [K]). In this text we will refer to this specific structure as a *signal*, e.g. GHG-signal. The presence of natural climate variability implies that this statistical isolation of relevant cause-effect relations is basically a signal-in-noise problem. Furthermore the signals have to be estimated reliably to obtain a meaningful detection variable in the second step.

Despite the fact that the CM has no linear components, such an estimation can be obtained by driving the CM in its final configuration, i.e. frozen weights, with one forcing time-series at a time, thus setting all other inputs to their mean.

The term *detection* in this context refers to the process of demonstrating that a simulated climate change is significantly different than can be explained by natural climate variability alone. To get a realistic estimation of this natural climate variability it is necessary to consider all potential natural and anthropogenic climate forcing mechanisms in the simulation at the same time. The unexplained parts of the simulations (residuals) are added to these statistically extracted causes of climate variability so that a potential signal is tested on a certain significance level against all other extracted signals plus residuals. This way we obtain what we call *climate noise*.

If the ratio between an anthropogenic greenhouse forcing signal at a given location x and time t $S_{anthr}(x,t)$ and the standard deviation of climate noise $s_{noise}(x)$ is denoted by a *detection variable* $d(x,t)$, it is possible to compute the space-time related probability of climate change on a certain significance level $a(Si)$

$$S_{anthr}(x,t) > a(Si) \cdot s_{noise}(x) \tag{10}$$

which leads to the definition of the detection variable $d(x,t)$

$$d(x,t) \equiv \frac{S_{anthr}}{s_{noise}}, \tag{11}$$

which is based on the signal-to-noise ratio (Von Storch and Zwiers, 1999).

The probability P of an anthropogenic climate change, i.e. the significance level, can thus be computed using

$$P(z \leq |d(x,t)|) = \mathrm{erf}\left(\frac{d(x,t)}{\sqrt{2}}\right), \tag{12}$$

where

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2)\, du \qquad (13)$$

is the error function which can be treated with numerical methods (Press et al., 1992). Equation (12) can be applied here because s_{noise} is sufficiently Gaussian distributed which was tested using a Kolmogorv-Smirnoff test (Press et al., 1992).

5. Preprocessing of the data

Due to the fact that with the exception of global or hemispheric mean temperature all spatial data sets represent variations in time and space a preanalysis using empirical orthogonal functions (EOF) (Preisendorfer, 1988) was performed for the area weighted time series of near surface air temperature. In this way we obtain 72 time-related principal components ranked according to their explained variance instead of 72 climate variable time series at 72 different areas. This transformation can be written as

$$z(x,t) = \sum_{j=1}^{m} \lambda_j EOF_j(x) PC_j(t), \qquad (14)$$

where $z(x,t)$ is the original space-time related data field transformed into m time-related principal components $PC_j(t)$ and a series of space-related principal components called empirical orthogonal functions $EOF_j(x)$. The factor λ_j is the eigenvalue and quantifies the amount of variance of the related principal component existent in the original data. The EOF_j provides the information about the weight of the corresponding PC existent at the related point of space. The $PC_j(t)$ serve further on as the target function of the analysis on the area weighted scale. In default of a objective criterion for how many $PC_j(t)$'s to use here the first four principal components, which explain well over 50% of the total variance of the original data field, have been selected as target functions for the investigations on the area weighted scale, see Fig. 4.

The dominant EOF 1 in Fig. 4 holds 32% of the total variance and represents approximately the global mean temperature series shown in Fig. 2 whereas EOF 2 roughly refers to internal climate variability which is likely to be caused by ENSO (Staeger, 1998). Higher $EOFs$ can not be identified that easy with processes in the climate system.

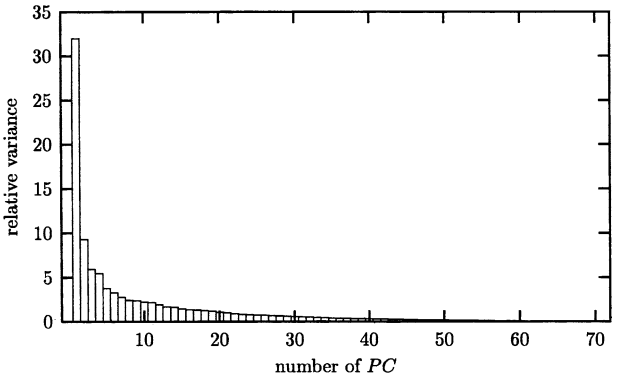

Fig. 4. Relative explained variance vs. PC number. The explained variance flattens after PC_4. The first four PC's explain well over 50% total variance of the original data field

6. Results and interpretation

Figure 5 shows the simulation and the corresponding anthropogenic signals obtained using the CM-FSA algorithm described in Section 3.1. The simulation as well as the plotted signals are the average signals of thirty model runs to reduce the (low) probability of falling into local minima of the cost function.

The simulation quality amounts to 81% explained variance (0.9 correlation). A similar Multiple Linear Regression Model (MLR) driven by the same forcing time-series ends up with an explained variance of 75% on the global scale

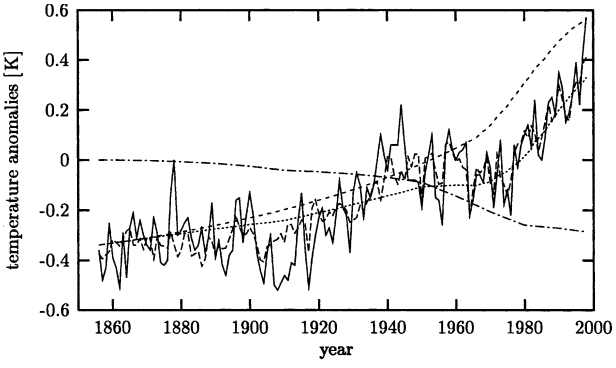

Fig. 5. Results of the global CM simulation (dashed) and the corresponding GHG-signal (short-dashed), SU-signal (dashed-dotted) and the combined anthropogenic signal (dotted). Also shown are observed global temperature anomalies (solid) provided by (Jones et al., 1994), (Jones, 1999a) respectively

(Walter, 2001), results not shown. An identical simulation using a BPN lead us to an explained variance of 84% (Walter and Schönwiese, 2002). Thus, if all parameters of the CM-FSA as well as the BPN have been chosen correctly, it can be concluded that at least for global surface temperature, there is only little nonlinearity effective. If the CM-FSA algorithm found the global minimum of the cost function it can furthermore be concluded that the BPN results mentioned suffered from a slight overfitting.

The signal amplitudes shown in Fig. 5 amount to 0.90 K for the GHG forcing, −0.28 K for the SU forcing and 0.66 K for the combined anthropogenic forcing (GHG and SU). The combined anthropogenic signal reflects the observed trend of 0.60 K (Jones, 1999a) for the analyzed time period rather well. Recent simulations by General Circulation Models (GCM), came up with similar results. Johns et al. (2003) used a coupled Atmosphere-Ocean GCM which included a representation of the anthropogenic sulfur cycle and both direct and indirect forcings from sulfate aerosols. For the historical period 1860 to present they obtain a GHG-signal of roughly 1.0 K which is very similar to our findings. Similar results were obtained by Roeckner et al. (1999) also using a coupled AOGCM. For the period 1860 to 2000 they obtained a GHG-signal of 0.9 K. In the case of an absent SU forcing their simulated temperature increase evolves too fast compared to the observational record. Beside this we find in accordance with the results of Roeckner et al. (1999) and Stott et al. (2000) a more pronounced warming due to anthropogenic GHG emissions over land than over ocean. Stott et al. (2000) give a range of about 0.2 K/decade warming for recent decades, which is close to our results. Furthermore the recent IPCC report (Houghton et al., 2001) gives similar signal amplitudes for the above forcing mechanisms.

We obtained the largest model-data discrepancy in the period 1900 to 1920, see Fig. 5. Two factors which could be responsible for this model-data difference in this period are (i) mid-latitude land clearance may have increased the albedo and caused slightly greater cooling than simulated (Bonan et al., 1992), and (ii) warming may be underestimated in the early stage of the instrumental record because of sparse data coverage (Jones et al., 1999).

A remarkable feature of the SU signal carried out is its time structure: a moderate cooling due to anthropogenic emissions of Sulfur dioxide until the 1940's, followed by a rather pronounced cooling effect for the time period 1940 to 1970 and a rather weakend cooling from there on. This time structure is consistent with federal environmental legislation for most industrialized nations.

Realistic simulations and signal amplitudes in a physical sense have been carried out for the northern and southern hemisphere as well (Walter, 2001), results not shown.

On the area related scale the CM-FSA simulated a maximal anthropogenic (GHG + SU) effect over northern America (1.13 K) and over Central Asia (1.17 K), (Walter, 2001), results not shown. These results are reasonable in a physical sense because of the enhanced heat capacity of these regions. In the northern Atlantic region a rather weak warming (0.11°K) is simulated by the CM-FSA. This effect is due to an enhanced downwelling of surface water (Smethie, 1993) and is shown by physical models as well, see again (Houghton et al., 2001).

To ensure that the CM has captured all relevant mechanisms, the residuals of the simulations have to be tested. This is usually done with testing for Gaussian distribution of the residuals e.g. Kolmogorov-Smirnov test (Press et al., 1992). In case of a nonlinear model the residuals undergo nonlinear transformations during the training of the network, therefore the residuals do not have to be Gaussian distributed. An alternative statistical method for testing this kind of residual is the autocorrelation function (Von Storch and Zwiers, 1999). Because the elements of a white noise process are independant it follows that their autocorrelation function $\rho(\tau)$ is

$$\rho(\tau) = \begin{cases} 1 : \tau = 0 \\ 0 : \tau \neq 0 \end{cases}, \tag{15}$$

where τ is the lag.

The autocorrelation function of the residual for the global scale is shown in Fig. 6, also shown is the 95% confidence interval. For $\tau = 1, 2$ there is a slight autocorrelation obvious. This is due to the inertia of the climate system. Within the climate system most temperature information is stored via the oceans SST, thus it is possible to model global near surface air temperature evolution as an autoregressive process. For lags greater

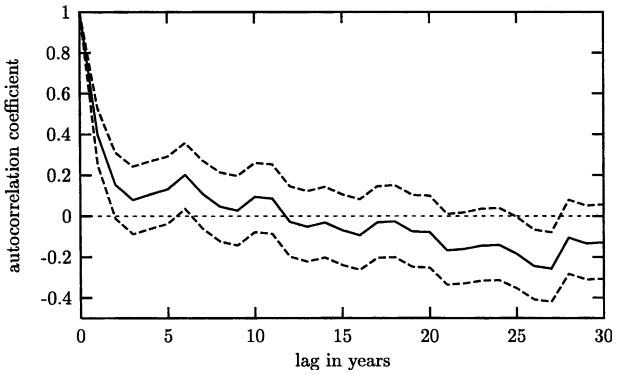

Fig. 6. Autocorrelation function (solid) with corresponding 95% confidence interval (dashed) of the residuals

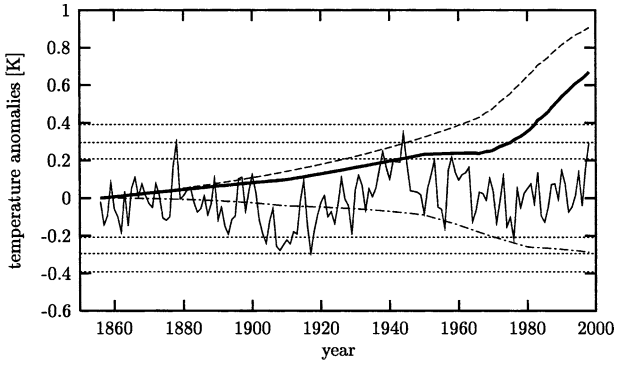

Fig. 7. Residual and obtained natural signal of the global CM-FSA simulation (solid), GHG-signal (dashed), SU-signal (dashed-dotted) and combined anthropogenic signal GHG + SU (thick solid). Also shown are the 95%, 99% and 99.9% significance levels (dotted)

than 2 the autocorrelation coefficient shows no significant deflection from zero. Thus it can be concluded that the CM-FSA captured all relevant time structures and that the remaining residual can be treated as noise. On the area weighted scale the autocorrelation function of the residuals show the same properties (Walter, 2001), results not shown.

To obtain a useful detection variable one has to look for a statistical relationship between the $PC_j(t)$'s introduced in Section 5 and the forcings of interest. The observed spatio-temporal variations of the near surface temperature can be written as

$$z(x,t) = S_{anthr} + S_{nat} + \varepsilon, \tag{16}$$

where $z(x,t)$ is the original data field, S_{anthr} is the combined (GHG + SU) anthropogenic signal, S_{nat} are the natural signal, i.e. effects due to volcanism, ENSO and solar variations, and ε is the unexplained residual. By testing the proportion between S_{anthr} against $S_{nat} + \varepsilon$ one obtains an assessment of the signal-to-noise ratio, and in consequence of the confidence level of the detection of an anthropogenic induced climate change.

In Fig. 7 the results for such a detection approach on the global scale are shown. Here the solid line represents the residual of the global CM-FSA simulation plus the natural signals (solar, volcanism and ENSO) obtained. This combination of unexplained variability and natural climate variations is what we call *climate noise*, see Section 4. The dotted lines represent the 95%, 99% and 99.9% significance levels, which can be computed using Eq. (10). The obtained anthropogenic signals (GHG-signal: dashed, SU-signal: dashed-dotted and combined anthropogenic GHG + SU: thick solid) do not agree with the assumption of a undisturbed climate system. The GHG-signal as well as the combined anthropogenic signal are detectable with a probability of $>99.9\%$. The GHG-signal exceeds this significance level in 1961. Because of the cooling effect of SO_2-emission the combined anthropogenic signal exceeds this level not until 1983. The SU-signal comes short of exceeding the 99% significance level and thus can only be detected with a probability of $>95\%$. From these results it can be concluded, that it is *virtually certain* that an increase in global near surface air temperature because of the anthropogenic emissions of GHGs has already happened. Furthermore it is *very likely* that in the analyzed time period SO_2 emissions contributed a substantial cooling effect to the global near surface air temperature evolution. Thus, it is again *virtually certain*, that for the time period 1856–1998 at least the low frequent trend in the observations of global near surface air temperature is caused by anthropogenic emissions of GHGs and SO_2.

Figure 8 shows the results from the detection approach introduced in Section 4 for a probability $p>90\%$ that an anthropogenic (GHG & SU) climate change has happened. Because of the fact that not only the signal but the signal-to-noise ratio determines the detection of the anthropogenic signal the detection strategy introduced above succeeds at a high confidence level where the overall standard deviation of climate noise, i.e. natural signals plus residual, is rather small, see Fig. 9. Because of a reduced year-to-year variability this is mainly the case for the tropics

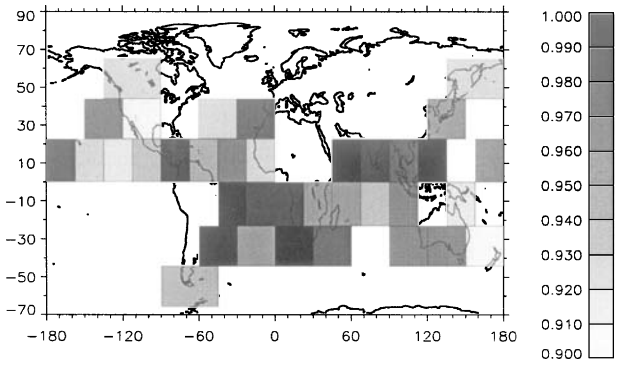

Fig. 8. Probability $P \geq 90\%$ for an anthropogenic induced climate change

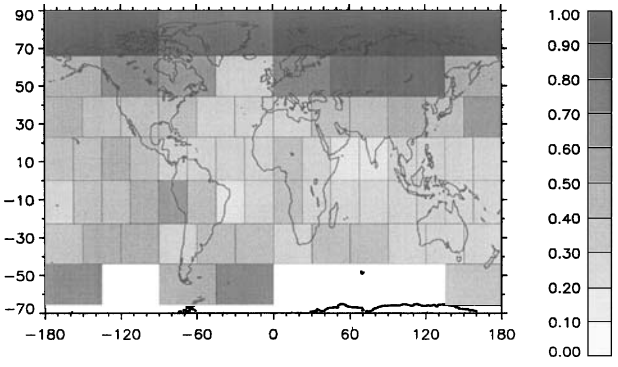

Fig. 9. Standard deviation of the residuals on the area-related scale

and oceanic regions due to their large heat capacity and missing orographic effects on the atmospheric circulation.

However, the detection succeeds on a high confidence level for the mid-latitudes of the northern and southern hemisphere as well; this time because of enhanced signal amplitudes compared to those in the tropical regions. A linear MLR did not come up with similar results (Walter, 2001), results not shown, which is mainly due to the enhanced simulation quality of the CM and thus, a reduced standard deviation of the remaining climate noise. Thus, regions with a low standard deviation correspond to regions with a high probability of an anthropogenic climate change. In 36 out of 72 considered areas this probability exceeds 90%.

It is worth mentioning that a similar approach using the standard BPN yielded better results in terms of the explained variance (Walter and Schönwiese, 2002), but as mentioned above this effect may be due to overfitting which the CM-FSA tries to prevent.

7. Conclusions

In this work the possibilities of NNM, especially the CM-FSA, for the purpose of attributing and detecting anthropogenic climate change have been studied. An NNM simulation represents nothing less than a nonlinear optimal fit. Therefore the results are highly sensitive to the choice of internal free parameters of the network architecture, i.e. learning constant ε in Eq. (8) and cooling schedule Eq. (9) respectively. However, if one carefully selects these parameters, the CM-FSA algorithm provides a strong nonlinear statistical tool for climatological data analysis with a minimized risk of over-/underfitting.

The results we obtained show a significant anthropogenic climate change for most regions of the globe, see Fig. 5, Fig. 7 and Fig. 8 respectively. On the global scale the best estimate of the effect of anthropogenic GHG and SO_2 emissions into the atmosphere amounts to 0.66 K warming for the time period 1854 to 1998, which is close to the observed trend (Jones et al., 1994), (Jones, 1999a).

Thus, in accordance with GCM results, e.g. (Roeckner et al., 1999), (Stott et al., 2000), we have to conclude that a significant anthropogenic climate change is *virtually certain* and already visible in the global near surface air temperature record.

References

Ackley DH, Hinton GE, Seijnowski TJ (1986) A learning algorithm for Boltzman machines. In: Anderson JA, Rosenfeld E (eds) Neurocomputing. Foundations of research. Cambridge, MA: MIT Press, pp 638–650

Adrian ED (1926) The impulses produced by sensory nerve endings: part I. J Physiol 61: 49–72

Anderson JA, Rosenfeld E (eds) (1986) Neurocomputing. Foundations of research. Cambridge, MA: MIT Press, 729 p

Bonan GB, Pollard D, Thompson SL (1992) Effects of boreal forest vegetation on global climate. Nature 359: 716–718

Charlson RJ, Schwartz SE, Hales JM, Cess RD, Coakley JA Jr, Hansen JE, Hoffman DJ (1992) Climate forcing by anthropogenic aerosols. Science 255: 423–430

Freeman JA, Skapura DM (1992) Neural networks. Algorithms, applications and programming techniques. Reading, MA: Addison-Wesley, 401 p

Geman S, Geman D (1986) Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. In: Anderson JA, Rosenfeld E (eds) Neurocomputing. Foundations of research. Cambridge, MA: MIT Press, pp 614–634

Grieger B, Latif M (1993) Reconstruction of the El Niño attractor with neural networks. Report des Max-Planck-Institut für Meteorologie Nr. 112, Max-Planck-Institut für Meteorologie, Hamburg, 1993

Grieser J, Schönwiese C-D (1998) Parametrization of spatio-temporal patterns of volcanic aerosol induced stratospheric optical depth and its climate radiative forcing. Atmósphera 12: 111–133

Grossberg S (1982) How does a brain build a cognitive code? In: Grossberg S (ed) Studies of mind and brain. Boston: D Reidel Publishing, pp 1–52

Hansen J, Lebedeff S (1987) Global trends of measured surface air temperature. J Geophys Res 92: 13345–13372

Hansen J, Sato M, Lacis A, Ruedy R, Tegen I, Matthews E (1998) Climate forcings in the industrial era. Proc Natl Acad Sci 95: 12753–12758

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (2001) Climate change 2001. The scientific basis. Cambridge: Cambridge University Press, 881 p

Hsieh WW, Tang B (1998) Applying neural network models to prediction and data analysis in meteorology and oceanography. Bull Amer Meteor Soc 79(9): 1855–1870

Hsu K-L, Gao X, Sorooshian S, Gupta HV (1997) Precipitation estimation from remotely sensed information using artificial neural networks. J Appl Meteorol 36: 1176–1190

Johns TC, Gregory JM, Ingram WJ, Johnson CE, Jones A, Lowe JA, Mitchell JFB, Roberts DL, Sexton DMH, Stevenson DS, Tett SFB, Woodage MJ (2003) Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emission scenarios. Climate Dynamics 20: 583–612

Jones PD (1999a) http://www.cru.uea.ac.uk:80/cru/data/temperat.htm

Jones PD (1999b) http://www.cru.uea.ac.uk/cru/data/soi.htm

Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its change over the past 150 years. Rev Geophys 37: 173–199

Jones PD, Wigley TML, Briffa KR (1994) Global and hemispheric temperature anomalies – land and marine instrumental records. In: Boden TA, Kaiser DP, Sepanski RJ, Stoss FW (eds) Trends'93: A compendium of data on global climatic change. ORNL/CDIAC-65. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, pp 603–608

Lean J, Beer J, Bradley R (1995) Reconstruction of solar irradiance since 1610: Implications for climate change. Geophys Res Lett 22: 3195–3198

Lean J, Rind D (1999) Climate forcing by changing solar radiation. J Climate 11: 3069–3094

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21: 1087–1092

Preisendorfer R (1988) Principal component analysis in meteorology and oceanography. Developments in Atmospheric Science, 17. Amsterdam: Elsevier, 425 p

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipies. Cambridge: Cambridge University Press, 934 p

Ramanathan V, Cicerone R, Singh H, Kiehl J (1985) Trace gas trends and their potential role in climate change. J Geophys Res 90: 5547–5566

Roeckner E, Bengtsson L, Feichter J, Lelieveld J, Rodhe H (1999) Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle. J Climate 12: 3004–3032

Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Review 65: 386–408

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Parallel distributed processing: explorations in the microstructure of cognition. Cambridge, MA: MIT Press, pp 318–362

Rumelhart DE, Mc Clelland JL (1986) Parallel distributed processing: explorations in the microstructure of cognition. Cambridge, MA: MIT Press, 2 Vols., 1158 p

Schönwiese C-D, Denhard M, Grieser J, Walter A (1997) Assessments of the global anthropogenic greenhouse and sulfate signal using different types of simplified climate models. Theor Appl Climatol 57: 119–124

Smethie WM (1993) Tracing the thermohaline circulation in the western North Atlantic using chlor–flourmethanes. Prog Oceanogr 31: 51–99

Staeger T (1998) Statistische Analyse des ENSO- und Vulkanismus-Signals in Klima-Zeitreihen. Master's thesis, Johann Wolfgang Goethe-Universität, Frankfurt, 84 p

Stevens MJ, North GR (1996) Detection of climate response to the solar cycle. J Atmos Sci 53: 2594–2608

Stone M (1974) Cross-validatory choice and assessments of statistical predictors. J Roy Stat Soc B36: 111–147

Stott PA, Tett SFB, Jones GS, Allen MR, Mitchell JFB, Jenkins GJ (2000) External control of 20th century temperature by natural and anthropogenic forcings. Science 290: 2133–2137

Sun D-Z, Trenberth KE (1998) Coordinated heat removal from the tropical Pacific during 1986–87 El Niño. Geophys Res Lett 25: 2659–2662

Szu H (1986) Fast simulated annealing. In: Denker JS (ed) Neural networks for computing. New York, NY: American Institute of Physics, pp 420–425

Trenberth KE, Branstator GW, Karoly D, Kumar A, Lau N-C, Ropelewski C (1998) Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. J Geophys Res 103: 14291–14324

Von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge: Cambridge University Press, 484 p

Walter A (2001) Zur Anwendung neuronaler Netze in der Klimatologie. PhD thesis, Johann Wolfgang Goethe-Universität, Frankfurt. Berichte des Deutschen Wetterdienstes No. 218, 168 p

Walter A, Denhard M, Schönwiese C-D (1998) Simulation of global and hemispheric temperature variations and signal detection studies using neural networks. Meteorol Z NF 7: 171–180

Walter A, Schönwiese C-D (2002) Attributution and detection of anthropogenic climate change using a back-propagation neural network. Meteorol Z 11(5): 335–343

Authors' address: Andreas Walter (e-mail: A.Walter@meteor.uni-frankfurt.de) and Christian-D. Schönwiese (e-mail: Schoenwiese@meteor.uni-frankfurt.de), J.W. Goethe-Universität, Institute for Meteorology and Geophysics, P.O. Box 11 19 32, D-60054 Frankfurt a.M., Germany.