Simulation of global temperature variations and signal detection studies using neural networks

A.Walter, M.Denhard and C.-D. Schönwiese

Institute for Meteorology and Geophysics, J.W. Goethe Universitiy, Frankfurt a.M., Germany

Quelle: Meteorologische Zeitschrift, N.F.7,pp. 171-180, August 1998

Abstract

The concept of neural network models (NNM) is a statistical strategy which can be used if a superposition of any forcing mechanisms leads to any effects and if a sufficient related observational data base is available. In comparison to multiple regression analysis (MRA), the main advantages are that NNM is an appropriate tool also in the case of non-linear cause-effect relations and that interactions of the forcing mechanisms are allowed. In comparison to more sophisticated methods like general circulation models (GCM), the main advantage is that details of the physical background like feedbacks can be unknown. Neural networks learn from observations which reflect feedbacks implicitly. The disadvantage, of course, is that the physical background is neglected. In addition, the results prove to be sensitively dependent from the network architecture like the number of hidden neurons or the initialisation of learning parameters.

We used a supervised backpropagation network (BPN) with three neuron layers, an unsupervised Kohonen network (KHN) and a combination of both called counterpropagation network (CPN). These concepts are tested in respect to their ability to simulate the observed global as well as hemispheric mean surface air temperature annual variations 1874 - 1993 if parameter time series of the following forcing mechanisms are incorporated : equivalent CO2 concentrations, tropospheric sulfate aerosol concentrations (both anthropogenic), volcanism, solar activity, and ENSO (all natural). It arises that in this way up to 83% of the observed temperature variance can be explained, significantly more than by MRA. The implication of the North Atlantic Oscillation does not improve these results. On a global average, the greenhouse gas (GHG) signal so far is assessed to be 0.9 - 1.3 K (warming), the sulfate signal 0.2 - 0.4 K (cooling), results which are in close similarity to the GCM findings published in the recent IPCC Report. The related signals of the natural forcing mechanisms considered cover amplitudes of 0.1 - 0.3 K. Our best NNM estimate of the GHG doubling signal amounts to 2.1K, equilibrium, or 1.7 K, transient, respectively.

1. Introduction

The recent IPCC report (Houghton et al. 1996) states that the global mean surface air temperature has increased by approximately 0.3 - 0.6 K since the late 19th century and that there are small differences between those reconstructions using different data bases. Based on the data set 1854 - 1995 Jones (1994, 1996) we find a linear trend of +0.6 K with a trend-to-noise ratio of 2.7 where noise is represented by the annual data standard deviation. The non-parametric Mann-Kendall test reveals a trend confidence of > 99% (error probability < 0.01). The main hemispheric peculiarities are that the amplitudes of both the year-to-year and the relatively low-frequent variations are much more pronounced in the northern than in the southern hemisphere including a considerably northern hemisphere cooling between the fourties and seventies of the recent century.

The global or hemispheric long-term warming is often seen as a consequence of the enhanced greenhouse effect, i.e. radiative (IR) forcing due to the atmospheric concentration increase of CO2, CH4, CFC, N2O, O3 and others (greenhouse gases, GHG). Another large-scale but more regional human influence on climate is the formation of sulfate (SO4- -) aerosol particles (SUA) due to SO2 emissions which leads to a cooling of the atmospheric surface layer. According to IPCC the radiative forcing so far (industrial time, roughly since 1850) amounts to + (2.1 - 2.8) Wm-2 in case of GHG and (0.3 - 0.9) Wm-2 in case of SUA. In this context IPCC sees a major progress in the fact that it was possible to simulate the combined GHG - SUA effect by means of coupled ocean - atmosphere general circulation models (AOGCM) leading to the result that the GHG effect in the global mean surface air temperature so far should have a magnitude of approximately +1K (Mitchell et al. 1995, Hasselmann et al. 1995). Moreover, the ?fingerprint" technique (comparison of regional patterns of model and observational data) gave rise to the IPCC statement (again Houghton et al. 1996) that ? the balance of evidence suggests that there is a discernible human influence on climate". Santer et al. (1996) who have implied also the climatic effects due to stratospheric O3 depletion and extended their ?fingerprint" studies to the three-dimensional thermal structure of the atmosphere state again a statistically similarity of observational with model data but say also that they have not quantified the relative magnitude of natural and anthropogenic climate effects. Similarly, also IPCC sees a need for such multiple (multi-forced) signal detection studies including the effects of volcanism and solar activity.

It is only this multi-forcing problem which, from an observational point of view, is addressed in the present paper. Our study is focussed on the global or northern or southern hemisphere, respectively, mean surface air temperature annual data reconstructions 1874-1993 (Jones 1994,1996). The forcing mechanisms considered and represented by related annual parameter time series are GHG (CO2 equivalents), SUA, volcanism, solar activity, and ENSO (El Niño - southern oscillation). However, due to negligible contributions of NAO to the explained temperature variance, this forcing was neglected in the detailed computations. As an improved technique alternative to multiple regression analysis (MRA; Schönwiese 1983, 1987, 1991; Bayer and Schönwiese, 1995) we used neural network models (NNM) for the simulations and signal detection studies. These statistically-derived signals are compared with radiative forcing and the results from OAGCM projections.

In the following, we describe the strategy and the particular modifications of NNM as used in this study (section 2), outline briefly our data base (section 3), present the NNM best simulation and signal assessments (section 4), including intercomparissons with the results from other modeling (especially AOGCM), discuss some NNM problems and caveats (section 5) and draw some conclusions.

2. Neural network strategy

The fundamental ideas of neurocomputing were first published in the fourties (Mc Culloch and Pitts 1943, Hebb 1949). More recent textbooks are available from e.g. Smith (1993) or Freeman and Skapura (1991). The fundamental concept defines a model neuron (a unit) to which an input vector x = xj is presented. The model neuron allocates to each input vector component xj a weight wi. If the relation

$$I = \text{Sum } \mathbf{wj} \cdot \mathbf{xj} > w0 \qquad (1)$$

where I is called neuron impulse and w0 is a defined threshold, is fulfilled the neuron is activated and produces a response y* (output). In analogy to biological reactions y* is defined as a sigmoid function of I, often tangens hyperbolic (tanh) with a slow reaction phase in case of small I, near-linear quick reaction phase for I near average and a saturation phase for large I.

Now, this unit has to be trained to produce a correct response. In case of controlled training this correct response, say y0 , is known so that it can be compared to the actual neuron response y*. Then, in a sequence of testing, according e.g. to Widrow and Hoff (1960), it can be tried to minimize (y0 - y*)2. As soon as this error is small and approximately stable for further testing, the neuron has learned to answer correctly or quasi-correctly and the training can be finished.

The simplest architecture of a neural network model (NNM) with supervised training and error minimization is a three layer backpropagation network (BPN), see Fig. 1, where an input layer of neurons receives the forcing vector. In weighted form as described above this information is transmitted to a so-called hidden layer of neurons whereas the output layer (one neuron in case of one variable considered, otherwise again several neurons) produces the response y* and, at the same time, receives the control value y0 for error minimization. In its application to climatology the vector x = xj is transformed into a matrice {xji} where xj represent different forcing mechanisms like GHG, SUA, volcanism and so on at each time step i (time series x1(i), x2(i), etc.); y = yi may be a temperature time series as response to the forcing {xji}. At each time step NNM, in this case BPN, is trained and corrected by (y*i - y0i)2 minimization.

It is reasonable to add a verification period to an initial training period. Thereby, in case of climate trends, it is problematical to use, for example, the first half of an observational period for training and the second half for verification because in the verification period y0 values may occur which were not existent in the training period. Therefore, we used a random selection of training and verification data, 75% or 25% respectively, of the total data set, distributed irregularly over the whole observation period. The choice of the number of processing units is performed by means of optimization criteria.

The number of processing neurons in the hidden layer has to be optimized to avoid that in addition to the desired data features noise is modeled as well, i. e. to limit the networks power. This so-called ?overfitting" problem can be prevented by limiting the number of processing units in the hidden layer. Therefore, after training is finished, the networks performance has to be tested using the validation data set. The error-function of the network output in dependence on the number of processing neurons decreases asymptotically to a minimum error. From this point the error-function will rise steeply, so that the network architecture which implies the smallest value of the error-function represents the optimal architecture and only this network architecture will be used for the simulations.

Another internal parameter which has to be optimized is the so-called learning constant g. This constant connects the weights wj at time step i-1 to the weights wj at time step i via the equation

$$\mathbf{wj}(i) = \mathbf{wj}(i-1)+\mathbf{x}(i)\ \text{gamma}(y0 - y*)**2 \qquad (2)$$

where **x** represents the input vector. Equation (2) is the so-called ?Delta Rule".

Therefore the optimal weights are highly sensitive to the choice of g. To achieve the best value for the learning constant, if not using so-called ?adaptive learning rates" (Smith, 1993), g has to be iterated through several network runs. The value of g can than be determined by comparison of the several achieved error-functions.

Kohonen (1982) proposed a self-organizing unsupervised neural network model called Kohonen network (KHN). In this case, the input (forcing) information is received by a two-dimensional field of neurons (Kohonen layer) which, unlike to the BPN case, all are able to interact. As a consequence of the input, one of the neurons implies a maximum impulse and neurons in a defined neighborhood are also stimulated but of minor impulses. At each training step i the weights and, in consequence, the impulses are changed until a particular structure of neuron impulses is established. It is possible to obtain something like ?neuron impulse charts" which show, e.g. for training and verification, which neurons in the two- dimensional Kohonenlayer are stimulated. Like in the BPN case, also the KHN strategy allows to choose the optimal number of (hidden) neurons for a given problem by means of optimization criteria.

However, as an alternative technique to BPN, we did not use KHN but the counterpropagation network (CPN) which is a combination of BPN and KHN (Hecht-Nielsen 1988). That means, the input layer is first connected with a Kohonen layer (like in Fig.2) which

transfers its output to a modified output layer similar to BPN (see Fig.1) called Grossberg layer (Grossberg 1980). Finally, similar to BPN, an error minimization is performed as described above. Again a choice of the number of processing neurons is possible and ? neuron impulse charts" are available.

In comparison to multiple regression analysis (MRA), outstanding advantages of neural network models (NNM) are that free non-linear cause-effect-relations as well as free interactions of the forcing mechanisms are allowed. (MRA is either linear or defined non-linear relations have to be prescribed; the forcings have to be independent and time series autocorrelations hampers MRA considerably.) In contrast to all deterministic climate models, i.e. variety from energy balance (EBM) to coupled ocean-atmosphere general circulation model (OAGCM), the governing physical equations, including all feedbacks, can be unknown. NNM need only observational information which reflects all feedbacks implicitly and learn by training. Disadvantages are, of course, the neglected physical background, the fact that the non-linear cause-effect relations learned by training are not explicitly available and a problematical sensitivity to the number of hidden neurons, of the values of initial training parameters and of the data period (caveats discussed in section 5). Moreover, signal analysis neglects the advantages of possible forcing interactions because, so far, we found that only BPN with only one forcing can be handled for this purpose. So, it is necessary to compare particularly the signal assessments, as far as possible, to these from OAGCM experiments. Nevertheless, in a first step, it has to be tested how well, based on a particular cause-effect observational data base, the NNM concept works to reproduce climate variation time series. The typical number of hidden neurons we used for this purpose was roughly 5-10 in case of BPN and 50-100 in case of CPN.

3. Data base

As already mentioned in the Introduction, concerning observed climate variation effects, we used 1874-1993 annual data time series of mean global or mean hemispheric, respectively, surface air temperature provided by Jones (1994, 1996), see solid lines in Figures 3-5. It is broadly accepted that this is a high-quality data set where urban heat island error contribution to the secular trend is < 0.05 K (Jones et al. 1990).

Anthropogenic GHG forcing is represented by CO2 equivalents also used by IPCC (Houghton et al. 1996) starting with a value of 293 ppm in 1874 and ending with 415 ppm in 1994 (only CO2, Mauna Loa value 1994: 358 ppm). Because anthropogenic sulfate aerosol particles (SO4- -) have an atmospheric life time of only a few days this forcing, on an annual time scale, may be directly proportional to the related SO2 emissions. Our related data base is again like reported by IPCC where it is important that the mean northern hemisphere emissions show first a relatively weak increase since the last century, after Second World War a sharp increase, in the seventies and eighties a relative decrease (due to air pollution reduction measures) and recently again an increase (due to enormous emissions in South East Asia). In addition to a first estimate, approximately digitized from IPCC publications, we used also a gridded SO2 emission data set from Charlson et al. (1992), same as in the OAGCM simulations by Hasselmann et al. (1995). For the recent years we refer to IPCC scenarios.

Explosive volcanism whose ejecta reach the stratosphere and form (gas-to-particle conversion) the climate relevant sulfate aerosol particles has been considered by many authors. On a short time scale (e.g. day-to-day) volcanic radiative forcing may be as large as 4.5 Wm-2 (Pinatubo, 1991; IPCC, Houghton et al. 1996) but related to annual data a magnitude of not more than roughly 1 Wm-2 may be supposed. In our context the long-term global or hemispheric time history of explosive volcanism is important where the US Smithsonian Institution chronology (Simkin et al. 1981, 1993) is a major data source. Following some earlier assessments (e.g. Schönwiese, 1996) we have evaluated an annual volcanic activity index time series 1700 - 1994 similar to Sato et al. (1993) or Robock and Free (1995) which, however, takes into account the formation and deposition processes of stratospheric sulfate (Grieser, 1996).

The role of solar phenomena and their influence on climate is a matter of controversial discussion. From a statistical point of view we (Schönwiese et al. 1994) have considered several hypotheses and found that the best confidence is implied with sunspot-related time series, especially concerning the assessments by Foukal and Lean (1990), which are used in this paper. An alternative assessment by Hoyt and Schatten (1993) leads to assumptions of solar irradiance forcing which seems to be significantly too large when compared to the IPCC statements: forcing 0.1 - 0,5 Wm-2. Finally, the ENSO mechanism is represented by the southern oscillation index SOI as suggested by Wright (1984, updates available from WMO). For a comparison of all forcing parameter time series see Fig. 6.

4. NNM simulations and signal results

First, it has to be tested to which extent the neural network is able to reproduce the observed global or hemispheric mean annual surface temperature variations 1874 - 1993 if the neural network model (NNM) is driven by greenhouse gases (GHG), tropospheric sulfate aerosols (SUA), volcanism, solar activity and ENSO (see section 3 for data sources). As outlined in section 2, we used the backpropagation (BPN) and counterporopagation (CPN) network, random selection of training (75%) and verification (25%) data, and optimization techniques for the appointment of the numbers of hidden neurons and values of training parameters.

Figures 3 - 5 show that this concept works well explaining 62% - 83% of variance (corresponding to .79 - .91 multiple correlation), maximum for global data and CPN, including a satisfactory reproduction of amplitudes (standard deviations) and trends, see Table 1. For comparison, a related multiple regression analysis (MRA) leads to 59% - 73% explained variance (plots not shown) so that it can be concluded that the relationship under consideration imply some non-linearities, however, not in dominating size. A residuum analysis (results not shown) confirms that no significant further forcing mechanisms are neglected because the model residuum data are Gaussian distributed and the autocorrelation functions differ not confidently from nought. A remarkable detail is that the northern

hemisphere observations which show both a larger year-to -year variability and more pronounced relatively low-frequent variations (including a cooling roughly 1940 - 1970) when compared to the southern hemisphere is better reflected by the BPN type of the neural networks.

This may be of some relevance concerning the signal analysis where we found that only the BPN is approximate for. For the assessments of these temperature signals which may be due to definitive isolated forcing mechanisms BPN was driven by all forcing parameter time series separately; in addition, we performed a combined GHG + SUA forcing. The results are summarized in Table 2 whereas the Figures 7 - 9 specify only the antropogenic signals, however, in its time series histories (Table 2 specifies only the maximum amplitudes of the signals).

In addition to the anthropogenic signal time series GHG, SUA, and GHG+SUA, Figures 7 - 9 show also the 10 yr Gaussian low pass filtered data as observed and modeled. It is not surprising that NNM fitting in this low-frequent domain works even better than in case of unsmoothed annual data. Fig. 6 , considering global mean temperature, reveals a GHG signal of approximately +1K, a SUA signal of -0.2K and a combined GHG+SUA signal of approximately +0.6K (due to the nonlinear model structure not simply the sum of the GHG and SUA signals). It is a very remarkable fact that these results concerning the GHG and the GHG+SUA signals are very similar to the results from OAGCM simulations from Mitchell et al. (1995) as well as Hasselmann et al. (1995) so that the present study verifies these OAGCM results on a statistical-observational basis. However, the NNM simulations can also reproduce the year - to - year variations by adding further forcing mechanisms. The SUA signal, as shown in Fig. 7, implies an interesting structure: weak cooling until the end of Second World War, significantly cooling subsequently until the seventies and again weak cooling afterwards. So, the SUA signal reflects reasonably the SO2 forcing mentioned in section 3.

Again reasonable are the interhemispheric differences of the NNM signal assessments. So, the SUA forced cooling signal is much more pronounced in the northern (Fig. 8) than in the southern (Fig. 9) hemisphere. Therefore, the observed northern hemisphere cooling between roughly 1940 - 1970 , absent in the southern hemisphere, seems to be an anthropogenic sulfate effect. Furthermore, our results suggest that the large - scale (global and hemispheric mean) low - frequent temperature variations 1874 - 1993, including the GHG forced non-linear trend, may be anthropogenic whereas the relatively high-frequent variations seem to be natural. This implies a hypothetical separation and quantification of anthropogenic and natural forcing in observed temperature variations. Note that also the (natural) signals due to volcanism, solar activity and ENSO as revealed by NNM analysis seem to have a reasonable magnitude when compared with IPCC (Houghton et al. 1996) statements, particularly radiative forcing, see again Tab. 2.

On conditions that scenarios of future atmospheric GHG concentrations prove to be reasonable and the physical as well as statistical cause-effect relations can be transferred into the future, predictions of the anthropogenic enhanced greenhouse effect are possible. To avoid thereby a strong dependence on the scenarios one could ask, how the global mean surface air temperature would react in the case of a atmospheric CO2 equivalent doubling compared to preindustrial values (circa 1800/1860). These predictions can than be obtained for equilibrium or, if taking the atmospheric delay time into account, transient. IPCC (Houghton et al., 1996) gives under the restrictions of AOGCM modeling values of 2.1 - 4.6 °C (equilibrium) and 1.3 - 3.8 °C (transient) receptively.

When applying neural networks for predictions there occur the problems that these models learn strictly out of past values and that future atmospheric CO2 equivalent concentration values leave the range of the training period. Therefore there is principally no prediction possible when applying neural networks to such kind of problems. One attempt to avoid this problem is to fit and extrapolate a function which corresponds broadly with the present signal time series. In Fig. 10 we used therefor a quadratic function which gives values of 2.1 °C (equilibrium) and 1.7 °C (transient). When comparing these values to the values given by IPCC (Houghton et al., 1996) it is obvious that they suggest a relatively low heating, but they are still in the range of uncertainties given.

Figures 11 and 12 shown the autocorrelation function of the unexplained variances of the CPN and BPN simulations respectively. These unexplained parts can be either represent external forcings which have not been taken into account during the simulations or they can be interpretated as noise. In the case of a non -linear model, such as the NNM concept, these differences undergo as well the non linear transformation, therefore they don?t have to be Gaussian distributed. For the Analysis of these residuums the autocorrelation function is the appropriate tool. For the CPN simulations (Fig. 11) as well as for the BPN simulations (Fig. 12) the autocorrelation function of the residuums show no significant deviation from nought, except for the CPN simulations on the northern and southern hemisphere respectively. Here at a time lag of about 28 years cyclic variances occur, although these cyclic variances are not evident in the global CPN simulation. An explanation of this phenomenon can here not be given.

5. Some NNM caveats

When applying neural networks to any kind of time series one has to be aware of numerous pitfalls. These potential reasons for dissatisfying NNM results include:

· The choice of the number of processing units in the hidden layer (see section 2).

· The value of the learning constant g (see equation 2).

· The division of the data sample into a training and a verification period.

· Problems occurring when applying the NNM to different lengths of the same time series.

Techniques to avoid problems 1. and 2. are briefly discussed in section 2.

The problems occurring with the division of the data into a training and a validation sample apply to the high sensibility of the NNM when changing the length of the training as well as the validation sample, e.g. modifying the training sample from 30 to 50 data points. It is not the length of the training sample itself which determines the quality of the results, but the appearance of ?learned" features in the validation phase. Therefore one has to assure that features similar to those learned occur during the validation period as well. This problem has to be addressed by a ?try and error" kind of method. We found that a division of 75% of the data for training purposes and 25% for validation guarantees best results, i.e. the lowest values of the error function.

The problems in context to different lenghts of the same time series are the most sophisticated one to deal with. During the optimization phase of our simulations we received results where even the sign of the signals changed if the simulation ran one year less. Mainly the SUA signal on the northern scale seemed to be highly sensible to the length of the simulation. This problem occurs, because NNM neglects the physical background of any given problem. A NNM simulation represents nothing less than a non-linear optimal fit, and these simulations are highly sensitive to any internal parameter.

To avoid this kind of dissatisfying results, especially when applying NNM for signal analysis, a separate model for the low-frequency fraction in the time series has to be integrated into the NNM. If such a model is established in addition to the vanilla NNM, the network will give stable results in its application to signal analysis. The development of such a model for the low frequencies fraction in the data will be our next step following the promising NNM concept.

References

[1] Bayer, D.; Schönwiese, C.-D.,1985: Some statistical aspects of anthropogenic and natural forced global temperature change. Atmósfera 8 , 3-22.

[2] Charlson, R.J. et al.,1992: Climate Forcing by Anthropogenic Aerosols. Science 255, pp. 423-430.

[3] Foukal, P.V.;Lean, J.,1990: An Empirical Model of Total Solar Irradiance Variation between 1874 and 1988. Science, 247, pp. 556-558.

[4] Freeman, J.A.; Skapura, D.M.,1991: Neural Networks. Algorithms, applications and programming techniques. Addison-Wesley, Reading, Ma.

[5] Grieser, J.,1996: Eine makroskopische Klimadynamik und ihre diagnostische sowie prognostische Anwendung auf globale Temperaturvariationen. Phd Thesis, Univ. Frankfurt/Main, Germany.

[6] Grossberg, S.,1980: How Does a Brain build a Cognitive Code? Psychological Review 87.

[7] Hasselmann, K. et al.,1995:Detection of Anthropogenic Climate Change using a Fingerprint method. Report No. 168, Max-Planck-Institut für Meteorologie, Hamburg.

[8] Hebb, D.O.,1949: The Organization of Behaviour. New York.

[9] Hecht-Nielsen, R.,1988: Applications of Counterpropagation Networks. Neural Networks, Vol.1.

[10] Houghton, J.T. et al.,1996: Climate Change. The IPCC Scientific Assessment (World Meteorological Organization / United Nations Enviroment Programme / Intergovernmental Panel on Climate Change) Cambridge 1990; Climate Change 1995. The Science of Climate Change (IPCC Second Assessment Report). Cambridge.

[11] Hoyt, D.V., Schatten, K.H.,1993: A discussion of plausible solar irradiance variations, 1700 - 1992. JGR, 98, 18, 895- 18, 906.

[12] Jones, P.D.,Wigley, T.M.L., Briffa, K.R.,1994: Global and hemispheric temperature anomalies - Land and marine instrumental records. pp. 603 - 608. In T.A. Boden, D.P. Kaiser, R.J. Sepanski and F.W. Stoss (eds.), Trends ?93: A compendium of data on global change. ORNL/CDIAC-65. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., U.S.A..

[13] Jones, P.D.,1996: priv. commm..

[14] Kohonen, T.,1982: Self-organized Formation of Topologically Correct Feature Maps. Biol. Cybernetics 43.

[15] Mc Culloch, W.S.;Pitts, W.,1943: A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophysics 5, 115-133.

[16] Mitchell, J.F.B. et al.,1995: Climate Response to Increasing levels of Greenhouse Gases and Sulphate Aerosols. Nature 376, pp. 501-504.

[17] Robock, A. and Free, M.P.,1993: Ice cores as an index of global volcanism from 1850 to the present. JGR, 100, D6, 11549 - 11567.

[18] Sato, M. Hansen, J.E., Mc Cormic, M.P. and Pollack, J.B.,1993: Stratospheric Aerosol Optical Depths, 1850 - 1990, JGR, 98, D12, 22987 - 22994.

[19] Schönwiese, C.-D.,1983: Northern hemisphere temperature statistics and forcing. Part A: 1881 - 1980 AD. Arch. Met. Geoph. Biocl., Ser. B, 32, 337 - 360.

[20] Schönwiese, C.-D.,1987: Observationla Assessments of hemispheric and Global Climate Response to increasing Greenhouse Gases. Contributions to Atmospheric Physics, 60, 1, 48 - 64.

[21] Schönwiese, C.-D. and Stähler, U.,1991: Multiforced statistical assessment of greenhouse gas-induced surface air temperature chnage 1890 - 1985,. Clim. Dyn.,6, 23 - 33.

[22] Schönwiese, C.-D. et al.,1994: Solar Signals in Global Climate Change. Clim.Change, 27, pp. 259-281.

[23] Siebert, L.,1993: priv. comm..

[24] Simkin, T. Siebert, L., Mc Clelland, L., Bridge, D., Newhall, C. and Latter, J.H., 1983: Volcanoes of the world. Hutchinson Ross Publishing Company,. Stroundsburg, Pennsylvania.

[25] Smith, M.,1993: Neural Networks for Statistical Modeling. New York.

[26] Widrow, M.; Hoff, M.E.,1960: Adaptive Switching Circuits. IRE WESCON Convention Record, New York.

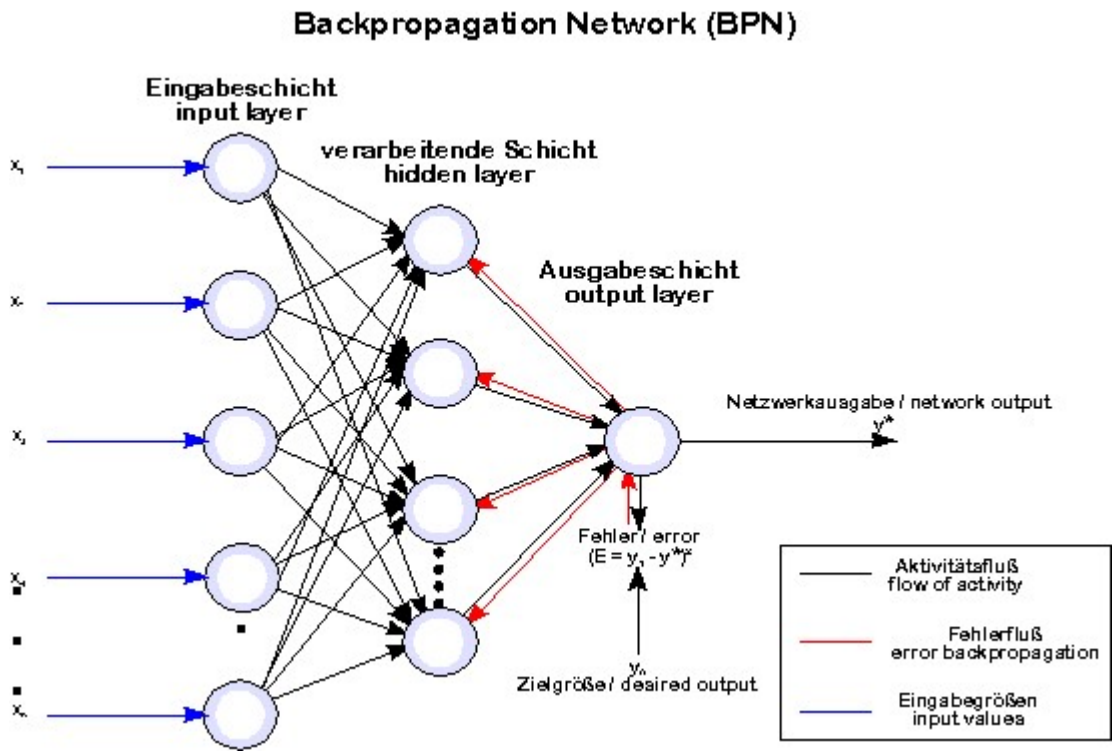
[27] Wright, P.B.,1985: Relationships between indices of the souther oscillation. Mon. weath. Rev. 112, 1913.

Table 1. Explained variance and standard deviations (modelled and observed) of global and hemispheric mean annual surface air temperaturee variations 1874 - 1993 using backpropagation (BPN) and counterpropagation (CPN) neural networks (NNM).

Temperature series	**NNM version**	**Explained variance**	**Standard deviation**		**Trend**	
			modelled	observed	modelled	observed
global	BPN CPN	79% 83%	0.1899 0.1876	0.2107 0.2107	0.60 K 0.55 K	0.58 K 0.58 K
northern hemisphere	BPN CPN	75% 62%	0.2014 0.1999	0.2332 0.2332	0.60 K 0.55 K	0.59 K 0.59 K
southern hemisphere	BPN CPN	78% 78%	0.1933 0.1854	0.2074 0.2074	0.61 K 0.53 K	0.58 K 0.58 K

Table 2. Comparison of radiative forcing (from IPCC, Houghton et al. 1996) with the related signals in global and hemispheric mean annual surface temperature variations 1874 - 1993 using the backpropagation (BPN) version of neural networks NNM.

Mechanism	Forcing in Wm**-2	Temperature signals (best estimate in parenthesis)		
		global	northern hem.	southern hem.
Greenhouse Gases (GHG)	(+) 2.1 - 2.8	0.9-1.3 (1.0)	0.6-1.3 (1.1)	0.4-0.7 (0.6)
Sulfate Aerosol (SUA)	(-) 0.3 - 0.9	0.2-0.4 (0.3)	0.1-0.6 (0.4)	0.1-0.3 (0.2)
Volcanism	(-) < 1	0.1-0.2 (0.2)	0.1-0.2 (0.2)	0.1-0.2 (0.2)
Solar activity	(+) 0.1 - 0.5	0.1-0.2 (0.1)	0.1-0.2 (0.1)	0.1-0.2 (0.1)
El Niño / Southern Oscillation	(+) -	0.2-0.3 (0.2)	0.1-0.3 (0.1)	0.2 - 0.4 (0.3)


Backpropagation Network (BPN)
Eingabeschicht
input layer
verarbeitende Schicht
hidden layer
Ausgabeschicht
output layer
Netzwerkausgabe / network output
Fehler / error
Zielgröße / desired output
Aktivitätsfluß
flow of activity
Fehlerfluß
error backpropagation
Eingabegrößen
input values

Counterpropagation Network (CPN)