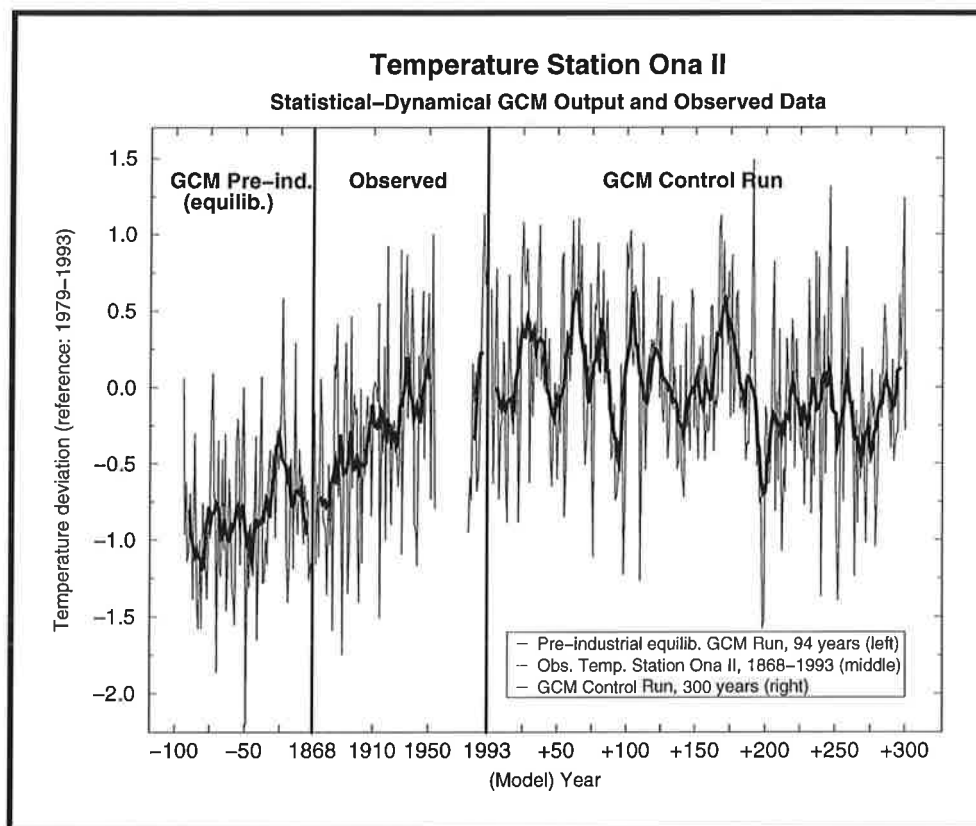




Max-Planck-Institut für Meteorologie

REPORT No. 274



A STATISTICAL-DYNAMICAL MODELING APPROACH FOR THE SIMULATION OF LOCAL PALEO PROXY RECORDS USING GCM OUTPUT

by

Bernhard K. Reichert • Lennart Bengtsson • Ove Åkesson

HAMBURG, August 1998

AUTHORS:

Bernhard K. Reichert,
Lennart Bengtsson

Max-Planck-Institut für Meteorologie
Hamburg
Germany

Ove Åkesson

Swedish Meteorological and
Hydrological Institute
Norrköping
Sweden

MAX-PLANCK-INSTITUT
FÜR METEOROLOGIE
BUNDESSTRASSE 55
D-20146 HAMBURG
F.R. GERMANY

Tel.: +49 - (0)40 - 411 73 - 0
Telefax: +49 - (0)40 - 411 73 - 298
E-Mail: <name>@dkrz.de

A Statistical-Dynamical Modeling Approach for the Simulation of Local Paleo Proxy Records using GCM Output

Bernhard K. Reichert¹⁾, Lennart Bengtsson¹⁾, Ove Åkesson²⁾

1) Max-Planck-Institut für Meteorologie

20146 Hamburg, Germany

2) Swedish Meteorological and Hydrological Institute

60176 Norrköping, Sweden

submitted to

Journal of Geophysical Research

ISSN 0937 - 1060

ABSTRACT

Recent proxy data obtained from ice core measurements, dendrochronology and valley glaciers provide important information on the evolution of the regional or local climate. General Circulation Models integrated over a long period of time could help to understand the (external and internal) forcing mechanisms of natural climate variability. For a systematic interpretation of in situ paleo proxy records, a combined method of dynamical and statistical modeling is proposed. Local 'paleo records' can be simulated from GCM output by first undertaking a model-consistent statistical downscaling and then using a process-based forward modeling approach to obtain the behavior of valley glaciers and the growth of trees under specific conditions. The simulated records can be compared to actual proxy records in order to investigate whether e.g. the response of glaciers to climatic change can be reproduced by models and to what extent climate variability obtained from proxy records (with the main focus on the last millennium) can be represented. For statistical downscaling to local weather conditions, a multiple linear forward regression model is used. Daily sets of observed weather station data and various large-scale predictors at 7 pressure levels obtained from ECMWF re-analyses are used for development of the model. Daily data give the closest and most robust relationships due to the strong dependence on individual synoptic-scale patterns. For some local variables, the performance of the model can be further increased by developing seasonal specific statistical relationships. The model is validated using both independent and restricted predictor data sets. The model is applied to a long integration of a mixed layer GCM experiment simulating pre-industrial climate variability. The dynamical-statistical local GCM output within a region around Nigardsbreen glacier, Norway is compared to nearby observed station data for the period 1868-1993. Patterns of observed variability on the annual to decadal scale and the mean temperature change due to pre-industrial climatic conditions are realistically simulated for this location. The local output produced by the described method will be used to force a process-based model for the production of 'synthetic' proxy data, e.g. the simulation of a valley glacier.

1. INTRODUCTION

Understanding spatio-temporal patterns and mechanisms of natural climate variability as well as the anthropogenic impact on climate requires the extension of instrumental records further back in time by the usage of paleoclimatic proxy data. Several attempts to reconstruct reliable temperature patterns over the last few centuries have been made [e.g. Landsberg et al., 1978; Groveman et al., 1979; Bradley and Jones, 1993; Barnett et al., 1996; Bradley, 1996; Mann et al., 1998]. Proxy records obtained from ice cores [e.g. Thompson, 1982], tree rings [e.g. Briffa et al., 1992], corals [e.g. Dunbar et al., 1994] as well as historical data [e.g. Pfister, 1992] and long instrumental records [e.g. Jones and Bradley, 1992] were used to reconstruct large-scale or global-scale patterns. Valley glaciers [e.g. Oerlemans, 1992; 1997] can also provide important information on the evolution of the regional or local climate.

How can proxy data best be interpreted? What are the underlying forcing mechanisms? Recent interpretation studies [Mann et al., 1998] have investigated the influence of external forcings, such as solar irradiance variations and explosive volcanism on northern hemispheric temperature variations. Whether pre-industrial climate variations are for the most part due to natural internal variations in the climate system can however not be excluded. The characteristic variability of a General Circulation Model (GCM) [e.g. Roeckner et al., 1996; Manabe and Stouffer, 1996] integrated over a long period of time could help to indicate credible explanations. However, a different methodology for a systematic evaluation of paleo proxy data is required as will be proposed in this paper.

The question we would like to address is whether we are able to simulate 'synthetic' paleo proxy records from GCM output for comparison with actual in situ proxy data. Our strategy is to perform a model-consistent statistical downscaling of the output of a GCM combined with a process-based forward modeling approach to simulate, for example, the behavior of valley glaciers and the growth of trees under specific conditions. Simulated records can be compared to actual in situ proxy records in order to investigate whether for example the response of glaciers to climatic change can be reproduced by models and to what extent climate variability obtained from proxy records (with the main focus on the last millennium) can be interpreted.

The growth of a valley glacier is mainly controlled by local temperature and precipitation [Paterson, 1981]. Such data are very difficult to obtain from gridpoint-scale GCM output

because of very large deviations due to local orographic conditions. As will be shown in this paper, a careful statistical model derived from the present climate can provide reliable local data which can be used to force the growth of such a valley glacier. A similar approach is required and can be developed for the evaluation of dendrochronological data.

Specific questions to be investigated concern the stability of the statistical model in general, as well as its performance on single seasons which are most important for a specific proxy indicator. Can we, for example, obtain from a GCM reliable local summer temperatures for the growing season of trees and which predictors do we need in order to do so?

Further questions concern the importance of horizontal GCM resolution and the time sampling for the determination of suitable predictors. It is found, for example, that daily data sets for the development of the statistical model give the closest and most stable relationships due to a strong dependence on individual synoptic patterns.

2. GENERAL STRATEGY

The general strategy proposed in this paper is the following (fig. 1). First we develop a statistical model between daily large-scale circulation patterns and corresponding local data observed by operational weather stations located near a proxy site to be investigated. large-scale patterns are represented by daily ECMWF (European Centre for Medium Range Weather Forecasts) re-analyses (ERA) for the period 1979-1993. We use daily data in order to include synoptic timescale variability and to achieve a physically robust relation. The obtained statistical relationships are applied to the daily coarse spatial gridpoint output of a GCM in order to achieve local GCM output (statistical downscaling). A forward modeling approach for a specific proxy, e.g. a glacier model [Oerlemans, 1996], can then be used to produce 'synthetic' (paleo-) proxy data which finally can be compared to actual in situ proxy data.

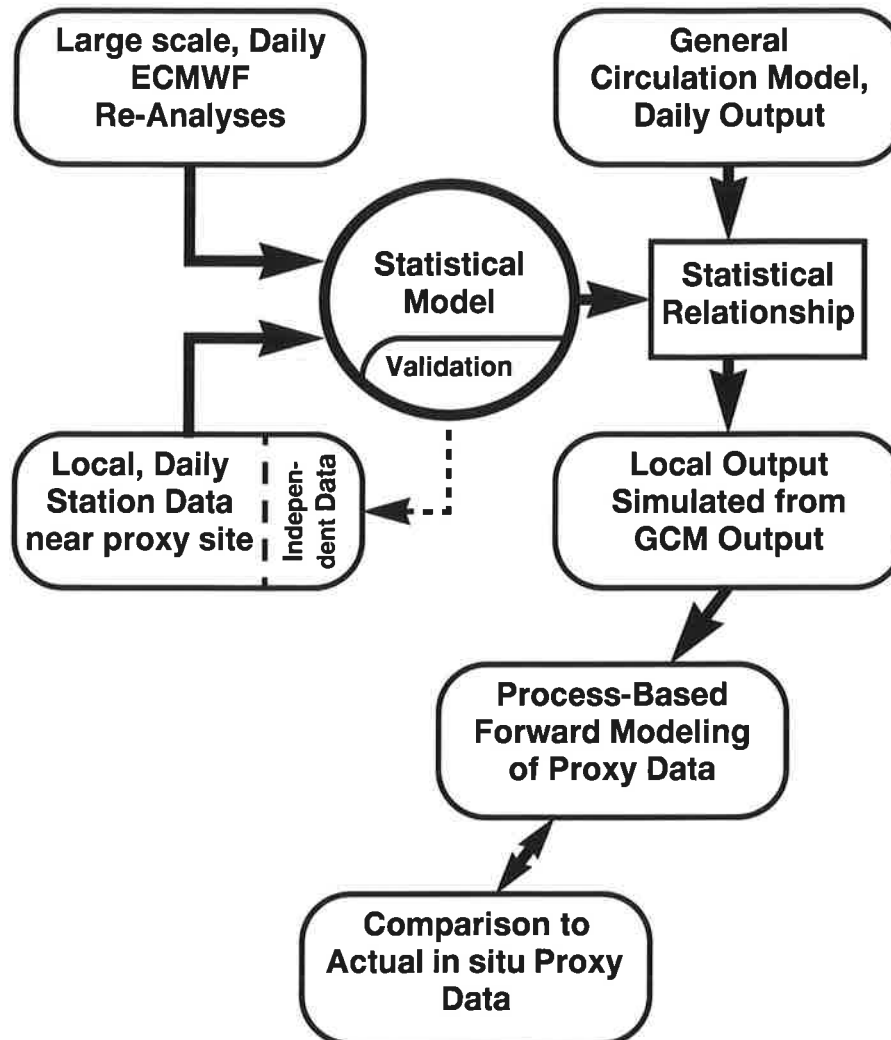


Fig. 1. General strategy for the interpretation and usage of in situ paleo proxy data. See text for further explanations.

3. STATISTICAL MODEL

The response of local weather to large-scale flow patterns of the atmosphere has been noted for a very long time, not only by meteorologists, but also by laymen interested in weather. The most common feature is perhaps the precipitation in mountainous regions which is particularly determined by orographic forcing generating enhanced precipitation on the windward and reduced precipitation on the leeward side. In most areas the local conditions (topography and land surface characteristics) have a major effect not only on precipitation but also on wind and temperature as well as on cloudiness and visibility. Bergeron [1930; 1981] proposed that a special climatology should be established classifying local climate in terms of the large-scale flow. We may notify such an approach as a dynamic climate classification. The significance of this approach became obvious as it became possible to predict the synoptic flow by numerical models. A dynamic climatology can be produced for any particular local weather parameter (predictand), e.g. local precipitation, cloud cover, cloud height, visibility, maximum or minimum temperature, by the use of different large-scale predictors, e.g. surface pressure, wind, geopotential thickness, vertical velocity, large-scale precipitation.

For the purpose of this analysis we use a multiple linear forward regression model in order to establish relationships between the large-scale flow and local weather parameters. For the $i = 1, 2, \dots, n$ values of an observed (dependent) quantity y_i (predictand) it takes the form of a linear combination

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i$$

where $x_{i0} = 1$ and $x_{i1}, x_{i2}, \dots, x_{ip}$ are the settings of the p corresponding (independent) quantities (predictors), $\beta_0, \beta_1, \dots, \beta_p$ are the regression parameters which are to be estimated and ε_i are unknown independent random errors (see for example von Storch and Zwiers [1998]). We use least-squares estimation which means that best estimates of the unknown regression parameters are calculated by minimizing $\sum_{i=1}^n \varepsilon_{ik}^2$ for each predictand $k = 1, 2, \dots, q$.

However, it is neither necessary nor desirable to include all potential predictors in the data set for the prediction of a specific observed local variable. The maximum number of predictors that may be used in the model in order to get a 'stable' solution which fits not only the developmental sample but works also on independent data sets, is a function of the sample size.

Furthermore, some predictors included in the model might be a linear or 'near-linear' combination of other predictors (collinearity or 'near-collinearity') which could cause unstable results. In order to address these problems we choose the following selection procedure for the large daily data set that is used in this study: The model is built up stepwise using an interactive forward selection procedure of independent variables. After having chosen a single predictor with maximum correlation, the next independent variable providing the best fit in conjunction with the first one is added and tested for near-collinearity. In a critical case the user may decide whether this variable should be included or not. Further variables are added in a recursive fashion until a saturation criteria (the correlation does not improve significantly) is reached.

4. DATA SETS: ECMWF RE-ANALYSES AND OBSERVED DATA

The development of the statistical model is based on ECMWF re-analyses (ERA) [Gibson et al., 1997] used for an area of about $11^{\circ} \times 11^{\circ}$ in Norway (covering the proxy site of Nigardsbreen glacier at $61^{\circ}43'N$, $7^{\circ}08'E$ to be investigated) and on local observational records for 22 synoptic weather stations within that area.

The ECMWF re-analysis project has produced a validated and reasonably consistent global data set of assimilated data for the period 1979-1993 [Gibson et al., 1997]. In this study, ERA data constitute the potential predictors for the development of the statistical model. We use ERA 24 hour forecasts for precipitation in order to address the spin-up problem and to have a consistent picture of precipitation [Stendel and Arpe, 1997]. For all other surface and pressure level variables (Table 1), 6 hourly initialized analysis are taken and daily averaged. We extract pressure level variables on the 1000, 925, 850, 700, 500, 400 and 300 hPa levels. In order to meet the resolution of the ECHAM GCM runs which we intend to use afterwards, the ERA output is interpolated to T30 resolution ($\sim 3.8^{\circ} \times 3.8^{\circ}$). Additional experiments with original T106 ($\sim 1.1^{\circ} \times 1.1^{\circ}$) resolution are also analyzed. For each location of the operational weather stations we compute weighted area means for an area covering roughly 1200×1200 km as input for the statistical model.

The predictands of the model consist of observational data for 22 operational weather stations (Fig. 2) in the surrounding of the Nigardsbreen glacier for the period 1979-1993. We interpolate missing values in the 6 hourly weather data before daily averaging. However, data

quality is adequate for most stations in that period and only few data are missing. Table 2 shows the available observed parameters which are used as predictands (dependent variables).

PRESS. LEV. DATA (1000, 925, 850, 700, 500, 400, 300 hPa)	COMPOSITED PREDICTORS
Temperature	Geopot. 925-1000 hPa
Dewpoint Temperature	Geopot. 850-1000 hPa
Wind-velocity (u, v)	Geopot. 700-1000 hPa
Vertical velocity	Geopot. 500-1000 hPa
Vorticity	Geopot. 500-850 hPa
Divergence	Geopot. 500-700 hPa
Geopotential height	Seasonal cycle: sin(day)
Relative humidity	Seasonal cycle: cos(day)
SURFACE DATA	Seasonal cycle: sin(2day)
Large scale precip.	Seasonal cycle: cos(2day)
Convective precip.	Large-scale+conv. precip.
Mean sea level press.	Sqrt. of total precip.
Total cloud cover	Vertic. integr. liquid water
Tot. column water vap.	Lapse rate (low. Trop.)
	K-Index (George, 1960)

Table 1. Potential large-scale predictors from ECMWF re-analyses used as input for the statistical model. For the model version excluding near-surface predictors (see text), only potential predictors on 850, 700, 500, 400 and 300 hPa levels, Mean sea level pressure and the seasonal cycle were used (shaded fields).

OBSERVED DATA	COMPOSITED
Pressure reduced	Sqrt of 24h precipitation
Wind direction	Cube of 24h precipitation
Wind velocity	Fourth root of precip.
Temperature	Logarithm of precip.
Dewpoint temperature	Exponential of precip.
24h precipitation	Sine of wind direction
Visibility	Cosine of wind direction
Weather code	U wind component
Total cloud amount	V wind component
Low and middle cloud	Relative Humidity
Cloud base	
Low level cloud	
Middle level cloud	
High level cloud	
Snow	

Table 2. Observed parameters from operational weather stations representing the predictands (dependent data) of the statistical model. Results only for a selection of these parameters will be shown below.

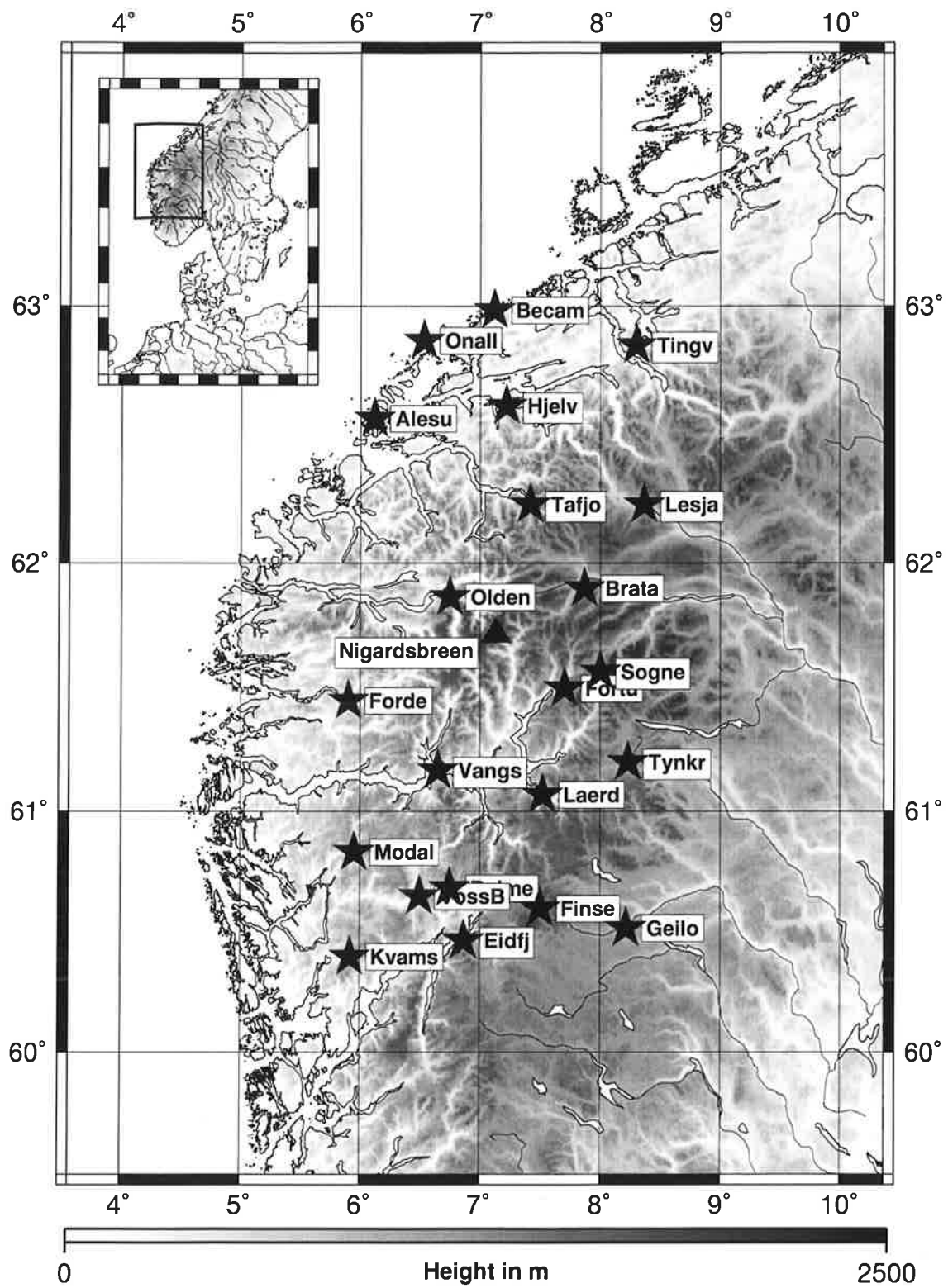


Fig. 2. Location of operational weather stations used for development of the statistical model. We anticipate to provide reliable local output for Nigardsbreen valley glacier (triangle at $61^{\circ}43'N$, $7^{\circ}08'E$).

5. RESULTS OF STATISTICAL MODELING

The complete data set that can potentially enter the statistical model consists of daily values of 82 large-scale potential predictors from the ECMWF re-analyses and 20 local observed predictands for the period 1979-1993, which makes about 5500 daily sets of 102 variables for each of the 22 stations.

5.a Role of near-surface Predictors

In order to investigate the role of near-surface predictors for the statistical model, we choose two different model versions. In the first one, we allow all predictors (complete table 1) to potentially enter the equations. The predictors which are finally selected by the model represent only a small subset of these variables, for local precipitation, for example, usually not more than five large-scale predictors play a significant role and are therefore actually used. In the second one we only use potential predictors above the 850 hPa level (850, 700, 500, 400, 300 hPa), mean sea level pressure and the seasonal cycle (shaded fields in table 1). Here, near-surface predictors are excluded in order to be able to get stable results even when the model is applied to various GCMs which might differ in the underlying topography and in the representation of surface processes. It is found that correlation between observed and predicted variables in this model only slightly decreases compared to the first model version which suggests that predictors above the 850 hPa level are already sufficient for the prediction of the desired local surface variables. The large-scale variability of the synoptic timescale flow is well determined from predictors above the 850 hPa level.

5.b General Model Performance

Figures 3.a-f show an example of statistical model output for the Norwegian station Kvamskogen (60°24'N, 5°55'E, 408 m asl). Here, we perform the statistical model on a daily basis with re-analyses at T30 resolution including predictors above 850 hPa only (second model version) for the period 1979-1992. The light gray curves show large-scale direct re-analyses for the station without statistical modeling, the black curves represent statistically simulated results, the dotted curves show observational data. The correlation coefficient r is shown in the top left corner of each graph.

The station is generally characterized by an exceptional high amount of observed precipitation (up to 4000 mm/year) which can naturally not be represented by the direct large-

scale re-analyses precipitation (about 1300 mm/year).

In order to achieve the best fit, the statistical model finally selects three large-scale predictors for the prediction of local precipitation. These can be found in the upper part of table 3.b together with their relative impacts in the final equations. The predictors enter the equations such that i) negative vertical velocity (upward air movement at 850 hPa, relative impact 60%), ii) positive u-wind velocity (westerly winds at 700 hPa, relative impact 20%) and iii) positive vorticity (cyclonic movement at 700 hPa, relative impact 20%) on the large-scale determine an increase of local precipitation, which for this station is strongly orographically enhanced. Using these three predictors only, the explained variance (r^2) between the statistical simulation and the observed record of precipitation is 47.7% for daily data and for the complete years 1979-1992 (lower part of table 3.b). Figure 3.a (daily values for February/March 1980) shows an example for periods where the model produces too little precipitation (e.g. around March 13th) and periods where actually non-existent precipitation is generated (e.g. within Feb 17th to Feb 27th no precipitation was observed). However, the explained variance for monthly means of these data is as high as 81.7% (fig 3.b for 1980-1983), annual means (fig 3.c) even show a more remarkable explained variance of 88.6%. The annual mean precipitation is extremely realistically simulated by the statistical model (fig 3.c) on the basis of daily input values. Further experiments with monthly and yearly values show that daily values are required to obtain such robust and physically reasonable couplings. Local climatic conditions can be modeled well using daily large-scale predictors. However, the orographic effect plays an important role and is perhaps the easiest to determine, especially in winter time with pronounced synoptic flow.

Fig. 3.d-f show the model results for annual means of temperature, cloud cover and relative humidity respectively. The predictors chosen for local temperature (table 3.a, upper part) are large-scale temperature at 850 hPa with the highest relative impact (65%), followed by the seasonal cycle (25%) and relative humidity at 850 hPa (10%). The explained variance of 91.3% for daily data (table 3.a, lower part) is not representative because it includes the seasonal cycle. However, monthly data can explain 85% of variance after having removed the seasonal cycle and yearly data as much as 91.9%. The explained variance of annual total cloud cover (67.2%; fig 3.e) is also improved compared to the direct large-scale re-analyses. The same is true for annual relative humidity (fig 3.f) compared to re-analyses relative humidity on 925 hPa.

This is just one example out of 22 stations which we use in this study. The statistical model may use different optimized predictors for the same observed variable at each station due to its local setting.

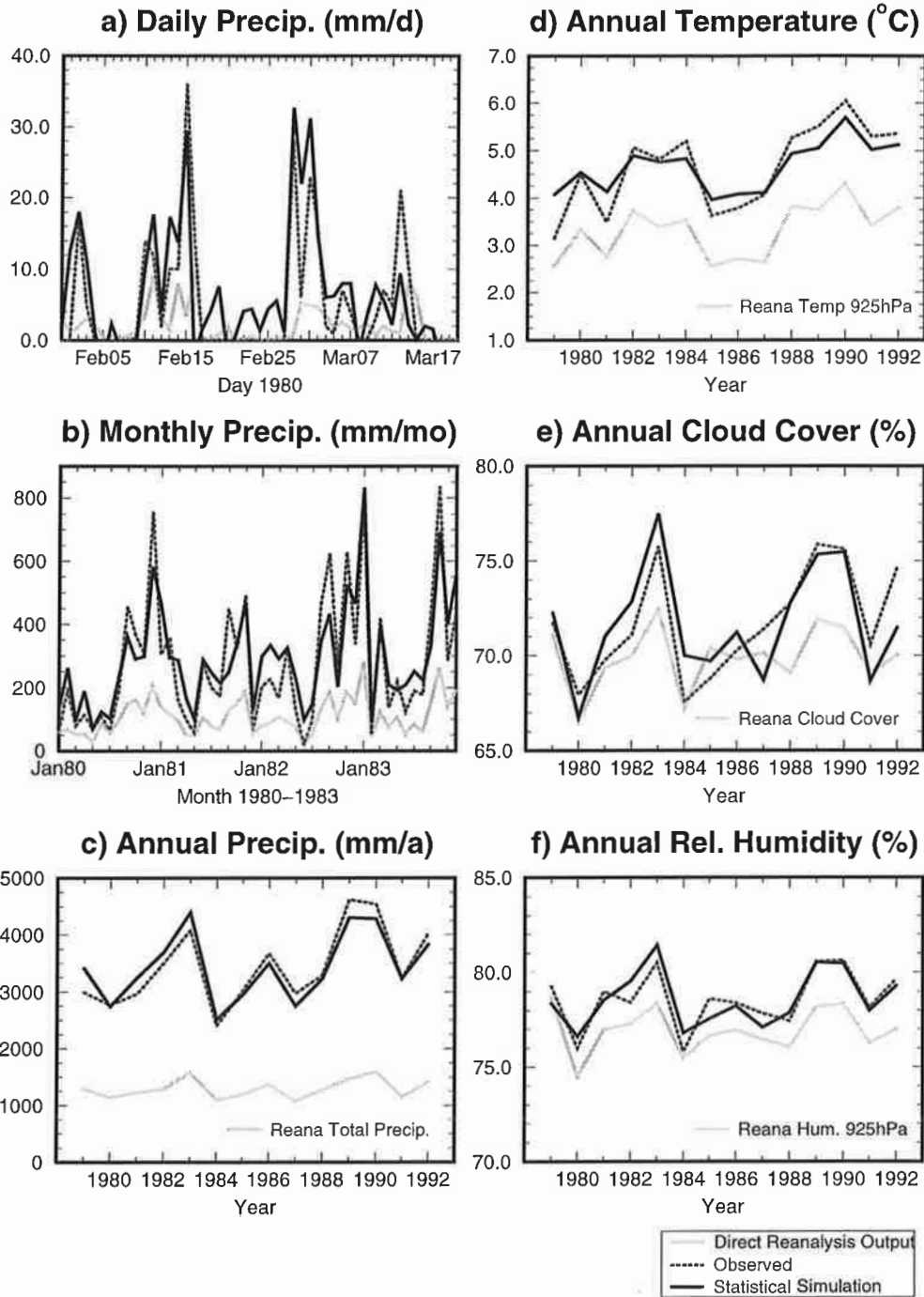


Fig. 3. Results of the statistical model for station Kvamskogen (60°24'N, 5°55'E, 408 m asl, see fig. 2). Local output for a) daily precipitation 1 Feb - 19 Mar 1980, b) monthly precipitation Jan 1980 - Dec 1983, c-f) annual precipitation, temperature, total cloud cover and relative humidity 1979-92. The black lines show statistically simulated results, the dotted lines represent observed data, the light gray lines are directly interpolated re-analyses without statistical modeling. The correlation coefficient r is shown in the top left corner of each graph (see text).

Predictand:	a) Local Temperature (for complete years 79-92)			b) Local Precipitation (for complete years 79-92)		
	Predictors		rel. imp.	Predictors		rel. imp.
	Temp. 850 hPa		65 %	Vert. Vel. 850 hPa		60 %
	Seas. Cycle		25 %	Vorticity 700 hPa		20 %
	Rel. Hum. 850hPa		10 %	Zon. Wind 700hPa		20 %
Expl. Var. for	Daily Data (A)	Monthly Data (B)	Annual Data (C)	Daily Data (A)	Monthly Data (B)	Annual Data (C)
i) Year	(91.3 %)	85.0 %	91.9 %	47.7 %	81.7 %	88.6 %
ii) JJA	69.8 %	79.4 %	77.5 %	29.6 %	57.2 %	46.0 %
iii) DJF	71.7 %	93.5 %	95.7 %	59.9 %	90.1 %	94.0 %

Table 3. Predictors selected by the statistical model (upper part of table) and percentage of observed variance explained (lower part of table). The predictors are shown with their individual relative impact in the final equations for a) local temperature and b) local precipitation for station Kvamskogen using daily data for complete years within the period 1979-92. The lower part of the table shows the percentage of observed variance explained by the statistical model for i) the whole year, ii) JJA only and iii) DJF only each for A) original daily data, B) monthly averaged output of the statistical model after removing the seasonal cycle and C) yearly means of statistical model output (which means seasonally averaged output in case of JJA and DJF).

5.c Model Runs for Specific Seasons of the Year

Is it possible that different predictors may be required for different seasons of the year? The simulation of proxy indicators may require realistic local output with a particular interest in specific seasons (e.g. the growing season of trees, the melting period of glaciers). In order to investigate the seasonal performance of the statistical model, we carried out further experiments allowing daily data for single seasons only as input (table 4.a-b). Compared to the model with full year daily input (table 3.a-b), the composition of predictors and their relative impacts may change.

1) JJA temperatures: If we restrict the input data to daily values of JJA (table 4.a) then local JJA temperatures are determined by large-scale zonal wind at 850hPa (relative impact 17%) and vertical velocity at 500hPa (relative impact 12%) in addition to 850 hPa temperature (relative

impact 71%). Here, the daily, monthly and seasonal explained variance for JJA is 81%, 91.8% and 89.1% respectively (lower part of table 4.a). Compared to the seasonal performance with the full year input (69.8%, 79.4% and 77.5%; table 3.a, bottom) the explained variances improve significantly and the predictors have changed.

2) DJF temperatures: The explained variance for northern hemispheric winter (76% in table 4.b with daily DJF input data only) is slightly improved compared to the full year input (72% in table 3.b) whereas monthly and seasonal values remain nearly constant. This means that the large-scale flow patterns in winter time are already reasonably well determined by the full year input data.

Experiments with daily data for specific seasons are also carried out for the prediction of local precipitation. However, the composition of predictors only slightly changes and the explained variance cannot be improved significantly.

We may conclude that the capability of the statistical model to simulate single seasons can be improved when a seasonal specific statistical model is developed for some local observed variables (e.g. temperature) as demonstrated above.

Predictand:	a) Local Temperature for JJA only (1979-92)			b) Local Temperature for DJF only (1979-92)		
	Predictors		rel. imp.	Predictors		rel. imp.
	Temp. 850 hPa		71 %	Temp. 850 hPa		60 %
	Zon. Wind 850hPa		17 %	Rel. Hum. 850 hPa		24 %
	Vert. Vel. 500 hPa		12 %	Zon. Wind 500hPa		16 %
Expl. Var. for	Daily Data (A)	Monthly Data (B)	Seas. Data (C)	Daily Data	Monthly Data	Seas. Data
	ii) JJA	81.0 %	91.8 %	89.1 %		
iii) DJF				75.8 %	94.5 %	95.6 %

Table 4. Predictors selected and performance of the seasonal specific statistical model for local temperature. The model was developed with re-analyses and weather station data from a) JJA only and b) DJF only within the period 1979-1992 (see text and table 3 for further explanations). The lower part of the table shows the percentage of observed variance explained by the statistical model for ii) JJA only and iii) DJF only each for A) original daily data, B) monthly averaged output of the statistical model after removing the seasonal cycle and C) seasonal means of statistical model output.

5.d Model Validation Experiments

Does the model work on independent data and is it transferable to other time periods with different climatic conditions? In order to address these questions two further experiments are carried out.

In the first experiment we develop the model for the second half of the ECMWF re-analyses time period (1985-1992) only (fig. 4.a-b). For the purpose of validation, the statistical relationship obtained is then applied to an independent validation sample, in this case the first half of the re-analyses time period (1979-1984). The annual means of precipitation and temperature for this experiment are shown in fig. 4.a-b. Although the model is not developed for 1979-1984 (i.e. it does not use any local observation in that period for fitting), it can still realistically simulate the local variables for this independent time period. The differences to fig 3.c-d respectively (where we have used the full time period 1979-1992 for development) are insignificant.

The second experiment addresses the question whether the model can produce realistic output for climatic conditions which differ from the conditions it is actually developed for. If the model is developed using present-day climatic conditions (represented by ECMWF re-analyses), is it then applicable to GCM output for pre-industrial times? The statistical distribution of daily local temperatures for station Kvamskogen after statistical modeling of ECMWF re-analyses for the period 1979-1992 is shown in fig. 5. The standard experiment (fig 5.a) includes all occurring temperature events for the development of the model. For the validation experiment (fig 5.b) events with temperatures less than -5°C are excluded prior to model development and the statistical relationships are calculated the same way. Although this model is not developed for events $< -5^{\circ}\text{C}$ it has still realistically simulated them, which can be clearly seen in the distribution (fig. 5.b upper graph) and the time series (fig. 5.b lower graph) of temperatures. The differences between standard and validation experiment (fig. 5.a and 5.b resp.) are insignificant.

The experiments demonstrate the ability of applying the statistical model to independent data sets and they also show that we can produce realistic output for climatic conditions different from present-day climate which we use for development of the model. However, the latter statement may only be true for climatic conditions which are not fundamentally different in their large-scale flow properties. This is to a large extent true for the pre-industrial output of the GCM

which we will apply below. Here, the general coupling between the large-scale flow and local weather parameters deduced from present-day climate are to a great extent maintained for pre-industrial climatic conditions.

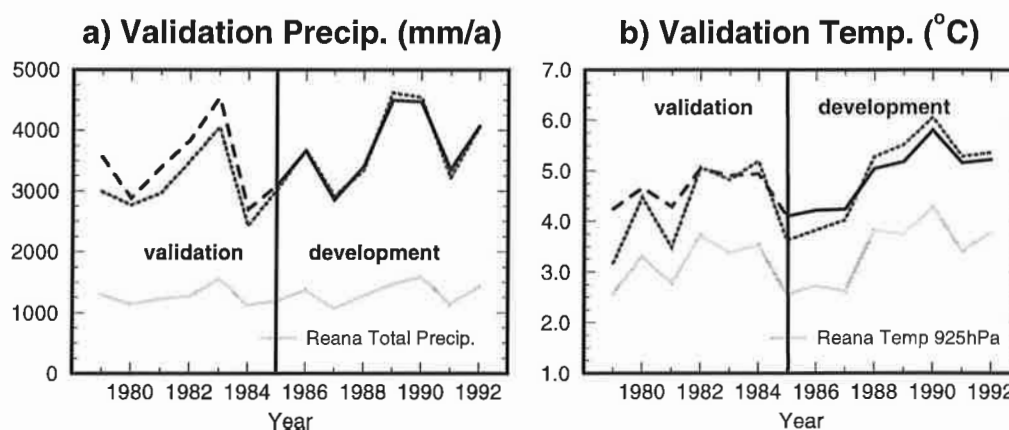


Fig. 4. Validation experiment for station Kvamskogen a) for local annual precipitation and b) for local annual temperature. The model is developed for the 1985-1992 period only (developmental sample). The statistical relationship obtained is then applied to independent re-analyses for 1979-1984 (validation sample). The dashed line shows the statistical model output for the validation sample, other lines like in fig. 3. The differences to fig 3.c-d (full time period 1979-1992 used for development) are reasonably small.

Statistical Distribution of Daily Local Temperatures
Station Kvamskogen, Stat. Modeled ECMWF Re-Analyses 1979–1992

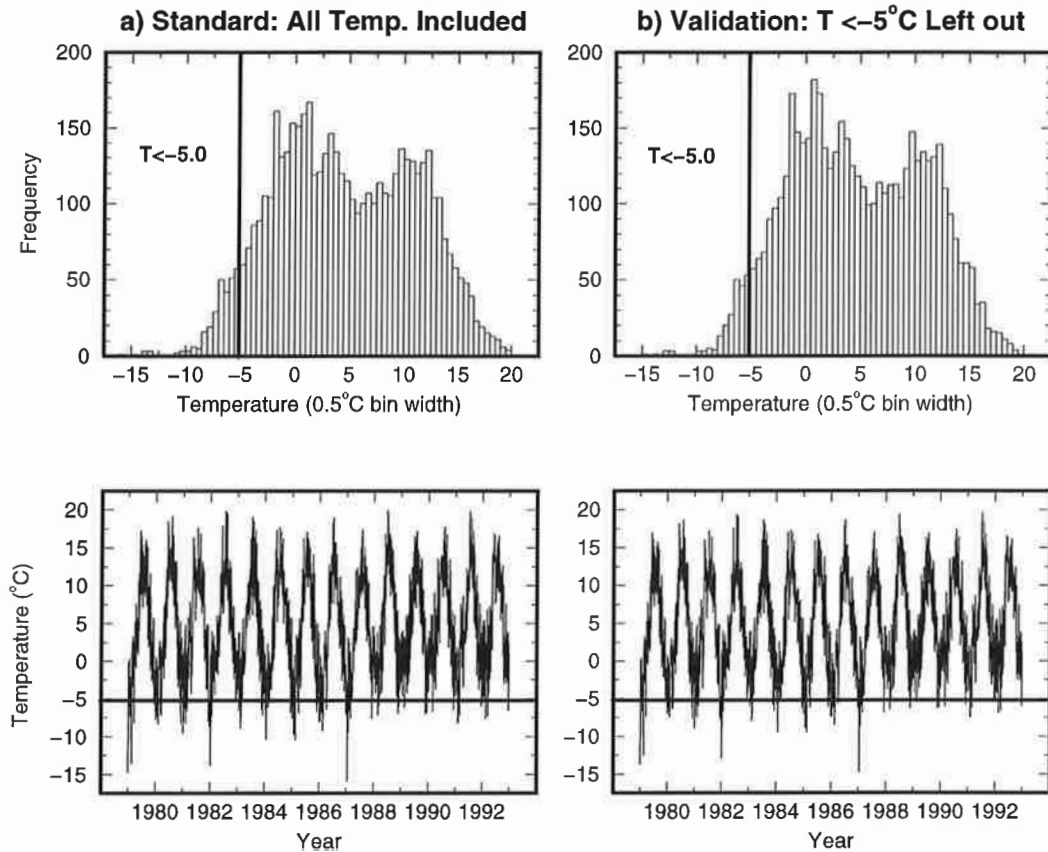


Fig. 5. Statistical distribution (upper graph) of daily local temperatures for station Kvamskogen after statistical modeling for the period 1979-1992 and time series of daily temperatures (lower graph) for a) all temperature events included for the development of the model and b) events with temperatures less than -5°C excluded from model development. Although the model in b) is not developed for events $< -5^{\circ}\text{C}$ it has still realistically simulated them, the differences in distribution between a) and b) are reasonably small. This gives some support for the assumption that we can produce realistic output for climatic conditions which differ from the conditions we use for development of the model.

5.e Impact of Spatio-Temporal Resolution of Predictors

Our intention is to produce reliable local monthly or annual mean output from re-analyses and GCMs comparable to corresponding local weather station data. Although we do not aim to produce perfect predictions on a daily basis, it turns out that our model requires daily predictor data as input in order to achieve the statistically closest and most robust relationships. Consequently, the output can best be averaged to monthly or annual means after the statistical calculations have been carried out using daily predictor data.

Monthly means as input are tested as well but the results are more unstable. The model sometimes attempts to use near-collinear predictors for least-squares estimation which particularly becomes a problem when being applied to independent data.

However, this also emphasizes the role of synoptic timescale variability in the close relationship between local weather and large-scale circulation patterns. Averaging the predictors means that this information is partly removed and the correlation therefore weakens.

The impact of spatial resolution of predictors was tested by producing statistical model output both with original re-analyses at T106 resolution ($\sim 1.1^\circ \times 1.1^\circ$) and with an interpolated data set (T30 res.; $\sim 3.8^\circ \times 3.8^\circ$) for comparison. Although the interpolated re-analyses might not really represent the output of an actual T30 model it can still give us some hints about the impact of model resolution. It is found that summer temperatures (JJA) are most sensitive. For the example given in table 3.a the explained variance of monthly data for JJA (seasonal cycle removed) is increased from 79.4% (T30 res.) to 84.5% (T106 res.); annual data improves from 77.5% (T30 res.) to 83.3% (T106 res.). The sensitivity for local precipitation is of the same order. Monthly means for JJA increase from 57.2% (table 3.b) to 62.1%, annual means from 46.0% to 51.7%. In spite of these improvements, we may conclude that the overall ability to produce satisfying statistical model output is already given using T30 predictor resolution, which is also the resolution of the ECHAM4 GCM runs which will be used below.

6. APPLICATION TO ECHAM4 GCM: CONTROL AND PRE-INDUSTRIAL RUN

The statistical relationships derived from re-analyses and local station data are applied to the output of the ECHAM4 GCM coupled to a Mixed Layer Ocean developed at the Max-Planck-Institut für Meteorologie (MPI) and the Deutsches Klimarechenzentrum (DKRZ) in Hamburg [Roeckner et al., 1996; Roeckner, 1997; Roeckner et al., 1998]. We use a long integration of a control run and a run with pre-industrial greenhouse gas forcing for our experiments.

The control experiment of the 19-layer ECHAM4/MLO GCM lasts for 600 years at T30 resolution. A 100-year equilibrium run with pre-industrial greenhouse gas forcing [IPCC, 1995] is also used. The global mean surface air temperature for the pre-industrial run is 1.0°C lower than the control run. Here, we take the daily large-scale pressure level and surface output of these models for the area of Nigardsbreen glacier, Norway and apply the statistical relations obtained as described above.

As an example, we show the statistically downscaled and yearly averaged GCM output of surface air temperature deviations for station Ona II (fig. 6). This station is chosen because we are able to compare simulated data to a long instrumental observed temperature record for the period 1868-1955 which is available from the Global Historical Climatology Network temperature data base [Peterson et al., 1997]. Additionally we include weather station data for 1979-1993 obtained from the SMHI, Sweden (Åkesson, pers. comm.) for this station.

Figure 6 shows the statistically corrected GCM surface air temperature output of the equilibrium run with pre-industrial greenhouse gas forcing along with the corrected GCM output for 300 years of the control run. For comparison we also show the observed temperature relative to the 1979-1993 reference period.

The first striking feature is the temperature increase of about 0.7°C between the pre-industrial GCM run and the control run for this station, which is about the same as the observed temperature difference. This means that the dynamical GCM output in combination with the statistical modeling procedure is able to realistically simulate pre-industrial to present-day temperature changes for this location.

Furthermore, it can be clearly seen that the annual variability in GCM output (both control and pre-industrial run) resembles very much the observed one and that 10-year running means show a comparable or even slightly higher variability than observed. Pronounced low frequency

fluctuations with even longer periods are clearly simulated by the downscaled GCM control experiment.

We may conclude that the dynamical-statistical modeling procedure can realistically simulate both the patterns of observed variability on the annual to decadal scale as well as temperature changes due to different climatic scenarios.

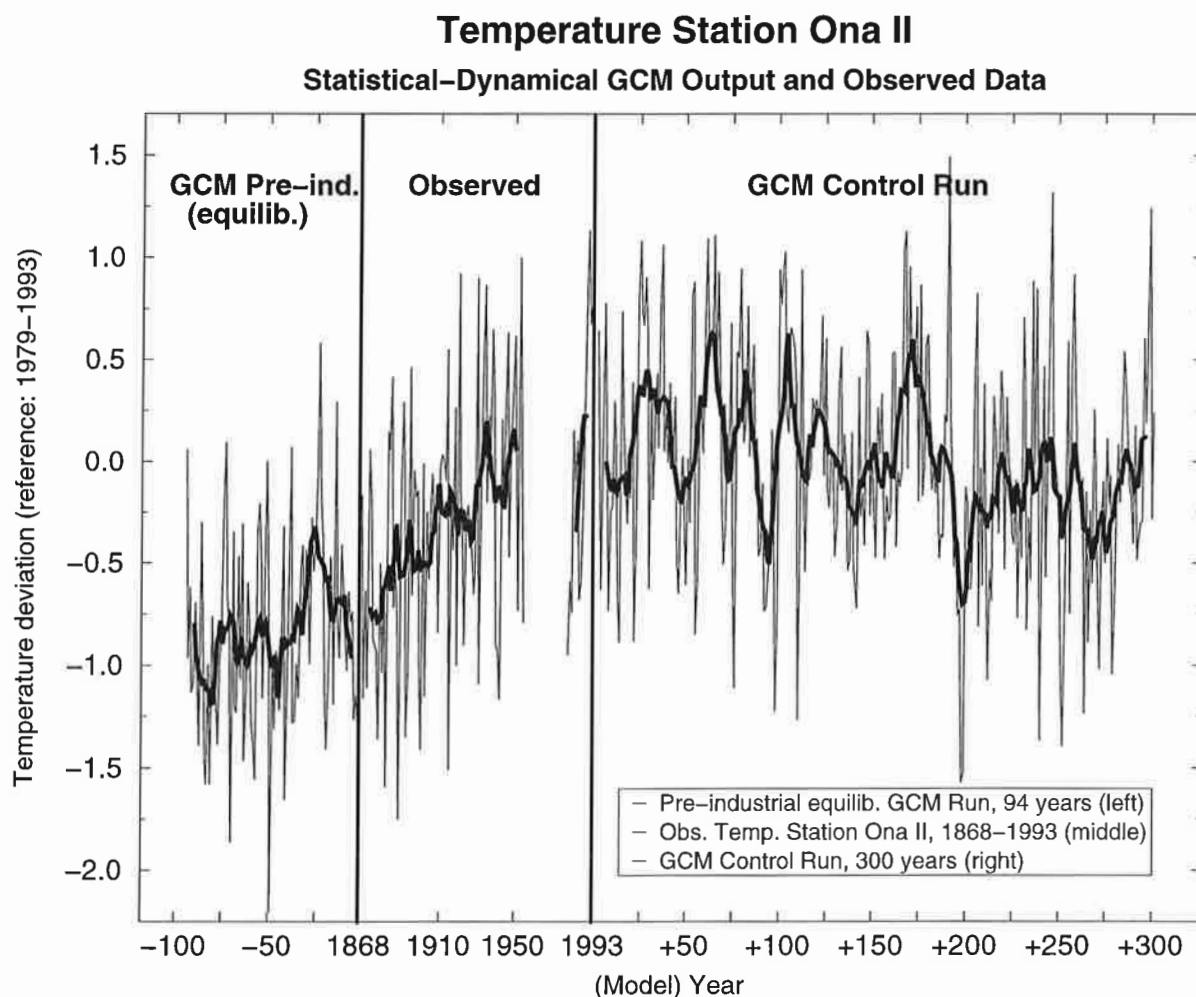


Fig. 6. Application of the statistical model to GCM output (ECHAM4 T30 L19 Mixed Layer Ocean) for station Ona II. Thin lines are annual means, thick lines are 10-year-running means. The left part shows the statistically corrected local output from the equilibrium GCM run with pre-industrial GHG forcing, the right part represents the 300 year control run of the same model. For comparison, the observed temperature relative to the 1979-1993 reference period is shown in the middle (see text).

7. CONCLUSIONS

Local output for temperature, precipitation and other parameters has been produced by a General Circulation Model in combination with a statistical downscaling model. Daily ECMWF re-analyses using predictors excluding near-surface predictors gave stable and physically reasonable relationships for statistical model development. The model showed a high performance using large-scale predictors from re-analyses and local surface observations for the area of Nigardsbreen glacier, Norway. Analyzing single seasons individually, it became clear that for some local surface variables it is useful to develop a specific set of predictors for seasons which might be most relevant for a specific proxy indicator (e.g. for the growing season of trees). Daily predictor data were required in order to achieve statistically the most stable and physically the most reasonable relationships. Satisfying results for the model could be achieved using T30 resolution ($\sim 3.8^\circ \times 3.8^\circ$) predictor data. We validated the model using separate developmental and validation intervals for the re-analysis time period and we carried out a validation experiment with a restricted predictor data set. The method has been applied to a long control integration of the ECHAM4 / Mixed Layer Ocean GCM and to an equilibrium run with pre-industrial greenhouse gas forcing. The output has been compared to patterns of observed station data in the area of Nigardsbreen glacier, Norway for the period 1868-1993. Patterns of observed variability on the annual to decadal scale and of mean temperature changes due to different climatic scenarios have been realistically simulated for this location.

The proposed dynamical-statistical modeling approach could help to improve a systematic interpretation of paleo proxy records and model-data intercomparisons for past climatic scenarios. A simulation of the growth of trees and the response of valley glaciers to specific climatic conditions is in preparation.

ACKNOWLEDGEMENTS

The authors would like to thank M. Giorgetta, M. Stendel and B. Machenhauer for helpful comments and discussions. The ECMWF provided the re-analyzed data set, the SMHI contributed with operational weather station data and NOAA made GHCN data available. The study was supported by the European Commission under contract ENV4-CT95-0072. The model simulations were performed at the Deutsches Klimarechenzentrum (DKRZ) in Hamburg.

REFERENCES

- Barnett, T.P., B. D. Santer, P. D. Jones, R. S. Bradley, and K. R. Briffa, Estimates of low frequency natural variability in near-surface air temperature, *The Holocene*, 6, 255-263, 1996
- Bergeron, T., Richtlinien einer dynamischen Klimatologie, *Met. Z.*, 47, 246, 1930 (English translation by H. C. Willett, Ground plan of a dynamic climatology, *Mo. Wea. Rev.*, 59, 219, 1931)
- Bergeron, T., Synoptic meteorology: an historical review, *Pure and Appl. Geoph.*, 119, 443-473, 1981
- Bradley, R. S., and P. D. Jones, 'Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends, *The Holocene*, 3, 367-376, 1993
- Bradley, R. S., Are there optimum sites for global paleotemperature reconstruction?, in *Climatic Variations and Forcing Mechanisms of the last 2000 Years*, edited by P. D. Jones, R. S. Bradley and J. Jouzel, NATO ASI Series, I, 41, pp. 603-624, Springer, 1996
- Briffa, K. R., P. D. Jones, T. S. Bartholin, D. Eckstein, F. H. Schweingruber, W. Karlen, P. Zetterberg, and M. Eronen, Fennoscandian summers from AD 500: temperature changes on short and long timescales, *Clim. Dyn.*, 7, 111-119, 1992
- Dunbar, R. B., G. M. Wellington, M. W. Colgan, and P. W. Glynn, Eastern pacific sea surface temperature since 1600 A.D., The delta 18O record of climate variability in the Galapagos corals, *Paleoceanography*, 9, 291-315, 1994
- George, J. J. (Ed.), *Weather forecasting for Aeronautics*, Academic Press, pp. 407-415, 1960
- Gibson, J. K., P. Kållberg, S. Uppala, A. Nomura, A. Hernandez, and E. Serrano, ERA Description, *ECMWF Re-Analysis Project Report Series*, 1, 1997

Groverman, B. S., Reconstruction of Northern Hemisphere temperature: 1579-1880, in *Meteorol. Progr. Publ.*, Maryland Univ., College Park, pp. 79-181, 1979

IPCC Intergovernmental Panel on Climate Change, *Climate Change 1994, Radiative forcing of climate change*, edited by J. T. Houghton, L. G. Meira Filho, J. Bruce, H. Lee, B. A Callander, E. Haites, N. Harris and K. Maskell, p. 80/194, Cambridge, 1995

Jones, P. D., and R. S. Bradley, Climatic variations in the longest instrumental records, in *Climate Since A. D. 1500*, edited by R. S. Bradley and P. D. Jones, pp. 246-268, Routledge, 1992

Landsberg, H. E., B. S. Groverman, and I. M. Hakkarinen, A simple method for approximating the annual temperature of the northern hemisphere, *Geophys. Res. Letters*, 5, 505-506, 1978

Manabe, S., and R. J. Stouffer, Low-frequency variability of surface air temperature in a 1000-year integration of a coupled atmosphere-ocean-land surface model, *J. Climate*, 9, 376-393, 1996

Mann, M. E., R. S. Bradley, and M. K. Hughes, Global-scale temperature patterns and climate forcing over the past six centuries, *Nature*, 392, 779-787, 1998

Oerlemans, J., Climate sensitivity of glaciers in southern Norway: application of an energy-balance model to Nigardsbreen, Helstugubreen and Alftobreen, *J. Glaciology*, 38, 223-232, 1992

Oerlemans, J., Modelling the response of valley glaciers to climatic change, in *Physics and chemistry of the atmospheres of the earth and other objects of the solar system*, edited by C. Boutron, pp. 91-123, ERCA, 2, 1996

Oerlemans, J., A flow-line model for Nigardsbreen: projection of future glacier length based

on dynamic calibration with the historic record, *Ann. Glac.*, 24, 382-389, 1997

Paterson, W. S. B. (Ed.), *The physics of glaciers*, 2nd edition, Pergamon press, 1981

Peterson, T. C., and R. S. Vose, An overview of the Global Historical Climatology Network temperature data base, *Bulletin of the American Meteorological Society*, 78, 2837-2849, 1997

Pfister, C., Monthly temperature and precipitation in central Europe from 1525-1979: quantifying documentary evidence on weather and its effects, in *Climate Since A. D. 1500*, edited by R. S. Bradley and P. D. Jones, pp. 118-142, Routledge, 1992

Roeckner, E., L. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dumenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida, The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate, *Max-Planck-Institut für Meteorologie, Report No. 218*, Hamburg, 1996

Roeckner, E., Climate sensitivity experiments with the MPI/ECHAM4 model coupled to a slab ocean, in *Euroclivar Workshop on Cloud Feedbacks and Climate Change*, edited by C. A. Senior and J. F. B. Mitchell, Hadley Centre for Climate Prediction and Research, Bracknell, England, 1997

Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld, and H. Rodhe, Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle, *Max-Planck-Institut für Meteorologie, Report No. 266*, Hamburg, (Journal of Climate, subm.), 1998

Stendel, M., and K. Arpe, Evaluation of the Hydrological Cycle in Re-Analysis and Observations, ECMWF Re-Analysis Validation Reports - Part 1, *ECMWF Re-Analysis Project Report Series, 6*, 1997

Thompson, L. G., Ice core evidence from Peru and China, in *Climate Since A. D. 1500*, edited by R. S. Bradley, and P. D. Jones, pp. 517-548, Routledge, 1992

von Storch, H., and F. W. Zwiers (Eds.), *Statistical Analysis in Climate Research*,
Cambridge University Press, 1998