

Statistical downscaling in European mountains: verification of reconstructed air temperature

Helen Kettle*, Roy Thompson

School of GeoSciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK

ABSTRACT: General circulation models constrained by past meteorological records have been run as part of a major community effort, using datasets from 1958 to the present date, to produce consistent gridded atmospheric databases known as reanalysis data. We derive linear regression models to transfer reanalysis gridded data to high elevation weather stations in Europe using daily observations from 1994 to 2001. The models are then used to reconstruct daily mean, minimum and maximum air temperatures since 1958 at the weather stations. The regression models use principal components of reanalysis temperature and pressure variables in addition to local (nearest grid point) temperature and pressure variables. An all-subsets regression technique in conjunction with cross validation is used to find the best model. The accuracy of the approach is verified using observed monthly data from 1980 to 1990 at 29 stations, and monthly data since 1958 at 8 stations. The verification results indicate that retrodiction to 1980 is good at all stations. However, validation at 3 alpine stations shows large differences between observed and reconstructed temperatures prior to 1970. Nevertheless, the basic spatio-temporal warming pattern we reconstruct for the European mountains has many similarities to that for the European 'lowlands'. We find regional climatic trends for the period 1958 to 2001 of typically 0.7°C per 100 yr for minimum temperatures and twice that for maximum temperatures. These trends are probably underestimated. Our reconstructions suggest that there has been an increase in the diurnal temperature range in European mountains in addition to the overall warming.

KEY WORDS: Downscaling · Climate · Mountains · 'Reanalysis data'

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

Mountain environments support a substantial proportion of the human population, contain a wide range of ecological variability and are the source of much of the world's water systems. Climatic change at high elevation sites can be dramatically amplified by feedback effects associated with the albedos of snow and ice. Consequently there is a pressing need to understand the impacts of interannual variability and extreme climate conditions on our mountain systems. To examine climatic change over the last century, long time series of meteorological observations are required. These records must be homogeneous such that variations are purely a result of the weather and climate and not caused by factors such as changes in instrumenta-

tion, instrument drift, urban warming or relocation of station. In many parts of the world, such high quality records are sparsely distributed. In recent years general circulation models (GCMs), constrained by historical meteorological records, have been run to produce consistent simulations of the state of the atmosphere over the past 4 decades. These assimilations are known as reanalysis models. One very useful aspect of the reanalysis modelling work is that it has the ability to transport information from data-rich to data-sparse regions. However, while GCMs can simulate large-scale upper air circulation fairly accurately, they are poor at reproducing surface variables on regional and local scales (Grotch & MacCracken 1991). Therefore, to examine climate at a sub-grid scale it is necessary to relate the gridded reanalysis data to an observed

*Email: h.kettle@ed.ac.uk

meteorological variable at a specific location, a process known as downscaling. Downscaling techniques can be summarised into 4 categories: regression methods; weather pattern-based methods; stochastic weather generators; and limited area modelling (Wilby & Wigley 1997). In this work we use regression methods to downscale air temperature. This involves the construction of a linear regression model relating reanalysis variables to observed surface air temperatures. Reconstruction of an observed variable outside its observation period relies on the assumption that the relationship between the large-scale circulation and the local climate does not change over time. As with all regression models, application of the model is limited by the range of data used in its construction. If observed data covering the full time period of the reanalysis data were used to construct the model, then it would be a robust (because it is built over many decades), but essentially redundant model as it is not providing any new information. To make real use of the reanalysis data we need to know whether models based on relatively short-term surface records can be used to reconstruct surface temperatures over previous decades.

Correlations between air temperatures recorded at several mountain weather stations in a given region are higher than correlations between air temperatures recorded at several low-lying stations in the same region (Weber et al. 1997), implying that mountain stations are less subject to local effects. Meteorological variables recorded at mountain stations present an ideal opportunity for examining climatic change as they are generally far from large cities and free from the warming associated with urbanisation. There is also evidence that the amplitude of temperature change this century at many high elevation sites is greater than the observed global change, implying that impacts of future climatic change will be greater at high elevations (Beniston et al. 1997). By reconstructing daily data at these locations it is possible to study changes in climate trends, variability and extremes.

The response of the local climate to the large-scale climate at a particular weather station can differ depending on the siting of the station. For example, if the station is on the side of a hill, the orientation of the slope may affect cloud cover and snow lie; stations set at the bottom of high valleys may experience temperature inversions; glaciers generate pronounced local effects and conditions at stations on high plateaus will differ from those on mountain tops. Small-scale influences such as local convective activity, orography, vegetation and soil characteristics can influence local wind systems, snow cover or energy exchange with the atmosphere, thus uncoupling the variability of individual weather elements from the large-scale circulation.

However, regional topography and land use at each station should be implicitly incorporated in the individual linear regression models used for downscaling. Due to the lack of human interference at mountain stations such factors should not have changed over the last 50 yr.

The aim of this study is to answer the following questions regarding downscaling in European mountains:

- (1) Is there a difference in the accuracy of downscaled minimum, maximum and mean daily air temperatures?
- (2) Is there a seasonal difference in downscaling accuracy?
- (3) Are temporal and spatial structures in the temperature series correctly downscaled?
- (4) Which reanalysis variables are the most important predictors?
- (5) Can we accurately reconstruct air temperatures between 1980 and 1990?
- (6) Can we accurately reconstruct air temperatures back to 1958?
- (7) How accurate are the long-term climatic trends derived from reconstructed temperatures?
- (8) How do the reconstructed trends compare with established climate trends?

2. DATASETS

In this work we focus on European mountain observatories where daily air temperatures since 1994 (which are used to build the regression models) and mean monthly temperatures from 1980 to 1990 (which are used for verification of the models) are available. The daily data provide daily minimum, maximum and mean temperatures. All of the stations used are well-maintained, quality-controlled World Meteorological Organisation (WMO) sites situated 1000 m above sea level (a.s.l.) in central and southern Europe, and above 700 m in Scandinavia. The height constraint is lower in Scandinavia because at this more northerly latitude, surface temperatures are lower so that conditions at these lower elevations are similar to those in the high Alps. To help ensure the regression models are fairly robust and not overfitted, the additional constraint that at least 5 yr of daily data must be available since 1994 was imposed. Details and locations of the 29 stations that satisfied our constraints are given in Table 1 and Fig. 1. For 8 of these stations, namely Sonnblick (Austria), Jungfrau and Säntis (Switzerland), Lomnický štít (Slovakia), Fokstua II (Norway), Mount Aigoual (France), Churanov (Czech Republic) and Navacerrada Pass (Spain) monthly data extend back several decades. The reanalysis data were obtained from the large assimilation datasets produced by the joint pro-

Table 1. Details of daily data (1994 to 2001) used in this study. **Bold** values indicate stations for which there is also long-term (from 1958) monthly data. m a.s.l.: metres above sea level

ID	Name	Country	Lat. (° N)	Long. (° E)	Altitude (m a.s.l.)	
1	012380	Fokstua II	Norway	62.15	9.28	0974
2	028010	Kilpisjarvi	Finland	69.05	20.78	0476
3	066800	Säntis	Switzerland	47.25	9.35	2500
4	067300	Jungfrau	Switzerland	46.55	7.98	3576
5	067500	Guetsch	Switzerland	46.65	8.62	2284
6	067530	Piotta	Switzerland	46.52	8.68	1016
7	067590	Cimetta	Switzerland	46.20	8.80	1648
8	067820	Disentis	Switzerland	46.70	8.85	1180
9	067910	Corvatsch	Switzerland	46.42	9.82	3299
10	067920	Samedan airport	Switzerland	46.53	9.89	1706
11	075600	Mount Aigoual	France	44.12	3.58	1565
12	082150	Navacerrada Pass	Spain	40.78	-4.02	1888
13	111460	Sonnblick	Austria	47.05	12.95	3107
14	111550	Feuerkogel	Austria	47.82	13.73	1621
15	112120	Villacheralpe	Austria	46.60	13.67	2160
16	112140	Preitenegg	Austria	46.93	14.92	1055
17	114570	Churanov	Czech Rep.	49.07	13.62	1126
18	117870	Lysa Hora	Czech Rep.	49.55	18.45	1327
19	119300	Lomnický Stit	Slovakia	49.20	20.22	2635
20	125100	Sniezka	Poland	50.73	15.73	1613
21	150520	Rarau	Romania	47.45	25.57	1541
22	151080	Ceahlau Toaca	Romania	46.93	25.92	1898
23	152800	Omu	Romania	45.45	25.45	2509
24	153020	Predeal	Romania	45.50	25.58	1093
25	160080	S. Valentino alla	Italy	45.75	10.53	1461
26	160210	Rolle Pass	Italy	46.30	11.78	2006
27	160220	Paganella	Italy	46.15	11.03	2129
28	161240	Cisa Pass	Italy	44.43	9.93	1040
29	161340	Cimone	Italy	44.20	10.70	2173

almost all of the local variance can be attributed to large-scale effects (Wigley et al. 1990), it is important that a circulation variable is included as a predictor in the regression models. Huth (2002) suggests the best downscaling models are constructed using 1 temperature-based and 1 circulation-based predictor. In this work, mean daily temperatures at 850 mb (T850) and mean daily sea level pressure (SLP) were extracted from a grid extending over Europe and the easternmost Atlantic Ocean to capture the large-scale circulations simulated by the GCM. The reanalysis data grid has 2.5° squares but as there is much spatial homogeneity we extract T850 and SLP from a sparser grid (5° lat. × 10° long.) extending from 30° to 80° N and from 30° W to 40° E. This provides 88 time series for both variables. This dimensionality is reduced and the linear dependency removed using principal component analysis (PCA) of the T850 and SLP fields for the summer and long winter separately. In summer the first 14 PCs of T850 (which explain 81% of the variance) and the first 9 of SLP (78% of the variance) are taken as potential predictors. In winter, the first

ject (Kalnay 1996) between the National Center for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR).

3. METHODS

3.1. Model construction

Long distance interactions between different parts of the atmosphere (e.g. high latitudes and low latitudes), known as teleconnection patterns, can vary with season (Huth 1997). Since the regression models incorporate information over a wide area, these patterns are important. To allow the model to cope with seasonal changes in teleconnections, we split the modelling data into summer (JJA) and a 'long winter' (the rest of the year). As regional climates are forced by global circulation, such that in some locations

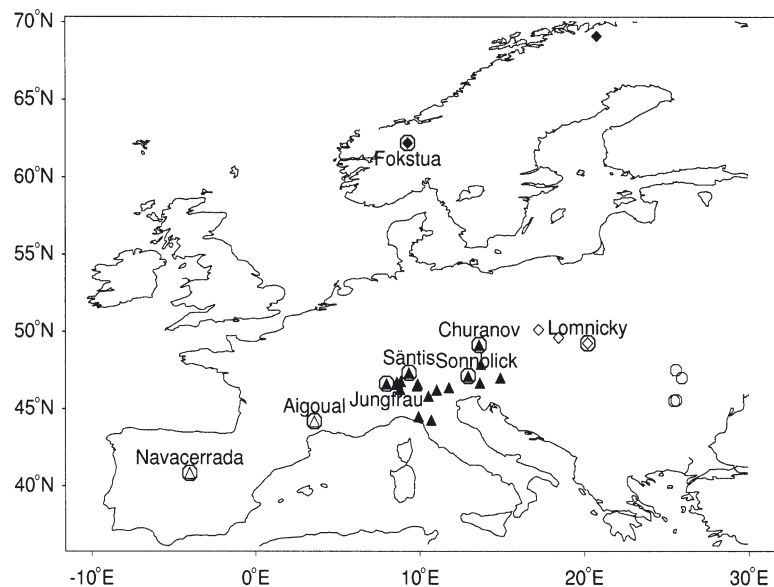


Fig. 1. Location of WMO stations. Stations are split into 5 regions: western Europe (Δ), central Alps (\blacktriangle), northeast Europe (\diamond), eastern Europe (\circ) and Scandinavia (\blacklozenge). See Table 1 for station details. Named stations with symbols surrounded by a circle mark those with long-term climate data

13 (89% of the variance) of T850 and first 10 (88% of the variance) of SLP are retained. Although it is desirable to retain more principal components, practical restraints on computing time limit the number of potential predictors. In addition to these circulation variables, at each surface site considered, we also use the following data from the nearest reanalysis grid point: air temperatures at pressure levels 200, 500 and 850 mb (T200, T500 and T850, respectively); minimum, maximum and mean surface temperatures and SLP. These combined with the PCA scores give a total of 30 potential predictors.

The downscaling models were derived by following a procedure that can be broken down into 6 steps:

Step 1. All time series were split into summer and long winter components. Steps 2 to 5 were then applied to data from the 2 seasons separately.

Step 2. All time series were normalised by subtracting the mean and dividing by the standard deviation following Huth (1999).

Step 3. A process known as ‘leaps and bounds’ or ‘all subsets regression’ was used to choose the best downscaling model in terms of cross validation root mean squared error (RMSE). For m predictors this method fits all 2^m possible subsets of predictors to the predicand. The fit criterion is Mallows’s C_p :

$$C_p = \frac{\text{RSS}}{\text{MSE}} - (n - 2p) \quad (1)$$

where n is number of observations, RSS is the sum of squared errors, MSE is the residual mean square error and p is the number of model parameters. A model that fits well will have a computed C_p value close to p . If there are several, the model with the smallest value of C_p is chosen. For each size of subset, the best subset of predictors (in terms of C_p) was retained leaving 30 possible subsets (models) to choose between.

Step 4. The best subset size was found using cross validation. Data from 1 yr are set aside from the model building process. The model is constructed using only the remaining data and then used to predict the year of data that had been held aside. This validation process is repeated for every year for which there are data. Thus for 8 yr of data there are 8 cross validation models built per subset size. The minimum of the average RMSE values over all of the validation periods is used to choose the best subset size.

Step 5. The final fit is produced using all of the data (no data set aside) with the number of predictors defined by the best subset size in the cross validation of Step 4. The best model (in terms of C_p) for this number of predictors was derived and used to reconstruct daily temperatures since 1958.

Step 6. The 2 reconstructions of summers/long winters were then combined to provide daily temperatures from 1 January 1958 to 31 December 2001.

This whole procedure (Steps 1 through 6) was repeated separately for the daily minimum, maximum and mean air temperature at each of the 29 stations.

Although the yearly cycle is not explicitly removed from the data, it is captured in the first principal components of the temperature and pressure datasets. It is thus handled naturally as an integral part of our modelling procedure.

3.2. Cross validation statistics

Average cross validation statistics were calculated for the best summer and winter models so that there were 2 RMSE, mean absolute error (MAE), bias and skill values for each station. The MAE shows the average error in the prediction without the disproportionate weighting the RMSE gives to occasional very large errors. The bias, which is the mean prediction error, is positive if the model is systematically overestimating air temperatures (and vice versa). The forecast skill is defined as:

$$\text{skill} = 100\% \times \left(1.0 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - \bar{x}_c)^2} \right) \quad (2)$$

by Lorenz (1956) in which x_i and \hat{x}_i are the actual and estimated air temperatures in the verification season for each data point i , and \bar{x}_c is the mean of the actual air temperatures in the calibration seasons. The closer the skill is to 100% the better the prediction.

4. EVALUATION OF MODELS

4.1. Predictors

At each station we have 30 variables (predictors) derived from the reanalysis assimilation datasets that can be used for downscaling. Some of these predictors will be used more frequently than others. To establish which are the important predictors for downscaling in European mountains, the number of stations using each predictor was calculated. The results, in Table 2, show that when downscaling to surface mean, maximum and minimum daily air temperatures the T850 at the local reanalysis grid point is the most used variable with the 1 exception of the minimum daily temperature in winter, which uses the local surface minimum temperature. In general, the summer models use local grid point data rather than the principal components of T850 and SLP. In contrast the winter models always select at least one of the principal component

Table 2. Frequency of use of predictors in daily air temperature regression models in order of usage with 1st indicating the predictor used by the largest number of stations, 2nd by the next largest number and so on. T850 and T500: mean daily temperatures at 850 and 500 mb, respectively; SLP: mean daily sea level pressure; PC: principal component

Variable	1st	2nd	3rd	Average no. of predictors
Mean (summer)	Local T850	Local SLP	Local T500	12
Mean (long winter)	Local T850	PC1 of T850	PC6 of SLP	24
Max. (summer)	Local T850	Local surface min.	Local T500	13
Max. (long winter)	Local T850	PC3 of SLP	Local surface min.	22
Min. (summer)	Local T850	Local surface min.	Local T500	12
Min. (long winter)	Local surface min.	T850	PC3 of SLP	22

Table 3. Cross validation statistics averaged over all the stations listed in Table 1. Skill: see Section 3.2 Eq. (2); RMSE: root mean squared error; MAE: mean absolute error

	Winter model			Summer model			Both models		
	Skill	RMSE	MAE	Skill	RMSE	MAE	Skill	RMSE	MAE
Mean	94.96	1.56	1.21	97.51	1.24	0.96	95.60	1.48	1.15
Min.	92.47	2.00	1.55	94.33	1.49	1.13	92.93	1.87	1.44
Max.	92.62	2.12	1.62	97.03	1.86	1.44	93.72	2.05	1.58

scores to be included in the top 3 predictors. This may reflect the increased strength of teleconnections in winter. The summer models typically require only half the number of predictors used by the winter models, which could be a function of the relatively small number of data points available in summer (3 mo compared to the 9 mo used in winter). Since all the models are thoroughly cross validated, there is no gross over-fitting even when a large number of predictors are retained.

4.2. Cross validation statistics

As described in Section 3.1, part of the model building process at each station involves cross validation. This allows the predictive accuracy of the model to be assessed without reducing the amount of data available for deriving the final model. The error statistics for the omitted seasons are averaged separately, providing summer and winter cross validation statistics at each station. Table 3 summarises these results for daily mean, minimum and maximum temperatures averaged over all of the stations. The cross validation statistics quoted for both models are averaged over the summer and winter models with a time weighting of 3/12 for summer and 9/12 for winter. The MAE and RMSE values indicate that the summer mean daily temperature and the winter daily maximum temperature are the most and least accurately downscaled vari-

ables respectively. In general, the skill of the downscaling appears much better in summer than winter. This is possibly because there is much greater natural variability in the winter months. For example, average standard deviations at Jungfrau mountain since 1994 are 3.1°C in summer and 4.4°C in winter. Because the cross validation skill takes into account the deviation of the observed data from its mean in the calibration period, it is possible for a high downscaling skill to be associated with a high RMSE (e.g. maximum temperatures in the summer model, Table 3). In our work, the cross validation skill cannot be used to compare between the summer and long winter models as the annual cycle in the data distorts the relative values. The skill is only useful for comparing between stations, but the RMSE can be used to compare between seasons.

4.3. Temporal and spatial structure

The accuracy of the temporal structure of the reconstructions is found by comparing the persistence (lag-1 autocorrelation) of the reconstructions with that of the observed time series. The persistence of the reconstructions is calculated over the time period of the surface observations. The results for the mean, maximum and minimum daily temperatures are shown in Fig. 2. The mean daily temperature reconstructions have a temporal structure closer to the observed series (bias

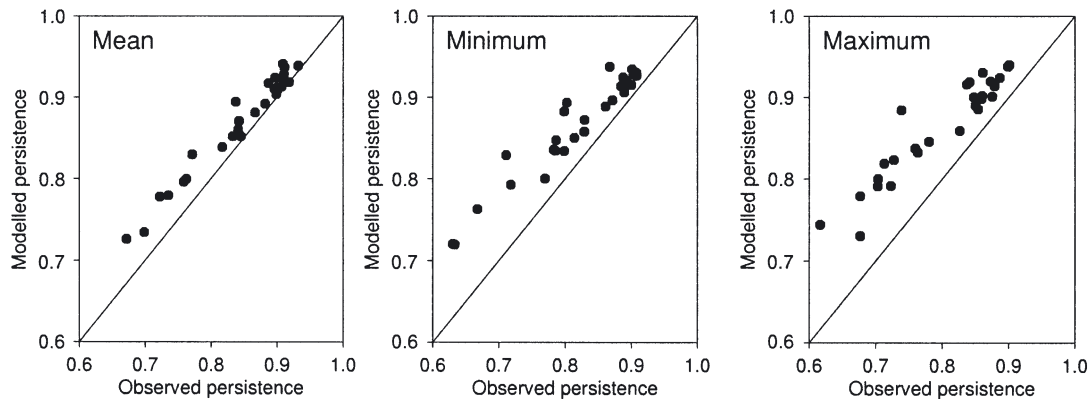


Fig. 2. Comparison of persistence (lag-1 autocorrelations) at all stations for mean, minimum and maximum daily air temperature. Straight lines: equal persistence in observed and reconstructed time series

0.025°C) than either the minimum or maximum temperature series (biases of 0.048 and 0.065°C, respectively). In all cases, the persistence of the modelled series is higher than that of the observed. This is unsurprising since the reconstructions are produced by linear regression models that have the effect of damping any extreme values (caused by local effects) and so increasing the autocorrelation. The mean daily temperature is less affected than the minimum or maximum since it has already been damped to some extent in the averaging process.

The accuracy of the spatial structure of the reconstructions is found by examining the inter-station correlations. This can be performed for all 29 stations for mean, minimum and maximum temperatures. The results are so similar that correlations with the mean temperatures at Lomnický Stit (identification no. 119300) serve as an indicative example. The contour plots in Fig. 3 show the correlations of mean air temperatures at each station with those at Lomnický Stit.

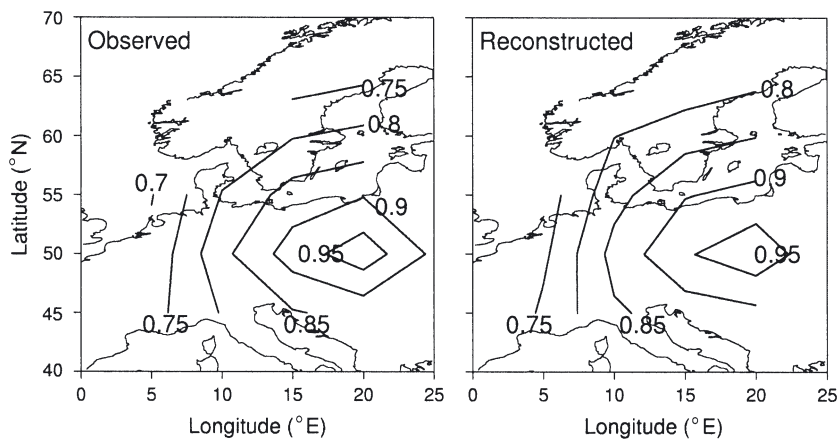


Fig. 3. Spatial correlations using the station Lomnický Stit (119300) for observations and reconstructions of daily mean temperatures

The form of the basic spatial pattern is the same for observed and reconstructed data, showing an east-west elongated maximum. However, the spatial correlations between the reconstructed time series are significantly overestimated so that the reconstructed series show higher correlations. For example the correlation contour that runs through the northern Gulf of Bothnia is 0.8 in our reconstructed series but only 0.75 for the original observations.

These results show that while the downscaling reconstructions may have small daily errors they have significant errors in persistence and spatial structure. This must be taken into account if these series are used to determine local climate regimes.

5. VERIFICATION OF MODELS

Although all our models are cross validated so that reconstructions with low predictive skill are treated with caution, there still remains the problem that for many stations daily data are only available from 1994. If the link between the large-scale circulation and the local climate has changed from 1958 to 2001, then our models will not necessarily be valid for this entire time period. Ideally, downscaling models should be built using several decades of data so that they cover most climatic situations (Zorita & von Storch 1999). To verify the accuracy of our technique the reconstructions have been compared with observed monthly data wherever they are available. The observed monthly mean data is given as the average of the mean minimum and mean maxi-

imum daily temperatures over the month rather than the average of the mean daily temperatures. This means that verification of reconstructions can only apply to the daily minimum and maximum temperatures and not to the daily mean as they are not used in the monthly mean calculation.

5.1. Verification (1980 to 1990)

Data from the beginning of 1980 to the end of 1990 at all 29 stations allow validation of our technique for reconstructing monthly means. For each station, the error statistics (RMSE, MAE and bias) are calculated for the summer and long winters separately. In addition to this, the skills of the reconstructions are calculated using Eq. (2) (in which the calibration mean is substituted with the value found using an estimation of the annual cycle) so that the skill represents the improvement the model affords with respect to the annual cycle. Unlike the daily cross validation skill, which includes the annual cycle, this criterion can be

used for intercomparison between the summer and long winter models. Error trends are also calculated using least squares linear regression. The results, in Table 4, show that our winter models have higher verification skills than our summer models, which is the opposite result to that found in the daily cross validation. The error trends, are either negligible compared with $0.01.5^{\circ}\text{C yr}^{-1}$ for the global average surface temperature rise since 1979 (IPCC 2001), or insignificant ($p > 0.05$) at all of these stations. There is a small lag-1 autocorrelation (~ 0.2) in the errors so that the significance of the trends may be slightly overestimated (von Storch & Navarra 1999). The autocorrelation in the errors implies a weakness in the model whereby local short-term weather events are not being reproduced by the downscaling.

During the 11 yr period 1980 to 1990 strongly increasing temperatures were observed (Fig 4). The trends of these surface observations range from 4°C per 100 yr (at Sniezka in Poland) to 39°C per 100 yr (at Villacheralpe in Austria) with an average of 23°C per 100 yr for the European mountains as a whole.

Table 4. Statistics of validation errors (reconstructed minus observed data) for monthly means. Trends significant at the 0.05 level are shown in **bold**. Skill: see Section 3.2 Eq. (2); MAE: mean absolute error; RMSE: root mean squared error

Stn name	Skill (%)		MAE ($^{\circ}\text{C}$)		RMSE ($^{\circ}\text{C}$)		Bias ($^{\circ}\text{C}$)		Trend in errors ($^{\circ}\text{C}$ per 100 yr)
	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer	
Fokstua II	85.9	80.9	0.69	0.47	0.85	0.59	-0.22	-0.27	-0.0007
Kilpisjarvi	75.9	82.8	1.11	0.57	1.45	0.67	0.66	0.38	-0.0027
Säntis	93.4	75.9	0.45	0.61	0.56	0.68	0.19	0.60	-0.0005
Jungfrau	88.8	68.5	0.51	0.51	0.67	0.61	0.14	0.49	0.0012
Guetsch	92.1	79.6	0.49	0.48	0.62	0.63	-0.02	0.07	-0.0014
Piotta	69.7	74.2	0.76	0.43	0.91	0.55	0.38	0.31	-0.0009
Cimetta	91.3	73.6	0.49	0.46	0.64	0.63	-0.04	-0.08	0.0008
Disentis	84.9	73.8	0.65	0.53	0.82	0.66	0.43	0.43	-0.0021
Corvatsch	87.3	85.4	0.55	0.40	0.71	0.49	-0.45	-0.27	0.0015
Samedan airport	59.8	53.7	1.02	0.68	1.20	0.78	0.38	0.57	-0.0025
Mount Aigoual	95.4	92.3	0.36	0.39	0.45	0.45	0.08	0.03	-0.0003
Navacerrada Pass	94.9	88.3	0.43	0.54	0.53	0.66	0.31	0.49	-0.0004
Sonnblick	93.7	79.7	0.44	0.44	0.54	0.54	0.04	0.41	-0.0005
Feuerkogel	96.1	88.0	0.36	0.36	0.44	0.44	0.08	0.22	0.0016
Villacheralpe	92.7	71.5	0.48	0.58	0.59	0.66	0.35	0.55	-0.0002
Preitenegg	82.7	73.2	0.70	0.49	0.87	0.57	0.52	0.45	-0.0013
Churanov	93.1	91.4	0.48	0.27	0.61	0.35	0.12	0.18	-0.0007
Lysa Hora	94.7	94.3	0.40	0.26	0.53	0.34	0.15	0.15	-0.0006
Lomnický Stit	93.8	76.1	0.40	0.53	0.53	0.62	0.11	0.50	-0.0001
Sniezka	95.8	95.1	0.38	0.24	0.48	0.29	0.11	0.06	0.0007
Rarau	91.2	70.0	0.53	0.74	0.67	0.80	-0.48	-0.74	-0.0001
Ceahlau Toaca	90.6	93.4	0.60	0.31	0.73	0.40	-0.38	-0.05	0.0001
Omu	87.7	83.7	0.60	0.45	0.78	0.56	-0.23	-0.39	0.0009
Predeal	93.3	87.7	0.42	0.42	0.53	0.49	-0.20	-0.36	-0.0007
S. Valention alla	78.3	54.3	0.71	0.61	0.91	0.73	-0.08	-0.45	-0.0021
Rolle Pass	89.6	70.9	0.53	0.53	0.66	0.65	0.24	0.45	-0.0014
Paganella	94.4	84.0	0.43	0.47	0.52	0.55	-0.04	0.35	-0.0011
Cisa Pass	71.3	75.2	0.74	0.48	0.92	0.63	0.59	0.32	0
Cimone	88.9	84.9	0.52	0.50	0.64	0.57	0.00	0.47	0
Mean	87.8	79.4	0.56	0.47	0.70	0.57	0.09	0.22	-0.0004

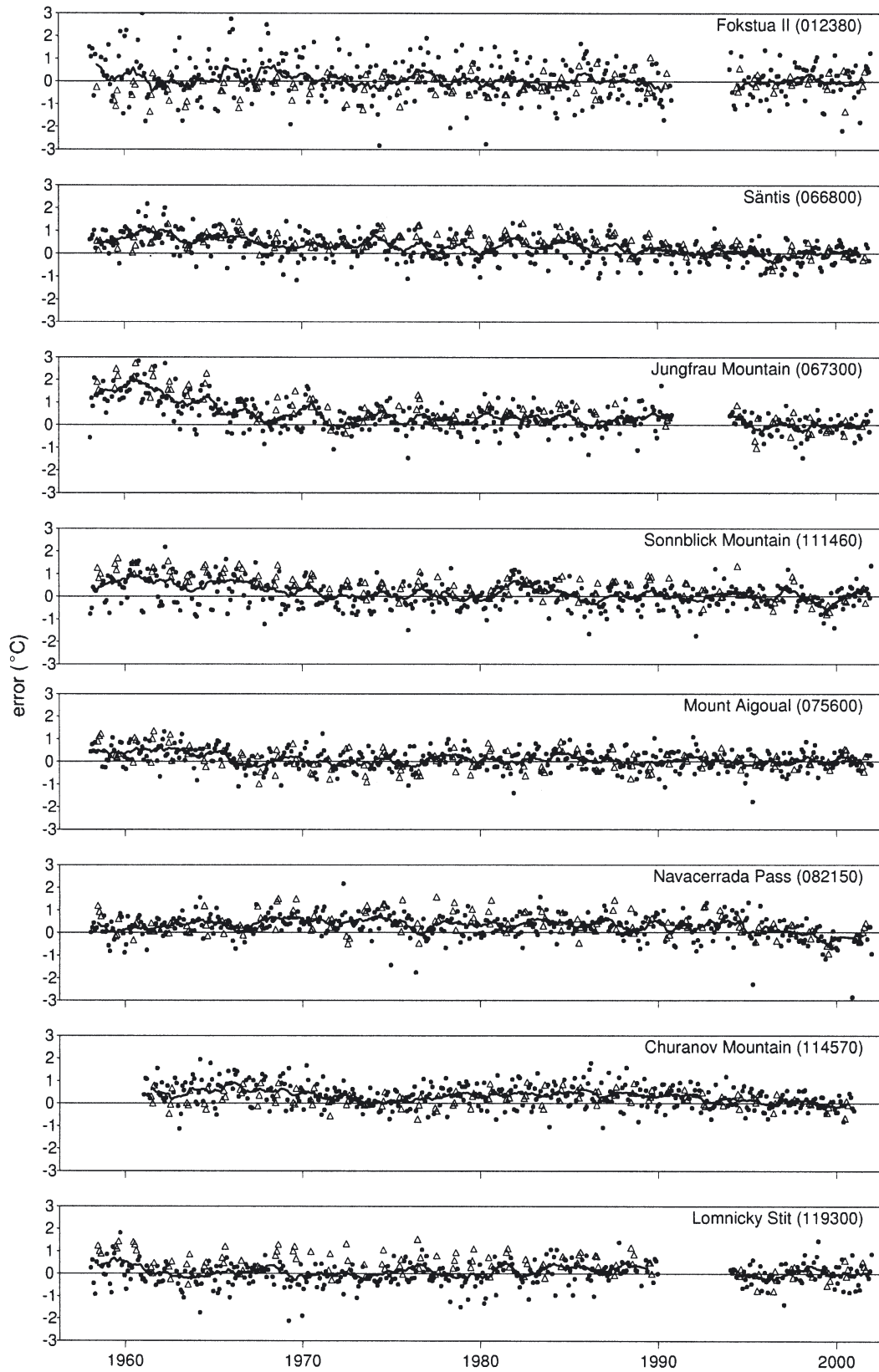


Fig. 5. Monthly mean validation errors (reconstructions minus observed data). (Δ) Summer months; (●) winter months; thick lines: yearly running mean

5.2.1. Poor choice of downscaling model

Reconstructions at Jungfrau mountain show the largest error trends. Therefore, to check that these error trends are not dependent on the form of the downscaling model, many regression models were examined. Rather than choosing the single best summer and winter models that gave the lowest cross validation error, 400 models were examined for Jungfrau mountain. These were random combinations of the 60 maximum temperature models (30 for summer and 30 for winter) and 60 minimum temperature models selected by the all-subsets regression. These models were then used to reconstruct mean monthly air temperatures since 1958 and compared with the observed data. The errors for all 400 models followed an almost identical pattern, showing large increases prior to 1970. The range of monthly error values for each model had a maximum value of 0.86°C and an average value of 0.38°C. These results serve to demonstrate that the downscaling is not dependent on the exact choice of predictors. We further note that even if the T850 and SLP predictors are erroneous the same trend in errors occurs when using just the surface minimum, maximum and mean temperatures from the nearest grid point. In summary, the errors shown in Fig. 5 are not simply caused by choosing a poor model.

5.2.2. Changes in the mode of circulation not captured by our 1994 to 2001 regression models

Alpine climate is strongly related to the North Atlantic Oscillation (NAO) (Wanner et al. 1997). Our model building period coincides with a period of high index NAO associated with zonal flow. However, the 1970s are a low index period, associated with meridional flow, so it seems feasible that changes in the NAO could have caused a change in the link between general circulation and local climate. On the other hand, the NAO is essentially a winter climate mode and errors in our reconstructions are, in most cases, higher in summer. To quantify the effect of the NAO on local climates the NAO index (normalised pressure difference between the Azores and Iceland) for winter (DJFM) is correlated with observed mean winter air temperatures at each station (Table 6). It is apparent that with the exception of Lomnicky Stit all the stations are highly correlated with the NAO. The next step is to establish whether the errors between our reconstructions and the observed data are correlated with the NAO. The results (Table 6) show that the largest correlations are with the errors at Lomnicky Stit and Mount Aigoual. Observed air temperatures at Lomnicky Stit are weakly affected by the NAO (correlation = 0.23) but the errors are

Table 6. Correlation coefficients of air temperatures and reconstruction errors with the North Atlantic Oscillation (NAO) for the winter months (DJFM). **Bold** values indicates correlations above 0.5

Stn name	NAO and observed air temperatures	NAO and errors
Fokstua II	0.65	-0.29
Säntis	0.49	-0.12
Jungfrau	0.53	+0.10
Mount Aigoual	0.59	-0.32
Navacerrada Pass	0.64	-0.15
Sonnblick	0.52	0.00
Churanov	0.66	-0.04
Lomnicky Stit	0.23	+0.43

strongly affected (correlation = 0.43), implying that our reconstructions at this station are too heavily influenced by the NAO. This is possibly due to the coarse grid in the original GCM simulation. However, at Mount Aigoual the air temperatures are highly correlated (0.59) and the errors also show an anti-correlation (-0.32), implying that when the NAO is in its negative phase our temperature reconstructions are too warm (and vice versa). To uncover any other links between our errors and large-scale circulation patterns, 12 other circulation indices, as defined by the Climate Prediction Center and available at www.cpc.ncep.noaa.gov/data/teledoc/telecontents.html, were correlated with the summer and winter retrodiction errors at each station. In winter all correlations were very low, the highest correlation out of over a hundred pairs examined being 0.34 (between Navacerrada and the Pacific Transition Pattern). This is surprising because the Pacific Transition Pattern is prominent between May and August, but it is aligned along the 40° N latitude circle on which Navacerrada lies. Correlations in summer were very low, with no correlations greater than 0.30. Overall, the occasional modest correlations between teleconnection patterns and the errors in our reconstructions are found. However, these results are unable to explain the large errors found in the central Alps and do not support the idea that a change in the influence of the large-scale circulation on the local climate has occurred.

5.2.3. Inaccuracies in the reanalysis output data pre-1970

Biases in the reanalysis data pre-1970 could be caused by paucity of observed data in this region to input to the reanalysis model during this time. Reanalysis assimilations are a composite of many different datasets such as land-based and ship-based measurements, upper air

data, satellite observations and numerical weather forecast output, but are obtained through the use of a consistent circulation model. Very different weightings can be given to these datasets. For example, surface measurements over land, with strong local biases, are often given little or no weight. Changes, through the reanalysis period, in the distribution, types and quality of the observations such as those of radiosonde and satellite can potentially lead to substantial inhomogeneities (Upala 1997). Trenberth et al. (2001) give an example for the tropics with jumps to warmer values below 500 mb in 1986 and 1989. Other analyses point to jumps around November 1978 between the pre- and post-satellite periods at the 100 mb level. Shifts in the reanalysis moisture fields (Trenberth & Guillemot 1998, Pocard et al. 2000) have similarly been noted in the mid-1970s and mid-1980s. Problems have also been recorded for surface pressure fields in the North Atlantic before 1968, where all observed pressure data below 1000 mb occurring during extra tropical cyclones were not input to the reanalysis model due to an error (Bob Kistler, <http://www.emc.ncep.noaa.gov/gmb/bkistler/psfc/psfc.html>). However, it is unlikely that these discrete events could cause the general error trend we observe for the Alpine stations.

5.2.4. Changes in measurement technique or changes to the type and location of the stations

Another cause of the large errors we find prior to 1970 could be changes in measurement technique or changes to the type and location of the mountain stations. According to Weber (1993), between 1978 and 1981 some of the Swiss stations were completely changed to automatic measurement systems with electronic thermometers and a nearly continuous recording of data every 10 min. In 1961 the station at Säntis changed from a small metallic shelter attached to the north wall of a building to a standard Stevenson screen (Weber 1993). The time of reading was also changed in Switzerland around 1970 (Weber 1993) with the big difference that the reading of the maximum temperature was changed from evening to morning. As described by Karl et al. (1986), this can have a large effect on daily minimum and maximum values. Thus, it is possible that a bias in the daily mean temperature (the arithmetic mean of the maximum and minimum) occurs around 1970 at the Swiss stations, which coincides with the errors found in our reconstructions. To investigate this source of error further we made 3 types of check. Firstly, we used Alexandersson's (1986) method to recheck the homogeneity of the long-term mountain station data. Secondly, we compared the trends in our reconstruction with those described by

Weber et al. (1997) and Beniston et al. (1994) for the Alps. Thirdly we compared our reconstructions with the observed monthly data from the gridded homogenised dataset, CRUTEM1 (Jones 1994).

Alexandersson's (1986) method reveals no major homogeneities, discrepancies, or sharp breaks in the records of the 3 Alpine stations of Jungfrau, Säntis or Sonnblick, when they are taken in comparison to one another and to a lowland reference series based on Basel, Geneva, Milan and Vienna. However, we note that the trends in mean temperature at Jungfrau (1958 to 1990) are the most extreme (most positive) of all 7 of these series, by a factor of almost 2. We find many similarities between our reconstructions and the trends reported by Weber et al. (1997) and Beniston et al. (1994). For example in the Alps the strongest warming trends in our reconstructions (Table 7) are for the winter season, especially in maximum temperatures. Weber et al. (1997, their Table IV) and Beniston et al. (1994, their Figs. 18 & 19) find exactly the same. Furthermore, Weber et al. (1997, their Table IV) report that the strongest cooling trends in the Alpine area took place in the eastern Alps, particularly in summer (JJA) and autumn (SON) months at mountain sites. We similarly reconstruct the most negative trends in the autumn at our mountain sites in the NE Alps (Table 7). Thus, even in the 1 region (Alps), and 1 time period (pre-1970), where the bias in our reconstructions is somewhat higher than might have been hoped for (see Säntis, Jungfrau and Sonnblick in Fig. 5), we nevertheless correctly deduce the main features not only of the inter-annual variability but also of the long-term trends. One difference between our mountain site reconstructions and the trends reported by other workers is for autumn temperatures in the central Alps. We deduce (Table 7) an average cooling trend of -1.33°C per 100 yr; however, in contrast Weber et al. (1997), Beniston et al. (1994) and CRUTEM1 all report a warming. Turning our attention more closely to the CRUTEM1 gridded homogenised dataset, the grid square which contains Jungfrau mountain does not contain any data but the adjacent square ($45-50^{\circ}\text{N} \times 10-15^{\circ}\text{E}$) contains nearly continuous monthly data over 1958 to 2001. The correlation between the grid square CRUTEM1 temperatures and our reconstructed temperatures at Jungfrau mountain is 0.65 (annual cycles removed), indicating that this comparison is valid. The difference between our reconstructions at Jungfrau mountain and the mean temperature in the grid square is shown in Fig. 6. The difference between these 2 time series remains approximately constant over the whole time period. The same result was found for both Sonnblick and Säntis. We note in Table 8 that the main negative temperature trends in CRUTEM1 (1958 to 2001) are for the autumn, in east Europe and

Table 7. Linear trends in reconstructed data (1958 to 2001). **Bold** values indicate trends significant at the 2-tailed 5% level. Annual trends are taken from yearly averages. Seasonal trends are taken from seasonal averages. Number of stations in each region is shown in parentheses. DTR: diurnal temperature range

	Summer (°C per 100 yr)	Winter (°C per 100 yr)	Autumn (°C per 100 yr)	Spring (°C per 100 yr)	Annual (°C per 100 yr)
Western Europe (2)					
Min.	3.06	2.79	0.89	2.11	2.16
Max.	4.42	4.09	1.69	3.24	3.34
DTR	1.37	1.31	0.80	1.13	1.18
Central Alps (18)					
Min.	0.95	2.95	-1.57	0.57	0.51
Max.	2.36	4.18	-1.10	1.41	1.57
DTR	1.41	1.23	0.47	0.84	1.06
Eastern Europe (4)					
Min.	1.40	2.03	-0.26	0.25	0.32
Max.	2.31	2.95	-1.11	0.87	0.93
DTR	0.90	0.91	0.15	0.62	0.61
NE Alps (3)					
Min.	1.7	4.12	-1.48	1.64	1.34
Max.	2.14	4.24	-1.58	2.29	1.63
DTR	0.44	0.14	-0.09	0.65	0.29
Scandinavia (2)					
Min.	0.50	4.26	-1.16	-0.67	0.83
Max.	0.98	3.18	-1.18	-1.10	0.63
DTR	0.48	-1.07	-0.03	-0.42	-0.20
All regions (29)					
Min.	1.18	3.02	-1.32	0.66	0.73
Max.	2.42	3.93	-0.96	1.39	1.56
DTR	1.24	0.91	0.36	0.73	0.83

Table 8. Linear trends in lowland mean monthly temperatures from CRUTEM1 (1958 to 2001). **Bold** values indicate trends significant at the 2-tailed 5% level. Annual trends are taken from yearly averages. Seasonal trends are taken from seasonal averages

	Summer (°C per 100 yr)	Winter (°C per 100 yr)	Autumn (°C per 100 yr)	Spring (°C per 100 yr)	Annual (°C per 100 yr)
Western Europe	4.24	1.68	1.52	4.28	2.93
Central Alps	2.56	1.25	0.43	1.56	1.45
Eastern Europe	2.35	1.11	-1.04	1.19	0.90
NE Alps	1.88	2.80	-0.69	1.66	1.41
Scandinavia	-0.26	4.25	0.17	0.77	1.23
All regions	2.10	2.27	-0.12	1.80	1.51

NE Alps. Otherwise, CRUTEM1 trends are positive, or negligible. Our one discrepancy with CRUTEM1 (1958 to 2001), and with Weber et al. (1997) and Beniston et al. (1994), is for the central Alps in autumn, where we incorrectly reconstruct a cooling that mainly occurred further to the east.

In summary, verification of our downscaling models is generally very good. However, high bias is found pre-1970, especially at Jungfrau mountain and to a lesser extent Sonnblick and Säntis. These discrepancies can be mainly explained by difficulties with the early observations at Jungfrau, and to a certain extent

by problems with our pre-1970 autumn downscaling in the central Alps.

6. LONG-TERM TRENDS ACROSS EUROPE

One of the main reasons to create long time series temperature data is to examine trends in regional climates. Table 7 summarises our reconstructed European trends in terms of 5 mountain regions. All 5 display rising temperatures. Rates of up to 4.42°C per 100 yr for maximum temperatures are found in the

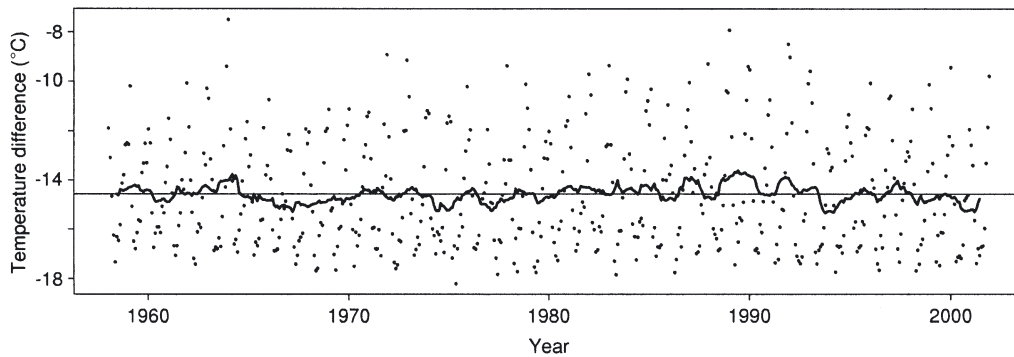


Fig. 6. Difference between monthly means of reconstructed temperatures at Jungfrau mountain and the average temperatures in grid square 45–50° N × 10–15° E in the CRUTEM1 dataset. (●) Monthly values; thick line: yearly running mean; thin (horizontal) line: mean difference between the 2 time series

westernmost region. In order to further check the validity of our reconstructions over this 44 yr time span we can compare the reconstructed trends with the observations at the 8 stations where long records exist. As noted above, a change occurs around 1970 at 3 alpine stations and so a comparison of the observed and reconstructed trends for post 1970 is also useful. The 1958 to 2001 and 1970 to 2001 trends are compared in Fig. 7. In general, the results indicate that the reconstructed trends are underestimated. The most extreme underestimation is Jungfrau mountain, which shows a reconstructed trend of only 1.8°C per 100 yr as opposed to the observed trend of 4.9°C per 100 yr between 1958 and 2001. The observed trend at Jungfrau mountain is higher than those observed at the other 7 long-term stations over the period 1958 to 2001 (Fig. 7). The differ-

ence in observed trends could imply that the observed trend is incorrect and that the errors found before 1970 are due to inhomogeneities in the observed data. Figs. 4 & 7 show that the trend in mean monthly air temperatures has increased with extremely high trends from 1980 to 1990. In this decade the NAO was predominantly in its high index phase, which is related to lower temperatures over the Atlantic Ocean and higher temperatures over the European continent (Hurrell 1995).

6.1. Comparison of reconstructed trends with established climatic trends in mountains

Our temperature reconstructions can be compared with published trends for the globe, for other mountain regions and for the European lowlands. Global climate trends are characterised by a faster rate of increase in minimum temperatures than maximum temperatures, leading to a decrease in the diurnal temperature range (DTR) (Karl et al. 1993, Easterling et al. 1997). This is equivalent to a day-night asymmetry since the daily minimum temperature usually occurs at night and the maximum during the day. It has been suggested that the cause of this asymmetry could be an increase in cloud cover (Karl et al. 1993). However, in some areas of the world the pattern has been different. According to Weber et al. (1994, 1997), one such area is the mountain region of central Europe where low-lying stations show the same pattern as the global trend but mountain top stations show maximum temperatures have increased at similar rates to minimum temperatures. Increased cloud cover concentrated at lower elevations could cause such an effect. When our downscaled reconstructions at all 29 stations are averaged to give an overall temperature trend for the European mountains in the period 1958 to 2001 (Fig. 8), the daily maximum temperatures appear to have risen faster than

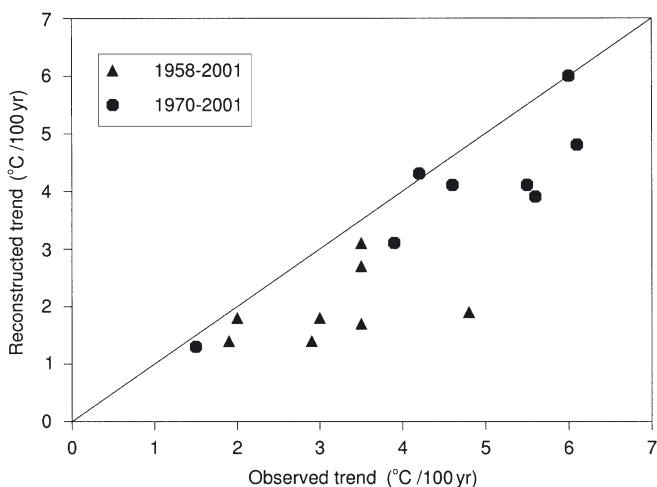


Fig. 7. Comparison of trends in monthly mean data for observed and reconstructed time series from 1958 to 2001 and 1970 to 2001. Straight line: perfect correspondence between observed and reconstructed

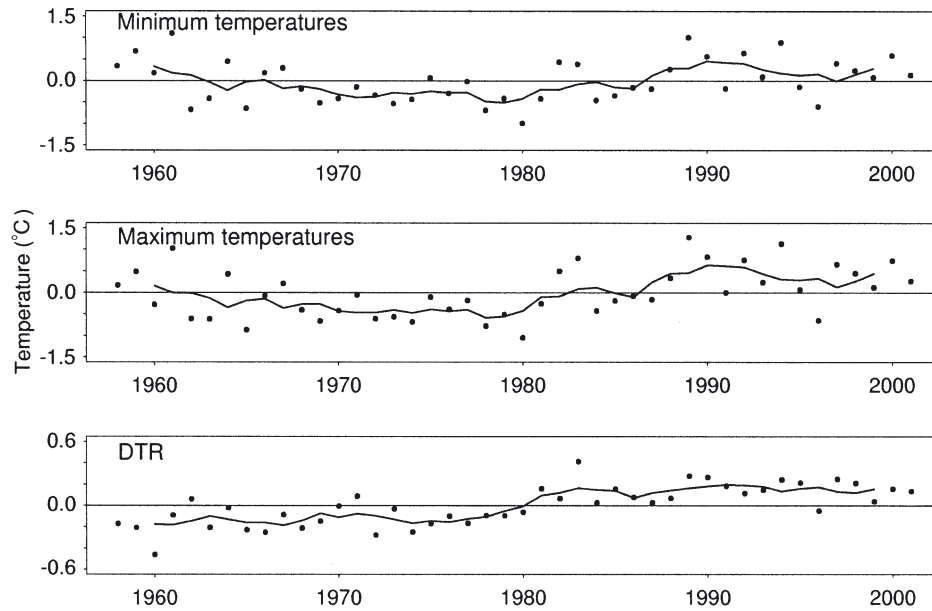


Fig. 8. Mean yearly maximum, minimum temperatures and diurnal temperature range (DTR) of reconstructions averaged over all 29 stations. Data are shown as anomalies from the mean value over the 1958 to 2001 period (thin horizontal lines). (●) Yearly anomalies; thick lines: 5 yr running mean

the daily minimum temperatures. Thus, there has been an overall increase in the DTR in European mountains, the opposite of the global trend.

As described in the verification work of Section 5.2.4, the main trends, with the exception of Jungfrau, that we reconstruct are generally in good agreement with those of other workers (e.g. Beniston et al. 1994, Weber et al. 1997). However, to examine the trends more thoroughly our reconstructions have been split into 4 seasons (DJF, MAM, JJA, SON) and 5 regions (western Europe, central Alps, eastern Europe, northeast Alps and Scandinavia). The locations of these regions are shown in Fig. 1. The trends of minimum and maximum temperatures and DTR for each season (and annually) in each region are listed in Table 7. Trend calculations using a regression approach are based on the assumption of independence (non-autocorrelated errors). Here the data have a yearly separation but examination of the error structure shows some autocorrelation (typically only 0.25 or less). This means that while the trend calculations are correct the significance of the trends may be slightly overestimated (von Storch & Navarra 1999).

In all regions there are no statistically significant ($p < 0.05$) trends in autumn. Most of the significant trends occur in winter with larger trends in maximum than minimum temperatures everywhere except Scandinavia. In one sense, this pattern is contrary to Weber et al.'s (1994) analyses which indicated that most of the warming in the central Alps

had taken place in the autumn. In western Europe (Navacerrada and Aigoual) we reconstruct significant increases in both maximum and minimum temperatures in summer and winter. The maximum temperature increases at a faster rate leading to increases in the DTR. This is in contrast to the observations for Pic du Midi in the French Pyrenees (Dessens & Bücher 1995), which is situated in between our stations. At Pic du Midi the minimum temperature increased at a much faster rate than the maximum. In the northeast Alps the only significant trends of maximum and minimum reconstructed temperatures are in winter, when they are roughly equal and very pronounced. However, Weber et al. (1997) reported a strong significant increase of minimum daily temperatures in spring for this region, a time when we have no significant trends. In the central Alps the significant trends in reconstructed maximum temperatures are in summer and winter, leading to an increase in DTR. Weber et al. (1997) found maximum and minimum temperatures increased mainly in winter and annual trends of maximum and minimum were roughly equal with only small changes to DTR over the period 1951 to 1990. In eastern Europe our reconstructed trends show a significant increase in maximum temperatures in summer and significant increases in DTR in all seasons except autumn. Weber et al. (1997) report very similar patterns. Our reconstructions in Scandinavia show different warming trends to the rest of Europe with no significant

seasonal trends, agreeing with Diaz & Bradley (1997), who found Scandinavia does not show the recent warming seen elsewhere.

Overall, our results for high elevation sites suggest that the strongest warming has occurred in western Europe with rather weaker warming in eastern Europe and Scandinavia. All regions except Scandinavia show a faster increase in maximum temperature than in minimum temperature, occurring mainly in winter. Furthermore our reconstructions indicate that on average the DTR has increased. These results, although contradicting some of Weber et al.'s (1994, 1997) analysis display the general spatial pattern observed by Diaz & Bradley (1997) regarding large warming trends in western Europe and small warming trends in eastern Europe and Scandinavia.

6.2. Comparison with lowland trends

It is also possible to compare our reconstructed trends with trends in mean temperatures determined from predominantly lowland data. To do this we averaged mean monthly data from the relevant grid squares in the CRUTEM1 (Jones 1994) dataset. The comparisons (Table 8) show the CRUTEM1 trends have much in common with our reanalysis downscaling. Indeed, the broad spatio-temporal patterns of temperature trends across Europe are very comparable. Here we focus on the similarities. First, the annual trends are highest for western Europe, and lowest for eastern Europe and Scandinavia. Secondly, the seasonal trends also show many common features. Trends are strong for the summer months particularly in western Europe, but with both our reconstructed mountain trends and the 'lowland' trends showing little or no trend in Scandinavia. In winter, trends have in general been lower. Once again Scandinavia is the exception with both our upland reconstructions and the 'lowland' trends being higher than elsewhere in Europe. Autumn trends are low everywhere in both datasets. Finally, in spring, western Europe and to a lesser extent NE Europe have experienced the high trends. However, while the spatio-temporal patterns match well, the magnitudes of the trends tend to be lower in our reconstructions for mountain regions. We have an average annual temperature increase, over all regions, of only 1.1°C per 100 yr in comparison to 1.5°C per 100 yr in the 'lowlands'. In general, the magnitudes of the trends in *maximum* temperature in the mountains are comparable with magnitudes of the trends in *mean* temperature in the lowlands. Once again Scandinavia breaks the general rule set by the normal European situation.

7. CONCLUSIONS

In this work linear regression models built on short-term daily data (1994 to 2001) are used to reconstruct daily mean, minimum and maximum air temperatures back to 1958 at 29 mountain weather stations. This study attempts to analyse the accuracy of these down-scaled air temperatures by answering the 8 questions raised in the 'Introduction'. Below is a summary of the answers:

(1) *Is there a difference in the accuracy of down-scaled minimum, maximum and mean daily air temperatures?* Yes, mean air temperatures are the most accurately downscaled. Minimum daily temperatures are the least accurate.

(2) *Is there a seasonal difference in downscaling accuracy?* Yes. The skill is higher in winter. However, the natural variability is larger in winter so the absolute errors are generally higher than in summer.

(3) *Are temporal and spatial structures in the temperature series correctly downscaled?* No, the persistence (lag-1 correlations) is consistently overestimated, particularly for the daily minimum and maximum air temperature reconstructions. The spatial correlations are also overestimated.

(4) *Which reanalysis variables are the most important predictors?* The air temperature at the 850 mb pressure level at the reanalysis grid point nearest to the station of interest is the most used variable. Principal components of SLP and T850 are more often used in the winter rather than summer models. In general the summer models require about half the number of predictors used by the winter models.

(5) *Can we accurately reconstruct air temperatures between 1980 and 1990?* Over this period there is very little bias (0.09°C) in the winter reconstructions but summer temperatures are slightly over estimated (0.22°C). Individual monthly mean temperatures can be reproduced to within ~0.5°C, and there are no trends in the errors.

(6) *Can we accurately reconstruct air temperatures back to 1958?* In some locations the reconstructions back to 1958 are reasonable, but at 3 alpine stations the apparent errors increase prior to 1970. Whether this change is due to inaccurate downscaling, inaccuracies in the reanalysis assimilation or inhomogeneities in the observed climate data remains to be determined. Our validation and verification studies point to caution in extrapolating these models further back in time than 1970.

(7) *How accurate are the long-term climatic trends derived from reconstructed temperatures?* In general the long-term trends are underestimated in the reconstructed data. This underestimation is very pronounced for the time periods 1958 to 2001 and 1970 to 2001, and occurs to a lesser extent in 1980 to 1990.

(8) *How do the reconstructed trends compare with established climate trends?* Overall, the reconstructions indicated that maximum temperatures are rising faster than minimum temperatures, resulting in an increase in the DTR from 1958 to 2001. This is in disagreement with some published trends (e.g. Weber 1994, 1997) which show maximum and minimum temperatures to be increasing at approximately equal rates in the European mountains with insignificant changes in DTR. However, our reconstructed trends for maximum temperatures are very similar to those for mean lowland temperatures.

Acknowledgements. Funding was provided by the European Union Framework V project EMERGE (European Mountain lake Ecosystems: Regionalisation, diagnostics & socio-economic Evaluation; contract no. EVK1-CT-1999-00032). Processed homogeneous gridded climate data are from the Climatic Research Unit (CRU) with source data from the NCEP/NCAR Reanalysis Project. The observed surface data are from World Meteorological Organisation (WMO) observatories obtained from the National Climatic Data Center (NCDC). The Jones CRUTEM1 gridded data are from CRU. None of our work would have been possible without these valuable climate resources.

LITERATURE CITED

- Alexandersson H (1986) A homogeneity test applied to precipitation data. *J Clim* 6:661–675
- Beniston M, Rebetez M, Giorgi F, Marinucci MR (1994) An analysis of regional climate change in Switzerland. *Theor Appl Clim* 49:135–159
- Beniston M, Diaz HF, Bradley RS (1997) Climatic change at high elevation sites: an overview. *Clim Change* 36: 233–251
- Dessens J, Bücher A (1995) Changes in minimum and maximum temperatures at the Pic du Midi in relation with humidity and cloudiness. *Atmos Res* 37:147–162
- Diaz HF, Bradley RS (1997) Temperature variations during the last century at high elevation sites. *Clim Change* 36: 253–279
- Easterling DR, Horton B, Jones PD, Peterson TC and 7 others (1997) Maximum and minimum temperature trends for the globe. *Science* 277:364–367
- Grotch S, MacCracken M (1991) The use of general circulation models to predict regional climate change. *J Clim* 4: 284–303
- Hurrell JW (1995) Decadal trends in North Atlantic Oscillation: regional temperatures and precipitation. *Science* 269:676–679
- Huth R (1997) Potential of continental-scale circulation for the determination of local daily surface variables. *Theor Appl Clim* 56:165–186
- Huth R (1999) Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Clim Res* 13: 91–101
- Huth R (2002) Statistical downscaling of daily temperature in central Europe. *J Clim* 15:1731–1742
- IPCC (Intergovernmental Panel on Climate Change) (2001) *Climate change 2001: the scientific basis*. Cambridge University Press, Cambridge
- Jones PD (1994) Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *J Clim* 7: 1794–1802
- Kalnay E, Kanamitsu M, Kistler R, Collins W and 17 others (1996) The NCEP/NCAR 40-year Reanalysis Project. *Bull Am Meteorol Soc* 77:437–471
- Karl TR, Williams CN, Young PJ (1986) A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States. *J Clim Appl Meteorol* 25:145–160
- Karl TR, Jones PD, Knight RW, Kukla G and 6 others (1993) Asymmetric trends of daily maximum and minimum temperature. *Bull Am Meteorol Soc* 74:1007–1023
- Lorenz EN (1956) Empirical orthogonal functions and statistical weather prediction. *Statistical Forecasting Scientific Rep. 1*. Department of Meteorology, Massachusetts Institute of Technology, Cambridge, MA
- Poccard I, Janicot S, Camberlin P (2000) Comparison of rainfall structures between NCEP/NCAR reanalyses and observed data over tropical Africa. *Clim Dyn* 16:897–915
- Trenberth KE, Guillemot (1998) Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. *Clim Dyn* 14:213–231
- Trenberth KE, Stepaniak DP, Hurrell JW, Fiorino M (2001) Quality of reanalyses in the Tropics. *J Clim* 14:1499–1510
- Uppala S (1997) Observing system performance in ERA. ECMWF Reanalysis Project Rep. 3. The National Academic Press, Reading
- von Storch H, Navarra A (1999) *Analysis of climate variability*, 2nd edn. Springer-Verlag, Berlin
- Wanner H, Rickli R, Salvisberg E, Schmutz C, Schuepp M (1997) Global climate change and variability and its influence on alpine climate—concepts and observations. *Theor Appl Clim* 58:221–243
- Weber RO (1993) Influence of different daily mean formulas on monthly and annual averages of temperature. *Theor Appl Clim* 47:205–213
- Weber RO, Talkner P, Stefanicki G (1994) Asymmetric diurnal temperature change in the Alpine region. *Geophys Res Lett* 21:673–676
- Weber RO, Talkner P, Auer I, Bohm R, Gajic-capka M, Zaninovic K, Brazdil R, Fasko P (1997) 20th-century changes of temperature in the mountain regions of central Europe. *Clim Change* 36:327–344
- Wigley TML, Jones PD, Briffa KR, Smith G (1990) Obtaining sub-grid-scale information from coarse-resolution general circulation model output. *J Geophys Res* 95:1943–1953
- Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. *Prog Phys Geogr* 21:530–548
- Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J Clim* 12:2474–2489

Editorial responsibility: Clare Goodess, Norwich, UK

*Submitted: February 17, 2003; Accepted: March 10, 2004
Proofs received from author(s): April 26, 2004*