SUMMER TEMPERATURE VARIATIONS IN THE EUROPEAN ALPS, A.D. 755–2004

ULF BÜNTGEN, DAVID C. FRANK, DANIEL NIEVERGELT, AND JAN ESPER

Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

(Manuscript received 18 August 2005, in final form 2 March 2006)

ABSTRACT

Annually resolved summer temperatures for the European Alps are described. The reconstruction covers the A.D. 755–2004 period and is based on 180 recent and historic larch [Larix decidua Mill.] density series. The regional curve standardization method was applied to preserve interannual to multicentennial variations in this high-elevation proxy dataset. Instrumental measurements from high- (low-) elevation grid boxes back to 1818 (1760) reveal strongest growth response to current-year June–September mean temperatures. The reconstruction correlates at 0.7 with high-elevation temperatures back to 1818, with a greater signal in the higher-frequency domain (r = 0.8). Low-elevation instrumental data back to 1760 agree with the reconstruction’s interannual variation, although a decoupling between (warmer) instrumental and (cooler) proxy data before ~1840 is noted. This offset is larger than during any period of overlap with more recent high-elevation instrumental data, even though the proxy time series always contains some unexplained variance. The reconstruction indicates positive temperatures in the tenth and thirteenth century that resemble twentieth-century conditions, and are separated by a prolonged cooling from ~1350 to 1700. Six of the 10 warmest decades over the 755–2004 period are recorded in the twentieth century. Maximum temperature amplitude over the past 1250 yr is estimated to be 3.1°C between the warmest (1940s) and coldest (1810s) decades. This estimate is, however, affected by the calibration with instrumental temperature data. Warm summers seem to coincide with periods of high solar activity, and cold summers vice versa. The record captures the full range of past European temperature variability, that is, the extreme years 1816 and 2003, warmth during medieval and recent times, and cold in between. Comparison with regional- and large-scale reconstructions reveals similar decadal to longer-term variability.

1. Introduction

For the European greater Alpine region (GAR), much progress in the last decade has been made in reconstructing climatic variations through studies of long instrumental observations (Auer et al. 2005, 2006, hereafter A06; Böhm et al. 2001; Camuffo and Jones 2002; Moberg et al. 2000), documentary evidence (Brázdil et al. 2005; Chouine et al. 2004; Glaser 2001; Le Roy Ladurie 2005; Menzel 2005; Pfister 1999), tree-ring data (Frank and Esper 2005b; Frank et al. 2005; Wilson and Topham 2004; Wilson et al. 2005), and multiproxy compilations (Casty et al. 2005c; Guiot et al. 2005; Luterbacher et al. 2004; Xoplaki et al. 2005). Atmospheric circulation patterns are also quite well documented for the European Alps (Wanner et al. 1997) and the North Atlantic/European sector (Casty et al. 2005a,b; Cook et al. 2002; Hurrell et al. 2003; Jacobfeit et al. 2003; Luterbacher et al. 1999, 2002; Pauling et al. 2006; Raible et al. 2006), with particular emphasis toward the winter half-year. Nevertheless, evidence is generally restricted to the recent centuries as there are few data for the medieval period. Longer-term understanding of European temperature variations is limited to a handful of records, such as the low-resolution evidence for a European Medieval Warm Period (MWP) and Little Ice Age (LIA) reported by Lamb (1965). Further evidence of the MWP has been derived from annually resolved tree-ring width (RW) and maximum latewood density (MXD) data (e.g., Briffa et al. 1990, 1992; Büntgen et al. 2005a; Grudt et al. 2002; Helama et al. 2002; Kalela-Brundin 1999; Schweingruber et al. 1988). However, due to paucity of data in both space and time, the occurrence of the LIA, and particularly the MWP, is still debated (e.g., Bradley 2003; Bradley and Jones 1993; Broecker 2001; Crowley 2000; Grove...
To date, tree-ring-based millennium-long temperature reconstructions (e.g., Büntgen et al. 2005a; Esper et al. 2003b; Luckman and Wilson 2005) (i) are key to understand local- to regional-scale climatic variations (Bradley 2000; Jones and Mann 2004), (ii) compile hemispheric-scale networks to assess spatial patterns of climatic change (Cook et al. 2004; D’Arrigo et al. 2006; Mann et al. 1998; Rutherford et al. 2005), and (iii) provide validation of the hindcast skill of climate model simulations (Houghton et al. 2001; Stainforth et al. 2005).

At the global scale, we know of only five MXD chronologies that stretch prior to A.D. 1000, that is, Lauenen from the Swiss Alps, Torneträsk from Swedish Lapland (Schweingruber et al. 1988), Polar Urals from Russia (Briﬁfa et al. 1995), Québec from Canada (Wang et al. 2001), and Columbia Iceﬁeld from Canada (Luckman and Wilson 2005). Herein, we present the ﬁrst MXD-based summer temperature reconstruction (A.D. 755–2004) that places the 2003 European heat wave (Chuine et al. 2004; Luterbacher et al. 2004; Menzel 2005; Schär et al. 2004) in a millennium-long context. In an effort to provide a reﬁned reconstruction of the timing and amplitude of past temperature variations, ecological disturbance signals are removed (Esper et al. 2006, manuscript submitted to Proc. Natl. Acad. Sci., hereafter EBFNL), age-related composite detrending techniques are applied (Briﬁfa et al. 1992, 1996), and wavelength-dependent calibration tests are performed (e.g., Osborn and Briﬁfa 2000). Results are compared to estimations of solar radiation (Crowley 2000; Usoskin et al. 2003), and inferences about the external forcing upon summer temperatures are made. Other GAR proxies are used to improve understanding of regional-scale temperature variations, and comparison with NH reconstructions is conducted to place these regional ﬁndings in a larger-scale context. In contrast to previous efforts (Büntgen et al. 2005a), this new analysis utilizing MXD data extends the existing Alpine record back by about 200 yr, updates the years 2003 and 2004, improves the growth/climate response signal, and enhances the “color” assessment and preservation in the reconstruction.

2. Data

a. Tree-ring data

The dataset consists of 180 MXD larch [Larix decidua Mill.] series from near timberline sites (86 recent samples) and subalpine construction timbers (94 historic samples) dating from 735–2004. Recent samples were collected in the Swiss Alps at elevations between 1900 and 2200 m asl. Historic buildings are located in an altitudinal belt of 1500–1900 m asl, with their construction wood often originating from higher elevations (Fig. 1): 110 samples derive from the Lôtschental (1258–2004), 39 from the Simplon region (735–1510), and 31 from the Aletsch region and Simmental (1681–1866). Samples were processed using a WALESCH 2003 X-ray densitometer with a resolution of 0.01 mm, and brightness variations transferred into g cm$^{-3}$ using a calibration wedge (Eschbach et al. 1995; Lenz et al. 1976). The mean segment length (i.e., the average number of rings per core or disc sample) is 264 yr, with means of 239 and 289 yr for the recent and historic subsamples, respectively (Fig. 2a). Average MXD is 0.87 g cm$^{-3}$, with little difference between the recent (0.90 g cm$^{-3}$) and historic (0.84 g cm$^{-3}$) material. The mean interseries correlation of the 180 MXD series is $r = 0.59$, calculated using COFECHA (Holmes 1983).

For regional tree-ring comparison, the GAR June–August temperature reconstruction for the A.D. 951–2002 period by Büntgen et al. (2005a) is used. This record combines 1527 subalpine larch and pine RW series from the Swiss and Austrian Alps (Fig. 1). A fraction of this rather large wood collection is used for the density measurements as utilized in this current study, that is, 120 of the 180 larch MXD series derive from the RW dataset.

b. Instrumental data

A revised version of homogenized instrumental temperature data including nine low- (high-) elevation 1$^\circ \times 1^\circ$ grid boxes, spanning the 1760 (1818)–2003 period, is considered (A06). This new release is a considerable upgrade in comparison to the original Böhm et al. (2001) data, that is, more series especially during the early instrumental period, and improved homogenization procedures and outlier corrections are considered. The low- and high-elevation grids cover the 45$^\circ$–47$^\circ$N and 6$^\circ$–9$^\circ$E area in the western-central Alps. The high grid meets the elevation criteria of the tree-ring sites, that is, >1500 m asl (Fig. 1). Interseries correlation between the single high- (low-) elevation grid points using June–September (JJAS) means is 0.99 (0.74). The higher interseries correlation for the high-elevation grid likely results from a mixture of factors, including the number of stations used for gridding, the greater spatial range of data used for interpolation of the low-elevation grid, and the greater common signal found at higher elevations. The regional mean of the high versus low-elevation grids, nevertheless, correlates at 0.94 over the 1818–2003 common period, indicating that elevational differences are reduced when many data over larger regions are combined (Böhm et al. 2001).
Since the GAR background climate is best captured by the high-elevation grid, and ideally preserved in the >1500 m asl tree-ring proxy data, these measurements are used for calibration. Low-elevation instrumental data back to 1760 are used for extra verification, and to address potential limitations in estimating the long-term temperature amplitude over the past millennium. Temperatures are expressed as anomalies from the twentieth-century mean (1901–2000).

A precipitation grid, similar to the low-elevation temperature dataset, covers the 1800–2003 period (Auer et al. 2005). The more clustered precipitation patterns within the GAR are expressed by a slightly lower grid box intercorrelation of 0.68, calculated for JJAS sums. Both the temperature and precipitation grids are used to assess the climatic signal preserved in the MXD chronology.

3. Methods

a. LBM correction

When analyzing the tree-ring data, negative MXD outliers induced by 8–9-yr cyclic larch budmoth (LBM) mass outbreaks were detected (EBFNL). The reason for these outliers is the defoliation of larch trees by LBM larvae during cyclic population peaks (Baltensweiler and Rubli 1999), causing exceptionally low MXD values (Schweingruber 1979). These patterns were used to detail a history of the frequency and magnitude of LBM population dynamics over the past millennium (EBFNL). For the current study, LBM effects are regarded as noise and removed from the MXD data, that is, 4649 LBM-affected tree rings were deleted and replaced with statistical estimates derived from the remaining, unaffected rings. In detail, this gap-filling procedure, for each year, comprises (i) averaging the MXD values of the remaining rings, (ii) adjusting the variance of the mean values of unaffected rings to the variance of the measurement series from which the ring was removed, (iii) replacing the gap with the variance-adjusted values obtained from unaffected rings, and (iv) calculating a mean chronology from the gap-filled single measurement series. Spectral analysis (Mann and Lees 1996; Percival and Walden 1993) using the multitaper method (MTM; Thomson 1982) was used to assess the power spectrum before and after LBM correc-
tion (not shown). Significant power at ~8–9 yr (peak at 8.9) diminished after the LBM correction was applied, and correlation with nearby nonhost fir and spruce chronologies (Lauenen and Tyrol; Schweingruber et al. 1988) increased from 0.34 and 0.36 to 0.54 for both records over the 1368–1975 common period (EBFNL).

Even though the removal of the LBM signal from the MXD data improves the calibration against instrumental data only slightly by about 0.03, we here use the LBM corrected series to avoid negative outliers due to insect population dynamics.

b. Chronology development

Tree-ring series were detrended using ARSTAN (Cook 1985) to remove nonclimatic, age-related growth trends (Fritts 1976). Since individual series detrending eliminates signals at wavelengths longer than about the mean series segment length (details in Cook et al.)

Fig. 2. The Alpine RCS–MXD chronology. (a) Temporal distribution of the 180 MXD samples. (b) The RCS–MXD chronology (red), and (c) the RCS–MXD chronology (red) with 95% bootstrap confidence limits (gray), and RCS chronologies after recent/historic data splitting (green/blue), all 20-yr low-pass filtered. (d) The temporal signal strength of the chronology is shown by EPS and RBAR statistics, with the horizontal line indicating the 0.85 EPS validity threshold, and (e) the 1 – NET parameter with high values indicating internal signal strength. The black line derives from 20-yr low-pass filtering, with both ends padded using the mean of the last 10 yr, somewhat comparable to the method described by Mann (2004).
1995), the regional curve standardization method (RCS; Briffa et al. 1992, 1996; Mitchell 1967) was applied to preserve low-frequency information in the resulting chronologies. RCS is a so-called age-related composite detrending method, where (i) all measurement series are aligned by cambial age; (ii) the mean of all age-aligned series, the so-called “regional curve” is smoothed—here, with a cubic spline of 10% the series length (Cook and Peters 1981); and (iii) the deviations of the individual measurements from this smoothed regional curve are calculated—here, as residuals (details in Esper et al. 2003a). The series were averaged using the biweight robust mean (Cook and Kairiukstis 1990), while the variance in the mean chronologies was stabilized using methods described by Osborn et al. (1997).

Resulting chronologies were truncated at a minimum sample replication of seven series. Bootstrap confidence limits of 95% were used to estimate uncertainty in the common signal represented in the MXD chronology (Efron 1987). To test for potential population differences along the last 1250 yr, and to assess the low-frequency information captured by the RCS chronology, the dataset was split temporally into the 86 recent and 94 historic subsamples, and two RCS runs were calculated. Results are compared with the chronology obtained using all data.

Signal strength of the RCS chronology is assessed using the interseries correlation (RBAR), the “expressed population signal” (EPS; Wigley et al. 1984), and the NET parameter (Esper et al. 2001). RBAR is a measure of common variance between single series, independent of the number of measurement series. EPS is an absolute measure of chronology error that determines how well a chronology, based on a finite number of trees, estimates the theoretical population chronology from which it has been drawn. EPS quantifies the degree to which this particular sample chronology portrays the theoretical population chronology. Both RBAR and EPS are calculated for 30-yr windows lagged by 15 yr along the chronology. NET combines the coefficient of variation (CV) and the Gleichläufigkeit (G)—the percentage of synchronous trends between single series—for each year of the mean chronology. The parameter shows high interannual variability between the single series related to the proportion of synchronous year-to-year changes. Since NET considers the relative variance between single series, it helps provide a signal strength estimate of the low-frequency component retained in RCS chronologies (Esper et al. 2001). To facilitate comparison with the RBAR and EPS statistics, we here show $1 - \text{NET}$, so that all metrics display increasing signal quality with increasing values.

c. Calibration and verification

Based on monthly correlation results, various calibration trials of the chronology were made against high-elevation JJAS, February—September (FS), and annual mean temperatures within the 1818–2003 period. The low-elevation grid is utilized for extra verification only. Split period calibration/verification (1818–1910/1911–2003) plus extra verification back to 1760 were undertaken to assess the model’s temporal robustness. To avoid loss of amplitude due to regression error (e.g., Esper et al. 2005a), simple scaling of the MXD–RCS chronology against instrumental targets, that is, adjusting the variance and mean, was applied.

The explained variance ($R^2$), reduction of error statistic (RE), coefficient of efficiency (CE), and the Durbin–Watson statistic (DW) were used to assess the reconstruction skill. Here, RE and CE are measurements of shared variance between target and proxy series, generally lower than the $R^2$ (Cook et al. 1994; Fritts 1976). The DW statistic tests for lag-1 autocorrelation in the model residuals. A DW value of 2 indicates no first-order autocorrelation in the residuals; values greater (less) than 2 indicate negative (positive) autocorrelation (Durbin and Watson 1951).

For better understanding of the modeled relationship between proxy and instrumental data, wavelength-dependent calibration was performed (e.g., Guiot 1985; Osborn and Briffa 2000; Rutherford et al. 2005; Timm et al. 2004), that is, predictor and predictand were decomposed into high- and low-pass components using a 20-yr smoothing spline. Linear regression was separately performed on both frequency bands. The regressed bands were then simply summed and the result scaled to the instrumental data to obtain the two-band reconstruction. Regression slope coefficients of the high- and low-pass components were found to vary systematically; however, after correction for lag-1 autocorrelation (Trenberth 1984) they were not statistically distinguishable. Osborn and Briffa (2000) suggest that these coefficients should be statistically distinguishable as a criterion for determining whether frequency-dependent calibration is appropriate. Nevertheless, we show results for the high- and low-pass and the combined two-band model to emphasize frequency dependence between predictor and predictand.

4. Results

a. Chronology characteristics

Replication, temporal distribution, and segment length of the 180 MXD series allow for the calculation of one composite RCS chronology (755–2004 after
truncation <7 series; Figs. 2a–c). Bootstrap confidence limits of 95% are reasonably narrow back to A.D. 755. They, however, increase between ~950 and 1450, indicating lower internal signal strength. After splitting the MXD data into the 86 recent and 94 historic subsamples, both datasets possess similar regional curves (not shown), and a correlation between the split chronologies of 0.50 for the 1544–1743 period of overlap. Interestingly, the recent and historic RCS chronologies lie during most periods of the past millennium within the 95% confidence limits estimated for the RCS chronology using all 180 series, indicating that the combination of living and historic material in one RCS run leads to similar results than obtained from the split approach.

The composite chronology using all data shows high MXD values in ~970, ~1150, ~1230, and ~1940, and during the most recent decade, with 2003 reaching the highest index value since A.D. 755. Within the early period of high index values, a distinct depression occurs during the eleventh century. A prolonged depression exists from ~1350 to 1820, with low MXD values in ~1460, ~1590, ~1680, and ~1820, and with 1816 showing the lowest value over the past 1250 yr.

Figures 2d–e denotes the temporal signal strength of the composite RCS chronology. Except for the 1194–1234 period, which is replicated by only 8–10 series, EPS, RBAR, and NET indicate internal consistency in common variance. Except for the period ~1200, EPS values clearly remain above the frequently applied threshold of 0.85 (e.g., Briffa and Jones 1990), indicating that the chronology closely represents a theoretical mean function of infinite replication (Wigley et al. 1984). Values of RBAR are particularly high prior to ~1100 and after ~1800, indicating more homogeneous data during the chronology’s early (Simplon) and late (Lötschental) periods. Low RBAR values in ~1200 reflect the overlap of particularly old and young material around that time. The step in ~1800 is likely influenced by the strong depression in MXD values in the early eighteenth century (see below), reported from many NH sites (e.g., Briffa et al. 2002).

The 20-yr low-pass-filtered NET values emphasize long-term internal signal changes dominated by the variance between single MXD series. This is key in evaluating the low-frequency component of time series generated using RCS (Esper et al. 2003a). Periods before ~1150 and after ~1750 possess generally high signal strength, with lower signal quality in between. Interestingly, lowest consistency occurs in the period ~1160–1200, followed by a strong increase in the ~1210s, a period during which EPS and RBAR indicate low signal strength, that is, the most problematic period of the MXD chronology in terms of replication, correlation, and variance between single measurement series is the depression centered around A.D. 1200.

b. Growth/climate response

High-elevation temperature and precipitation data are used to assess the proxy’s climate response back to 1818 (Fig. 3). Correlation analysis using previous-year April to current-year October monthly data indicates the significance of growing season conditions during June, and particularly July, August, and September on MXD formation (Fig. 3a). No correlation with previous-years temperatures is significant at p < 0.05, and weak or negative correlations with precipitation are likely influenced by the cross correlation with temperature (Briffa et al. 2002; Frank and Esper 2005a; Schweingruber 1996). JJAS and FS seasonal means reveal highest correlations of 0.69 and 0.57, respectively. These correlations are also significant at p < 0.01 after splitting the 1818–2003 period into two 93-yr subperiods, indicating temporal stability of the growth/climate relationship, when calibrating against high-elevation instrumental data (Table 1).

Moving 31-yr correlation analysis considering July, August, September, and JJAS indicates variable relationships for the individual monthly mean temperatures (Fig. 3b). Weaker correlations are obtained for July before ~1900, and for September after ~1940. Results for JJAS are temporally more stable and persistently significant at p < 0.01.

c. Calibration/verification trials

The R², RE, CE, and DW calibration and verification statistics against JJAS mean temperatures using the RCS chronology, its 20-yr high-pass-filtered component, and the 20-yr two-band model indicate reconstructive skill against high-elevation instrumental temperature data back to 1818 (Table 1). The RE and CE values gleaned from the 1818–2003 period, range between 0.20 and 0.59, demonstrating some useful information preserved by the model (Cook et al. 1994), with DW values ranging between 0.88 and 2.11. Extra verification against early, low-elevation instrumental data over the 1760–1817 period indicates that the RE and CE statistics are negative for both the RCS and two-band models, even though R² values (0.8) are exceptionally high during this period (Table 1). These results, together with the positive RE and CE values obtained for the 20-yr high-pass component, point to a misfit between instrumental and proxy data during the early extra verification period that is restricted to the lower-frequency component of these time series.
The low-pass component misfit is further detailed for the RCS chronology in comparison with differing seasonal (JJAS, FS) and annual temperature means (Fig. 4). Accordingly, the tree-ring data show highest correlations with warm season temperatures in the higher-frequency domain, but tend to be more statistically similar to the annual data in the lower-frequency domain, that is, residuals between smoothed proxy and instrumental data are slightly larger for the JJAS season. Overall, the residuals between target and proxy data are negative before ~1840, positive until ~1960, and negative again until 2004, indicating that there is no centennial-scale trend offset that could potentially arise from the application of RCS (Melvin 2004; see discussion below).

Even though, trend differences between warm season and annual temperature data (e.g., Hansen et al. 1999; Jones et al. 2003; Luterbacher et al. 2004) make the statistical differentiation and selection of the proper target season particularly challenging (Esper et al. 2005a), the monthly and seasonal correlation results, as well as the calibration and verification trails indicate that JJAS temperatures are best captured by the MXD data. For final calibration and transfer, the mean and variance of the RCS chronology is scaled to JJAS temperatures derived from high-elevation instrumental data (Fig. 5a). The model’s explained variance over the 1818–2003 calibration period is >70% in 41 yr, >50% in 84 yr, and 30%–50% in the remaining 32 yr, with moving 31-yr correlations demonstrating temporal stability (Fig. 3b).

While the reconstruction explains ~50% of summer temperature variability, comparison of the high- and low-pass components (Figs. 5b,c) reveals that particularly the interannual variations are rather well preserved. Calibration results obtained from the combined 20-yr two-band model are shown to highlight persisting misfit with early instrumental data used before 1818 (Fig. 5d). The two-band approach caused a slight rota-

### Table 1. Calibration and verification statistics against JJAS mean temperatures using (a) the simple scaling, (b) the 20-yr high-pass, and (c) the 20-yr two-band model.

| Period       | Calibration | | Verification | | |
|--------------|-------------|---|-------------|---|
|              | R² | DW | Period      | R² | RE | CE |
|              |    |    | (a)         |    |    |    |
| 1911–1903    | 0.53 | 1.12 | 1818–1910   | 0.32 | 0.45 | 0.27 |
| 1818–1910    | 0.32 | 1.50 | 1911–2003   | 0.53 | 0.56 | 0.47 |
| 1818–2003    | 0.47 | 0.88 | 1760–1817   | 0.64 | −1.75 | −2.33 |
|              | (b)         |    |              |    |    |    |
| 1911–2003    | 0.59 | 2.01 | 1818–1910   | 0.60 | 0.58 | 0.58 |
| 1818–1910    | 0.59 | 2.11 | 1911–2003   | 0.60 | 0.59 | 0.59 |
| 1818–2003    | 0.59 | 2.10 | 1760–1817   | 0.58 | 0.53 | 0.53 |
|              | (c)         |    |              |    |    |    |
| 1911–2003    | 0.55 | 1.10 | 1818–1910   | 0.58 | 0.39 | 0.20 |
| 1818–1910    | 0.52 | 1.38 | 1911–2003   | 0.48 | 0.42 | 0.31 |
| 1818–2003    | 0.51 | 1.12 | 1760–1817   | 0.64 | −0.68 | −1.04 |

Fig. 3. Climate response of the RCS-MXD chronology. (a) Correlations with high-elevation monthly mean temperatures (black) and precipitation (gray) of the previous and current year (1818–2003). [b, inset to a] Moving 31-yr correlations with current-year July (black), August (dashed), September (gray), and JJAS (black bold, labeled J-S) mean temperatures. Horizontal lines denote 95% and 99% significance levels, corrected for lag-1 autocorrelation (Trenberth 1984) using the month with the highest autocorrelation.
tion of the chronology, yielding to an insignificant reduction of the early offset, however, accompanied with a shift of the warmest year on record, from 2003 to 1928. Furthermore, use of the two-band model for reconstruction was not justified, because the regression slope coefficients of the high- and low-pass components are statistically not distinguishable.

d. Temperature history

Figure 6a shows the reconstructed Alpine temperature for the 755–2004 period, with its mean being 0.73°C colder than the 1901–2000 instrumental reference period. Warmest summers are in 2003 (+1.9°C), 970, and 1928 (both +1.7°C). Coldest summers are in 1816 (−4.5°C) and 1046 (−3.9°C). We emphasize that these results are derived from the proxy data without instrumental extension into the twenty-first century (e.g., Cook et al. 2004; Jones and Mann 2004).

Evidence for a pronounced MWP, LIA, and recent warmth is found. For the MWP, significant interdecadal fluctuations are recorded, with high temperatures in the 960s–80s and 1200s–20s, and low temperatures in the 1040s–60s. The reconstruction shows strong interdecadal fluctuations through a generally cooler period between ~1350 and 1820, coinciding with the LIA. Low temperatures are recorded during 1580–1710, and relatively high temperatures during ~1500 and ~1800. Since ~1710, an analog to the end of the Late Maunder Minimum (Eddy 1976; Luterbacher et al. 2001; Shindell et al. 2001; Wanner et al. 1995), temperatures discontinuously increased with notable depressions in ~1820 and ~1970. Reconstructed interannual- to multidecadal-scale variations of the last century show a first warming episode from the early 1910s to the end of the 1940s, and a second from the late 1960s to present. This course, and the most recent warming including the summer of 2003, is in line with temperature variations reported from high-elevation Alpine instrumental obser-

Fig. 4. Calibration (1818–2003) and extra verification (1760–1817) of the MXD–RCS chronology (red) against high-elevation (black) and low-elevation (thin) temperatures of different target season windows, using simple scaling. Residuals between 20-yr smoothed predictor and predictand indicate their relationship through time, and target season dependence. Correlations refer to unsmoothed series. Temperatures are expressed as anomalies from the twentieth-century mean.
5. Discussion

a. Proxy/target relationship

Unexplained variance in proxy data is found in the interannual- to decadal-scale frequency domain, with a superimposed misfit between colder tree-ring and warmer instrumental data before ~1840 and after ~1960 (Figs. 4, 5). An overall trend difference or centennial-scale discrepancy that would refer to potential limitations in applying RCS (Esper et al. 2003a; Helama et al. 2005; Melvin 2004) is, however, not revealed. Evidence for similar decadal-scale differences is seen in several Alpine studies that used different tree-ring and instrumental data, and applied varying detrending and calibration methods (e.g., Büntgen et al. 2005a; Frank and Esper 2005b; Wilson et al. 2005). Potential index inflation toward the chronology’s recent end, the so-called “end-effect” problem (Cook and Peters 1997), is herein excluded through calculating residuals rather than ratios. Proxy/target misfits are also found in different seasonal calibrations (Fig. 4), with maximum offset revealed for the summer months, and minimum offset for annual means. However, calibration and verification statistics of the 20-yr high-pass component show highest agreement with the JJAS season (Table 1).

Potential reasons for the unexplained variance in the reconstruction typically include (i) nonlinearity in the growth/climate response (Fritts 1976), with (ii) possible response shifts between precipitation and temperature, as reported from the European Alps (Büntgen et al. 2005b); (iii) growth response to maximum rather than mean temperatures (Wilson and Luckman 2003); (iv) changes in the growing season length including slow.
ecological shifts (Frank and Esper 2005b); and (v) methodological uncertainty in the detrending and calibration techniques performed (Cook and Kairiukstis 1990).

It seems interesting, however, that the observed decoupling between proxy and target data coincides with the timing of homogenization changes applied to the instrumental temperature data (warming before ~1840 and after ~1960, with cooling in between), with meteorological observations generally providing higher quality during the late twentieth century (A06; Böhm et al. 2001). Figure 5 denotes the decoupling between (warmer) early instrumental and (cooler) proxy data particularly before 1818, which could be affected by the less replicated, more error prone, and therefore more intensively homogenized early measurements, generally recorded by urban stations (Böhm et al. 2001). Although central Europe sets the standard for instrumental measurements (Jones and Moberg 2003), quality and quantity of early observations—16 (36) stations within the GAR provide data prior to 1800 (1850)—are incomparable with the modern network (Jones et al. 1997). The annual rate, magnitude, and frequency distribution of measured outliers increase before ~1840 (A06). Compared to the twentieth century, this early observational period is characterized by a slight variance increase, likely related to nonsystematic meteorological measurements composed of short sequences of

Fig. 6. (a) Alpine summer temperature reconstruction with the orange and blue boxes denoting the 10 warmest and coldest decades, respectively, and the smoothed red line being a 20-yr low-pass filter. Temperatures are expressed as anomalies w.r.t. 1901–2000. (b) High-frequency comparison between the MXD (red; this study) and RW (blue; Büntgen et al. 2005a) RCS chronologies. Records were normalized over the 951–2002 common period. The 51-yr moving correlations (black) indicate their temporal relationship, with the horizontal line denoting the 95% significance level, corrected for lag-1 autocorrelation (Trenberth 1984). (c) Length fluctuation (m) and 50-yr average mass balance (gray) of the Great Aletsch glacier (Haeberli and Holzhauser 2003).
sporadic observations (Brázdil et al. 2005; Camuffo and Jones 2002; Moberg et al. 2000; Parker and Horton 2005). Further uncertainty derives from the impact of urban artificial heating on temperature trends (Damon and Kunen 1976), with its quantification yet debated (Jones and Lister 2004; Kalany and Cai 2003; Klingbjer and Moberg 2003; Moberg et al. 2003; Parker 2004; Parker and Horton 2005).

The elevation difference between tree-ring sites and early instrumental stations could further contribute to the decoupling seen before 1818 (and latter in the calibration period), because vertical components in climatic variations are not static in nature. For the European Alps, elevation differences diminish the correlation between instrumental stations more than their horizontal separation does (Böhm et al. 1996), with its quantification yet debated (Jones and Lister 2004; Kalany and Cai 2003; Klingbjer and Moberg 2003; Moberg et al. 2003; Parker 2004; Parker and Horton 2005).

The sun is considered as the most important driving force of the earth climate system (Bard et al. 2000; Beer et al. 2000; Eddy 1976; Lean and Rind 1998), thus compared with the Alpine temperature reconstruction, using estimates of solar radiation (Crowley 2000) and sunspot numbers (Usoskin et al. 2003; Fig. 7). Correlations between the low-frequency solar activity and sunspot number records and the 40-yr smoothed temperature reconstruction are 0.64 and 0.58 over their common period, respectively. Even though correlations are not significant at $p < 0.05$ after correction for lag-1 autocorrelation, records share high values during the twelfth and thirteenth centuries (great solar maximum;
Eddy 1976), a prolonged depression during ~1350–1700, and increasing values toward the twentieth century. The prominent interdecadal solar minima—Oort, Wolf, Spörrer, Maunder, Dalton, and Damon (Stuiver and Braziunas 1989)—as well as the corresponding maxima are superimposed upon this secular trend.

We here provide tree-ring evidence for the Oort solar depression in ~1050, with magnitude comparable to that of the Late Maunder Minimum during ~1675–1715 (Eddy 1976; Luterbacher et al. 2001; Shindell et al. 2001; Wanner et al. 1995). Discrepancies between regional-scale summer temperatures and large-scale solar activity, such as in ~1180, ~1580, and ~1900, are likely caused by unexplained variance in the proxy records, and superimposed clusters of volcanic eruptions (see below). The offset in ~1970 likely refers to a cooling due to industrial sulfate aerosol emissions (Anderson et al. 2003), with the sun’s contribution to the recent warmth remaining an open question (Crowley 2000; Damon and Peristykh 2005; Foukal et al. 2004; Hansen 2000; Meehl et al. 2005; Solanki et al. 2004; Usoskin et al. 2003; Wild et al. 2005). Common evidence for a large-scale solar forcing upon longer-term preindustrial temperature variations derives from other regional tree-ring studies (e.g., Briffa et al. 1990, 1992, 1995; Esper et al. 2003; Luckman and Wilson 2005).

Comparison of our reconstruction with volcanic eruptions reveals no systematic relationship, likely related to the regional character of both the Alpine temperature and forcing data. Detailed analysis, however, suggests a cooling of several years following primarily tropical events with a volcanic eruption index (VEI) >4. Examples include Hekla in Iceland (1300), Pinatubo in the Philippines (1450), Kuwae in Vanuatu (1452), Cayambe in Ecuador (1570), Fuego in Guatemala (1581), Huaynaputina in Peru (1600), Momotombo in Nicaragua (1604), Colima in Mexico (1606), Vesuvius in Italy (1631), Parker in the Philippines (1641), Guagua-Pichincha in Ecuador (1660), Teon in the Banda Sea (1660), Katla in Iceland (1660), Gamkonora in Indonesia (1673), Taal in Indonesia (1754), Agung in Indonesia (1963), and Fuego in Guatemala (1974). For details on the location, intensity, type, and timing of these eruptions, see Simkin and Siebert (1994). Other regional (Gervais and MacDonald 2001; LaMarch and Hirschboeck 1984; Luckman and Wilson 2005) and larger-scale studies (Briffa et al. 1998; D’Arrigo and Jacoby 1999) support a posteruption summer cooling response to selected events.

We assume that periods of intensified eruption events, such as during the eleventh century, between ~1170 and 1300, during the second half of the fifteenth century, between ~1560–1700, ~1800–20, ~1900, and ~1960–70, forced summer temperature depressions on decadal time scales. However, several prominent volcanic eruptions reported for the past millennium (Simkin and Siebert 1994) did not leave their fingerprint. This is likely because of the regional scope of this study (Raible et al. 2006; Shindell et al. 2004), intensity and location of some eruptions (Robock 2000), and/or dating uncertainty of earlier events (Mann et al. 2005). Dating uncertainty is, for example, reported for the “unknown” mid-thirteenth-century eruption that caused the highest peak in ice core sulfate—referred to as atmospheric aerosol loading—during the last millennium (Oppenheimer 2003b; Zielinski 2000). For this current study, the ±1258 event likely caused a depression of ~1.5°C two years prior to the eruption date A.D. 1259 used by Crowley (2000; see also Luckman and Wilson 2005).

For the European Alps, most pronounced radiative forcing arises from the Tambora in Indonesia event in April 1815 (Sigurdsson and Carey 1989; Oppenheimer 2003a), causing a mean temperature depression of ~4.5°C in 1816, known as “the year without summer” (Harrington 1992; Robock 1994). This period is also characterized by a series of tropical eruptions (1808–15), which likely resulted in an aerosol-accumulated summer cooling effect (Chenoweth 2001; Dai et al. 1991), along with cooler conditions due to low solar activity in the Dalton minimum (Wagner and Zorita 2005).

Overall, reconstructed low-frequency summer temperature variations appear to mimic solar activity, with higher-frequency variations partly matching the timing of volcanic eruptions that can mask the sun–climate relationship (Donarummo et al. 2002). Interestingly, solar minima often coincide with periods of pronounced volcanic eruption activity (e.g., the Dalton Minimum), making both variables not clearly distinguishable. Recent anthropogenic impact further diminishes the proportion of natural forcing agents during the industrial period (Anderson et al. 2003; Bauer et al. 2003; Crowley 2000; Meehl et al. 2003; Stott et al. 2000, 2001, 2004).

c. Regional- to large-scale comparison

To assess the temperature history presented here, comparisons with regional- and large-scale millennium-long proxy records are performed (Figs. 6b–c, 7b–c). This current study and the chronology by Büntgen et al. (2005a), combining RW measurements from 1500+ larch and pine series from the European Alps, show reasonably high coherency on interannual to multicen-
tennial scales (Fig. 6b). Over the 951–2002 common period, correlations are 0.57 and increase to 0.66–0.71 after 20–80-yr low-pass filtering. The RW June–August Alpine temperature proxy (Büntgen et al. 2005a) shows warm summer conditions from before a.d. 1000 into the thirteenth century, followed by a prolonged cooling with lowest temperatures in the 1820s, and the recent warmth. The 51-yr moving correlations (mean $r = 0.46$) show weak coherency in $\sim 1180–1350$, $\sim 1550$, and $\sim 1650–1790$ between this record and the MXD-based reconstruction presented here. Enhanced interannual climate response of the MXD data refers to the lower biological memory of MXD data compared to their RW counterparts (Cook and Kairiukstis 1990; Fritts 1976; Frank and Esper 2005a), which show more decadal-scale variability (e.g., $\sim 1470$ and $\sim 1820$).

Interestingly, after dividing the larch/pine composite data into subsets of 1100 (Swiss) larch and 417 (Austrian) pine series, their RCS chronologies correlate at 0.60 and 0.31 with the new MXD reconstruction (951–1997), respectively. Correlations increase to 0.71–0.73 and 0.39–0.50, after 20–80-yr low-pass filtering.

Comparison over the past 1300 yr with regional length and mass balance reconstructions of the Great Aletsch glacier (Haeblerli and Holzhauser 2003; Hoelzle et al. 2003) indicates some lower-frequency coherency with our record (Fig. 6c). Evidence for multidecadal-scale fluctuations during the MWP, an early LIA, with a first extension in the fourteenth century, followed by a retreat phase from $\sim 1450$ to 1600, and again two advances in the seventeenth and nineteenth centuries, is provided. These ups and downs during the LIA likely refer to modifications in atmospheric circulation patterns during the Wolf, Maunder, and Dalton solar minima (Luterbacher et al. 2001, 2002), with the latter likely expressing the most extended Alpine glacier advance of the Holocene (Grove 1988). Glacier retreats comparable to the most recent one are reconstructed for three periods between the eighth and thirteenth centuries, interrupted by two advances, with the latter likely reflecting the Oort solar minimum.

Although cumulative tongue length fluctuations of the Great Aletsch and several other Alpine glaciers (e.g., Holzhauser 2002; Nicolussi and Patzelt 2001; Oerlemans 2005) resemble longer-term temperature variations and contribute to the understanding of past amplitude ranges, uncertainty remains. This is related to the glaciers’ time lag in response, reaching several decades, and the complex climatic signal including temperature, precipitation, and solar irradiation changes (Haeblerli and Holzhauser 2003). Visual comparison with the tree-ring proxy, therefore, must consider that reconstructed length changes do not account for higher-frequency temperature variations, and that the most recent warming trend is not picked up yet (Haeblerli and Holzhauser 2003).

Nevertheless, the Great Aletsch glacier is the largest and best documented glacier in the European Alps possessing a multimillennial-long history of advances and retreats supported by radiocarbon and tree-ring dating, moraine investigations, and annual measurements since 1892 (Holzhauser 2002), and it allowed for the estimation of an averaged 50-yr mass balance model (Hoelzle et al. 2003). Comparison with large-scale temperature records considers the tree-ring-based Esper et al. (2002), D’Arrigo et al. (2006), and the multiproxy-based Moberg et al. (2005) reconstructions (Fig. 7c). Correlations between this study and Esper et al. (2002), D’Arrigo et al. (2006), and Moberg et al. (2005), computed over the 951–1979 common period are 0.20, 0.28, and 0.37, and increase to 0.27, 0.40, and 0.56 after 40-yr smoothing, respectively. Correlations between the Alpine MXD reconstruction and other NH reconstructions by Briffa (2000), Jones et al. (1998), and Mann et al. (1999), computed over the 1000–1979 common period, are 0.18, 0.29, and 0.22, respectively, and increase to 0.26, 0.43, and 0.36 after 40-yr smoothing. No data overlap between the Alpine and NH reconstructions exists. All proxy records reveal an overall centennial to longer-scale common signal but show relative level differences for the MWP, LIA, and recent warmth (Esper et al. 2005b).

The MXD reconstruction generally portrays lower temperatures around the MWP, particularly during the Oort solar minimum compared to the NH reconstructions (and the Alpine RW record; Büntgen et al. 2005a), although higher temperatures are reconstructed in the twelfth and thirteenth centuries. The most recent temperature depression in $\sim 1970$ is distinct for the European Alps, and less pronounced in the large-scale reconstructions. Reasons for similarities and dissimilarities between the NH reconstructions are discussed elsewhere (Esper et al. 2005b; Mann et al. 2005b; Rutherford et al. 2005). All large-scale reconstructions, however, include only a few annually resolved proxies around a.d. 1000 (Esper et al. 2004), and generally share more data back in time, for example, the Torneträsk chronology from Swedish Lapland is used in all records. This data overlap makes the correlation results mentioned above not fully independent.

6. Conclusions

We utilized 180 recent and historic LBM corrected MXD series that span the 735–2004 period. The RCS
method was used to preserve both low- to high-frequency information from the data. Instrumental measurements from nine high- (low-) elevation grid boxes back to 1818 (1760) were used for comparison and reveal the proxy’s summer temperature response. For calibration, scaling models of different wavelength and seasonality were tested. The record correlates at 0.69 with high-elevation JJAS temperatures back to 1818, and the signal is weighted toward high-frequency variations. Extra verification using low-elevation temperatures back to 1760 shows the reconstruction’s interannual skill, but divergence with (warmer) instrumental data. Similar divergences also exist during more recent periods, with potential reasons being discussed.

High temperatures are recorded in the late tenth, early thirteenth, and twentieth century. A prolonged summer cooling from ~1350 to 1700 is followed by increasing temperatures, with distinct depressions during the ~1810–20s, the 1910s, and 1970s. Without instrumental extension, the 1250-yr-long record indicates warmest summer temperatures in 2003, and coldest temperatures in 1816, known as the “year without summer.”

Longer-term temperature variations match reasonably well with solar activity, and some annual to decadal-scale downturns appear to coincide with volcanic eruptions. Regional-scale comparison with an RW-based temperature reconstruction and a length and mass balance reconstruction of the Great Aletsch glacier indicates higher- and lower-frequency similarities. Large-scale comparison with reconstructed NH temperatures shows related decadal-scale variability superimposed on longer-term trends. These findings suggest that Alpine summer temperatures are somewhat synchronous with NH variations.

Since the new Alpine MXD record calibrates better with summer temperatures, and is longer than the recent RW reconstruction by Büntgen et al. (2005a), it improves our understanding of past Alpine temperature variations. Nevertheless, an updated RW/MXD hybrid including the year 2005, data from a newly developed millennial-long spruce chronology from Austria, and numerous of summer temperature sensitive tree-ring chronologies throughout the GAR compiled by Frank and Esper (2005a,b), would likely draw a superior Alpine temperature history.

Main uncertainty within the new reconstruction is likely related to the reduction of sample size ~1200. Wood prior to this “transition” originates from the Simplon region. Afterward, most samples were collected in the Lötschental. Limited site control, for example, elevation and exposition, stand characteristics, and ecology, is a characteristic feature of the historic material utilized. The “shape” of the RCS chronology is partly insecure, since various implications of data and methodology are not yet fully quantified (Esper et al. 2003a; Helama et al. 2005; Melvin 2004). Decadal-scale differences between (warmer) early instrumental measurements and (colder) proxy data reveal uncertainty in the longer-term temperature amplitude. Annual extremes, such as the warm and cold summers of 2003 and 1816, respectively, remain less pronounced than the instrumental target requires. Although, reconstructed temperature variations mimic natural forcing agents reasonably well, their quantification is still vague, and the twentieth-century contribution of anthropogenic greenhouse gases and aerosol remains insecure.

Acknowledgments. We thank M. Schmidhalter for providing historic wood from the Simplon region, R. Böhm for instrumental data, and R. J. S. Wilson, J. Luterbacher, and K. Treydte for comments and discussion. Supported by the EU project ALP-IMP, and the SNF projects NCCR Climate and EURO-TRANS (#200021-105663).

REFERENCES


Briffa, K. R., 2000: Annual climate variability in the Holocene:


—, D. C. Frank, and R. J. S. Wilson, 2004: Low frequency am


Kalela-Brundin, M., 1999: Climatic information from tree-rings of Pinus sylvestris L. and a reconstruction of summer temperatures back to AD 1500 in Femundsmarka, eastern Norway, using partial least squares regression (PLS) analysis. Holocene, 9, 55–77.


Luterbacher, J., —, R. Schmutz, D. Gyalistras, E. Xoplaki, and H.


—, T. Bartholin, E. Schär, and K. R. Briffa, 1988: Radiodensitometric-dendroclimatological conifer chronologies from Lapland (Scandinavia) and the Alps (Switzerland). *Boreas*, 17, 559–566.


Simkin, T., and L. Siebert, 1994: *Volcanoes of the World: A Re-


