SM9: “Greater Urals” large-regional analysis

GU1. Evaluating uncertainties in chronologies that represent different regional scales
In many statistical applications, there is a balance between minimising random sampling error and minimising systematic error (i.e. a trade-off between variance and bias). This issue arises in tree-ring chronology construction too, balancing the inclusion of more data to reduce the noise (i.e. the sampling error) against the inclusion of data from too large an area such that the signal becomes ambiguous or even incompatible. In this section we briefly evaluate some of the key aspects associated with this balance, as they apply to the Yamalia chronologies.

We have constructed four different chronologies to illustrate some of the issues associated with chronology sampling error and bias, and to compare these between a single-site chronology and a chronology developed from a much larger region. The single-site chronology is constructed using one-curve, signal-free RCS to standardise the living and sub-fossil tree-core measurement data from the Polar Urals site (i.e. the larch data for this region used by Esper et al. 2002). We produce two versions of this chronology, one using all the data including root-collar samples (125 cores from Pou_la.raw and Polurula.raw) and a second using only the stem sample data (104 samples remaining after removing Pou_la_root.raw and polu_root.raw).

The larger-region, Greater Urals, chronology combines this Polar Urals measurement data with living and sub-fossil tree-ring data from Yamal (specifically, the data of Hantemirov and Shiyatov 2002, also used by Briffa 2000), together with data from living Larch (*Larix sibirica*) trees collected from a further 14 sites. The site locations and characteristics of the ring-width data are listed in Tables GU1 and GU2. These sites span 22° of longitude and 3° of latitude, with maximum separation distance of more than 1000 km. For comparison, the correlation decay distance of instrumental summer temperature observations is between 1000 and 1500 km in this region (Figure 2d of Briffa and Jones, 1993). Cross correlations between all the individual site chronologies (Table GU3) range from 0.24 to 0.88; the lowest correlation being between the two sites with the greatest separation distance. Although not as strong as the common signal present within a smaller region (such as Polar Urals and Yamal – i.e. “Yamalia”), there is clearly a common signal in tree growth across this wide region. The correlations mostly reflect annual to decadal variations, however, and longer-term changes are not as coherent in some cases. There are also notable differences in the mean growth rates between these sites, partly reflecting different tree age/size distributions in the available samples, but also suggestive of different growth behaviour that could influence the longer-term changes recorded in tree growth.

As with the single-site chronology used in this section of our work, two chronologies are constructed from the larger region, one including and one excluding the Polar Urals root-collar samples. In each case, a single RCS curve is used to standardise the measurement data. The resultant chronologies are compared with those obtained from the single-site, in terms of both the central chronology estimates and also some aspects of the statistical uncertainty associated with each chronology. This latter aspect includes the uncertainty in each RCS curve that arises from having a finite sample of measurements from which to determine the expected growth rate of larch in this region as a function of a tree’s age. We determine this via a bootstrap procedure: for a case with N measurements series, we randomly select N samples with replacement to derive an alternative RCS curve and subsequently an alternative chronology. Note that this differs from some other applications of bootstrapping in dendrochronology because we apply the sampling with replacement to each complete measurement series (e.g. from one tree or from one tree core) rather than to each individual ring measurement (e.g. Guiot 1991). We repeated this 1000 times and we used this distribution of 1000 alternative values to estimate the 95% confidence (from the 2.5 to the 97.5 percentiles) interval for each value in the RCS curve and for each value in the chronology range.

The first comparison we show is between the chronologies obtained with and without the root-collar samples. Inclusion of the root-collar samples has a significant impact on the centennial variations in the single-site chronology (Figure GU1a), and we have argued in the main paper that this is an unwanted bias due to the much higher mean growth rate of these samples compared with those from tree stems. For the larger-region chronology (Figure GU1b), the impact of including these samples is still evident though considerably reduced (note that prior to the modern, living samples, the sub-fossil data in the "larger-region" chronology come entirely from the much narrower region represented only by the Yamal and Polar Urals chronologies). This illustrates that building chronologies using data from multiple sites can reduce potential systematic errors that arise from biases in one group of data that do not occur in other groups. Of course, the bias is still present, and where it can be identified and avoided (by adjustment or removal of the non-homogeneous data) then this is to be recommended rather than attempting to dilute the bias through the incorporation of additional data.

The single-site and larger-region chronologies show some similar variations (Figure GU1a,b), due to the common data and due to the common signal across the group of chronologies. There are some differences, however, such as the relative strength of the upward slope over the last two centuries, and these arise partly from differences in the RCS curves (Figure GU1e,f). The bootstrap confidence limits on the RCS curve for the larger region is notably smaller, suggesting that the use of additional data has reduced chronology uncertainty and perhaps led to a more reliable chronology (note that the bootstrap confidence limits on the chronology values are also reduced – Figure GU1c,d). This would be a questionable conclusion, however, because the main reason for the apparently more accurately estimated RCS curve is the big increase in the number of modern samples from living trees (see replication in Figure GU1a,b). Although this big data sample apparently reduces RCS error, it is at the expense of introducing bias into the RCS curve. Rather than the RCS curve being determined by tree growth across a range of calendar year periods, and thus climate conditions, the dominant numbers of modern samples mean that the younger ages in the RCS curve are overly influenced by ring widths from 1600–1800 CE, while the older ages in the RCS curve are overly influenced by 20^{th} century ring widths. The need for a relatively even coverage of ring-width measurements over a long period of time for defining an RCS curve has been noted before (e.g. Briffa et al 1996), and the present example is an illustration of this issue. The larger-region RCS curve may have an apparently smaller sampling error, but it may be less applicable to the standardisation of measurement data in the sub-fossil period and may suppress part of the trend during the living-tree period (Briffa and Melvin, 2011).

A further disadvantage of building a single chronology from a network of sites across a larger-region, despite apparently reducing the size of sampling errors, is that the temporal consistency may be reduced. In the current example, the modern part of the larger-region chronology represents a much larger region than the sub-fossil part, with the latter based only on the Polar Urals and Yamal datasets that lie within about 5° of longitude. If the larger-region chronology is intended to represent average tree-growth for the whole of the larger region, then the bootstrap confidence limits on the pre-1600 CE part of the chronology (Figure GU1b) will be underestimates because the cross-correlations (and hence the common signal) during this period are overestimated when the data are taken from only a small part of the large region.

A related temporal inconsistency is that the variability of the larger-region chronology is reduced during the modern period relative to the earlier period because of the larger sample of more widely spaced data that are averaged together. The inconsistency carries over to the creation of a temperature reconstruction from such a chronology (though this step is not undertaken here). The calibration during the overlap with instrumental data will optimise the fit between the reduced-variance modern part of the chronology, but the calibration equation will not be applicable to the earlier, sub-fossil part of the chronology that has higher variance (because it is based on a smaller sample from a smaller region).

Some advantages (such as reduced sampling error) may arise from building an RCS-standardised chronology from ring-width data gathered from a larger region, but the examples presented here illustrate some potential drawbacks. For the current study of the Yamalia region, the problems of a biased RCS curve and a temporally inconsistent chronology clearly outweigh the benefits of reduced sampling error, and therefore this strategy is not adopted in the main part of this study.

Table GU1. Details of the 17 Greater Urals sites from which we used data in this inter-comparison, including site name, altitude, grid locations, years spanned by data and species.

FILE	ALT	NORTH	EAST	START	END	SPECIES
NONBLASI.raw	70	6536	5038	1603	1990	LASI
KEDVLANO.raw	70	6415	5334	1674	1991	LASI
SHCHLA.raw	60	6613	5620	1630	1990	LASI
Russ002.rwl	550	6440	6000	1691	1969	LASI
KOZHLASI.raw	400	6527	6035	1588	1990	LASI
MUZYLASI.raw	30	6519	6439	1828	1994	LASI
Russ001.rwl	150	6650	6545	1541	1968	LASI
POU_LA.raw	250	6652	6538	914	1990	LASI
POLURULA.raw	250	6652	6538	778	1846	LASI
PLL.rwl	20	6545	6893	1557	1992	LASI
YamalAD.raw	-999	6500	6900	-200	1996	LAGM
KHADYTLA.raw	90	6712	6950	1782	1990	LASI
PLR.rwl	12	6536	6952	1609	1992	LASI
PDP.rwl	12	6535	6952	1505	1992	LASI
PLO.rwl	12	6535	6952	1684	1992	LASI
NADILASI.raw	80	6613	7140	1740	1990	LASI
KHEYLANA.raw	100	6523	7252	1767	1990	LASI

Table GU2. Some basic statistics for the 17 sites.
Indices created by standardising with a 30-year spline.
Corr - the mean correlation of each index series with the chronology (excluding that series).
Rbar - the mean inter index-series correlation.
MnRaw - the mean value of all measurements for the site.
File extension ".raw" and ".rwl" indicate ring-width measurements

File	Cores	Start	End	Years	Rings	Corr	Rbar	MnRaw
NONBLASI.raw	22	1603	1990	388	6822	0.77	0.63	0.537
KEDVLANO.raw	29	1674	1991	318	7487	0.78	0.63	0.608
SHCHLA.raw	26	1630	1990	361	5542	0.82	0.68	0.498
russ002.rwl	20	1691	1969	279	3789	0.79	0.66	0.328
KOZHLASI.raw	32	1588	1990	403	9514	0.85	0.73	0.422
MUZYLASI.raw	26	1828	1994	167	3410	0.80	0.67	1.079
russ001.rwl	21	1541	1968	428	4901	0.75	0.61	0.328
pou_la.raw	93	914	1990	1077	13929	0.78	0.66	0.476
polurula.raw	32	778	1892	1115	5361	0.68	0.54	0.655
PLL.rwl	27	1557	1992	436	8646	0.77	0.62	0.345
yamalad.raw	252	-202	1996	2199	40893	0.79	0.63	0.615
KHADYTLA.raw	34	1782	1990	209	3872	0.82	0.68	0.686
PLR.rwl	38	1609	1992	384	11508	0.76	0.60	0.452
PDP.rwl	11	1505	1992	488	4564	0.77	0.64	0.429
PLO.rwl	20	1684	1992	309	3984	0.75	0.59	0.408
NADILASI.raw	34	1740	1990	251	4998	0.82	0.70	0.474
KHEYLANA.raw	29	1767	1990	224	5369	0.79	0.64	0.736
polaresp.raw	125	778	1990	1213	19290	0.76	0.61	0.526
polaresp_nor.raw	104	872	1990	1119	15873	0.77	0.64	0.457
urals.raw	746	-202	1996	2199	144461	0.69	0.44	0.536
urals_nor.raw	725	-202	1996	2199	141060	0.69	0.44	0.529

Table GU3. Cross-correlations (*100) between the 17 site chronologies.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
1		74	59	45	55	52	33	30		34	31	39	40	41	36	33	24	Nonblasi
2	74		50	46	50	46	31	27		29	34	39	41	42	33	30	28	Kedvlano
3	59	50		51	61	61	50	50		52	46	48	51	48	57	51	44	Shchla
4	45	46	51		60	49	59	55		30	52	53	36	40	29	41	36	russ002
5	55	50	61	60		63	59	54		55	54	51	57	56	54	46	43	Kozhlasi
6	52	46	61	49	63		67	62		62	67	62	63	58	63	58	45	Muzylasi
7	33	31	50	59	59	67		82		49	80	80	58	58	57	68	55	russ001
8	30	27	50	55	54	62	82		76	48	74	68	54	52	56	63	46	Pou_la
9								76			69			40				Polurula
10	34	29	52	30	55	62	49	48			55	49	77	74	82	67	67	PLL
11	31	34	46	52	54	67	80	74	69	55		88	61	60	62	69	54	yamalad
12	39	39	48	53	51	62	80	68		49	88		57	58	53	69	50	Khadytla
13	40	41	51	36	57	63	58	54		77	61	57		85	83	80	67	PLR
14	41	42	48	40	56	58	58	52	40	74	60	58	85		75	77	73	PDP
15	36	33	57	29	54	63	57	56		82	62	53	83	75		68	61	PLO
16	33	30	51	41	46	58	68	63		67	69	69	80	77	68		71	Nadilasi
17	24	28	44	36	43	45	55	46		67	54	50	67	73	61	71		Kheylana

Chronologies were created using 30-year spline, signal-free standardisation. The period used for each correlation is the common period between that pair of sites where both sites have 5 or more samples (see Utab1.prn for details). Note: Polurula has no living-tree data and has little overlap with most other sites.

Figures GU1 to GU3 compare bootstrap confidence limits for chronologies created from the Greater Urals data (Urals.raw, Urals_nor.raw) and the Esper et al. (2002) larch data (Polaresp.raw and Polaresp_nor.raw) both with and without the root-collar samples (where "_nor" indicates files without root-collar samples). The four data sets were standardised using simple RCS (i.e. a single RCS curve and not signal-free RCS). The process was repeated 1000 times, in each case using random selection with replacement of the number of cores in the original file. The 2.5% lowest and 2.5% highest values of the 1000 bootstrap chronologies are used as the upper and lower 95% bootstrap confidence limits for each chronology.

Figure GU1. Shows the RCS curve created from the full chronology (red) and the bootstrap confidence limits (cyan) for each chronology: a) Greater Urals including root-collar samples, b) Greater Urals excluding root-collar samples, c) Polar Urals including roots, and d) Polar Urals excluding roots. Sample counts for each are shown with grey shading. The error range is small in all cases, though smallest for the Greater Urals because the sample counts are large. The inclusion/exclusion of the root-collar samples makes little difference to either the RCS curve or its bootstrap range (compare (a) and (b)), because of the large number of samples. The Polar Urals RCS curve without root-collar samples (d) has slightly narrower confidence limits that with root-collar samples (c), and the RCS curve itself is also affected (e.g. lower values for ring ages from 30 to 250 years without root-collar samples). This is a systematic bias that is not evident in the Greater Urals RCS curves.

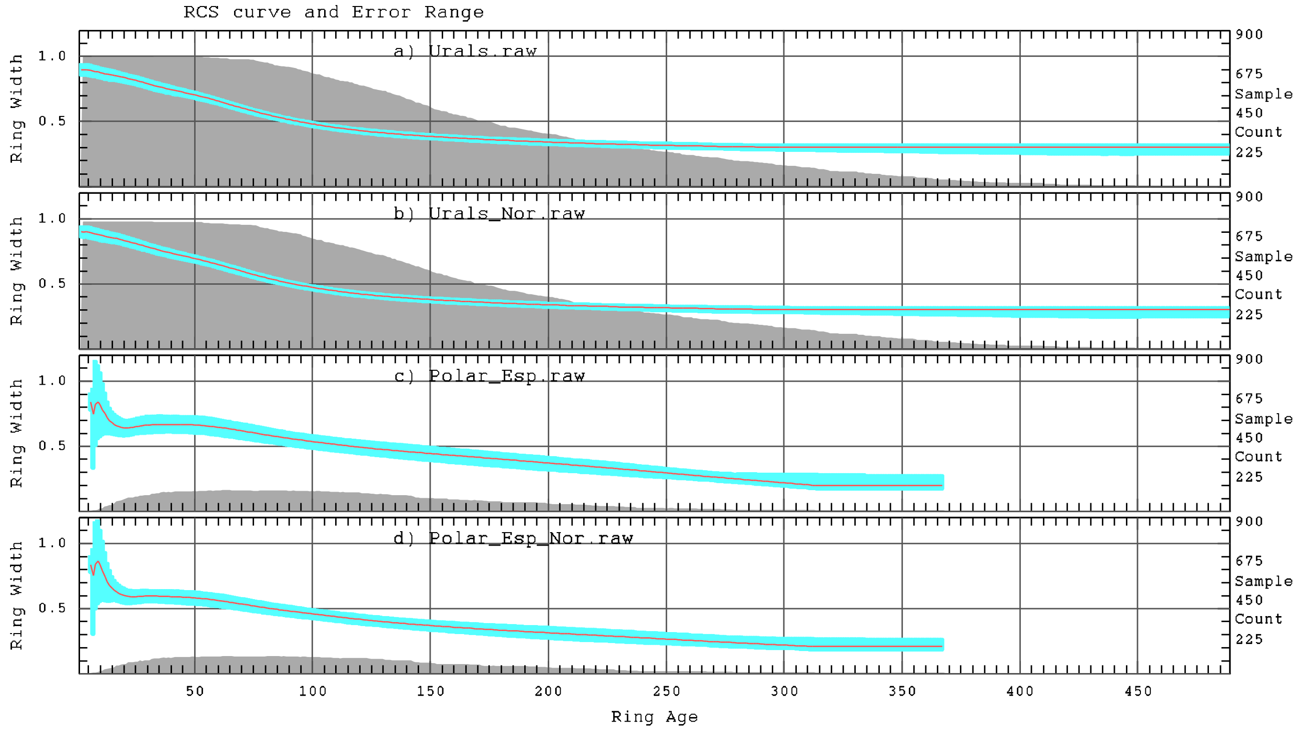


Figure GU2. Shows the full chronology (red) and the 95% bootstrap confidence limits (cyan) for the four chronologies: a) Greater Urals including root-collar samples, b) Greater Urals excluding root-collar samples, c) Polar Urals including roots, and d) Polar Urals excluding roots. Sample counts for each are shown with grey shading. The error range is much larger prior to 1600 in all cases, reflecting reduced sample counts. The bootstrap confidence limits for the Polar Urals chronology are larger than those for the Greater Urals chronologies due to the greater number of samples in the Greater Urals chronology. The inclusion/exclusion of the root-collar samples has a greater influence on the chronology and its bootstrap confidence limits for the smaller Polar Urals dataset, though there are still some small differences even for the Greater Urals chronology.

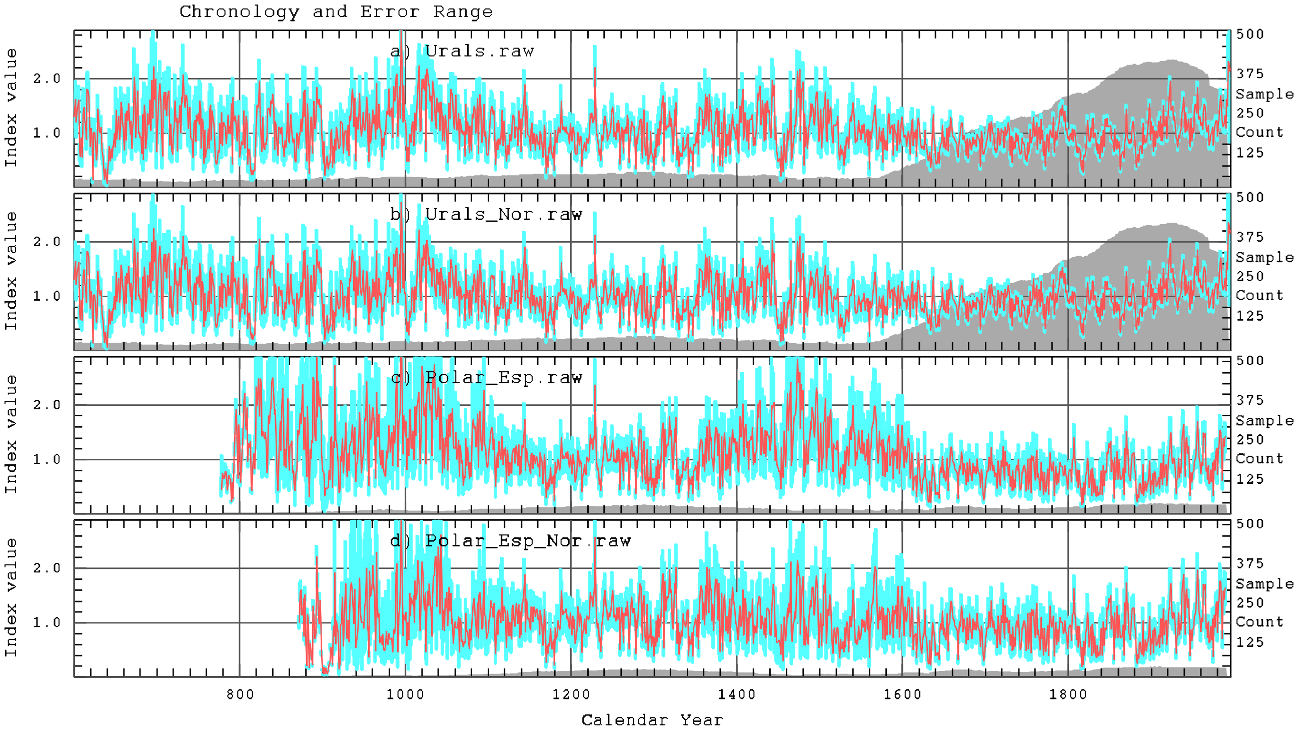


Figure GU3. Shows the same chronologies and bootstrap confidence limits as Figure GU2, but the confidence limits are now plotted about the mean of 1.0. For the two Polar Urals chronologies the uncertainty range is mostly greater than the range of chronology values prior to 1100 and between 1400 to 1600 in (c) because of the root-collar material and in (d) because of reduced sample counts. The Greater Urals chronology has smaller confidence limits because of the presence of the Yamal trees (and of course the many additional modern samples after 1600).

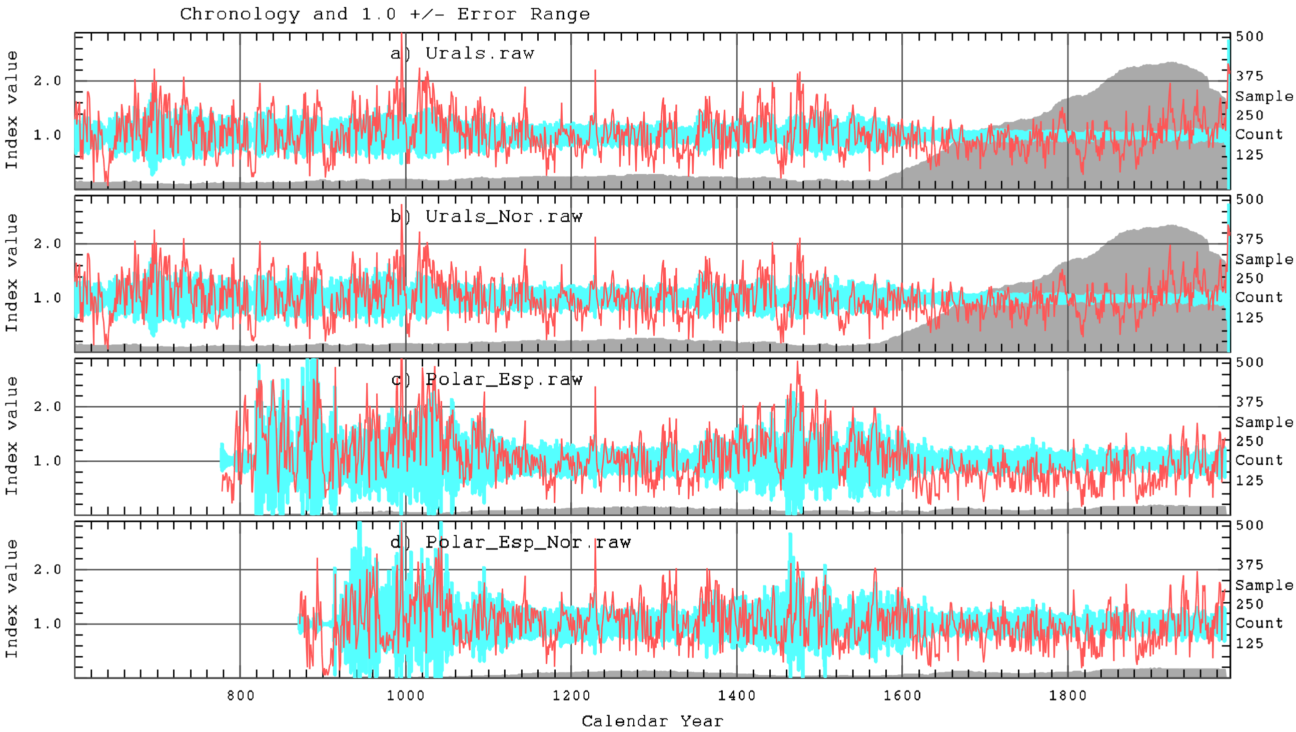


Figure GU4. For Figures GU4 to GU5, the data in the four files (Polaresp.raw and Polaresp_nor.raw, where "_nor" indicates files without root-collar samples) were standardised using one-curve, signal-free RCS. Chronologies were normalised (subtract mean and divide by standard deviation) over the period 1600 to 1990 CE. The chronologies with root-collar samples (blue) and without root-collar samples (red) are shown in a). The chronologies, smoothed with a 50-year spline, are shown in b). Annual values of two standard errors for each chronology are shown in c) where standard error is calculated as the standard deviation of indices for that year scaled by the square root of sample count. The presence of root-collar samples makes a large difference to the part of the chronology derived from sub-fossil samples.

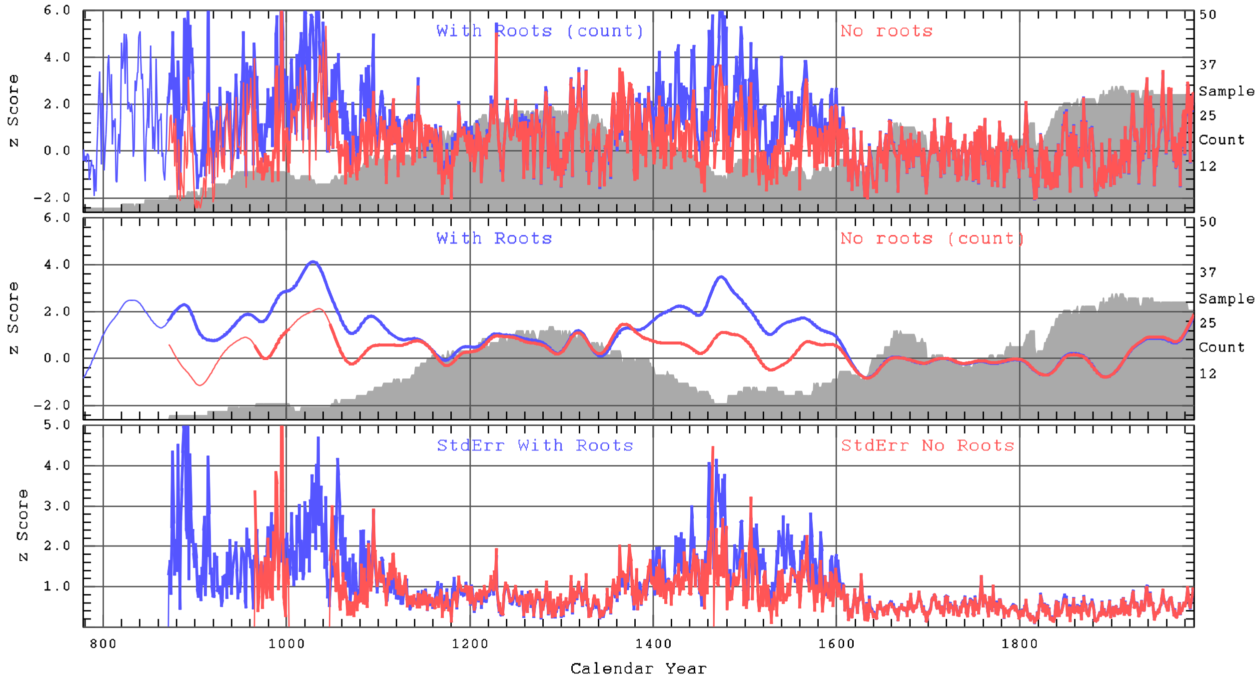


Figure GU5 as for GU4 but using Urals.raw and Urals_nor.raw data sets. There is a difference between chronologies created with and without root-collar samples.

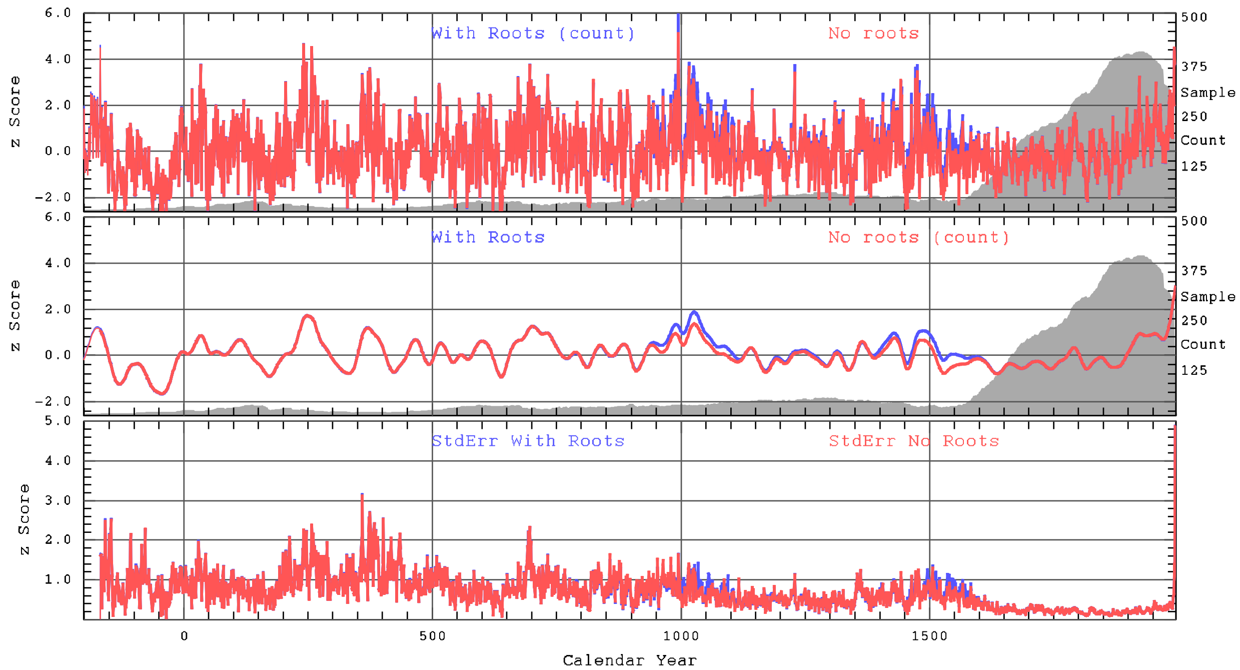


Figure GU6 shows rbar and adjusted, effective EPS for Greater Urals larch samples with the root-collar samples, processed using one-curve, signal-free RCS.

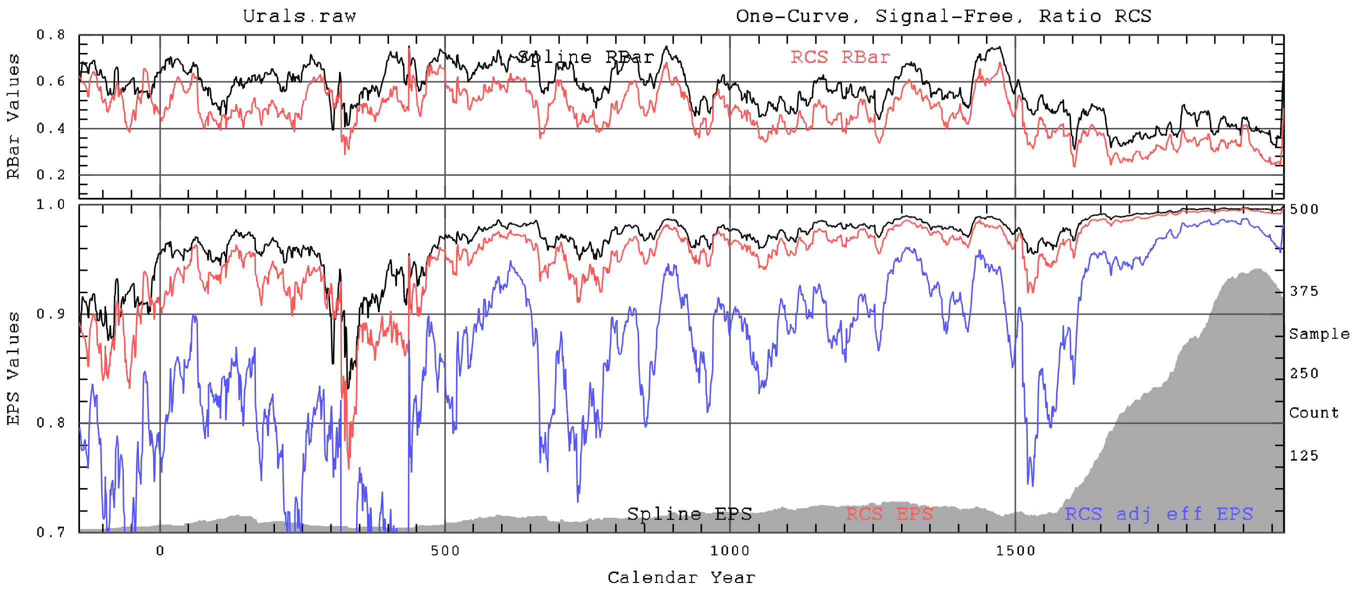
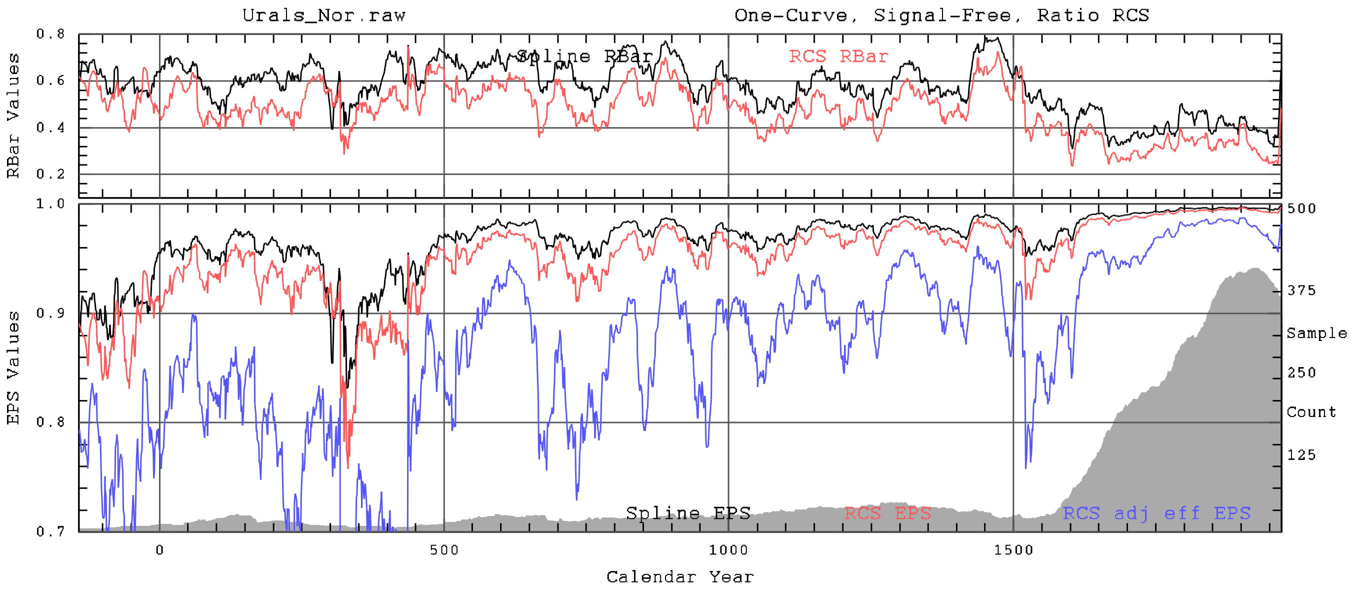


Figure GU7 shows rbar and adjusted, effective EPS for Greater Urals larch samples without the root-collar samples, processed using one-curve, signal-free RCS.



Comment: For the Greater Urals samples the adjusted effective EPS dips below 0.85 prior to 1100 and in the 17th century for both with- and without-root-collar data sets. Removing the root-collar samples has little effect on the rbar and slightly improves the EPS. The rbar is lower after 1600, when samples come from the very wide Greater Urals region, but because counts are extremely high in this period the EPS is maintained at about 0.95.

Figure GU8 shows rbar and adjusted, effective EPS for Esper et al. (2002) Polar Urals larch samples with the root-collar samples, processed using one-curve, signal-free RCS.

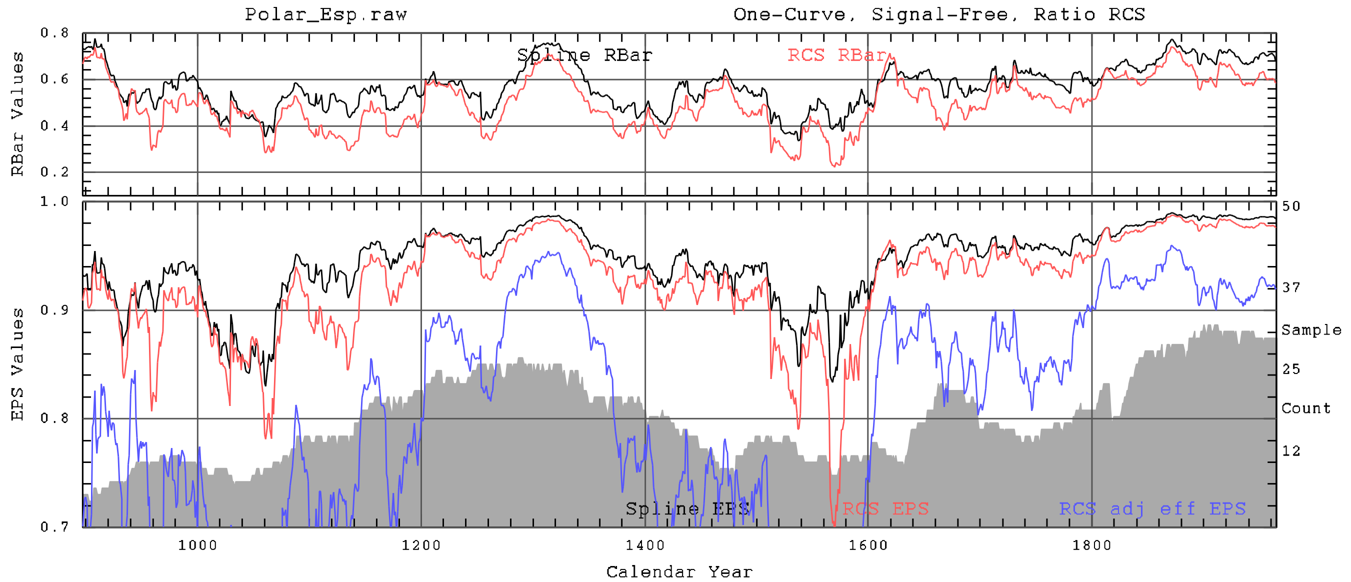
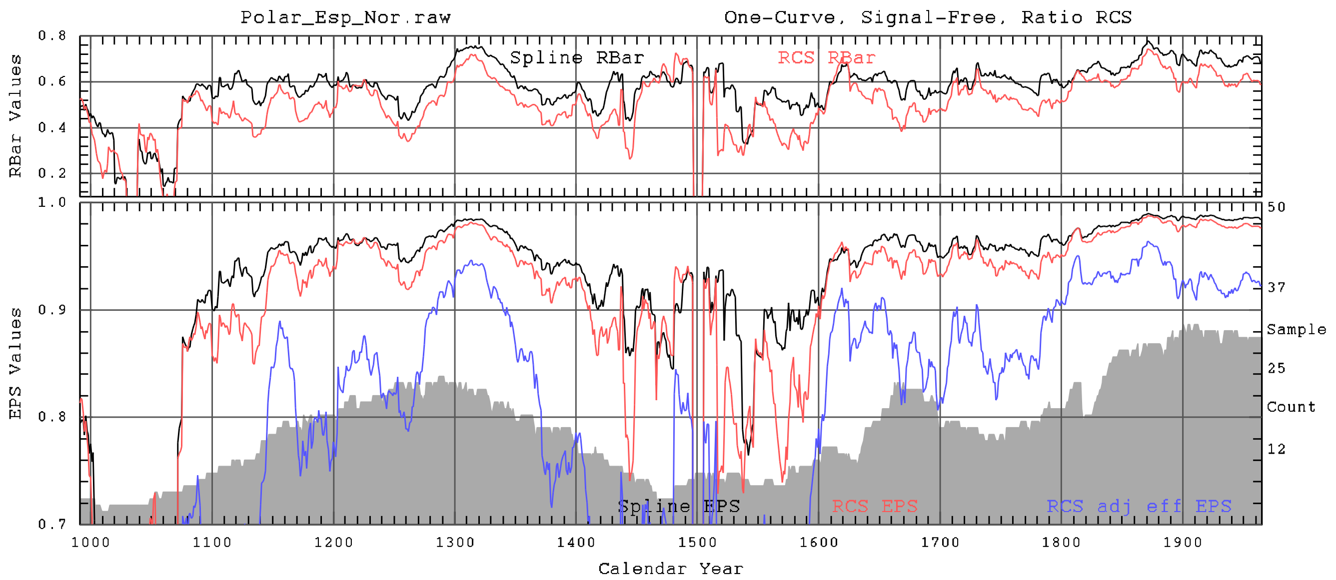


Figure GU9 shows rbar and adjusted, effective EPS for Esper et al. (2002) Polar Urals larch samples without the root-collar samples, processed using one-curve, signal-free RCS.



Comment: For the Esper et al. (2002) larch samples the adjusted effective EPS is well below 0.85 prior to 1150 and between 1350 and 1600 for both with and without root-collar data sets. Removing the root-collar samples improves the rbar slightly but the reduction of the sample count has a contrary effect. The EPS of the Esper et al. (2002) Polar chronology is poor prior to 1600.