Tracheids vs. Tree Rings as Proxies for Dendroclimatic Reconstruction at High Altitude: The Case of *Pinus sibirica* Du Tour

Mikhail S. Zharkov 1,*, Bao Yang 2, Elena A. Babushkina 1,3, Dina F. Zhirnova 1,3, Eugene A. Vaganov 1,4 and Vladimir V. Shishov 1,*

Abstract: Siberian pine (*Pinus sibirica* Du Tour) is a widespread and long-lived species in the northern hemisphere, which makes it a good potential proxy for climatic data. However, the tree-ring growth of this species weakly correlates with climatic conditions, which prevents its use in dendroclimatic reconstruction. It was proposed to use the measurements of tracheid characteristics as model predictors to reconstruct the smoothed temperature of the key periods in tree growth. In this study, algorithms for preprocessing tracheids and temperature data, as well as for model cross-validation, were developed to produce reliable high-resolution (weekly-based) temperature reconstructions. Due to the developed algorithms, the key time periods of Siberian pine growth were identified during the growing season—early June (most active cell development) and mid-July (setting new buds for the next growing season). For these time periods, reliable long-term temperature reconstructions ($R^2 > 0.6, p < 10^{-8}$) were obtained over 1653–2018. The temperature reconstructions significantly correlated ($p < 10^{-8}$) with independent reanalysis data for the 19th century. The developed approach, based on preprocessing tracheid and temperature data, shows new potential for Siberian pine in high-resolution climate reconstructions and can be applied to other tree species that weakly respond to climate forcing.

Keywords: wood anatomy; cell measurements; radial cell diameter; cell wall thickness; tracheidograms; tree-ring response; poorly sensitive to climate; temperature reconstruction

1. Introduction

Siberian pine forests are the most complex ecosystems in the Siberian taiga, characterized by regenerative-age dynamics, stability, spatial and temporal structure, and biodiversity [1]. The distribution range of Siberian pine (*Pinus sibirica* Du Tour), a forest-forming species of the “cedar” forest formation, extends from the northeast of European Russia to the south of East Siberia, reaching Mongolia in the southern part (http://agroatlas.ru/ru/content/related/Pinus_sibirica/map/, accessed on 27 June 2023).

The observed global warming of recent decades has contributed to a shift in the timespan of the active growing season and its heat availability [2–4]. The consequences of these changes are most acute at the edges of the growing range, where trees often grow at the limits of physiological endurance [5,6]. Moreover, the responses of tree plants to climate change are mixed, from range expansion to suppression or death [7–11]. Under current climate changes, it is therefore critical not only to understand the response of tree vegetation as an important part of the planetary carbon cycle [12–15] but also to obtain a quantitative assessment through robust reconstructions of long-term climatic fluctuations.
The use of long-term chronologies of the anatomical parameters of tree rings allows one to obtain a more accurate and detailed understanding of past and recent variability in principal climatic factors compared to simply analyzing tree-ring width (TRW) [16–19]. The seasonal development of xylem anatomical structure significantly determines tree productivity and survival; hence, forests’ vulnerability to climate change and their ability to fix carbon dioxide [20,21]. This approach also enhances opportunities to study the physiological mechanisms of plant adaptation to changing environmental conditions, due to changes in both hydraulic and mechanical functions of woody tissue [22–24]. Therefore, understanding how and at which intervals of the growing season the principal climatic factors (temperature or precipitation) modify the tree-ring structure turns out to be important both for reconstructing past climate-tree ring relationships and developing adequate prediction models of climatic factors influencing the anatomical structure of tree rings [24–29]. The application of the quantitative wood anatomy approach is particularly relevant for the species (i.e., *Pinus sibirica* Du Tour) for which the study of climate response is extremely difficult due to the limited climate sensitivity of their radial growth [30,31].

A unique 495-year chronology of wood anatomical characteristics for Siberian pine has recently been obtained for the timberline in the Western Sayan, which has made it possible to estimate the climatic response of tree-ring widths and integral anatomical characteristics (namely, TRW indices, mean and maximum radial cell diameters, and cell wall thickness) [19].

However, the dataset of detailed anatomical measurements used in this study, namely the radial cell diameter and cell wall thickness of *Pinus Sibirica* tracheidograms, allows us to estimate the century-old records of climate factors with much higher time resolution. In this work, we exploit the full potential of the anatomical structure of tree rings in a detailed analysis of their climatic response, with a focus on dendroclimatic reconstructions, using previously developed [32] and new approaches.

In this study, we employ tracheid measurements with the objective of revealing the potential of *Pinus sibirica* Du Tour for high-resolution temperature reconstructions. Algorithms for tracheid and temperature preprocessing and accurate model evaluation were developed. Additionally, two hypotheses were tested: (1) key periods of tree growth can be determined by the quality of the smoothed daily temperature reconstruction models, and (2) the use of TRW for modeling would not achieve the same reconstruction reliability as using tracheid data.

2. Materials and Methods

2.1. Climate Data

Temperature and precipitation data were obtained from the Tashtyp meteorological station (WMO 29956), 52.8 N, 89.9 E, 455 m a.s.l., and they included mean daily temperature and cumulative daily precipitation records from 1929 to 2016. Mean daily NOAA/CRES/DOE 20th Century Reanalysis V3 temperatures were obtained from the Climate Explorer web platform (http://climexp.knmi.nl/select.cgi?field=c3t2m_daily, accessed on 29 August 2023), using the grid box region: longitude from 89.5 to 90.5 E, latitude from 52.5 to 53.5 N (Figure 1).

To create a dependent variable for the reconstruction from the raw temperature data, a two-way sequential smoothing algorithm was applied (Figure 2).

To suppress high-frequency fluctuations in the mean daily temperature, intra-annual smoothing with varying lengths $w$ of the sliding window (from 1 up to 14 days) was applied (Figures 2 and S1). The optimal length of the sliding window was determined by optimization of the temperature reconstruction models as described below.

In addition to the intra-annual smoothing, an inter-annual smoothing procedure for temperature was applied to develop reconstruction models of the smoothed temperature based on anatomical characteristics. The second smoothing was done to reduce annual variance in temperature. We used a moving average with a sliding window of $W$ years (from 1 to 11 years) (Figures 2 and S2), using the same criteria to choose the length of $W$. 
Figure 1. The locations of the study plot (brown pinecone for *Pinus sibirica* Du Tour) and the climate station Tashtyp (red triangle) (A, B), and the average monthly mean temperature (T, °C) and monthly total precipitation (P, mm) in Tashtyp from 1929 to 2016 (C).

Figure 2. The two-way sequential smoothing algorithm for obtaining the dependent temperature variable $T_{\text{inter}}(w, W, \text{doy})$, where $w$ is the length of the intra-annual sliding window; $W$ is the length of the inter-annual sliding window; $\text{doy}$ is the day of the year for which the reconstruction model is developed.
For each pair \((w, W)\) of intra- and inter-annual sliding windows, the set \(DOY_{sig}(w, W) \subseteq \{152, \ldots, 243\}\) was created as a subset of tree-ring growing days (DOYs, growing season). \(DOY_{sig}(w, W)\) is the set of DOYs for which the Pearson correlation between inter-annual and intra-annual smoothed temperatures was highly significant \((p < 0.001)\).

As a result of the two-way sequential smoothing algorithm, we obtained dependent variables \(T_{inter}(w, W, \text{doy})\) of inter-annual smoothed temperatures for the model.

### 2.2. Tree Data Collection and Processing

Wood samples were collected from seven trees of *Pinus sibirica* Du Tour located on the border of the Republic of Khakassia and the Republic of Tyva (51°42.8’ N 89°51.9’ E, 1970–2020 m a.s.l.) (Figure 1). Tracheidogram measurements of tree rings over 1653–2018, obtained earlier (see [19] for more details), were used in this work (Figure S3).

For each tree ring, lumens (Ls) and double cell wall thickness (DCWT) were measured for 5 rows of cells (Figure S3). Then, the radial cell diameter (Cell Diameter, D) was calculated as the sum of lumen and double cell wall, and cell wall thickness (CWT) was determined as half of DCWT [33]. The obtained measurement were verified by an independent cell measuring tool [34].

The resulting tracheidograms of averaged D and CWT from the five rows of the measured series (Figure S4A,C) were then standardized to 15 cells (mean seasonal cell production over 1653–2018), resulting in standardized sD and sCWT series for each year of each tree [35]. Absolute tree ring width (TRW) values were obtained by summarizing DMeans of the corresponding rings.

The TRW values of all trees in a year were averaged to obtain the site tree-ring chronology. A standardized chronology was developed by bi-weight robust averaging individual tree indices, which were obtained by removing the age-related trends using cubic smoothing splines with a 50% frequency response at 67% of the series length [36].

The standardized tracheidograms of individual trees were year-to-year averaged (Figure S5). As a result, a single 30-dimensional object was obtained for each year (growing season), consisting of the values of cell diameters (D1–D15) and cell wall thicknesses (CWT1–CWT15) of the corresponding standardized curves (Figure S6).

Finally, 15 radial-cell-diameter and 15 cell-wall-thickness chronologies were obtained over 1653–2018 (Figure S7). Figure S8 shows examples of smoothed (“inter-seasonal smoothing”) chronologies of radial cell size and cell wall thickness. All the smoothed time series had significant Pearson correlations \((p\text{-value} < 10^{-16})\) with D1–D15, CWT1–CWT15 series for all the sliding windows.

Principal component analysis (PCA) was applied to the radial-cell-size and cell-wall-thickness chronologies to reduce the dimensionality of the resulting 30-dimensional objects. It was shown that 4 principal components explained 95% of the variance of the series, while 9 principal components explained 99% of the variance (Figure S10). In addition, the application of PCA allows one to avoid multicollinearity in the obtained series, as noted above [37].

The first \(P\) principal components, smoothed with the inter-annual window \(W\) (see the example for \(P = 5\), Figure S9):

\[
PC_i(W) = \{pc_i^W(y) | y \in \{1653 \ldots 2018\}\}, \ i \in [1, \ldots, P]
\]

were used as predictors (independent variables) in the model development process \((P \in [4, \ldots, 9])\). The principal component transformation matrix is presented in Table S1.

### 2.3. Reconstruction Models Development

For the best model fit, a triplet of hyperparameters \((w, W, P)\) was varied as follows: \(w \in \{1, \ldots, 14\}, W \in \{1, \ldots, 11\}, P \in \{4, \ldots, 9\}\).
For each triplet \((w, W, P)\), we obtained the set of independent variables \(\{PC_1(W), \ldots, PC_P(W)\}\) (Figure 3) and the set of dependent variables \(T_{\text{inter}}(w, W, \text{doy}), \text{doy} \in \text{DOY}_{\text{sig}}(w, W)\) (Figure 2).

For each \(\text{doy} \in \text{DOY}_{\text{sig}}(w, W)\), a separate multiple linear regression (MLR) model \(\text{MLR}_{w, W, P, \text{doy}}(\text{year})\) was developed.

The final MLR models were considered as ensembles of individual MLR models obtained in a new rolling leave-one-out cross-validation (RLOO CV) procedure:

\[
\text{MLR}_{w, W, P, \text{doy}}(\text{year}) = k_0 + \sum_{l=1}^{P} k_l \cdot PC_l^W(y),
\]

where \(k_l = k_l(w, W, P, \text{doy})\) is the \(l\)th coefficient of the final MLR model \((l = 0, P)\), estimated as \(k_l = \frac{\sum_{w=W} k_l^0}{N}\).

We note that \(k_l^0 = k_l^0(w, W, P, \text{doy})\) is the \(l\)th coefficient of the \(\theta\)th individual MLR model, and \(N = 2016 - 1929 + 1 = 88\) is the total number of individual MLR models obtained in the RLOO CV procedure.

**Figure 3.** The reduction of tracheidograms to smoothed principal components algorithm for obtaining independent variables from the raw tracheid data (Diameters (Ds) and Cell Wall Thicknesses (CWTs)).
Each \( \theta \)th individual MLR model can be described by this formula:

\[
MLR_{w, W, P, \text{doy}}(\text{year}) = k_0^\theta + \sum_{i=1}^{P} k_i^\theta \cdot pc_i^W(y), \quad \theta = 1929, 2016
\]  

(3)

To obtain the \( \theta \)th individual MLR model, \( T_{\text{inter}}(w, W, \text{doy}) \) was split into calibration \( (T_{\text{cal}}^{\text{cal}}(w, W, \text{doy})) \) and verification \( (T_{\text{inter}}^{\text{ver}}(w, W, \text{doy})) \) sets by the rules of the RLOO CV procedure:

1. Each \( \theta \)th element (year) of \( T_{\text{inter}}(w, W, \text{doy}) \) is considered as a verification set.
2. The elements from \( \left[ \theta - \left\lfloor \frac{W}{2} \right\rfloor, \theta \right) \cup \left( \theta, \theta + \left\lfloor \frac{W}{2} \right\rfloor \right] \) are omitted (\( \left\lfloor \frac{W}{2} \right\rfloor \) is the floored division). This is done to prevent the data from the \( \theta \)th element from getting into the calibration set due to smoothing with the \( W \) inter-annual sliding window and affecting the elements from \( \left[ \theta - \left\lfloor \frac{W}{2} \right\rfloor, \theta \right) \cup \left( \theta, \theta + \left\lfloor \frac{W}{2} \right\rfloor \right] \). All indices from \( \theta - \left\lfloor \frac{W}{2} \right\rfloor < 1929 \) or \( \theta + \left\lfloor \frac{W}{2} \right\rfloor > 2016 \) are omitted.
3. All other elements are considered as a calibration set.

In this study, the RLOO CV procedure was developed as an extension of the LOO CV procedure [38] for smoothed data.

After obtaining the calibration and verification sets, the coefficients \( k_i^\theta \) of the \( \theta \)th individual MLR model are obtained by training the model on the calibration set.

To evaluate the individual models on the calibration sets, the coefficient of determination \( (R_{\text{cal}}^2, \beta) \) and the Root Mean Squared Error \( (\text{RMSE}_{\text{cal}}, \beta) \) were calculated between \( T_{\text{inter}}(w, W, \text{doy}) \) and \( \{ MLR_{w, W, P, \text{doy}}(\text{year}) \mid \text{year} \in T_{\text{cal}}(w, W, \text{doy}) \} \).

After training the \( N = 88 \) models, one for each year, the chronology of the verification values was obtained as:

\[
\text{CRN}^{\text{ver}}_{w, W, P, \text{doy}} = \left\{ MLR_{1929}^{1929}(w, W, P, \text{doy}), \ldots, MLR_{2016}^{2016}(w, W, P, \text{doy}) \right\}
\]  

(4)

and the mean metrics \( R_{\text{cal}}^2 = \frac{\sum_{w=1929}^{2016} R_{\text{cal}}^2}{} \), \( \text{RMSE}_{\text{cal}} = \frac{\sum_{w=1929}^{2016} \text{RMSE}_{\text{cal}}}{N} \) were calculated to evaluate the total quality of the individual models on the calibration set.

To evaluate the individual models on the verification set, \( R_{\text{ver}}^2 \) and \( \text{RMSE}_{\text{ver}} \) were calculated between \( \text{CRN}^{\text{ver}}_{w, W, P, \text{doy}} \) and \( T_{\text{inter}}(w, W, \text{doy}) \).

After individual evaluation, the final MLR model \( MLR_{w, W, P, \text{doy}}(\text{year}) \) was developed by averaging the coefficients of the individual models.

To evaluate the final model, \( R_{\text{sim}}^2 \) and \( \text{RMSE}_{\text{sim}} \) (\( \text{sim—simulated} \)) were calculated between \( T_{\text{inter}}(w, W, \text{doy}) \) and \( \{ MLR_{w, W, P, \text{doy}}(\text{year}) \mid \text{year} \in T_{\text{inter}}(w, W, \text{doy}) \} \).

All data processing algorithms were implemented in Python and can be downloaded from: https://github.com/mikewellmeansme/dendroclimatic-reconstructor/ (accessed on 9 January 2024).

3. Results

3.1. Reconstruction of Temperature Dynamics

We chose \( R^2 > 0.5 \) as a threshold for the selection of qualitatively reconstructed periods on the calibration and verification sets to be sure that the model explained most of the variance. After varying the triplet of hyperparameters \((w, W, P)\), the next heatmap was obtained (Figure 4). Obviously, reliable reconstructions appeared when the annual smoothing windows \( W \) was not less than 8 years.
The approach of choosing the optimal triplet \((w, W, P)\) of hyperparameters is debatable, and different combinations of the triplet values sometimes result in the reconstruction of different days of the growth season. Mainly, temperatures at the beginning of June (weeks with DOY 152–154 at the center) and the middle of July (weeks with DOY 195-198 at the center) are adequately simulated by the tracheidograms.

There are also a couple of models for late June \((P = 4, \text{DOY 175})\) and August \((\text{DOY 220 and 232})\), but these results are not sustainable because they do not appear with other triplets of hyperparameters.

As an example, we chose the triplet \(w = 7, W = 9, P = 5\) to demonstrate the applicability of our approach to dendroclimatic reconstruction of summer temperatures (Figure 5).

\[
\text{Figure 4. Heatmap showing the number of models for which } R^2 \text{ on the calibration and verification sets was greater than 0.5 for different pair of hyperparameters } W \text{ (inter-annual sliding window) and } P \text{ (number of first principal components), and } w = 7 \text{ (7-day intra-annual sliding window of temperatures).}
\]

\[
\text{Figure 5. The values of } R^2 \text{ metric per day of the year (DOY) on the calibration (black line) and verification (red line) sets for the models of temperature reconstruction with } w = 7, W = 9, P = 5. \text{ The dotted gray line shows the threshold for selecting acceptable reconstructed periods } (R^2 > 0.5 \text{ on both calibration and verification sets). The blue lines show DOYS with acceptable metrics. The gaps in the curves show DOYS where the intra-annual temperatures did not correlate with their smoothed curves (see Figure S1).}
\]
We compared the new reconstructions with the observed and reanalyzed temperatures for two main sustainable periods: A—DOY 154 (the mean daily temperature from 30 May to 5 June) and B—DOY 197 (the mean temperature of the week from 12 July to 18 July) (Figure 6).

![Smoothed (W = 9, 1 year step) mean Tashtyp temperature of two periods: DOY 154 (A) and DOY 197 (B). Observed (red line), reconstructed (black line), reanalysis V3 temperatures (blue line).](image)

**Figure 6.** Smoothed (W = 9, 1 year step) mean Tashtyp temperature of two periods: DOY 154 (A) and DOY 197 (B). Observed (red line), reconstructed (black line), reanalysis V3 temperatures (blue line).

### 3.2. Reconstruction Model Evaluation

For each model, we obtained $R^2$ and RMSE metrics for the calibration set—performance of the models on the data on which they were trained; and for the verification set—performance of the models on the data that were not used in the training process. We also calculated simulated metrics—final performance of the models on the whole dataset. The obtained metrics are shown in Table 1.

<table>
<thead>
<tr>
<th>Period</th>
<th>DOY</th>
<th>$R^2$ Calibration</th>
<th>$R^2$ Verification</th>
<th>$R^2$ Simulated</th>
<th>RMSE Calibration</th>
<th>RMSE Verification</th>
<th>RMSE Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>154</td>
<td>0.74 ± 0.03</td>
<td>0.60</td>
<td>0.74</td>
<td>0.48 ± 0.15</td>
<td>0.60</td>
<td>0.48</td>
</tr>
<tr>
<td>B</td>
<td>197</td>
<td>0.79 ± 0.03</td>
<td>0.60</td>
<td>0.78</td>
<td>0.46 ± 0.15</td>
<td>0.63</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Since different $R^2$ and RMSE were obtained for each iteration of the rolling cross-validation for the calibration set, Table 1 summarizes the mean values of the statistics on the calibration sample (±s.d.).

The obtained metrics show a reasonably high quality of the models, with more than 74% of the explained variance with RMSE less than 0.49 °C on the calibration set, and 60% of the explained variance with RMSE less than 0.63 °C on the verification set.

Figure 6 shows that, over the period of instrumental observations (1929–2016), the modeled (black line) and observed (red line) temperatures for both periods (A and B) have an extremely high degree of synchrony ($R > 0.8, p < 10^{-26}$, Table 2), as do the observed and reanalyzed temperatures (blue line) ($R > 0.96, p < 10^{-45}$, Table 2), but the reanalyzed data are significantly underestimated relative to the observed temperature (>1 °C) (see Figure 7 for more detail).
Table 2. The Pearson correlations (Rs) and their significance (p) between the observed, reanalyzed, and simulated temperatures.

<table>
<thead>
<tr>
<th>Years Period</th>
<th>Observed</th>
<th>Reanalyzed Observed</th>
<th>Simulated</th>
<th>Simulated Reanalyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pearson R</td>
<td>p-Value</td>
<td>Pearson R</td>
<td>p-Value</td>
</tr>
<tr>
<td></td>
<td>Pearson R</td>
<td>p-Value</td>
<td>Pearson R</td>
<td>p-Value</td>
</tr>
<tr>
<td></td>
<td>Pearson R</td>
<td>p-Value</td>
<td>Pearson R</td>
<td>p-Value</td>
</tr>
<tr>
<td>1836–1928 A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1836–1928 B</td>
<td>0.67</td>
<td>&lt;10^{-12}</td>
<td>0.57</td>
<td>&lt;10^{-8}</td>
</tr>
<tr>
<td>1929–2016 A</td>
<td>0.97</td>
<td>&lt;10^{-51}</td>
<td>0.86</td>
<td>&lt;10^{-26}</td>
</tr>
<tr>
<td>1929–2016 B</td>
<td>0.96</td>
<td>&lt;10^{-45}</td>
<td>0.89</td>
<td>&lt;10^{-29}</td>
</tr>
</tbody>
</table>

Due to their strong correlation with the observed temperature, despite the underestimation, the reanalyzed data can be taken as a reliable source of information on temperature trends. In this regard, the significant correlation (R > 0.56, p < 10^{-8}, Table 2) between simulated and reanalyzed temperatures for the years with no direct climate observations (1836–1928) indicates that the models are sufficiently correct in indicating temperature trends for the modeled periods (A and B). In this paper, reanalyzed data are used for additional assessment of model quality beyond the metrics from Table 1.

The reanalyzed data are consistently underestimated (from 1.35 to 1.78 °C) relative to the observed data. This can be explained by the altitude difference between Tashtyp and the stations on the basis of which the reanalyzed data were constructed. As the altitude increases by 100 m, the temperature may proportionally decrease by a value close to 0.5 °C [39], so it can be assumed that the reanalyzed data are constructed on the basis of the data from the stations located 300–400 m above Tashtyp.

We noted that the obtained tree-ring width chronology did not correlate (p > 0.05) with the smoothed observed temperature of the Tashtyp weather station for intra-annual intervals up to 14-day smoothing.

Figure 7. Scatterplots for the simulated/reanalyzed temperatures versus the observed temperatures obtained for two periods: DOY 154 (A,C) and DOY 197 (B,D) (1929 to 2016).
Moreover, the enumeration of inter-annual sliding windows (Ws) for the standard TRW chronology from the range of 1 to 11 years did not allow one to obtain any models with $R^2 > 0.5$.

4. Discussion

It is known that both the width of tree rings and their anatomical structure are controlled by internal and external factors [35,40–43]. However, when comparing the variance and sensitivity coefficient of tree-ring width chronologies for different species under the same growing conditions, Siberian pine is shown to have rather low values of growth variability indices, even under climatically limited growing conditions [44,45]. There is no doubt that the low variability in Siberian pine growth is the result of a greater contribution of internal factors to its seasonal and perennial growth. These factors include adaptations of its physiology to the slow accumulation and utilization of nutrients [46,47], adaptations to wetter growing conditions (average variability of growth in other similar species of Siberian taiga, Siberian spruce, and Siberian fir [48,49]), and other ecophysiological features [50]. It is possible that the minor response of Siberian pine growth to low temperatures at high elevations is related to its genome size, which is larger than that in most tree species [51,52].

However, the long lifespan (up to 800–900 years [50,53]), wide distribution, and high economic and ecological value of Siberian pine have determined the ongoing attempts of dendrochronologists to obtain centuries-long annals of its tree rings and decode the contribution of the dynamics of environmental conditions, including climate forcing, on the formed rings. One of the approaches, due to a greater number of low-frequency fluctuations compared to annual fluctuations in tree-ring growth, is the use of smoothed time series with a window width of several years. The effectiveness of this approach was demonstrated in this work, and corresponds well with previously published results [30,54].

Another approach is to search more sensitive indicators of climate variability among the parameters of annual rings, including the use of quantitative wood anatomy. The use of long-term chronologies of tree-ring anatomical characteristics to reconstruct long-term changing climates over centuries is a relatively new and rapidly developing research direction [27,28,55,56].

We hypothesized that the ability to reconstruct seasonal, weekly-based temperature changes could be enhanced by using cell chronologies of Siberian pine anatomical characteristics, which was confirmed by both previously published results [19] and the new temperature reconstructions developed in the current research. Since the effects of internal factors are fairly uniform from year to year, by special treatment and combining the available cell measurements, we can accumulate the effects of external climate forcing and potentially reveal “hidden” correlations with climatic variables that cannot be observed based on traditional dendroclimatology techniques. To realize this approach, we developed several procedures to treat initial climate and anatomical data (Figures 2 and 3). We also developed a new reconstruction procedure of seasonal temperatures (Figure 4) based on the principal component decomposition of cell measurements, specifically radial cell sizes and cell wall thickness.

The search for optimal windows for temperature generalization within a season and on a long-term scale allowed us not only to obtain reconstruction models that had high convergence with instrumental and reanalysis data but also established physiologically based regularities. The high first-order autocorrelation within a decade, accounting for the increased contribution of low-frequency oscillations and the possibility of preferential reconstruction of smoothed climatic series, can, at least partly, be attributed to the perennial needles of the species. It is known that needles of different species can persist for up to 14 years (e.g., spruce at the limit of distribution in the mountains; [57]). Siberian pine has a needle life span of 4–6 years at lower elevations and up to 9–10 years at higher elevations [58,59]. The age of needles of a related species, European cedar Pinus cembra, can reach 9–12 years [60,61].
The temperature intervals during the season, for which the reconstruction models showed the highest and most reliable statistical estimates, are important indicators of xylogenesis phenology. Thus, the first period in the first week of early June indicates the time interval of the most active growth of Siberian pine under certain local conditions, with sufficient humidity and increasing daylight but low heat availability. The second period (the week in the middle of July) corresponds to the interval of the end of primary growth and the budding for the development of shoots and needles for the next season [62–64]. The resulting models had quite high metrics. The metrics of the models based on tracheid data \((R > 0.86)\) are much higher than those of models based on tree ring widths \((R \in [0.2, 0.6])\), both in the current work and compared to standard metric values of models based on climate-insensitive trees [19,65]. The obtained metric values are comparable to those from recent works on temperature reconstruction based on quantitative wooden anatomy of climate-insensitive species \((R^2 > 0.74\) on the calibration set) [65], which suggests that the cellular indices of these species are robust in explaining 60%–75% of the variance in temperature reconstruction.

In the current work, to test the quality of the models, in addition to the verification sample, whose data did not participate in the model training process and therefore can be considered independent, we also used V3 reanalysis grid data, with the modeled temperature also significantly correlating with these data, both during the years with instrumental measurements \((R > 0.79)\) and outside them \((R > 0.57)\).

Nevertheless, the potential of the proposed approach to modeling short-term climatic fluctuations requires further research. Various transformations or standardization of anatomical measurements may, for example, suppress a part of the non-climatic signal (including size-age) or the climatic signal inherited from previous stages of xylogenesis [66]. Different approaches to selecting relevant predictors among PCs and identifying promising intervals for intra- and inter-annual averaging, such as wavelet analysis, are worth exploring. The proposed approach can also be widely tested on new data from different forest stands with the same or other tree species that have low sensitivity to climate, as well as in different climatic zones.

5. Conclusions

The use of tracheid measurements has made it possible to realize the potential of Pinus sibirica Du Tour for high-resolution temperature reconstruction.

Using the developed algorithms, reliable models for temperature reconstruction were obtained. The seasonal, weekly-based intervals for which the models were obtained correlated with the key growth periods of Pinus sibirica Du Tour—the period of the highest cellular activity (early June) and the period of budding for the next season (mid-July), which confirmed our first hypothesis. Also, the smoothing windows, for which the first adequate reconstructions appeared, correlated with the lifespan of Pinus sibirica needles.

The use of tree ring widths as predictors did not allow one to obtain reliable reconstruction models for any seasonal time interval or with a single smoothing window due to their poor correlation with the temperature data. At the same time, the results of reconstructions derived from tracheid data did not only show good metric values for the verification set \((R^2 > 0.6)\) but also had a significant correlation with the independent V3 reanalysis data for the 19th century \((p < 10^{-8})\), which also confirmed the second hypothesis of this work.

Further modifying the developed algorithms can improve the quality of temperature reconstruction, but it is already evident at this stage that, despite the greater complexity of obtaining tracheidogram measurements compared to annual rings, quantitative wooden anatomy allows one to reconstruct temperature with much higher accuracy and resolution than classical approaches.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f15010167/s1, Figure S1. Example of daily (or intra-annual) smoothing. The raw (red dashed) and smoothed \((w = 7, 1\) day step) daily Tashtyp temperature; Figure S2. Example of unsmoothed and 9-year (or inter-annual) smoothing of temperature characteristics.
mean values of the 1st week of May (red dotted curve), June (green dotted curve), and July (blue dotted curve) and their smoothed analogs (solid thick lines), respectively; Figure S3. Example of cell measurements for the year 1653 of Tree №2; Figure S4. Dmean (A) and CWTmean (C) tracheidograms for the year 1653 of Tree №1, and their standardization to 15 cells (B,D); Figure S5. Example of the mean standardized tracheidograms (black curves) for the 1653 year: radial cell diameter (A) and cell wall thickness (B); Figure S6. Obtained tracheidogram objects; Figure S7. The obtained 30 cell chronologies: 15 mean standardized cell diameters (A) and corresponding 15 cell wall thicknesses (B); Figure S8. Example of 9-year smoothed cell chronologies: mean standardized cell diameters (A) and corresponding cell wall thicknesses; Figure S9. First five principal components (PCs) of the smoothed (9-year sliding window) tracheid chronologies; Figure S10. Cumulative explained variance of the tracheidogram objects; Figure S11. Visualization of the Rolling Leave-One-Out Cross Validation procedure for the data with the sliding windows $w = 7$, $W = 9$, for the $doy = 152$. The red cells are considered as a verification set for the corresponding model, the gray cells are omitted, and the white cells are considered as a calibration set for the corresponding model. Table S1. Examples of the temperature time series $T_{\text{inter}}(w, W, \text{doy})$ for $w = 7$, $W = 9$, $1 \text{DOY}_{\text{sig}}(7,9) = 63$; Table S2. Examples of thirty ($2n = 30$) mean tracheidogram chronologies; Table S3. Example of inter-annual smoothed tracheid chronologies for $W = 9$; Table S4. Example of PC chronologies for $W = 9$; Table S5. PCA transformation matrix for first five principal components; Table S6. Example of the calibration and verification sets for $w = 7$, $W = 9$, $\text{doy} = 152$, $\theta = 2000$; Table S7. Model coefficients.

**Author Contributions:** Conceptualization, M.S.Z., E.A.V. and V.V.S.; methodology, M.S.Z. and V.V.S.; software, M.S.Z.; validation, B.Y., E.A.V. and V.V.S.; formal analysis, M.S.Z., B.Y., E.A.V. and V.V.S.; investigation, M.S.Z., B.Y., E.A.V. and V.V.S.; resources, E.A.B.; data curation, M.S.Z. and V.V.S.; writing—original draft preparation, M.S.Z., B.Y., D.F.Z., E.A.V. and V.V.S.; writing—review and editing, M.S.Z., B.Y. and V.V.S.; visualization, M.S.Z.; supervision, E.A.V. and V.V.S.; project administration, V.V.S.; funding acquisition, E.A.V. and V.V.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Russian Science Foundation, grant no. 23-44-00067 (data analysis, reconstruction models, software development), and the National Natural Science Foundation of China, grant no. 42261134537, in the framework of a joint Russian–Chinese project (state-of-the-art analysis and international collaboration expenses), and by the Russian Ministry of Science and Higher Education, grant number FSRZ-2023-0007 (climate and wooden data collection and measurements).

**Data Availability Statement:** The data presented in this study are available upon reasonable request from the authors. All data processing algorithms were implemented in Python and can be found at: https://github.com/mikewellmeansme/dendroclimatic-reconstructor/ (accessed on 9 January 2024).

**Acknowledgments:** The authors express their deep gratitude to Yulia Olegovna Zharkova, Krasnoyarsk Science Center, Siberian Branch of the Russian Academy of Sciences, for linguistic and stylistic corrections of the text.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

**References**

1. Danchenko, A.M.; Beh, I.A. Cedar Forests of Western Siberia; Tomsk State University: Tomsk, Russia, 2010.


64. Silvestro, R.; Brasseur, S.; Klisz, M.; Mencuccini, M.; Rossi, S. Bioclimatic Distance and Performance of Apical Shoot Extension: Disentangling the Role of Growth Rate and Duration in Ecotypic Differentiation. *For. Ecol. Manag.* 2020, 477, 118483. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.