Supplementary Information of Boreal Env. Res. Vol. 28: 169–180, 2023

Supplementary Information of

Trends in thermal growing season length from years 1955–2020 — A case study in hemiboreal forest in Estonia

Joonas Kollo *et al*.

Correspondence to: Steffen M. Noe (*steffen.noe@emu.ee*)

Table S1. Changes in the growing season start, end and length in different countries across the globe over the period from 1920 to 2020 using different methods. P - phenological data, N - normalized difference vegetation index (NDVI) and C – climatological data based on temperature.

Type	Country	Time span	Change (days)			Reference
			Spring	Autumn	Length	
P	Lithuania/Latvia	1971-2000	6.4–15.4 earlier	8–23.9 earlier	7	Kalvane et al., (2009)
P	Estonia	1920-2000	8 earlier			Ahas, 1999
P	Germany	1951-1996	9.2 earlier		9.2	Menzel et al., (2001), Chmielewski et al., (2004)
P	North West Russia	1930-1998	2.5 earlier	5 later		Kozlov & Berlina (2002)
N	Fennoscandia	1982-1998	14–21 earlier	7–21 later		Hogda et al., (2001)
C	China	1951-2007	5.8 earlier	3 earlier	8.8	Song et al., (2010)
C	South East Estonia	1955-2020			14.8	This work
C	United Kingdom	1980-2000	5.5 earlier			Sparks et al., (2005)
C	United Kingdom	1991-2000	4.5 earlier			Fitter & Fitter (2002)
C	Germany	1950-2000	6–12 earlier	12 later	5.5–24.5	Menzel et al., (2003)

Table S2. Statistic parameters of the regression models (Fig.3). All tests at significance level $p = 0.05$. Normality of the residuals was tested using Jarque-Bera (JB) test, a value of JB ≈ 3 and p(JB) close to one indicate normal distribution in the residuals, i.e., low skewness and kurtosis in the data. Higher values indicate fewer outliers. Akaike's information criterion (AIC) gives a handle to decide which model to choose in case of different models given.

Model	R^2	p	AIC	JB	p(JB)
effective linear	0.207	<0.05	544.9	1.214	0.545
effective logarithmic	0.207	<0.05	544.9	1.205	0.547
active linear	0.188	<0.05	575.3	1.155	0.561
active logarithmic	0.189	<0.05	575.3	1.155	0.561

Fig. S1. Location of 11 meteorological stations and Järvselja across Estonia.

S. 2 Testing the influence of the measurement protocol on the temperature readings

To simulate the influence of sampling protocols on the temperature data we used the years from 2015 to 2020 measured at the SMEAR Estonia station. These data are available in 10-minute intervals and we defined three protocols according to the historical used methods. We named them meantSMEAR, which are daily mean temperatures of the 30 minute averaged data, the base case. However, we have no information how many data points were averaged during the period 2001 to 2014 and therefore we use a minimum maximum approach that assumes there was one defined maximum and minimum reading and that was used as daily mean. We named that scenario meantMiMa. Finally, the third scenario is mapping the data into three time slices, from 21;00 to 9:00, 9:00 to 15:00, and 15:00 to 21:00, just as the measurements were conducted during 1955 to 2000 and this was named meanNMA. In all three time slices, the maximum and minimum temperature was taken to calculate the mean temperature of the period. Daily averages are then calculated from these data.

Because we used the same base data eventually occurring deviations are due to the different sampling protocols. meantSMEAR will have the largest number of input data for the daily mean estimation and according to the law of the largest number:

Let $X_1, X_2, \dots, X_n$ be random variables with a finite expected value $EX_i = \mu < \infty$. Then, for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|\bar{X} - \mu| \geq \epsilon) = 0,$$

we can assume that the estimated daily mean is most accurate in this case.

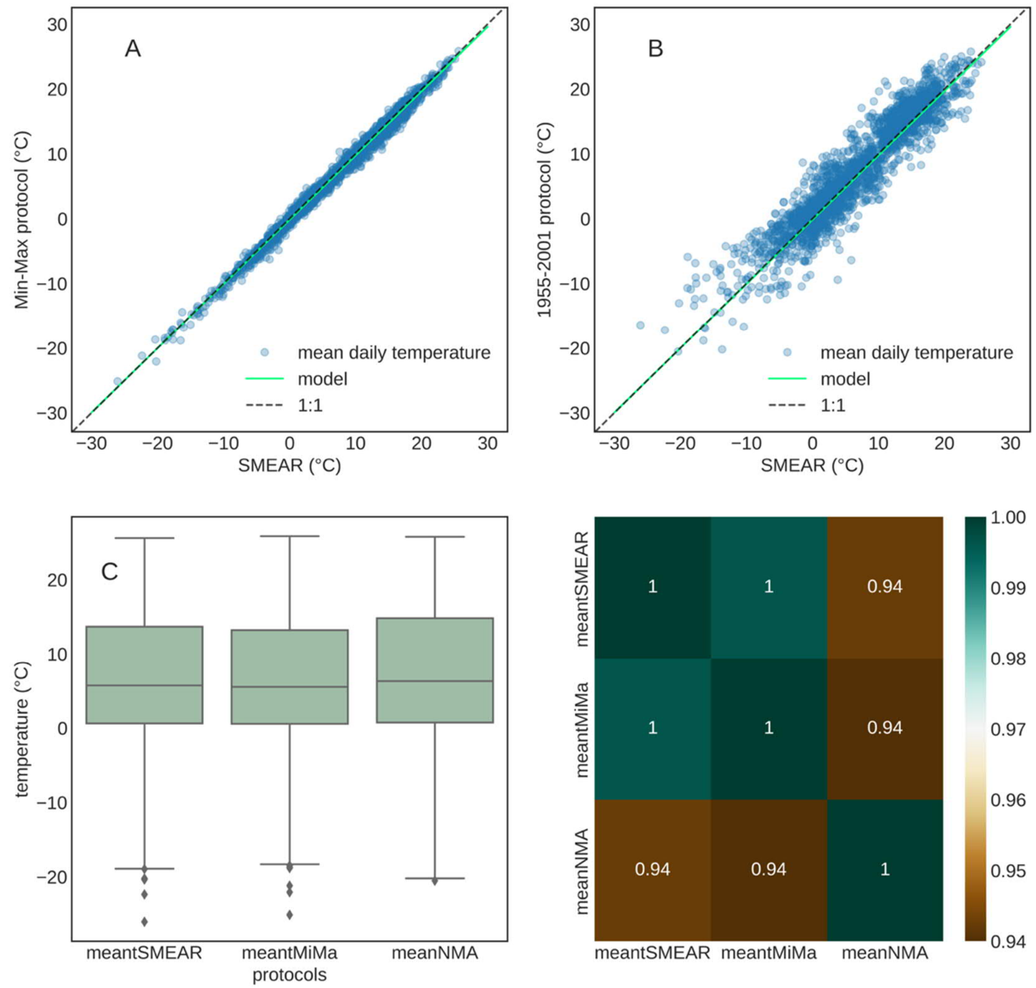

Fig. S2. On the top panel the linear relation between the meantSMEAR and meantMiMA and the meantSMEAR and meanNMA protocols are presented. In both cases, the linear model reached nearly the identity (1:1) and we obtained with $y = 0.992x - 0.175$,

$p < 0.05$ similar parameters for both. The boxplot on the lower panel indicates how similar the sampled medians, spread and quantiles are located. Only the meanNMA protocol led to a slightly larger spread in the data which is also indicated by the covariance matrix.

Given the results presented in figure S2, we can conclude that any possible effect by the differences in the sample protocol do not lead to biases in temperature readings. This is even more evident given the fact that the readings from 1955 to 2014 are rounded to either full or half degrees. Comparing the medians, quantiles, and means of the three time series generated by the protocol simulation shows that deviations remain below this limit given through the rounding criteria.

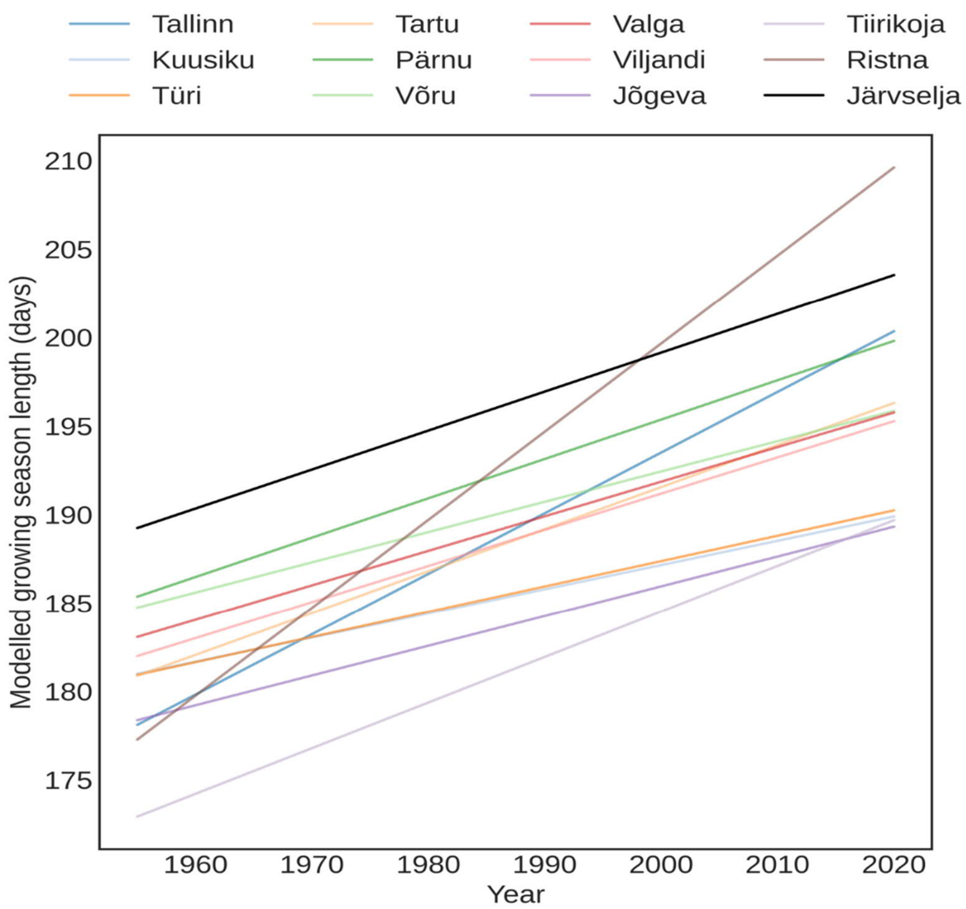

Fig. S3. Modelled linear relations between the year and the growing season length. The biggest changes in the growing season length occurred in Tiirikoja, Ristna and Tallinn where the slopes are steepest. Järvselja has a similar slope as most of the Estonian weather service's measurement stations.

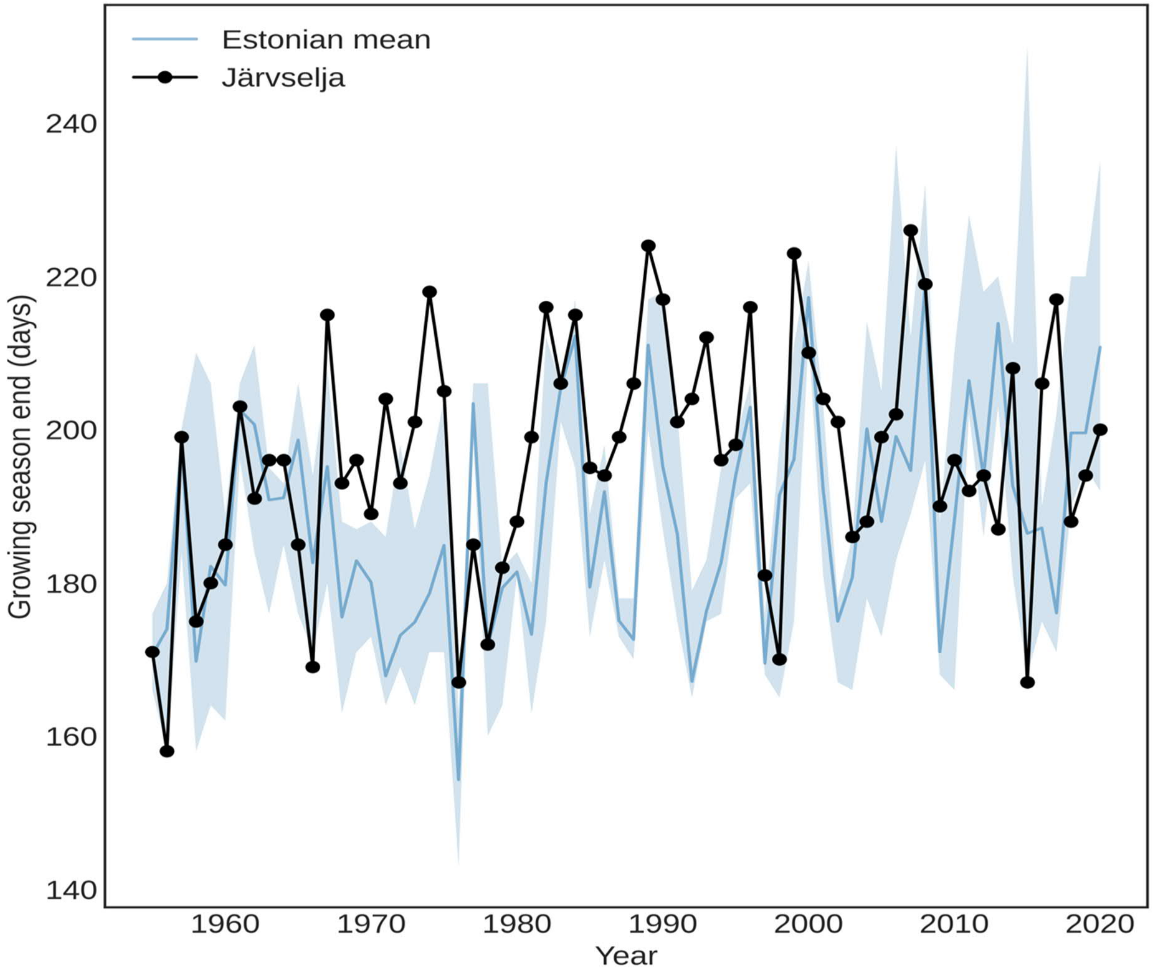

Fig. S4. Comparison between the effective thermal growing season length in Järvselja (black) and mean of 11 stations of the Estonian weather service's network (blue). The blue shaded area denotes the standard deviation of the averaged data.

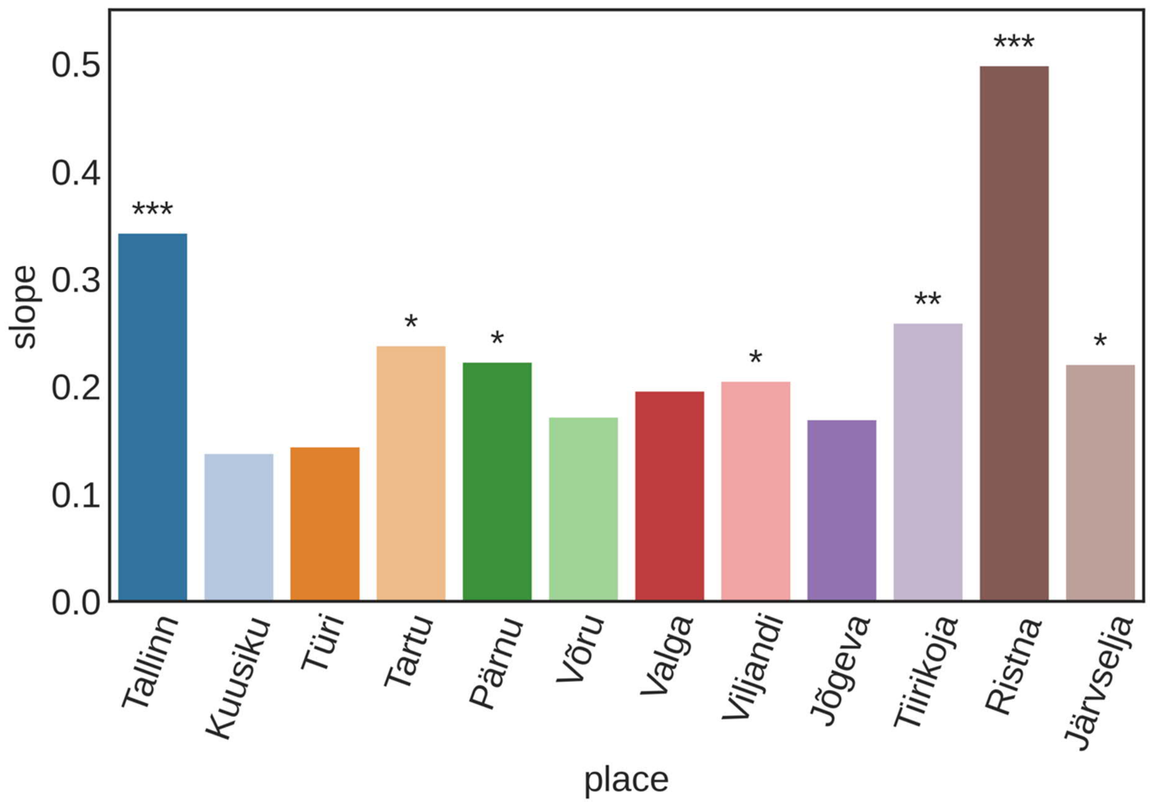

Fig. S5. Significance of the slopes in the change of the GSL for the Estonian area. Highest statistical significance was found for Tallinn and Ristna followed by Tiirikoja. Tartu, Pärnu, Viljandi and Järvselja show a weaker statistical significance in the estimation of the linear relation in GSL change and the most inland stations linear relations were found to be non-significant. Significance levels were chosen to $p<0.005$ with three asterisks, $p<0.01$ with two and $p<0.05$ with one asterisk and $p>0.05$ was left unmarked.

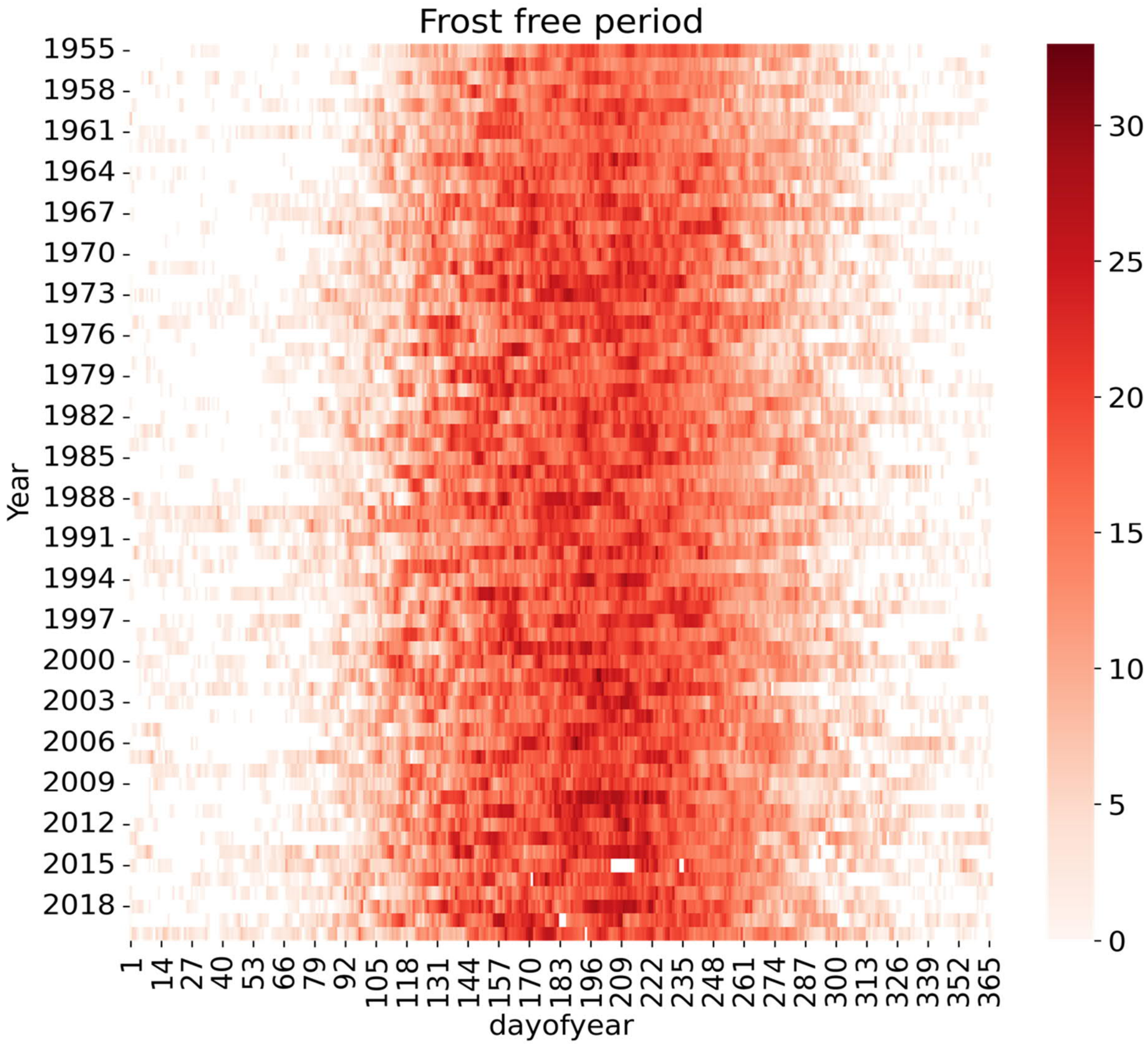

Fig. S6. Visualisation of the frost-free days in Järvselja. It is seen that number of frost-free days have been increasing during whole period from 1955 to 2020 and especially in the most recent decade frost-free days in winter became abundant