Re-calibration of Arctic sea ice extent datasets using Arctic surface air temperature records – Supplementary Information

Ronan Connolly [1,*], Michael Connolly [1] and Willie Soon [2]

[1] *Independent scientists, Dublin, Ireland.* E-mail: ronanconnolly@yahoo.ie
[2] *Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA*
* *Corresponding author*

Last updated: 28th December 2016. (Submitted to Hydrol. Sci. J.)

1. DISCUSSION OF UNADJUSTED AND MW-09 ADJUSTED TEMPERATURE DATASETS

As discussed in the paper, the source we use for determining surface air temperature trends is version 3 of the Global Historical Climatology Network (GHCN) monthly dataset, which was downloaded from http://www.ncdc.noaa.gov/ghcnm/ *[Accessed 11/1/2016]*. This dataset of monthly-averaged weather station temperature records is compiled and maintained by the US-based NOAA National Centers for Environmental Information (Lawrimore et al., 2011). They provide two versions of the dataset. The first dataset contains the raw monthly station records with only some minor quality control adjustments, but the records for the second dataset have been adjusted by the automated homogenisation algorithm of Menne & Williams (2009) in an attempt to remove/reduce any non-climatic biases which may exist in the raw station records. We will refer to the former as the "unadjusted dataset" and the latter as the "MW09-adjusted dataset".

Venema et al. (2012) and Williams et al. (2012) have found that this automated homogenisation algorithm is quite effective at identifying and removing certain types of artificially-introduced biases from synthetic temperature records derived from climate models. On this basis, they have argued that the "MW09-adjusted" dataset is more reliable. However, two of us (RC and MC) have noted that the algorithm can lead to considerable "blending" of biases in datasets where multiple neighbouring stations are simultaneously affected by similar biases (Connolly and Connolly, 2014a), including networks affected by urbanization bias (Connolly and Connolly, 2014b). Moreover, we also found that the adjustments applied by the algorithm are quite inconsistent, and the algorithm frequently yields either false positive or false negative results (Soon et al., 2015).

We recognise that the "unadjusted" dataset is affected by numerous non-climatic biases. Indeed, we argue that the extent of these biases has been substantially underestimated (Connolly and Connolly, 2014a; Connolly and Connolly, 2014b; Soon et al., 2015). However, we argue that the "MW09-adjusted" dataset is *not necessarily* an improvement, and in many cases may make individual station records *less* reliable. Instead of solely relying on automated homogenisation algorithms, we have recommended that the compilation of detailed station histories documenting potential non-climatic biases associated with individual station records should be a priority.

Nonetheless, from Figure 2 in the paper, we saw that the long-term temperature trends for the Arctic region are quite similar for both the unadjusted and the MW09-adjusted datasets. That is, both datasets imply that the Arctic went through a period of warming (~1900s-1940s), followed by a period of cooling (~1940s-1970s), followed by another period of warming (~1970s-present).

The net effect of the MW09-adjustments is to slightly reduce the apparent warmth of *both* the present warm period and the early-20th century warm period, and so the relative warmth of both warm periods is comparable for both datasets. The timing of the warming and cooling periods is also the same. Therefore, the results of our analysis in this paper should be quite similar, regardless of which dataset we use. With this in mind, in the paper, we presented the results using the unadjusted dataset. However, for interested readers, the equivalent analysis using the MW09-adjusted dataset is also provided here in the Supplementary Information.

2. INTERPOLATION OF MAHONEY ET AL. (2008) SIBERIAN SEA ICE EXTENTS

For some years (particularly in the earlier period), Mahoney et al. (2008; Mahoney, 2008) did not have seasonal estimates for all four of the Siberian seas, i.e., Kara, Laptev, East Siberian and Western Chukchi. However, as discussed in Table 1, the average sea ice extent for each of the four seas is closely related to the total Siberian Arctic sea ice extent for a given season. With this in mind, for those years when we did not have data for all four seas in a given season, we estimated the total Siberian sea ice extent by applying the relationships in Table 1 to those seas which did have data, and obtaining the average extent.

Relationship of each sea to total Siberian sea ice extent

Sea	***Equation Of Line***	r^2	p
Kara Sea	y = 2.11 x + 1,886,000	0.80	**<0.001**
Laptev Sea	y = 2.45 x + 1,071,000	0.80	**<0.001**
E. Siberian Sea	y = 3.91 x – 1,495,000	0.66	**<0.001**
W. Chukchi Sea	y = 5.62 x + 1,427,000	0.73	**<0.001**

Table 1. Linear relationships between the seasonal sea ice extents of each of the four Siberian seas (km^2) to the total sea ice extent for the Siberian Arctic, determined from linear least squares fitting for those years where data was available for all four seas.

3. SEA ICE EXTENT: TEMPERATURE RELATIONSHIPS

Table 2 and Table 3 list the results of linear least squares regression fits for the satellite era between regional, seasonal Arctic sea ice extents and the corresponding temperature trends using the unadjusted and MW09-adjusted (respectively) temperature datasets. The linear fits for Table 2 (i.e., unadjusted temperature dataset) are also plotted graphically in Figure 1.

The first point to note is that for all but three of the regions and seasons, the *p* values of the linear fits are much less than 0.05. The *p* value of a statistical fitting indicates the probability that the relationship is **not** significant. In other words, the lower the *p* value is, the more confidence we can have that the relationship is not just coincidental. By convention, a *p* value of less than 0.05 is usually taken as being statistically significant (>95% significance), although this threshold is rather arbitrary (e.g., see Briggs, 2016). Therefore, during the satellite era, there was a statistically significant linear relationship between surface air temperatures and sea ice extent for winter, summer and autumn in the North American Arctic; for all four seasons in the Nordic Arctic; and for summer and autumn in the Siberian Arctic.

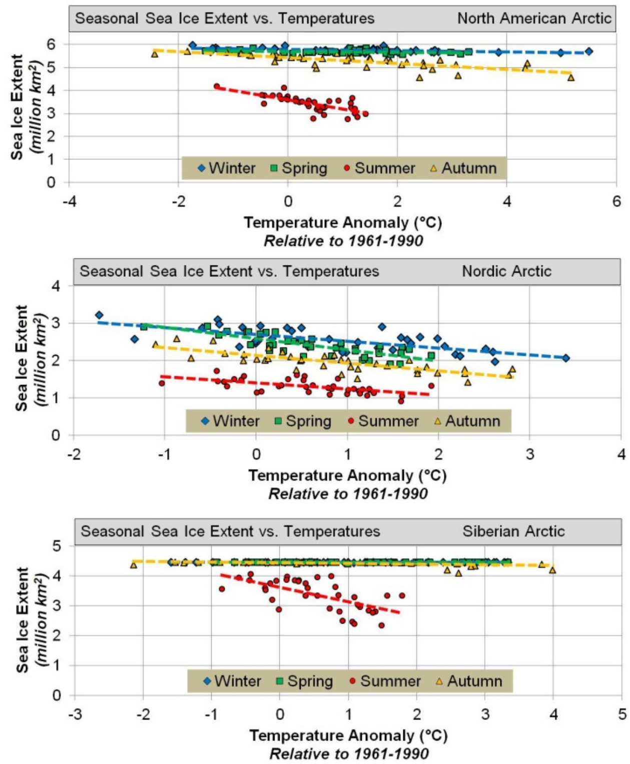

Figure 1. Linear least square regression fits between seasonal sea ice extent and temperature anomalies (unadjusted) for the three Arctic regions during the satellite era.

The *p* values for the linear fits for winter and spring in the Siberian Arctic are not statistically significant, suggesting that for this region, during those seasons, the relationship between surface air temperatures and sea ice extent breaks down. This is not surprising because, as discussed in the paper, during the winter and spring the Siberian Arctic is essentially ice-packed, and the sea ice extent is almost constant. The *p* value for the North American spring is greater than 0.05, but it is still less than 0.1, suggesting that there is still a weak relationship between temperatures and sea ice extent (>90% significance).

The equations of the linear fits are listed in terms of $y = mx + c$, where m and c are the slopes and intercepts of the lines respectively, y is the mean seasonal sea ice extent (in km^2) and x is the seasonal temperature anomaly (°C), relative to the 1961-1990 baseline period. All of the slopes are negative meaning that, as expected (e.g., Zakharov, 1997; Alekseev et al., 2016), when the regional temperature increases, its average sea ice extent decreases. The magnitude of the slope indicates the magnitude of this relationship, and this varies with region and season. Therefore, if the average summer air temperature for the North American Arctic rose by 1°C, we would expect the sea ice extent to decrease by 408,000 km^2, while an equivalent temperature rise during the winter should only decrease the winter sea ice extent by about 28,000 km^2.

As an aside, the intercept values in the equations of the lines correspond to the expected sea ice extent for a seasonal temperature anomaly of exactly 0°C. Since we calculated our temperature anomalies relative to the 1961-1990 baseline (when station coverage is greatest), this means that the intercepts for a given season and region correspond to the expected sea ice extent for the 1961-1990 average.

Finally, the r^2 values in Table 2 correspond to the coefficients of determination, and indicate how well the relationships between temperatures and sea ice extents are described by the linear fit. They can range from a value of 0 ("not at all") to 1 ("perfect"). Generally the r^2 values are largest for the regions and seasons where the slopes are most negative, i.e., when there is a strong relationship between sea ice extent and temperature. For the Siberian Arctic during the winter and spring, the r^2 values are negligible, again because there is almost no variability in the sea ice extent for this region in these seasons.

Satellite era, 1979-2015 (Unadjusted data)

Season	***Equation Of Line***	r^2	***p***
North American Arctic			
Winter	y = -28,000 x + 5,788,000	0.28	**<0.001**
Spring	y = -17,000 x + 5,702,000	0.08	0.08
Summer	y = -408,000 x + 3,595,000	0.52	**<0.001**
Autumn	y = -129,000 x + 5,441,000	0.55	**<0.001**
Annual	y = -123,000 x + 5,153,000	0.58	**<0.001**
Nordic Arctic			
Winter	y = -182,000 x + 2,702,000	0.49	**<0.001**
Spring	y = -302,000 x + 2,591,000	0.62	**<0.001**
Summer	y = -164,000 x + 1,404,000	0.36	**<0.001**
Autumn	y = -206,000 x + 2,139,000	0.59	**<0.001**
Annual	y = -247,000 x + 2,233,000	0.64	**<0.001**
Siberian Arctic			
Winter	y = -100 x + 4,459,000	0.00	0.77
Spring	y = -600 x + 4,458,000	0.01	0.67
Summer	y = -486,000 x + 3,616,000	0.40	**<0.001**
Autumn	y = -23,000 x + 4,442,000	0.19	**0.005**
Annual	y = -110,000 x + 4,276,000	0.54	**<0.001**

Table 2. Linear least square regression fits between seasonal sea ice extent (y, km^2) and temperature anomalies (x, °C) for the three Arctic regions during the satellite era, i.e., 1979-2015. Fits that are significant to >95% (i.e., p<0.05) are highlighted in bold.

Satellite era, 1979-2015 (MW09-adjusted data)

Season	***Equation Of Line***	r^2	***p***
North American Arctic			
Winter	y = -27,000 x + 5,779,000	0.24	**0.002**
Spring	y = -18,000 x + 5,702,000	0.08	0.09
Summer	y = -454,000 x + 3,629,000	0.57	**<0.001**
Autumn	y = -131,000 x + 5,438,000	0.56	**<0.001**
Annual	y = -129,000 x + 5,148,000	0.53	**<0.001**
Nordic Arctic			
Winter	y = -184,000 x + 2,699,000	0.48	**<0.001**
Spring	y = -308,000 x + 2,589,000	0.61	**<0.001**
Summer	y = -162,000 x + 1,396,000	0.33	**<0.001**
Autumn	y = -195,000 x + 2,115,000	0.53	**<0.001**
Annual	y = -257,000 x + 2,228,000	0.62	**<0.001**
Siberian Arctic			
Winter	y = -10 x + 4,459,000	0.00	0.98
Spring	y = -700 x + 4,458,000	0.01	0.62
Summer	y = -436,000 x + 3,557,000	0.26	**0.001**
Autumn	y = -22,000 x + 4,437,000	0.17	**0.01**
Annual	y = -111,000 x + 4,263,000	0.46	**<0.001**

Table 3. Linear least square regression fits between seasonal sea ice extent (y, km^2) and temperature anomalies (x, °C) for the three Arctic regions during the satellite era, i.e., 1979-2015. Fits that are significant to >95% (i.e., p<0.05) are highlighted in bold.

Table 4 and Table 5 list the equivalent results for the pre-satellite era sea ice estimates before calibration.

Pre-satellite era, 1901-1978 (Unadjusted data)

Season	***Equation Of Line***	r^2	***p***
North American Arctic (Walsh dataset)			
Winter	y = -4,000 x + 6,192,000	0.03	0.15
Spring	y = -9,000 x + 6,146,000	0.02	0.25
Summer	y = -117,000 x + 4,246,000	0.06	**0.04**
Autumn	y = -1,000 x + 6,047,000	0.00	0.79
Annual	y = -26,000 x + 5,660,000	0.07	**0.02**
Nordic Arctic (Walsh dataset)			
Winter	y = -57,000 x + 3,001,000	0.19	**<0.001**
Spring	y = -124,000 x + 2,970,000	0.13	**0.001**
Summer	y = -201,000 x + 1,845,000	0.16	**<0.001**
Autumn	y = -51,000 x + 2,590,000	0.13	**<0.001**
Annual	y = -122,000 x + 2,605,000	0.24	**<0.001**
Siberian Arctic (Walsh dataset)			
Winter	y = 4,581,000, i.e., constant	N/A	N/A
Spring	y = -200 x + 4,581,000	0.02	0.24
Summer	y = -93,000 x + 4,138,000	0.06	**0.04**
Autumn	y = 4,581,000, i.e., constant	N/A	N/A
Annual	y = -7,000 x + 4,470,000	0.01	0.41
Siberian Arctic (Russian datasets)			
Winter	y = -11,000 x + 4,152,000 *(over 1940-1978 period)*	0.04	0.22
Spring	y = -29,000 x + 4,073,000 *(over 1940-1978 period)*	0.15	**0.01**
Summer	y = -97,000 x + 3,455,000 *(over 1900-1978 period)*	0.05	**0.04**
Autumn	y = +16,000 x + 3,715,000 *(over 1943-1978 period)*	0.01	0.59
Annual	y = -44,000 x + 3,827,000 *(over 1943-1978 period)*	0.09	0.08

Table 4. Linear least square regression fits between seasonal sea ice extent (y, km^2) and temperature anomalies (x, °C) for the three Arctic regions during the pre-satellite era, i.e., 1901-1978. Fits that are significant to >95% (i.e., p<0.05) are highlighted in bold.

Pre-satellite era, 1901-1978 (MW09-adjusted data)

Season	***Equation Of Line***	r^2	***p***
North American Arctic (Walsh dataset)			
Winter	y = -3,000 x + 6,191,000	0.03	0.17
Spring	y = -10,000 x + 6,143,000	0.02	0.18
Summer	y = -142,000 x + 4,211,000	0.10	**0.005**
Autumn	y = -2,000 x + 6,047,000	0.00	0.56
Annual	y = -30,000 x + 5,653,000	0.10	**0.004**
Nordic Arctic (Walsh dataset)			
Winter	y = -52,000 x + 2,991,000	0.16	**<0.001**
Spring	y = -114,000 x + 2,950,000	0.11	**0.003**
Summer	y = -217,000 x + 1,808,000	0.17	**<0.001**
Autumn	y = -42,000 x + 2,579,000	0.09	**0.008**
Annual	y = -111,000 x + 2,579,000	0.20	**<0.001**
Siberian Arctic (Walsh dataset)			
Winter	y = 4,581,000, i.e., constant	N/A	N/A
Spring	y = -200 x + 4,581,000	0.02	0.21
Summer	y = -126,000 x + 4,127,000	0.09	**0.009**
Autumn	y = 4,581,000, i.e., constant	N/A	N/A
Annual	y = -11,000 x + 4,469,000	0.02	0.25
Siberian Arctic (Russian datasets)			
Winter	y = -10,000 x + 4,151,000 *(over 1940-1978 period)*	0.03	0.27
Spring	y = -30,000 x + 4,070,000 *(over 1940-1978 period)*	0.15	**0.01**
Summer	y = -113,000 x + 3,444,000 *(over 1900-1978 period)*	0.06	**0.03**
Autumn	y = +17,000 x + 3,716,000 *(over 1943-1978 period)*	0.01	0.56
Annual	y = -41,000 x + 3,823,000 *(over 1943-1978 period)*	0.07	0.12

Table 5. Linear least square regression fits between seasonal sea ice extent (y, km^2) and temperature anomalies (x, °C) for the three Arctic regions during the pre-satellite era, i.e., 1901-1978. Fits that are significant to >95% (i.e., p<0.05) are highlighted in bold.

3.1 Analysis of residuals from linear fits

Table 6 lists some statistics describing the *residuals* of the regional sea ice data for each of the seasons for the satellite era after subtracting the linear fits from Table 2.

If the residuals from a statistical fit are independent and identically distributed (i.i.d.) random variables with a normal distribution, we would expect the skewness and excess kurtosis statistics both to be close to zero. For most of the seasons and regions, these values are reasonably close to zero, indicating that the linear fit regression model based on temperature anomalies captures much of the variability in the sea ice data. However, for the Siberian Arctic the winter, spring (and to a lesser extent autumn) statistics are not. As discussed above and in the paper, this is a consequence of the Siberian Arctic sea ice being mostly landlocked for most of the year up until the summer. That is, there is very little variability in Siberian Arctic sea ice extent during the winter and spring (and to a lesser extent autumn), and hence there is little influence from the regional temperature anomalies.

If the residuals from the linear fits are substantially affected by autocorrelation, then this could mean that the *p* values reported in Table 2 are overly optimistic. This is because if the residuals are affected by autocorrelation, then the data points are **not** wholly independent and the effective sample size of the data (N_{eff}) is less than the total sample size (N=2015-1979=36 in this case).

A simple test for autocorrelation is the Durbin-Watson test. This test returns a value between 0 and 4. If the value is exactly 2 then there is no evidence of autocorrelation. If the value is much greater than 2 then the data may be affected by negative autocorrelation (relatively rare), while if the value is much less than 2, the data may be affected by positive autocorrelation.

More specifically for our case, for N=36, if the Durbin-Watson statistic is less than 1.21, then there is a 99% chance of some positive autocorrelation. If the Durbin-Watson statistic is less than 1.41, then there is a 95% chance of some positive autocorrelation. On the other hand, if the Durbin-Watson statistic is greater than 1.52, then there is no statistical evidence for positive autocorrelation.

We can see from Table 6 that most of the Durbin-Watson statistics are greater than 1.52, and therefore there it is not necessary to correct those *p* values for autocorrelation. However, for some of the fits, the Durbin-Watson statistics are a bit low, suggesting the possibility of some autocorrelation. In particular, the Durbin-Watson statistics for North American spring and Nordic spring are only 1.12 and 1.17 respectively, which is a bit below the 99% significance lower threshold of 1.21. Also, the North American winter and Siberian summer values (1.24 and 1.41) are below (or at) the 95% significance lower threshold of 1.41.

So, it is possible that some of the *p* values for these fits are slightly too optimistic, i.e., slightly too low. However, with the exception of North American spring, the uncorrected *p* values for these cases are all several orders of magnitude less than 0.01. Therefore, the statistical significance of these fits still seems robust.

As discussed earlier, the *p* value for the North American spring is relatively high (*p*=0.081), and the fact that the Durbin-Watson statistic is only 1.12 suggests that the true *p* value should be even higher. So, the apparent relationship between temperature and sea ice for this region/season should be treated with caution.

Satellite era, 1979-2015 (Unadjusted data)

		Statistics of residuals		
Season	***p (of fit)***	***Skewness***	***Excess Kurtosis***	***Durbin-Watson***
North American Arctic				
Winter	6.3×10^{-4}	0.45	1.35	1.24
Spring	0.081	0.77	0.16	1.12
Summer	4.4×10^{-7}	-0.14	1.26	**1.62**
Autumn	7.5×10^{-8}	-0.84	0.86	**1.93**
Nordic Arctic				
Winter	1.6×10^{-6}	0.02	-0.98	**1.66**
Spring	5.6×10^{-9}	0.61	-1.17	1.17
Summer	7.5×10^{-5}	0.80	-0.49	**1.66**
Autumn	2.1×10^{-8}	0.67	-0.12	**2.00**
Siberian Arctic				
Winter	0.77	-5.61	32.80	**2.19**
Spring	0.67	-5.81	34.60	**2.10**
Summer	2.5×10^{-5}	-0.25	-0.75	1.41
Autumn	0.0058	-2.09	4.99	**1.60**

Table 6. Analysis of the residuals from the linear fits in Table 2. The *p* values from Table 2 are repeated here at a higher precision for comparative purposes. Durbin-Watson statistics that are greater than the 95% upper bound threshold test against AR1 autocorrelation are shown in bold (N=36, upper bound = 1.52).

Table 7 reports the equivalent statistics for the pre-satellite era fits for the uncalibrated data derived in Table 4. As discussed above and in the paper, these fits are much less robust statistically. The Durbin-Watson thresholds for possible autocorrelation are more strenuous because of the larger sample size for the pre-satellite era (N=1978-1901=77 years for most of the regions and N=1978-1940=38 years for some of the Russian datasets).

Pre-satellite era, 1901-1978 (Unadjusted data)

		Statistics of residuals		
Season	***p (of fit)***	***Skewness***	***Excess Kurtosis***	***Durbin-Watson***
North American Arctic				
Winter	0.15	0.24	1.77	1.56
Spring	0.25	-0.08	-0.81	1.11
Summer	0.04	-0.31	-0.80	0.81
Autumn	0.79	-0.57	4.37	1.27
Nordic Arctic				
Winter	5.4×10^{-5}	1.25	1.75	0.89
Spring	1.3×10^{-3}	-0.31	-0.39	1.06
Summer	2.9×10^{-4}	0.68	0.43	0.96
Autumn	9.4×10^{-4}	1.50	7.14	1.31
Siberian Arctic (Walsh dataset)				
Winter	-	-	-	-
Spring	0.24	-8.59	75.08	**1.98**
Summer	0.038	-0.06	0.44	1.14
Autumn	-	-	-	-
Siberian Arctic (Russian datasets)				
Winter	0.27	-5.89	40.31	**1.71**
Spring	0.01	-7.76	61.05	**1.77**
Summer	0.03	-0.33	0.37	1.50
Autumn	0.56	-5.80	39.89	**1.85**

Table 7. Analysis of the residuals from the linear fits in Table 4. The *p* values from Table 4 are repeated here at a higher precision for comparative purposes. Durbin-Watson statistics that are greater than the 95% upper bound threshold test against AR1 autocorrelation are shown in bold (N=77 for Walsh dataset and Russian Summer, upper bound = 1.65; N=38 for other Russian datasets, upper bound = 1.54).

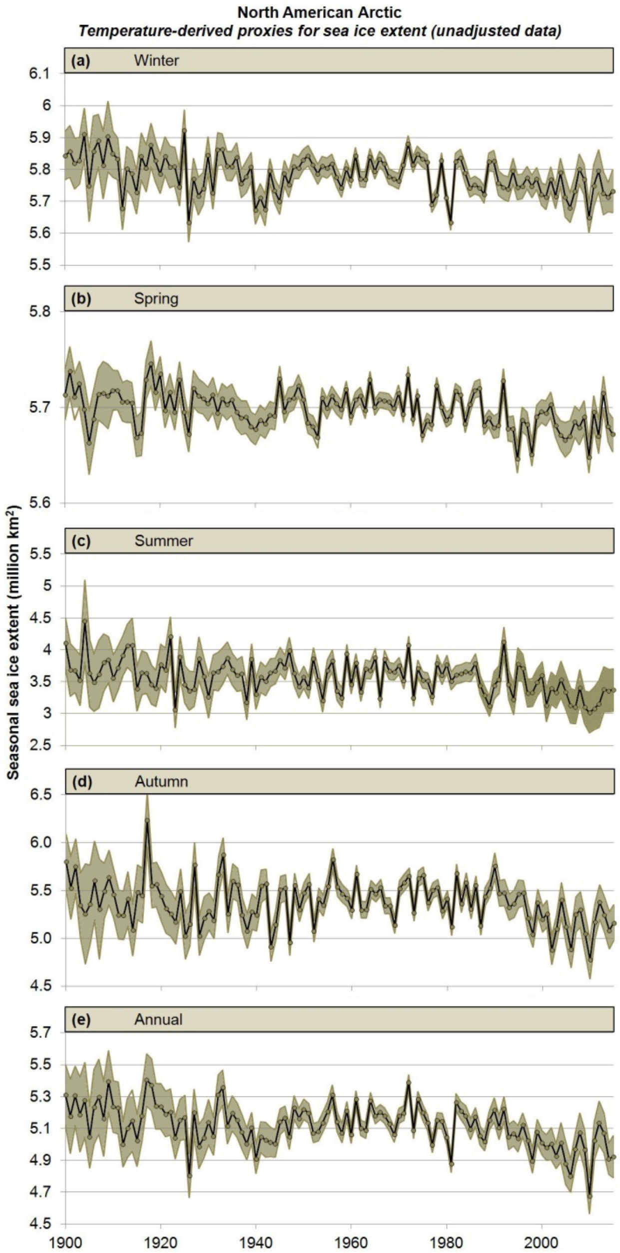

Figure 2. Temperature-derived proxies for sea ice extent for the North American Arctic using the unadjusted temperature data. Shaded bands correspond to twice the standard errors of the means. Note that the y-axes are different for each panel.

4 TEMPERATURE-DERIVED PROXIES FOR SEA ICE EXTENT

4.1 North American Arctic

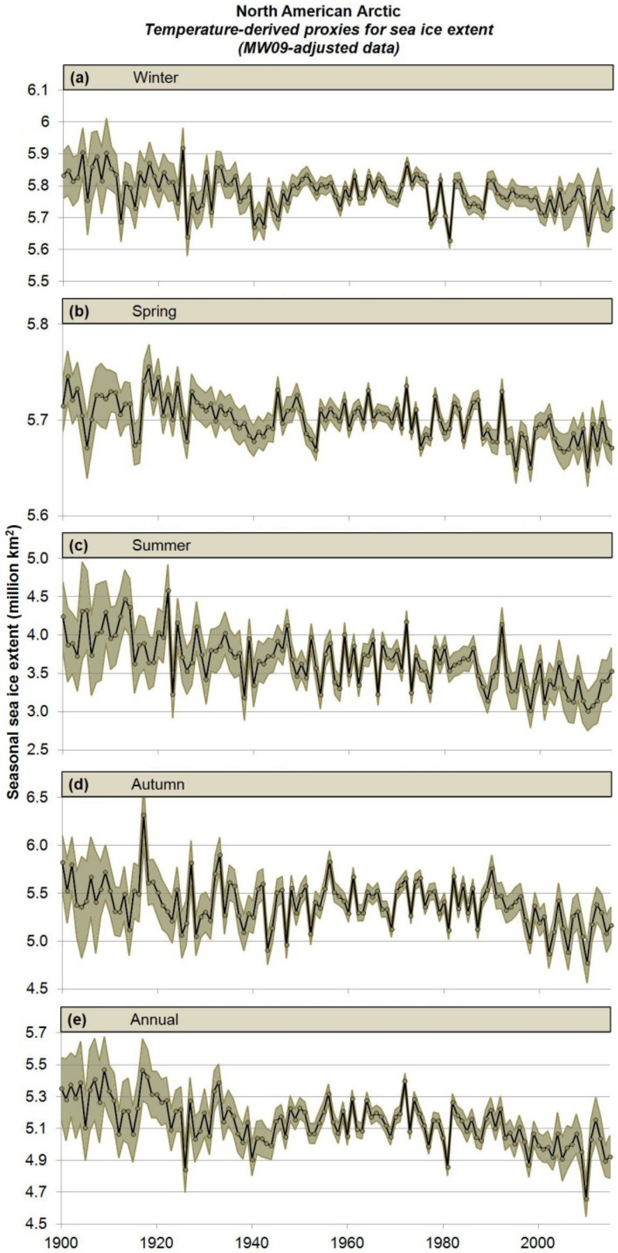

Figure 3. Temperature-derived proxies for sea ice extent for the North American Arctic using the MW09-adjusted temperature data. Shaded bands correspond to twice the standard errors of the means. Note that the y-axes are different for each panel.

4.2 Nordic Arctic

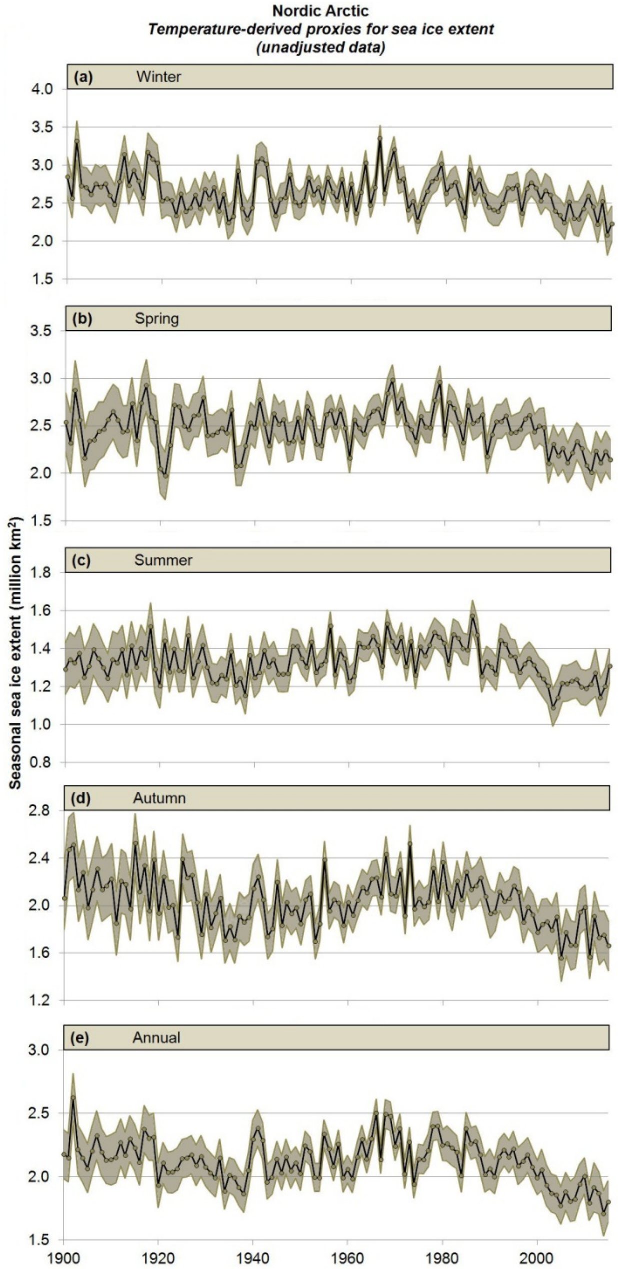

Figure 4. Temperature-derived proxies for sea ice extent for the Nordic Arctic using the unadjusted temperature data. Shaded bands correspond to twice the standard errors of the means. Note that the y-axes are different for each panel.

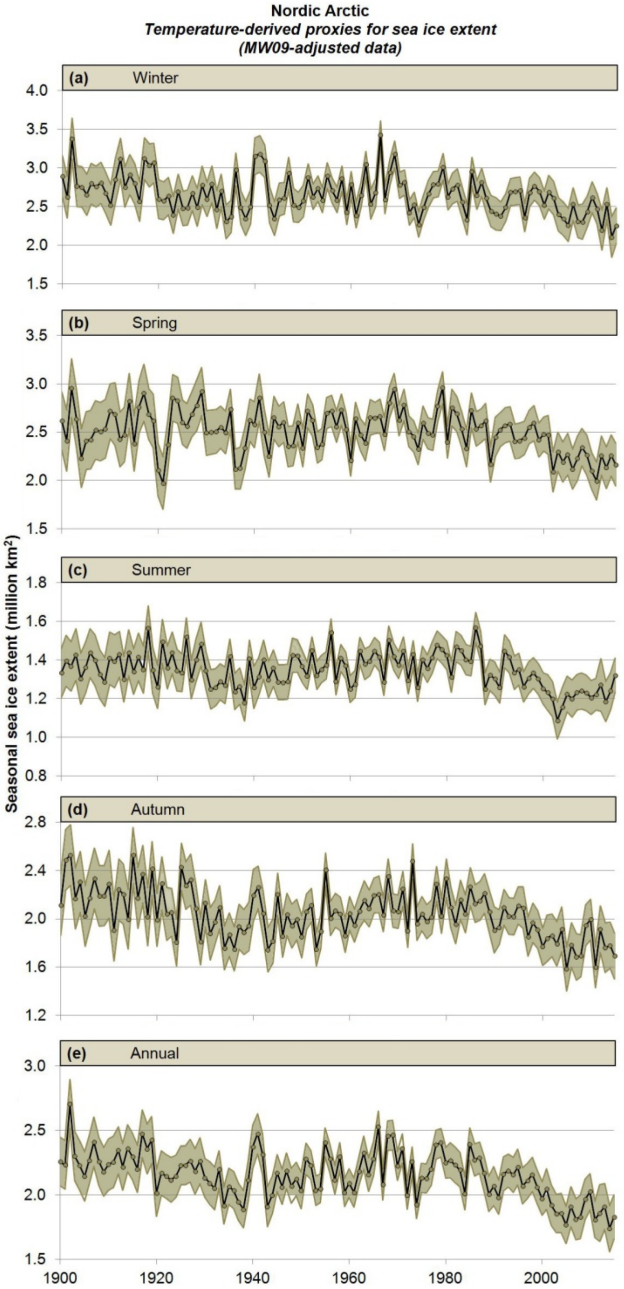

Figure 5. Temperature-derived proxies for sea ice extent for the Nordic Arctic using the MW09-adjusted temperature data. Shaded bands correspond to twice the standard errors of the means. Note that the y-axes are different for each panel.

4.3 Siberian Arctic

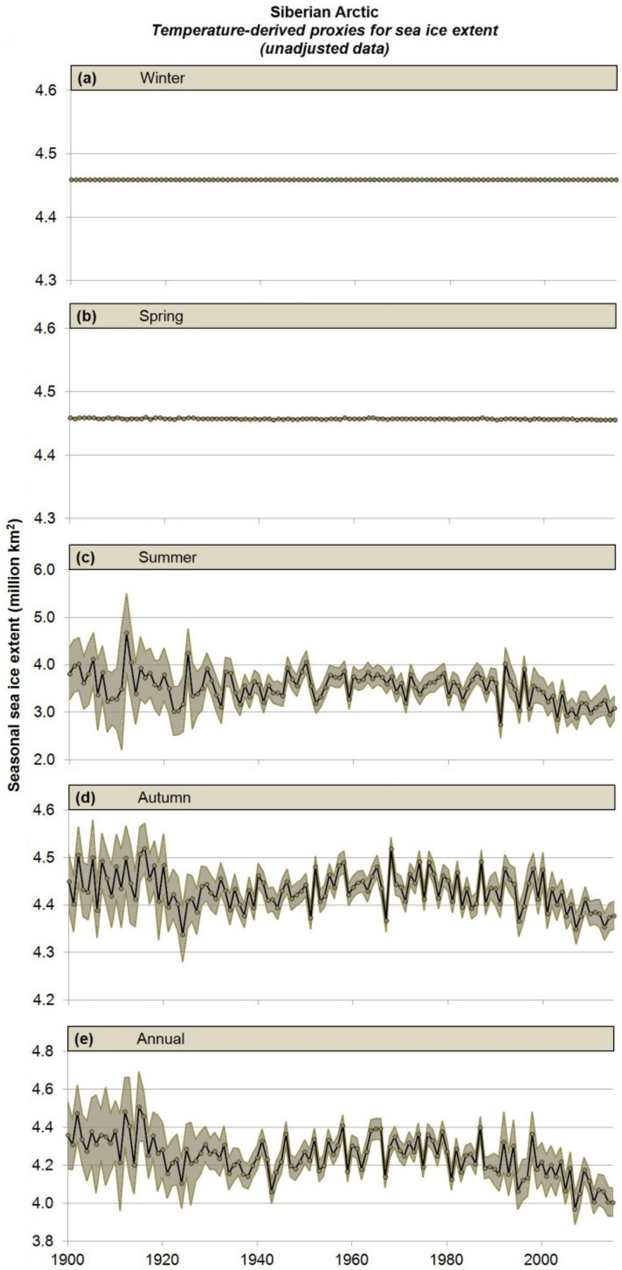

Figure 6. Temperature-derived proxies for sea ice extent for the Siberian Arctic using the unadjusted temperature data. Shaded bands correspond to twice the standard errors of the means. Note that the y-axes are different for each panel, and that there was almost no variability in the winter and spring.

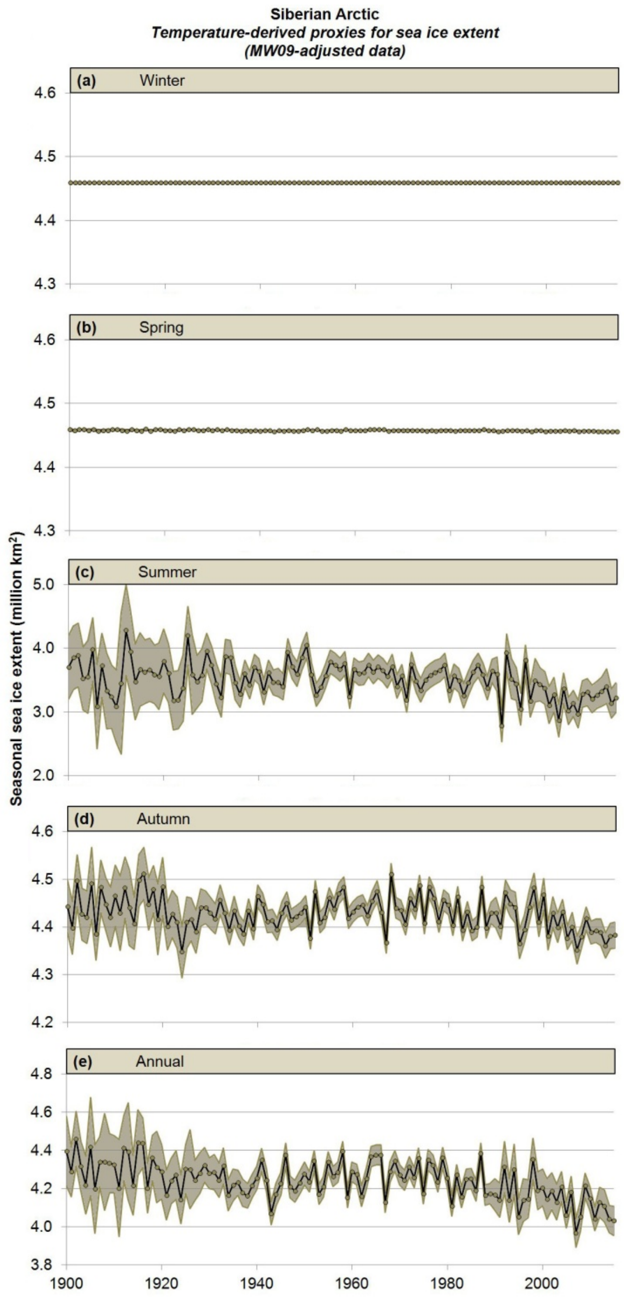

Figure 7. Temperature-derived proxies for sea ice extent for the Siberian Arctic using the MW09-adjusted temperature data. Shaded bands correspond to twice the standard errors of the means. Note that the y-axes are different for each panel, and that there was almost no variability in the winter and spring.

5. COMPARISON OF PRE-CALIBRATED SEA ICE DATA WITH TEMPERATURE-BASED PROXIES

North American Arctic (unadjusted data)

Season	***Temperature-based***		***Walsh dataset***	
	Mean	S.D. (σ)	Mean	S.D. (σ)
1901-1939				
Winter	5,808,000	67,000	6,197,000	34,000
Spring	5,706,000	18,000	6,158,000	72,000
Summer	3,663,000	251,000	4,386,000	171,000
Autumn	5,406,000	268,000	6,063,000	3,000
Annual	5,168,000	132,000	5,701,000	57,000
1940-1945				
Winter	5,715,000	20,000	6,186,000	0
Spring	5,693,000	21,000	6,124,000	0
Summer	3,586,000	169,000	3,778,000	0
Autumn	5,320,000	187,000	6,063,000	0
Annual	5,024,000	70,000	5,538,000	0
1946-1952				
Winter	5,806,000	28,000	6,191,000	23,000
Spring	5,701,000	10,000	6,193,000	79,000
Summer	3,657,000	287,000	4,609,000	120,000
Autumn	5,350,000	174,000	6,064,000	5,000
Annual	5,156,000	87,000	5,765,000	39,000
1953-1971				
Winter	5,795,000	47,000	6,173,000	53,000
Spring	5,704,000	16,000	6,119,000	67,000
Summer	3,593,000	229,000	4,129,000	228,000
Autumn	5,449,000	157,000	6,011,000	90,000
Annual	5,163,000	92,000	5,608,000	78,000
1972-1978				
Winter	5,801,000	70,000	6,232,000	88,000
Spring	5,699,000	15,000	6,130,000	36,000
Summer	3,589,000	96,000	3,948,000	398,000
Autumn	5,515,000	190,000	6,022,000	87,000
Annual	5,178,000	128,000	5,583,000	130,000
	Temperature-based		***Satellite dataset***	
1979-2015				
Winter	5,752,000	47,000	5,752,000	90,000
Spring	5,687,000	19,000	5,687,000	67,000
Summer	3,427,000	246,000	3,427,000	341,000
Autumn	5,286,000	219,000	5,285,000	295,000
Annual	5,038,000	124,000	5,037,000	163,000

Table 8. Comparison of mean and standard deviations for North American Arctic seasonal sea ice extents during the different periods, according to our temperature-derived sea ice proxies (using the unadjusted dataset), the gridded Walsh dataset, and satellite estimates.

North American Arctic (MW09-adjusted data)

Season	***Temperature-based***		***Walsh dataset***	
	Mean	S.D. (σ)	Mean	S.D. (σ)
1901-1939				
Winter	5,808,000	66,000	6,197,000	34,000
Spring	5,713,000	20,000	6,158,000	72,000
Summer	3,898,000	311,000	4,386,000	171,000
Autumn	5,454,000	279,000	6,063,000	3,000
Annual	5,227,000	148,000	5,701,000	57,000
1940-1945				
Winter	5,710,000	19,000	6,186,000	0
Spring	5,694,000	22,000	6,124,000	0
Summer	3,668,000	191,000	3,778,000	0
Autumn	5,324,000	189,000	6,063,000	0
Annual	5,022,000	74,000	5,538,000	0
1946-1952				
Winter	5,798,000	29,000	6,191,000	23,000
Spring	5,703,000	10,000	6,193,000	79,000
Summer	3,732,000	308,000	4,609,000	120,000
Autumn	5,353,000	177,000	6,064,000	5,000
Annual	5,158,000	92,000	5,765,000	39,000
1953-1971				
Winter	5,785,000	45,000	6,173,000	53,000
Spring	5,704,000	17,000	6,119,000	67,000
Summer	3,636,000	255,000	4,129,000	228,000
Autumn	5,447,000	159,000	6,011,000	90,000
Annual	5,159,000	97,000	5,608,000	78,000
1972-1978				
Winter	5,792,000	66,000	6,232,000	88,000
Spring	5,700,000	16,000	6,130,000	36,000
Summer	3,618,000	113,000	3,948,000	398,000
Autumn	5,513,000	194,000	6,022,000	87,000
Annual	5,174,000	130,000	5,583,000	130,000
	Temperature-based		***Satellite dataset***	
1979-2015				
Winter	5,752,000	44,000	5,752,000	90,000
Spring	5,686,000	19,000	5,687,000	67,000
Summer	3,427,000	257,000	3,427,000	341,000
Autumn	5,286,000	220,000	5,285,000	295,000
Annual	5,038,000	124,000	5,037,000	163,000

Table 9. Comparison of mean and standard deviations for North American Arctic seasonal sea ice extents during the different periods, according to our temperature-derived sea ice proxies (using the MW09-adjusted dataset), the gridded Walsh dataset, and satellite estimates.

Nordic Arctic (unadjusted data)

Season	***Temperature-based***		***Walsh dataset***	
	Mean	S.D. (σ)	Mean	S.D. (σ)
1901-1939				
Winter	2,669,000	260,000	2,936,000	73,000
Spring	2,482,000	220,000	2,960,000	216,000
Summer	1,317,000	81,000	1,808,000	272,000
Autumn	2,077,000	217,000	2,565,000	19,000
Annual	2,140,000	148,000	2,567,000	128,000
1940-1945				
Winter	2,688,000	298,000	2,902,000	0
Spring	2,552,000	157,000	2,618,000	0
Summer	1,308,000	50,000	1,445,000	0
Autumn	1,979,000	215,000	2,568,000	0
Annual	2,135,000	171,000	2,383,000	0
1946-1952				
Winter	2,651,000	161,000	2,914,000	49,000
Spring	2,455,000	164,000	3,071,000	101,000
Summer	1,351,000	70,000	1,857,000	220,000
Autumn	1,948,000	137,000	2,567,000	5,000
Annual	2,112,000	94,000	2,602,000	80,000
1953-1971				
Winter	2,734,000	268,000	3,194,000	229,000
Spring	2,582,000	187,000	3,012,000	210,000
Summer	1,380,000	87,000	1,835,000	198,000
Autumn	2,091,000	168,000	2,599,000	259,000
Annual	2,215,000	173,000	2,660,000	180,000
1972-1978				
Winter	2,654,000	250,000	2,926,000	190,000
Spring	2,587,000	214,000	2,745,000	222,000
Summer	1,405,000	75,000	1,552,000	109,000
Autumn	2,133,000	207,000	2,567,000	205,000
Annual	2,215,000	163,000	2,447,000	124,000
	Temperature-based		***Satellite dataset***	
1979-2015				
Winter	2,539,000	211,000	2,539,000	302,000
Spring	2,404,000	219,000	2,403,000	277,000
Summer	1,307,000	111,000	1,306,000	184,000
Autumn	1,951,000	200,000	1,944,000	251,000
Annual	2,049,000	182,000	2,048,000	228,000

Table 10. Comparison of mean and standard deviations for Nordic Arctic seasonal sea ice extents during the different periods, according to (a) our temperature-derived sea ice proxy (using the unadjusted dataset) and (b) the gridded Walsh dataset.

Nordic Arctic (MW09-adjusted data)

Season	*Temperature-based*		*Walsh dataset*	
	Mean	S.D. (σ)	Mean	S.D. (σ)
1901-1939				
Winter	2,669,000	260,000	2,936,000	73,000
Spring	2,482,000	220,000	2,960,000	216,000
Summer	1,317,000	81,000	1,808,000	272,000
Autumn	2,077,000	217,000	2,565,000	19,000
Annual	2,140,000	148,000	2,567,000	128,000
1940-1945				
Winter	2,688,000	298,000	2,902,000	0
Spring	2,552,000	157,000	2,618,000	0
Summer	1,308,000	50,000	1,445,000	0
Autumn	1,979,000	215,000	2,568,000	0
Annual	2,135,000	171,000	2,383,000	0
1946-1952				
Winter	2,651,000	161,000	2,914,000	49,000
Spring	2,455,000	164,000	3,071,000	101,000
Summer	1,351,000	70,000	1,857,000	220,000
Autumn	1,948,000	137,000	2,567,000	5,000
Annual	2,112,000	94,000	2,602,000	80,000
1953-1971				
Winter	2,734,000	268,000	3,194,000	229,000
Spring	2,582,000	187,000	3,012,000	210,000
Summer	1,380,000	87,000	1,835,000	198,000
Autumn	2,091,000	168,000	2,599,000	259,000
Annual	2,215,000	173,000	2,660,000	180,000
1972-1978				
Winter	2,654,000	250,000	2,926,000	190,000
Spring	2,587,000	214,000	2,745,000	222,000
Summer	1,405,000	75,000	1,552,000	109,000
Autumn	2,133,000	207,000	2,567,000	205,000
Annual	2,215,000	163,000	2,447,000	124,000
	Temperature-based		*Satellite dataset*	
1979-2015				
Winter	2,539,000	210,000	2,539,000	302,000
Spring	2,403,000	216,000	2,403,000	277,000
Summer	1,306,000	105,000	1,306,000	184,000
Autumn	1,944,000	182,000	1,944,000	251,000
Annual	2,048,000	180,000	2,048,000	228,000

Table 11. Comparison of mean and standard deviations for Nordic Arctic seasonal sea ice extents during the different periods, according to (a) our temperature-derived sea ice proxy (using the MW09-adjusted dataset) and (b) the gridded Walsh dataset.

Siberian Arctic (unadjusted data)

Season	*Temperature-based*		*AARI dataset*	
	Mean	S.D. (σ)	Mean	S.D. (σ)
1900-1932 (Frolov et al., 2008)				
Winter	4,459,000	0	N/A	N/A
Spring	4,458,000	700	N/A	N/A
Summer	3,622,000	375,000	3,563,000	222,000
Autumn	4,441,000	42,000	N/A	N/A
1933-1952 (Mahoney et al., 2007)				
Winter	4,459,000	0	4,120,000	141,000
Spring	4,458,000	600	4,062,000	88,000
Summer	3,562,000	255,000	3,512,000	314,000
Autumn	4,421,000	28,000	3,696,000	337,000
1953-1971				
Winter	4,459,000	0	4,167,000	24,000
Spring	4,458,000	600	4,065,000	71,000
Summer	3,635,000	195,000	3,242,000	256,000
Autumn	4,444,000	34,000	3,712,000	199,000
1972-1978				
Winter	4,459,000	0	4,179,000	4,000
Spring	4,458,000	400	4,071,000	66,000
Summer	3,610,000	141,000	3,336,000	130,000
Autumn	4,455,000	32,000	3,830,000	148,000
	Temperature-based		*Satellite dataset*	
1979-2015				
Winter	4,459,000	0	4,459,000	3,000
Spring	4,457,000	800	4,458,000	9,000
Summer	3,359,000	326,000	3,359,000	513,000
Autumn	4,415,000	38,000	4,414,000	86,000

Table 12. Comparison of mean and standard deviations for Siberian Arctic seasonal sea ice extents during the different periods, according to (a) our temperature-derived sea ice proxy (using the unadjusted dataset) and (b) the AARI datasets.

Siberian Arctic (MW09-adjusted data)

Season	*Temperature-based*		*AARI dataset*	
	Mean	S.D. (σ)	Mean	S.D. (σ)
1900-1932 (Frolov et al., 2008)				
Winter	4,459,000	0	N/A	N/A
Spring	4,458,000	700	N/A	N/A
Summer	3,622,000	375,000	3,563,000	222,000
Autumn	4,441,000	42,000	N/A	N/A
1933-1952 (Mahoney et al., 2007)				
Winter	4,459,000	0	4,120,000	141,000
Spring	4,458,000	600	4,062,000	88,000
Summer	3,562,000	255,000	3,512,000	314,000
Autumn	4,421,000	28,000	3,696,000	337,000
1953-1971				
Winter	4,459,000	0	4,167,000	24,000
Spring	4,458,000	600	4,065,000	71,000
Summer	3,635,000	195,000	3,242,000	256,000
Autumn	4,444,000	34,000	3,712,000	199,000
1972-1978				
Winter	4,459,000	0	4,179,000	4,000
Spring	4,458,000	400	4,071,000	66,000
Summer	3,610,000	141,000	3,336,000	130,000
Autumn	4,455,000	32,000	3,830,000	148,000
	Temperature-based		*Satellite dataset*	
1979-2015				
Winter	4,459,000	0	4,459,000	3,000
Spring	4,457,000	800	4,458,000	9,000
Summer	3,359,000	326,000	3,359,000	513,000
Autumn	4,415,000	38,000	4,414,000	86,000

Table 13. Comparison of mean and standard deviations for Siberian Arctic seasonal sea ice extents during the different periods, according to (a) our temperature-derived sea ice proxy (using the MW09-adjusted dataset) and (b) the AARI datasets.

6. RECONSTRUCTIONS USING MW09-ADJUSTED DATASET

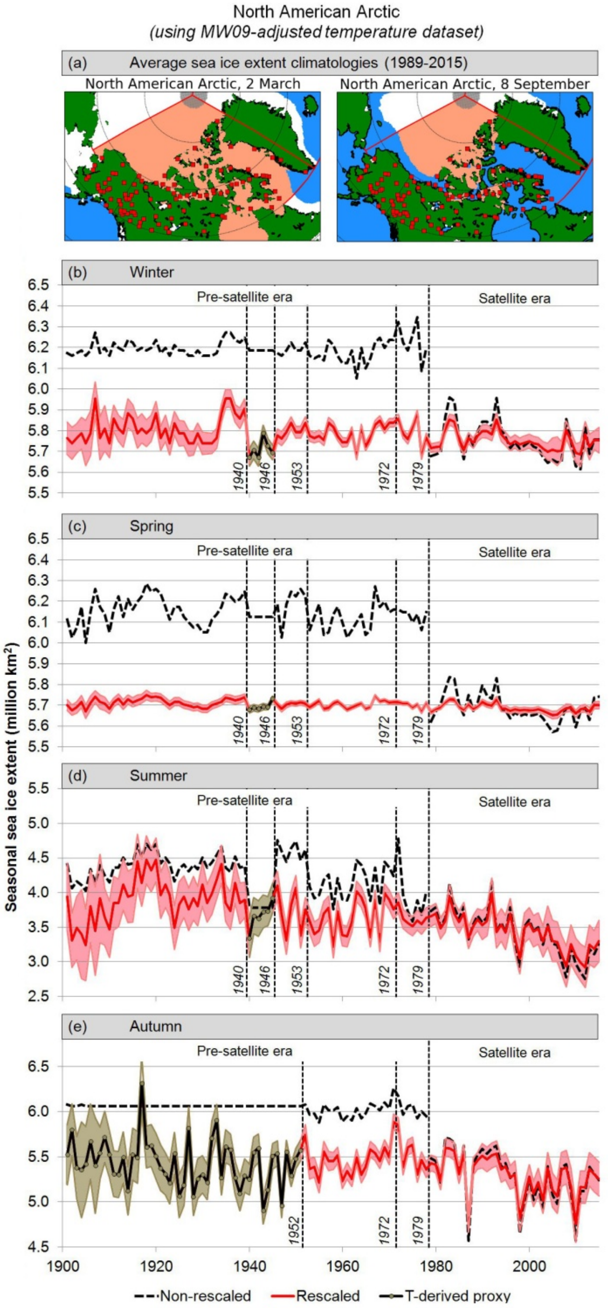

Figure 8. (a) The mean sea ice extents, and station locations for the North American Arctic region. (b-e) Arctic sea ice extent trends before and after rescaling for each season. Note that the y-axes are different for each season.

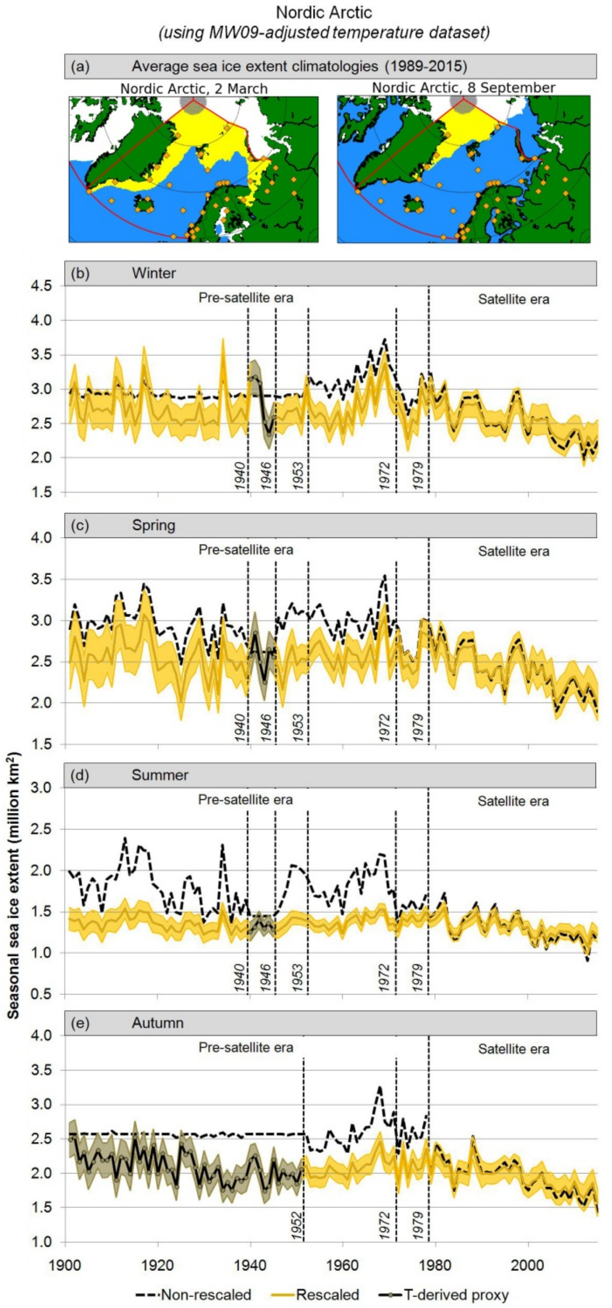

Figure 9. (a) The mean sea ice extents, and station locations for the Nordic Arctic region. (b-e) Arctic sea ice extent trends before and after rescaling for each season. Note that the y-axes are different for each season.

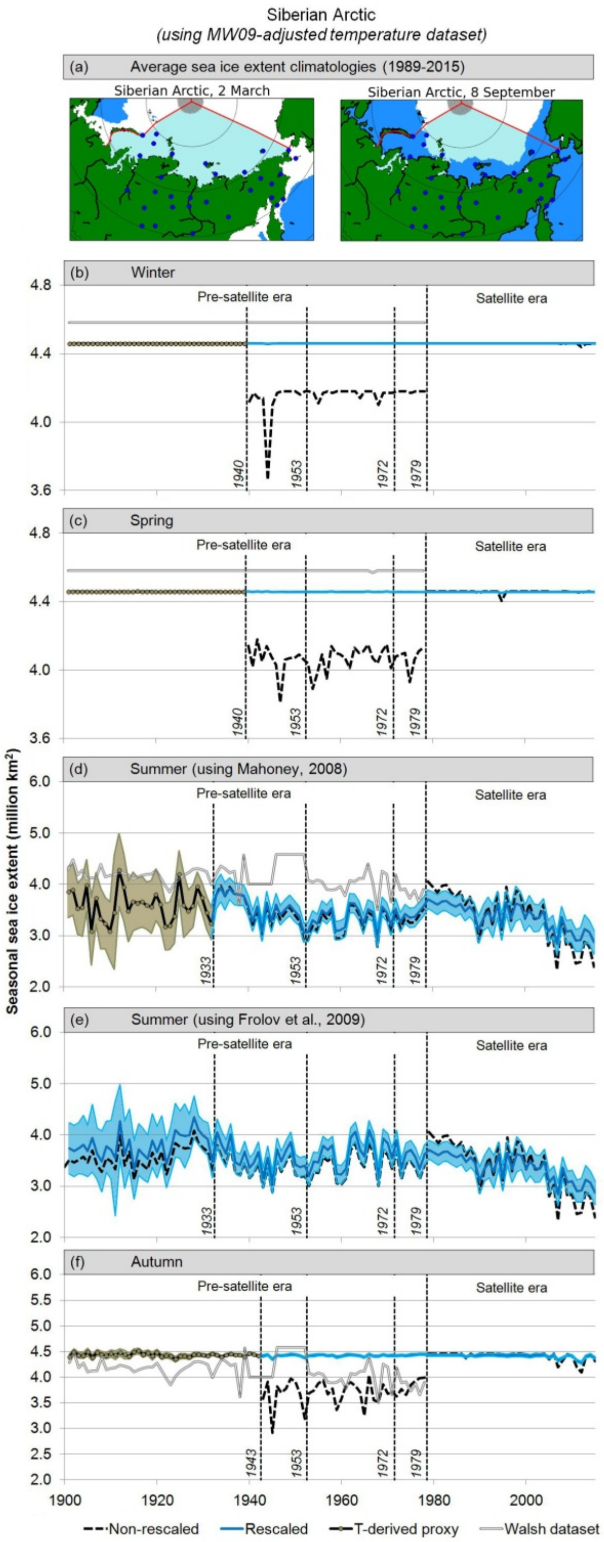

Figure 10. (a) The mean sea ice extents, and station locations for the Siberian Arctic region. (b-e) Arctic sea ice extent trends before and after rescaling for each season. Note that the y-axes are different for each season. Two rescaled estimates for summer trends are provided.

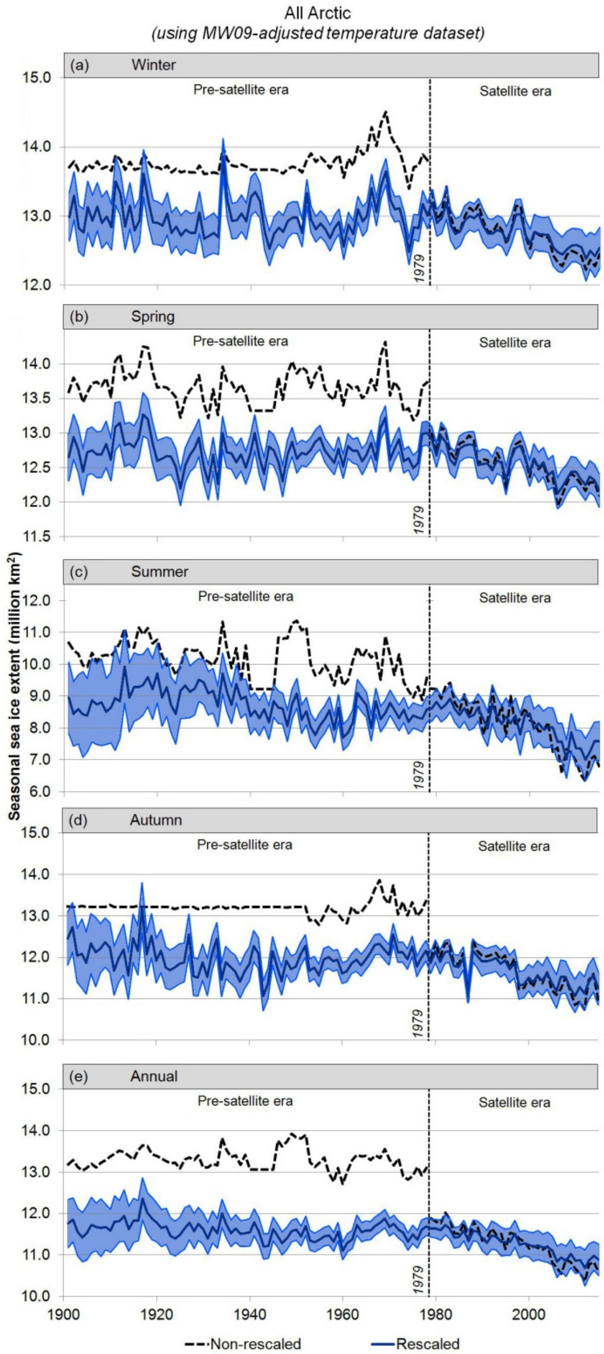

Figure 11. All-Arctic seasonal and annual trends.

REFERENCES

Briggs, W.M., 2016. Uncertainty: the soul of modeling, probability & statistics. Springer International Publishing, Switzerland. 255pp. ISBN: 978-3-319-39756-6. doi: 10.1007/978-3-319-39756-6

Connolly, R., and Connolly, M., 2014a. Has poor station quality biased U.S. temperature trend estimates? *Open Peer Review Journal*, 11 (Climate Science), version 0.1. url: http://oprj.net/articles/climate-science/11 [Accessed 9/2/2016]

Connolly, R., and Connolly, M., 2014b. Urbanization bias III. Estimating the extent of bias in the Historical Climatology Network datasets. *Open Peer Review Journal*, 31 (Climate Science), version 0.1. url: http://oprj.net/articles/climate-science/34 [Accessed 9/2/2016]

Lawrimore, J.H., Menne, M.J., Gleason, B.E., et al., 2011. An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *Journal of Geophysical Research*, 116, D19121. doi:10.1029/2011JD016187

Mahoney, A., 2008. Sea ice edge location and extent in the Russian Arctic, 1933-2006, Boulder, CO, USA. National Snow and Ice Data Center. doi:10.7265/NSW37T8Z [Accessed 16/1/2016]

Mahoney, A.R., Barry, R.G., Smolyanitsky, V., and Fetterer, F., 2008. Observed sea ice extent in the Russian Arctic, 1933-2006. *Journal of Geophysical Research*, 113, C11005. doi:10.1029/2008JC004830

Menne, M.J., and Williams, C.N., 2009. Homogenization of temperature series via pairwise comparisons. *Journal of Climate*, 22, 1700-1717. doi:10.1175/2008JCLI2263.1

Soon, W., Connolly, R., and Connolly, M., 2015. Re-evaluating the role of solar variability on Northern Hemisphere temperature trends since the 19th century. *Earth-Science Reviews*, 150, 409-452. doi: 10.1016/j.earscirev.2015.08.010

Venema, V.K.C., Mestre, O., Aguilar, E., et al., 2012. Benchmarking homogenization algorithms for monthly data. *Climate of the Past*, 8, 89-115. doi:10.5194/cp-8-89-2012

Williams, C.N., Menne, M.J., and Thorne, P.W., 2012. Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. *Journal of Geophysical Research*, 117, D05116. doi:10.1029/2011JD016761