Definitions of the Timing of the Seasons

Here we illustrate two different definitions of the timing of the seasons that are found in the literature. The first framework for looking at timing of the seasons is the threshold-based definition. In this framework, a season begins when a given, unchanging, threshold temperature is passed. This is a useful framework for thinking about seasons from a biological perspective [8, 10, 11, S1] since organisms are, in most cases, more sensitive to the absolute temperature than to the temperature relative to the annual mean. The second framework is the one used in this paper: the phase of the yearly period sinusoidal (Fourier) component. This is a useful framework for understanding the physical response of a system to forcing.

Suppose that the annual cycle in surface temperature at a given location is described by a mean temperature plus a sinusoid $T(t) = \overline{T} + A\cos\left(2\pi t - \theta\right)$. Figure S1a shows such an annual cycle where: $A = 20^{\circ}C$, $\theta = 210$ days, and $\overline{T} = 0^{\circ}C$.

In our threshold-based definition of the annual cycle, we will define the time when the temperature is above zero degrees Celsius as the "warm season" and the time when the temperature is below zero degrees Celsius as the "cold season". We will define "threshold-based-spring" as the day of the year when the the temperature rises above zero, and "threshold-based-fall" as the day of the year when the temperature drops below zero.

In our base-case (Fig. S1a, blue line), the warm season starts on 01 April, the cold season starts on 01 October, and the warm and cold seasons are of equal length (182.5 days).

Now consider what happens if the temperature increases, such that every day is warmer by the same amount. Figure S1b shows the case where the phase (θ) and amplitude (A) have not changed, but the mean temperature ($\overline{T}$) has increased by $6^{\circ}C$. In the threshold-based framework, "spring" now comes 18 days earlier and "fall" comes 18 days later and the warm season is now 71 days longer than the cold season. But from a sinusoidal perspective, nothing has changed about the timing. The date of the peak temperature has not changed, the date of the minimum temperature has not changed, and the dates of most rapid transition have not changed. Only the mean has changed.

Next consider the case where there is a shift in the timing of the seasons, but no change in the mean temperature. Figure S1c shows the case where the annual mean temperature ($\overline{T}$) and amplitude (A) have not changed, but the phase has shifted earlier by 20 days. Only in this case will the two frameworks for timing of the seasons give the same result. In both frameworks the timing of spring has shifted earlier by 20 days, and the timing of fall has also shifted earlier by 20 days. In the threshold framework, the lengths of the warm and cold seasons are equal.

Finally, we can consider the case where both the mean and the phase are changing at the same time. Figure S1d shows the case where the annual mean temperature ($\overline{T}$) has increased to $6^{\circ}C$ and the phase

has shifted earlier by 20 days. In the threshold framework we would say that spring has shifted earlier by 38 days and fall has shifted earlier by 3 days, lengthening the warm season at the expense of the cold season. In the sinusoidal (Fourier) perspective used in this paper we would say that the annual cycle has shifted earlier by 20 days, and the mean temperature has increased.

The Structure of Natural Variability

The land-sea contrast plays a dominant role in setting the climatological distribution of the annual cycle [17, 20-22]. The strength of this relationship is particularly clear when viewed in terms of lag (λ) and gain ($\mathcal{G}$) (e.g. after we have accounted for the variable insolation forcing with latitude).

Figure 2a of the accompanying article shows long-term-mean λ plotted against $\mathcal{G}$ for extratropical grid boxes. The grid points fall on an arc typical of what is seen anytime two sinusoids of different amplitudes interact. At one end of the arc (high $\mathcal{G}$, small λ) are continental points that are very isolated from the ocean. By "very isolated" we mean that one has to go a long way to the west (counter to the prevailing wind direction in midlatitudes) before one encounters any ocean. At the other end of the arc (low $\mathcal{G}$, large λ), are oceanic points that are far from land in the sense that one has to go a long way to the west before one encounters a continent. As discussed in the main body of the article, the position of an individual grid box along this continuum (SRI) is well predicted by the distance between a grid point and the coast to its west (R=0.88). Visually inspection (Fig. S2) shows the strength of the relationship between SRI and "westward distance from the coast".

The curvature of the mixing line between the land and ocean is a robust feature of any seasonally driven model that contains interacting land and ocean regions and can be understood as the natural consequence of interacting sinusoids. We present an analytical relationship for the nonlinear interaction between amplitude and phase in the main text. Some intuition into this relationship can be gained by looking at the shape of weighted averages of two sine waves. Figure S3 presents visually the effect of averaging two sine waves of different amplitude and phase. Here the apexes of each sine wave (red dots) indicate the amplitude and phase of the sinusoid and trace out the shape of the non-linear relationship between amplitude and phase. The result is that when weighted averages of a large and small sine wave of different phases are taken, the large amplitude sine wave is more stable against perturbations in phase, and the small amplitude sine wave is more stable against perturbations in amplitude.

This nonlinear relationship between amplitude and phase also gives insight into the differences in temporal variability seen at different locations. The pointwise relationship between climatological λ and $\mathcal{G}$ and their variability can be visualized as covariance ellipses. In Figure S4a, each ellipse represents the long-term-mean and the variability at a single Northern Hemisphere extratropical grid box. Each ellipse is centered at the long-term-mean λ and $\mathcal{G}$ for the particular grid box. The semi-major axis of the covariance ellipse represents the first eigenvector (scaled by the first eigenvalue) of the pointwise covariance matrix between λ and $\mathcal{G}$. Likewise, the semi-minor axis represents the second eigenvector (scaled by the second eigenvalue) of the same covariance matrix. Color represents westward distance from the coast as given in figure 2a, where positive values (red tones) represent land, and negative values (blue tones) represent

ocean. Ellipses are thus elongated in proportion to their variability in a particular direction. If the semi-major and semi-minor axes are parallel to the x- and y-axis, then λ and $\mathcal{G}$ variability are independent (their covariance matrix is diagonal), and the length of the semi-major and semi-minor axes represent the standard deviation of λ and $\mathcal{G}$ in each direction (times a constant of 1.178, which arises because these are bivariate covariance ellipses). The black line represents the nonlinear relationship between λ and $\mathcal{G}$.

Temporal $\mathcal{G}$ variability is greatest on land and correlates strongly with long-term-mean $\mathcal{G}$ (R=0.83). We thus expect to observe larger amplitude trends where the amplitude is large due to natural variability alone. Conversely, temporal variability in λ is largest on the ocean, and is correlated with the inverse of $\mathcal{G}$ (R=0.62). This structure can be understood as related to the fact that a random perturbation of some finite size will more readily alter the phase of a small amplitude sinusoid than a large sinusoid. We can think of the ocean phase as being less stable due to the simple fact that the annual cycle is not as big over the ocean.

The addition of randomly-phased noise with the same magnitude everywhere produces λ variability $\sim \mathcal{G}^{-1}$. This can be understood by viewing variability in the complex plane, where perturbations are represented by a a radial component and an azimuthal component (figure S4b). The radial component of the noise alters the gain ($\Delta \mathcal{G}$). The azimuthal component (ΔAz) of the noise alters the phase, but the observed change in phase is also a function of $\mathcal{G}$ ($\Delta \lambda = \frac{\Delta Az}{\mathcal{G}}$). The observed azimuthal component of temporal variability scales with the long-term-mean $\mathcal{G}$ (R=0.83). Note that this is the same scaling relationship we find between $\mathcal{G}$ variability and long-term-mean $\mathcal{G}$. Thus, despite the fact that the variance in phase (($\Delta\lambda)^2$) is larger on the ocean, the variance in the temperature perturbations associated with this variability (($\Delta Az)^2$) is larger on land because the phase perturbations are acting on a larger amplitude sinusoid. The variability appears to have a simple, understandable (and surprisingly linear) structure if viewed in the proper framework. An argument could be made that we should be studying variability in azimuthal excursion, not variability in λ, because it is the energetically relevant quantity. However, the same conclusions are reached if the trend analysis is done using azimuthal excursion, rather than λ.

If the pointwise variability is of random phase, the ratio of radial to azimuthal perturbations should be 1. In fact the average ratio is 1.14 ($\sigma = 0.17$) over land and 1.15 ($\sigma = 0.18$) over the ocean. The component of noise that maps onto amplitude variability of the annual cycle is $\sim 15\%$ stronger than the component that maps onto phase variability, perhaps because wintertime variability is larger than in any other season.

Correlation of Lag and Gain Variability with Standard Climate Indices

We test a number of climate indices against timeseries of the spatial average land and ocean lag (λ) and gain ($\mathcal{G}$). (See section 3.6 of [6] for definition and discussion of each index.) We find a number of cross-correlations significant, the most significant being the relationship between λ_{land} and the northern annular mode (NAM) index (R=-0.50, p<0.001) (Table S1). The NAM also exhibits significant cross-correlation with $\mathcal{G}_{\mathrm{land}}$ (R=-0.42, p=0.007). The North Atlantic Oscillation (NAO), a closely related index [S2], also exhibits a significant relationship with λ_{land} (R=-0.42, p=0.007). The Pacific/North American (PNA) index shows a significant relationship with $\mathcal{G}_{\mathrm{ocean}}$ variability (R=0.30, p=0.04), and the Southern Oscillation Index (SOI) exhibits significant cross-correlation with λ_{ocean} (R=-0.32, p=0.02). No significant relationships are found between annual cycle parameters and the Pacific Decadal Oscillation (PDO) index or the southern annular mode (SAM) index.

It is also of interest that λ_{land} and $\mathcal{G}_{\mathrm{land}}$ are significantly correlated (R=0.43, p=0.009), as are λ_{ocean} and $\mathcal{G}_{\mathrm{ocean}}$ (R=0.33, p=0.004), emphasising the intimate relationship between amplitude and phase.

All cross-correlations are done between 1950 and 2007, except for correlations with the SAM which begin in 1979. Significances are assessed using a monte-carlo approach which accounts for the temporal autocorrelation of the timeseries. Phase of the climate index and annual cycle parameter are randomized and correlations are computed (100,000 iterations) to build up a distribution to test (2-tailed) significance against.

The significance of the relationships between northern annular mode variability and λ_{land} raises the question of whether the entire λ_{land} excursion found in the recent record is associated with the excursion in annular mode indices observed at the end of the 20th century [S3, S4]. To test this hypothesis, we regress the 1950-2007 average land timeseries against the NAM index for the same time period. The 1950-2007 average λ_{land} timeseries shows a trend of -1.7 days/(57 years), which is significant at the 99.5% level. Regressing against the NAM index for the same time period, and subtracting the regression timeseries, we are left with a residual trend of -1.0 days/(57 years), which remains significant at the 98.7% confidence level (Fig. S5). Here significance of linear trends is assessed by the phase-randomization monte-carlo method described in the methods, where the 1950-2007 trends are assessed against the autocorrelation structure for the same time period (1950-2007).

Land and ocean spatial average timeseries are constructed by separately averaging λ and $\mathcal{G}$ timeseries from the 152 land locations and the 227 ocean locations with complete temporal coverage from 1950-2007.

We also construct an average λ_{land} timeseries for the smaller set of locations that have complete

temporal coverage from 1850-2007, in order to put the 1954-2007 trend in the context of the variability seen in the preceding 104 years (Fig. S6). As described in the main text, the 1954-2007 trend is found significant (p=0.006) when assessed against the 1850-1953 trends. No similar 54-year period in the early (1850-1953) record shows a significant variability.

Index values for the NAM, the NAO, the PNA pattern, and the SAM are taken from the National Oceanic and Atmospheric Administration's Climate Prediction Center [S5, S6]. Index values for the PDO are taken from values derived by the University of Washington's Joint Institute for the study of the Atmosphere and Ocean [S7]. Values for the SOI are taken from the National Center for Atmospheric Research's Climate & Global Dynamics Climate Analysis Section [S8].

Structure of Anomaly Annual Cycle

Table S2 shows the distribution of variance between the yearly sinusoidal component (that is, the 1 cycle per year Fourier component) and other frequencies for four 27-year periods.

AMIP2 simulated trends

We compare the observed λ_{land} and $\mathcal{G}_{\text{land}}$ trends to those simulated by the Atmospheric Model Intercomparison Project [30-32]. For calculating trends, we restrict ourselves to extratropical locations where there are at least 18 full years of observations over the 23 year AMIP2 simulation period, and apply the same spatial mask to the model simulations as to the data. Figure S7 compares the observed and simulated trends. Of the 25 model runs, 13 incorrectly simulate λ_{land} trends towards later seasons. Of the 12 model runs which simulate a shift in the correct direction (earlier), all predict trends which are less than one third the observed shift. The agreement between observed and simulated $\mathcal{G}_{\text{land}}$ is better, though 13 of 25 of the runs also simulate the incorrect sign $\mathcal{G}_{\text{land}}$ trend over the AMIP2 period.

The Central England Timeseries

Thomson [1] introduced the problem of the large-scale trends in the phase of the annual cycle of surface temperatures by consideration of the phase history of the annual cycle in the Central England Timeseries, the longest duration monthly thermometer record we possess [S9]. *Thomson* observed a steady trend towards later seasons from 1669 to 1940 of 51.1 arc seconds per year (4 days over the 280 years). He attributed this long-term trend in phase to the precession of the equinoxes, claiming to have detected the signal of the anomalistic year (the time from perihelion to perihelion) in this instrumental temperature record. This was quite unexpected, and produced some controversy [35], as our prior intuition about the climate system would have led us to predict that the annual cycle in local surface temperature was dominantly controlled by the tropical year. This observation remains an important poorly understood property of the instrumental record.

More directly relevant to the work at hand, *Thomson* observed in the later part of the record (starting around 1950) an excursion in phase that is anomalous relative to the rest of the record. He attributed this excursion to the observed warming induced by the rapid rise in atmospheric carbon dioxide concentrations [6]. This is a particularly interesting case for our purposes because the excursion noticed by *Thomson* is towards *later* seasons, while we argue for a large-scale shift towards *earlier* seasons over land. As can be seen in Fig. 1e of the main text, England lies in the transition zone between the high North Atlantic that has experienced fairly strong trends towards later seasons and the interior of Europe, that has experienced similarly strong trends towards earlier seasons.

With that in mind, we revisit the Central England thermometer record with an additional seventeen years of data. Figure S8 shows the phase of the annual cycle in surface temperature at Central England. Here we have fit the annual cycle over 21 year periods (sliding window) because we are interested in the low-frequency behavior and corrected for the 3-day offset associated with England's transition to the Gregorian calendar [1]. The centennial-scale trend towards later seasons, interpreted by *Thomson* as precession, is clear (black dotted-line). The phase trend for the 327 year period is 0.016 days per year (57.3 arc seconds per year, least-squares fit). Likewise we see the excursion away from the long-term slope towards later seasons beginning around 1950. However, with the additional data we find that the phase excursion identified by *Thomson* has fully recovered to the long-term trend line. It appears that the phase excursion at Central England is best interpreted in terms of the observed enhanced variability in phase trends [1].

A note on the 3-day calendrical correction. This correction was proposed by *Thomson* [1] to explain the large shift in the phase of the annual cycle in surface temperature in Central England in 1752, the year that England changed from the Julian to the Gregorian calender. *Manley* [S10,S11], who developed

the Central England Timeseries, had only weekly averaged temperature data for some parts of the early record, rather than daily values. It is not known exactly how *Manley* corrected for this calendrical change (his original paper on the record does not give details, stating "This paper is long, and would be far longer if all the stages of a compilation which has taken nearly five years to complete were argued in detail" [S10]), but it is believed that he shifted those records made prior to 14 September 1752 later by two weeks (14 days). This leaves a 3-day (11 days minus 14 days) spurious shift associated with the change in calendar. The belief that Manley corrected for the calendrical shift in this manner is supported by the fact that a 3-day offset in the phase of the annual cycle is indeed found in 1752 [1]. *Manley* could have instead corrected by one week (7 days), but this would have produced a larger, 4-day (11 days minus 7 days), spurious shift, and would imply a large discontinuity in the phase of the annual cycle.

Supplementary Methods

Linear trends from spatially averaged timeseries: Significances of linear trends are computed using a monte-carlo method that accounts for the full autocorrelation structure of the timeseries. We take the Fourier transform of the record segment whose spectral structure we are testing against, randomize the phase of each component, transform back to the time domain, and then calculate the linear trend. This process is repeated many times to build up a distribution of trends under the null hypothesis that the observed trends are a result of random variability as characterized by the autocorrelation structure of the control period.

Composite land annual cycle: Land 27-year composite annual cycles are formed by first averaging within the 27-year averaging window at a given location, and then averaging different locations together. The local 27-year average is formed using only those years where all 12 monthly temperature estimates are available, and we restrict ourselves to the 160 locations where 20 full years of data are available for each of the four 27-year record periods. Southern Hemisphere annual cycles are rotated by 6 months before averaging in space.

Comparison with model simulations: We calculate λ_{land} and $\mathcal{G}_{\mathrm{land}}$ in model output using the same method as used for the observations. We regrid model output to the CRU grid and calculate λ and $\mathcal{G}$ trends for each grid point. We then apply the same mask we use for comparing observed 1900-1953 and 1954-2007 land trends to the model output and spatially average the trends to recover model λ and $\mathcal{G}$ trends. Because the models do not separate land and ocean temperatures within individual grid boxes, we compare model trends to trends derived from the HadCRUT3 dataset, as opposed to the CRUTEM3 dataset. 1900-1953 trends are renormalized to units of days/(47 years) and $°\mathrm{C}/\frac{\mathrm{kW}}{\mathrm{m}^2}/(47$ years$)$ for comparison to 1954-2000 trends. We exclude from analysis four "Climate of the 20th Century" runs (one because less than 40 years of model output exist, and three because they extend into the future and we are unsure where the transition from observed to projected forcing occurs).

References

[S1] Sparks, T. & Menzel, A. Observed changes in seasons: An overview. *International Journal of Climatology* **22**, 1715–1725 (2002).

[S2] Wallace, J. North Atlantic Oscillation/annular mode: Two paradigms - one phenomenon. *Quarterly Journal of the Royal Meteorological Society* **126**, 791–805 (2000).

[S3] Hurrell, J. Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science* **269**, 676 – 679 (1995).

[S4] Thompson, D. & Wallace, J. The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophysical Research Letters* **25**, 1297–1300 (1998).

[S5] Barnston, A. & Livezey, R. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review* **115**, 1083–1126 (1987).

[S6] Mo, C. Relationships between low-frequency variability in the Southern Hemisphere and sea surface temperature anomalies. *Journal of Climate* **13**, 3599–3610 (2000).

[S7] Zhang, Y., Wallace, J. & Battisti, D. ENSO-like interdecadal variability: 1900–93. *Journal of Climate* **10**, 1004–1020 (1997).

[S8] Trenberth, K. Signal versus noise in the Southern Oscillation. *Monthly Weather Review* **112**, 326–332 (1984).

[S9] Parker, D., Legg, T. & Folland, C. A new daily Central England temperature series, 1772-1991. *International Journal of Climatology* **12**, 317–342 (1992).

[S10] Manley, G. The mean temperature of Central England, 1698-1952. *Quarterly Journal of the Royal Meteorological Society* **79**, 242–261 (1953).

[S11] Manley, G. Central England temperatures - monthly means 1659 to 1973. *Quarterly Journal of the Royal Meteorological Society* **100**, 389–405 (1974).

Supplementary Tables

Table S1:   **Correlations between annual cycle parameters and standard climate indices 1950-2007:** P-values are generated by using a 2-tailed, monte-carlo approach. Relationships that are significant at more than 95% confidence are in bold.

	land lag (λ_{land})	ocean lag (λ_{ocean})	land ($\mathcal{G}_{\text{land}}$)	ocean gain ($\mathcal{G}_{\text{ocean}}$)
λ_{land}	-	R=0.04 (p=0.8)	**R=0.43 (p=0.009)**	R=0.00 (p=1)
λ_{ocean}	R=0.042 (p=0.8)	-	R=-0.11 (p=0.6)	**R=0.33 (p=0.004)**
$\mathcal{G}_{\text{land}}$	**R=0.43 (p=0.009)**	R=-0.11 (p=0.6)	-	R=0.13 (p=0.3)
$\mathcal{G}_{\text{ocean}}$	R=0.00 (p=1)	**R=0.33 (p=0.004)**	R=0.13 (p=0.3)	-
land temp	R=-0.33 (p=0.12)	**R=0.45 (p=0.009)**	**R=-0.56 (p=0.002)**	**R=0.39 (p<0.001)**
ocean temp	R=-0.18 (p=0.6)	R=0.34 (p=0.09)	R=-0.24 (p=0.5)	**R=0.44 (p<0.001)**
time	**R=-0.47 (p=0.005)**	R=0.31 (p= 0.2)	**R=-0.46 (p=0.04)**	R=0.19 (p=0.2)
PDO	R=-0.2 (p=0.3)	R=0.12 (p=0.6)	R=-0.33 (p=0.1)	R=-0.02 (p=0.9)
PNA	R=-0.00 (p=1)	R=0.18 (p=0.2)	R=-0.16 (p=0.2)	**R=0.30 (p=0.04)**
NAO	**R=-0.42 (p=0.007)**	R=0.10 (p=0.6)	R=-0.31 (p=0.1)	R=-0.20 (p=0.1)
NAM	**R=-0.50 (p<0.001)**	R=0.12 (p=0.5)	**R=-0.42 (p=0.007)**	R=-0.28 (p=0.06)
SAM	R=-0.01 (p=1)	R=-0.02 (p=0.9)	R=0.03 (p=0.9)	R=0.11 (p=0.6)
SOI	R=-0.01 (p=1)	**R=-0.32 (p=0.02)**	R=0.13 (p=0.4)	R=-0.10 (p=0.5)

Table S2: **Energetics of 27-year anomalies from 108-year mean annual cycle** Partitioning of 27-year anomaly annual cycle (Fig. 2d) energy between the yearly period sinusoidal components and other frequencies. **% anomaly energy at 1 cycle per year** is the ratio of variance explained by the annual component to total variance in the timeseries. **total anomaly energy at 1 cycle per year** is the total variance explained by the annual sinusoid component (in $^\circ\text{C}^2$). **total anomaly energy not at 1 cycle per year** is the total variance not explained by the annual sinusoid component (in $^\circ\text{C}^2$).

	(1900-1926)	(1927-1953)	(1954-1980)	(1981-2007)
% of anomaly energy at 1 cycle per year	24%	88%	19%	80%
total anomaly energy at 1 cycle per year	0.2	2.6	0.4	7.4
total anomaly energy *not* at 1 cycle per year	0.7	0.4	1.9	1.8
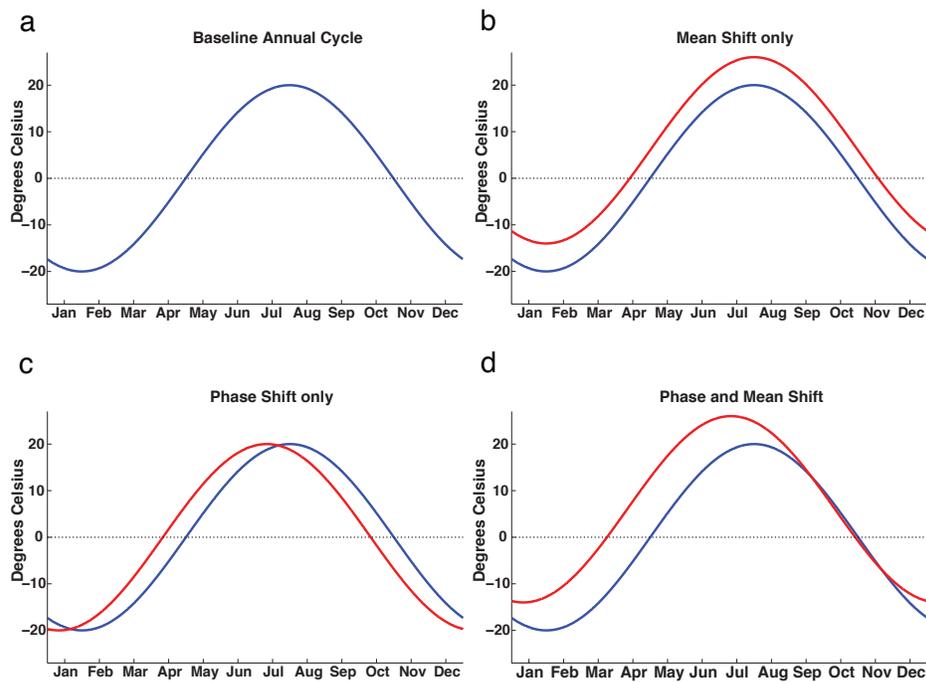
Figure S1: **Cartoon of 2 different frameworks for changes in timing of seasons.** (a) baseline annual cycle (blue in all subpanels) (b) annual mean warming with no phase shift (red line) (c) phase shift with no change in mean (red line) (d) phase shift and change in annual mean temperature (red line)
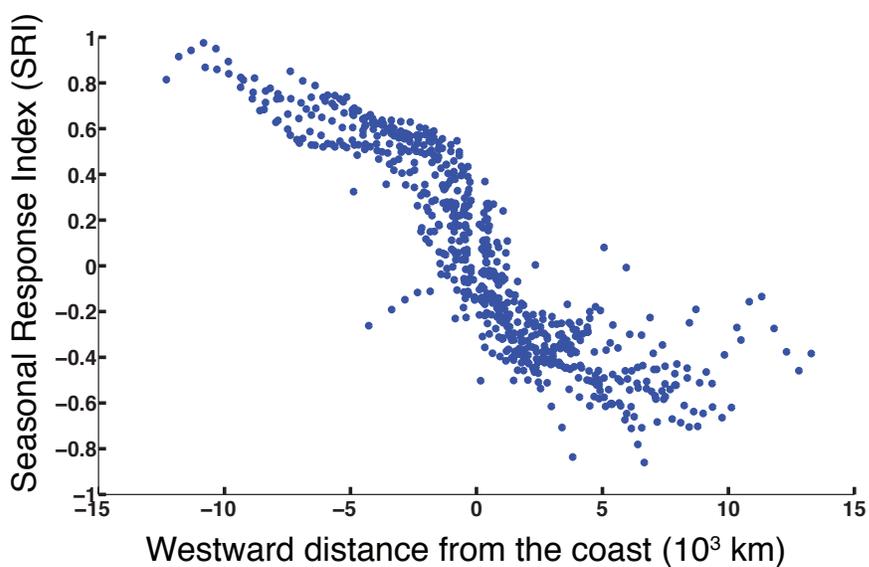
Figure S2: **Seasonal Response Index (SRI) vs westward distance from the coast:** Seasonal Response Index (SRI) versus westward distance from the coast for extratropical Northern Hemisphere grid boxes (R=0.88). Westward distance is defined as the distance one must travel due west from the grid box before encountering a grid box of the opposite surface cover type, and is positive for land grid points and negative for ocean grid points.
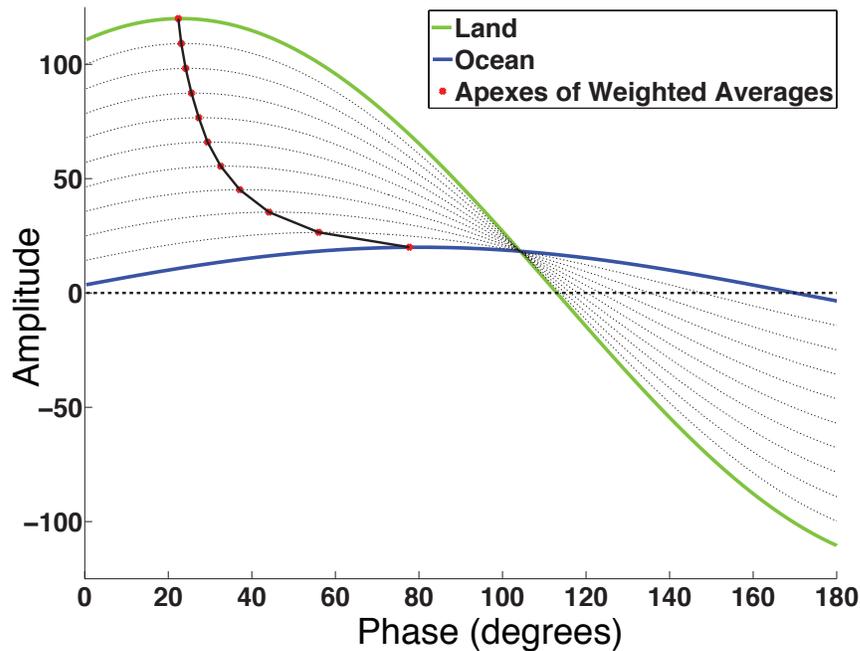
Figure S3: **Nonlinear mixing relationship between sinusoids:** Visualization of effect of averaging sine waves on amplitude and phase. Green and blue lines represent sine waves conceptually meant to represent the annual cycles of land and ocean. Thin black lines give sine waves that represent weighted averages of the land and ocean annual cycles. Red dots indicate the apexes (and thus the amplitude and phase) of each endmember and of the intermediate (thin black lines) sine waves. The thick black line indicates the form of the non-linear relationship between amplitude and phase, and is a discrete form of the black line given in Fig. 2a of accompanying article.
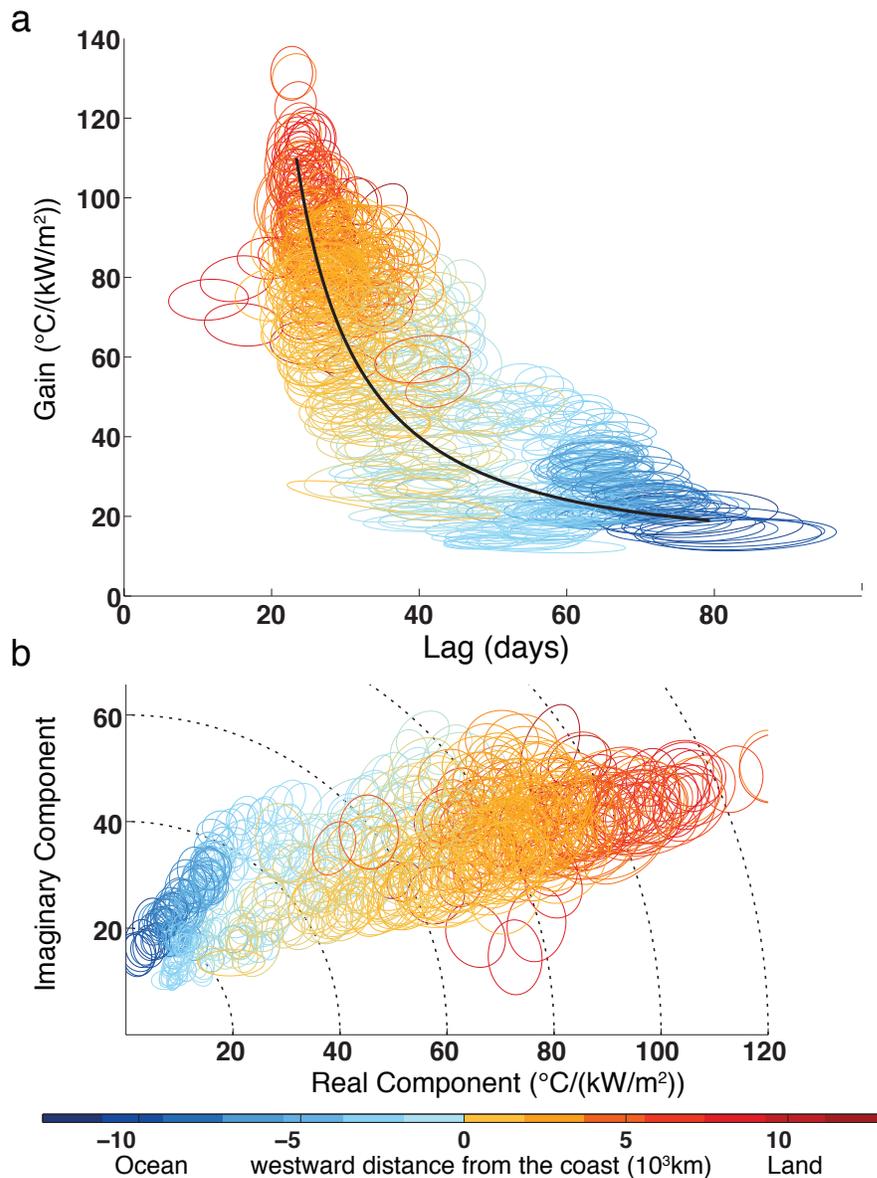
Figure S4: **Extratropical Northern Hemisphere relationship between seasonal parameters: Upper Panel:** Observed relationships between local gain ($\mathcal{G}$) and lag (λ) for Northern Hemisphere Extratropical locations. Center of ellipse represents long term mean λ and $\mathcal{G}$ for an individual grid box. Ellipses represent variability in each direction; the semi-major and semi-minor axes are the eigenvectors (scaled by their eigenvalues) of the local $\lambda/\mathcal{G}$ temporal covariance matrix. Alternatively, they can be thought of as level curves of the bivariate probability density functions. Color represents distance one must travel to the west from the grid point to reach the coast (positive for land, negative for ocean). Outliers with $\lambda < 20$ days are from the Indian Subcontinent and presumably reflect monsoon dynamics. The black line shows the non-linear relationship between amplitude and phase for weighted averages of two endmember sine waves **Bottom Panel:** Same, but parameters are plotted in the complex plane ($\mathcal{G}(\cos\lambda + i\sin\lambda)$)
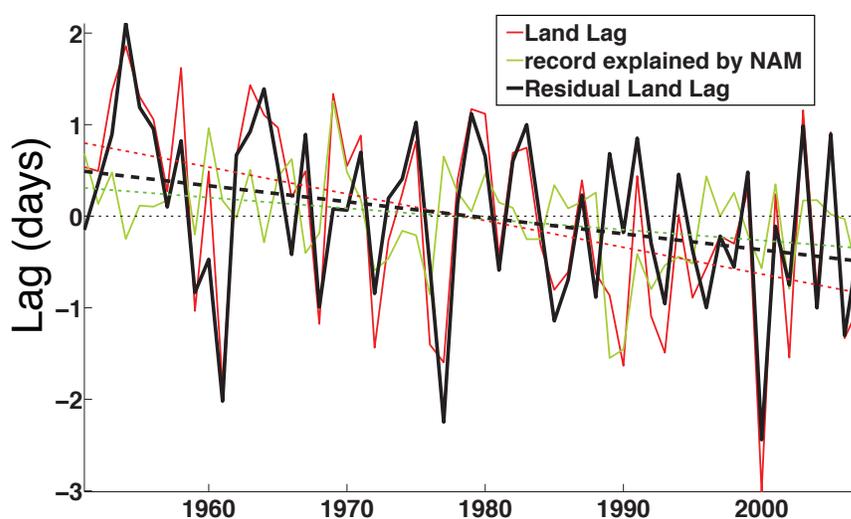
Figure S5: **Regression of land phase lag variability against northern annular mode (NAM) index:** Red solid line is the average λ_{land} anomaly timeseries obtained by averaging phases from grid boxes with complete temporal coverage from 1950-2007. The green solid line is the portion of the timeseries that is explained by regressing against the NAM index. The black solid line shows the residual after subtracting the portion of the record explained by regression against NAM. The residual still exhibits a significant trend. Dotted lines give the linear trends associated with each solid line of the same color.
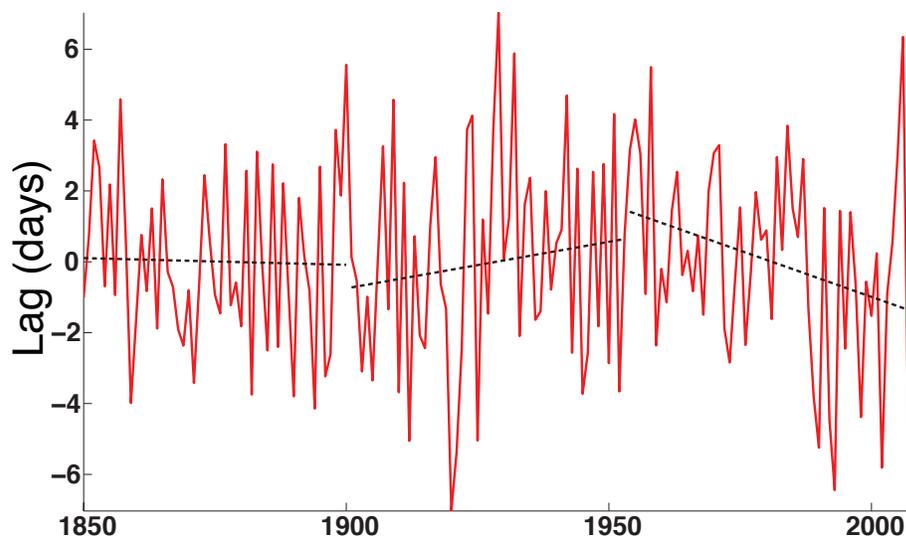
Figure S6: **Land phase lag anomalies 1850-2007:** Spatial average λ_{land} anomaly timeseries, 1850-2007 from 22 grid boxes with complete temporal coverage from 1850-2007. Black dotted lines represent 1850-1900, 1901-1953, and 1954-2007 trend lines. The 1954-2007 trend is significant when compared to variability 1850-1953 variability. No similar length (54-year) record segment in the 1850-1953 period shows a significant trend as assessed by this metric.
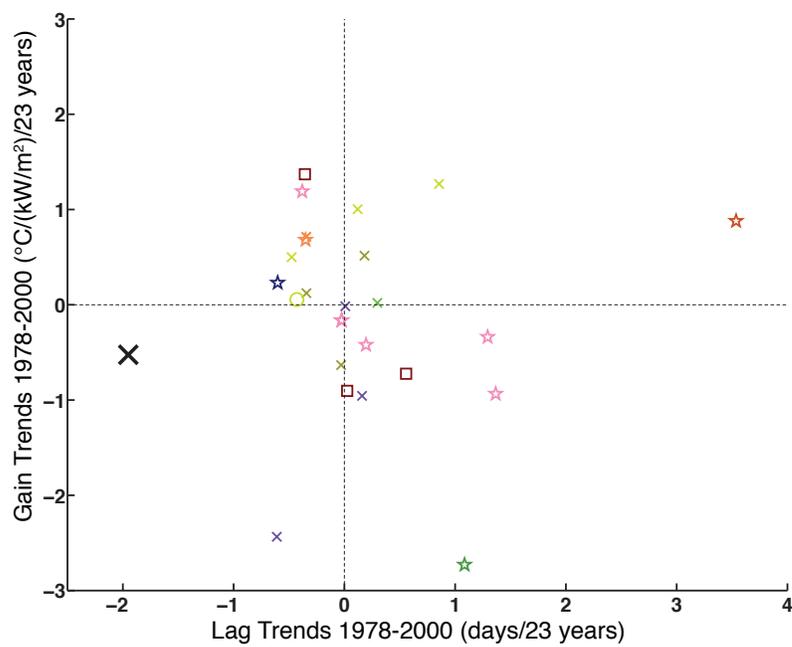
Figure S7: **AMIP model simulated trends vs. observations:** Colored symbols are 1978-2000 trends in land lag (λ_{land}) and gain ($\mathcal{G}_{\text{land}}$) from the AMIP2 simulations (which are forced with observed historical SST fields). Large black X is the observed λ_{land} and $\mathcal{G}_{\text{land}}$ trends for the same period. Shapes and colors for model simulations correspond to the model key in Fig. 3 of main text.
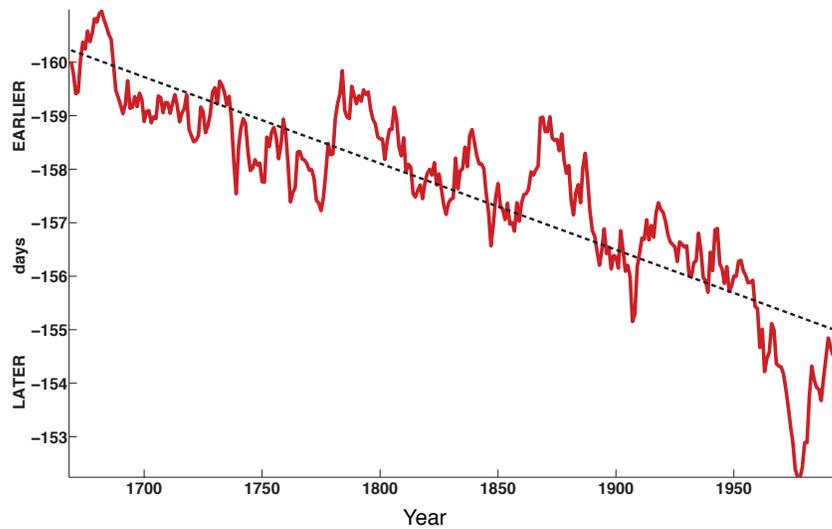
Figure S8: **Phase of Central England Timeseries:** Phase of the annual sinusoidal component in surface temperature calculated from the Central England thermometer timeseries in 21-year record pieces (red line). The black dotted line is the linear trend (0.016 days per year).