

Contribution of changes in atmospheric circulation patterns to extreme temperature trends

Daniel E. Horton^{1,2}, Nathaniel C. Johnson^{3,4,5}, Deepti Singh¹, Daniel L. Swain¹, Bala Rajaratnam^{1,2,6} & Noah S. Diffenbaugh^{1,2}

Surface weather conditions are closely governed by the large-scale circulation of the Earth's atmosphere. Recent increases in the occurrence of some extreme weather phenomena^{1,2} have led to multiple mechanistic hypotheses linking changes in atmospheric circulation to increasing probability of extreme events^{3–5}. However, observed evidence of long-term change in atmospheric circulation remains inconclusive^{6–8}. Here we identify statistically significant trends in the occurrence of atmospheric circulation patterns, which partially explain observed trends in surface temperature extremes over seven mid-latitude regions of the Northern Hemisphere. Using self-organizing map cluster analysis^{9–12}, we detect robust circulation pattern trends in a subset of these regions during both the satellite observation era (1979–2013) and the recent period of rapid Arctic sea-ice decline (1990–2013). Particularly substantial influences include the contribution of increasing trends in anticyclonic circulations to summer and autumn hot extremes over portions of Eurasia and North America, and the contribution of increasing trends in northerly flow to winter cold extremes over central Asia. Our results indicate that although a substantial portion of the observed change in extreme temperature occurrence has resulted from regional- and global-scale thermodynamic changes, the risk of extreme temperatures over some regions has also been altered by recent changes in the frequency, persistence and maximum duration of regional circulation patterns.

Although most land regions show robust warming over the past century¹³, the pattern of change has not been spatially uniform¹⁴. This heterogeneity results from regional differences in the response of the climate system to increasing radiative forcing, and from the background noise of climate variability. Together, these factors substantially increase the challenge of climate change detection, attribution and projection at regional and local scales^{14–16}.

The spatial pattern of changes in extreme weather events has generated arguments that global warming has caused dynamic and/or thermodynamic changes that have differentially altered extreme event

probabilities^{1,17}. Thermodynamic arguments are well understood and observed. For example, the accumulation of heat in the atmosphere has resulted in upward trends in hot extremes, downward trends in the majority of cold extremes, and more intense hydroclimatic events^{1,2}. Dynamic arguments have greater uncertainties^{15–19}. Changes in the large-scale atmospheric circulation—for instance, an increase in the occurrence or persistence of high-amplitude wave patterns—could alter the likelihood of extreme events²⁰. Recent extremes in the Northern Hemisphere mid-latitudes^{1,2,17} have motivated hypotheses of a dynamic linkage between ‘Arctic amplification’, altered atmospheric circulation patterns, and changes in the probability of mid-latitude extremes^{3–5,17}. Despite divergent views on the causal direction of this linkage¹⁷, altered atmospheric dynamics are consistently invoked. Although trends in mean-seasonal mid-atmospheric geopotential heights have been identified (fig. 2.36 of ref. 21; Fig. 1), evidence of changes in the occurrence of sub-seasonal atmospheric patterns remains equivocal, as does their contribution to extreme event probabilities^{6–8}.

Previous efforts to detect trends in atmospheric circulation may have been hampered by narrowly defined, spatially sensitive, and/or non-standardized metrics^{3,6–8,17}. We therefore employ a large-scale spatial characterization approach—self-organizing map (SOM) cluster analysis—to track the occurrence of highly generalized mid-atmospheric circulation patterns. We use 500 hPa geopotential height anomaly fields to describe daily circulation, and group each day's pattern into one of a predefined number of SOM clusters based on a measure of pattern similarity^{9–12} (Methods). The number of clusters is largely dependent on the degree of specificity/generalization required to test a particular hypothesis^{9–12}. To facilitate generalized large-scale mid-atmospheric classification, we use four clusters per domain. Using three reanalyses (Methods), we calculate linear trends (yr^{-1}) in the time series of annual values of (1) the total number of days in each season on which each SOM pattern occurs (‘occurrence’ (d yr^{-1})); (2) the mean length of consecutive occurrence (‘persistence’ (d event^{-1})); and (3) the longest consecutive occurrence (‘maximum duration’ (d event^{-1})). We consider trends in each metric to be robust if

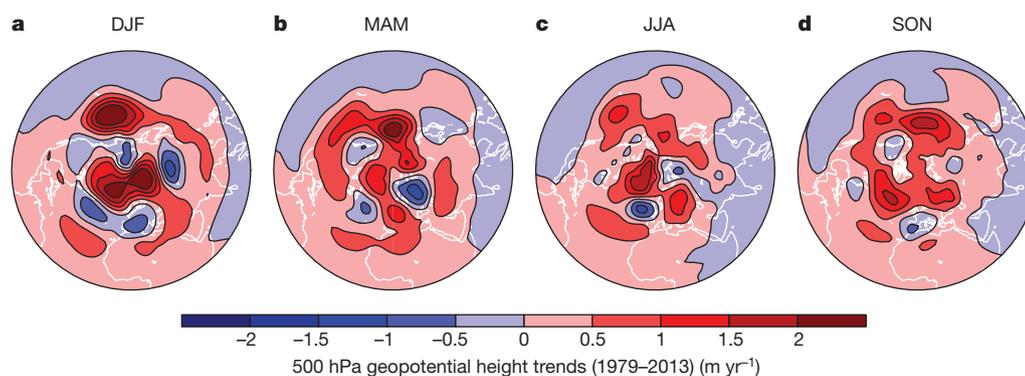


Figure 1 | Trends in mid-atmospheric geopotential heights. Northern Hemisphere polar projections of 1979–2013 seasonal trends (m yr^{-1}) in 500 hPa geopotential heights. Trends are computed for winter (a; December, January, February (DJF)), spring (b; March, April, May (MAM)), summer (c; June, July, August (JJA)) and autumn (d; September, October, November (SON)) seasons. Geopotential height fields are sourced from NCEP–DOE–R2.

¹Department of Earth System Science, Stanford University, Stanford, California 94305, USA. ²Woods Institute for the Environment, Stanford University, Stanford, California 94305, USA. ³International Pacific Research Center, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA. ⁴Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA. ⁵Cooperative Institute for Climate Science, Princeton University, Princeton, New Jersey 08540, USA. ⁶Department of Statistics, Stanford University, Stanford, California 94305, USA.

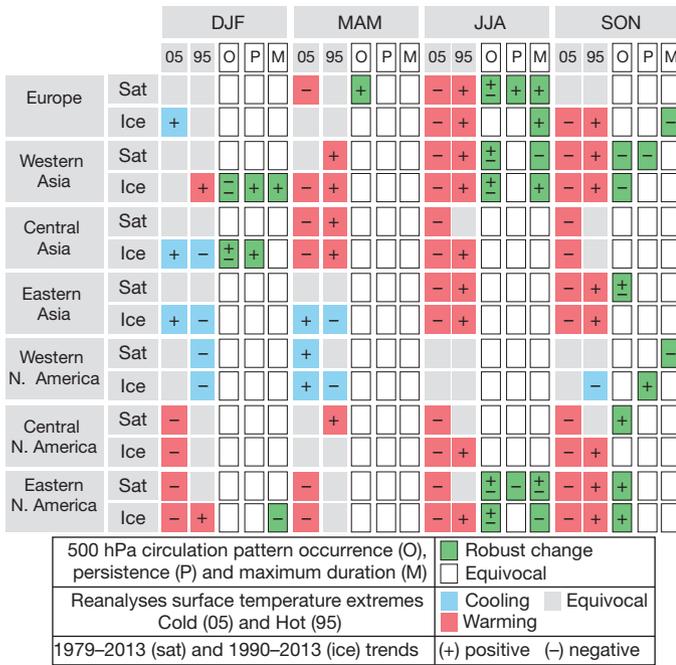


Figure 2 | Trends in surface temperature extremes and atmospheric circulation patterns. Trends are calculated for each Northern Hemisphere season (December, January, February (DJF), winter; March, April, May (MAM), spring; June, July, August (JJA), summer; September, October, November (SON), autumn) for two periods: 1979–2013 (satellite era) and 1990–2013 (ice era). Regional domains (see Fig. 3a) in which one or more of the four SOM circulation patterns demonstrate robust trends in mid-atmospheric circulation pattern occurrence (O), persistence (P), or maximum duration (M) are shown in green (Extended Data Figs 1 and 2). Positive (+) and negative (–) symbols are displayed when all three reanalyses show statistically significant trends in a particular circulation pattern (5% significance level; Methods), and agree on the sign of those trends. Multiple symbols within a box indicate multiple robust pattern trends. White boxes without symbols indicate no statistically significant trends and/or reanalysis disagreement (see Methods). Regional domains with positive and/or negative trends in cold (05) and/or hot (95) extremes receive (+) or (–) symbols when the three reanalyses agree on the sign of the area-weighted trend. Red and blue boxes indicate that the extreme temperature trend results in warming and cooling, respectively, while grey boxes indicate reanalysis disagreement.

matching circulation patterns from all three reanalyses have statistically significant trends of the same sign. We assess the robustness of trends for seven mid-latitude regions (Figs 2 and 3a), over both the era of satellite observation (1979–2013; ‘satellite era’) and the era of rapidly diminishing Arctic sea ice²² (1990–2013; ‘ice era’). We report circulation patterns that pass these robustness criteria, but also discuss results in the context of (1) comprehensive multiple hypothesis testing; (2) removal of the assumption of linear time-series relationships; (3) use of fewer/more clusters; and (4) addition of atmospheric thermal dilation controls (Methods; Extended Data Table 1 and Extended Data Figs 1–3).

Of the 112 total circulation patterns analysed in each period (Methods), the three reanalyses exhibit statistically significant trends in pattern occurrence for a total of 17, 16 and 16 patterns in the satellite era, and 15, 13 and 14 patterns in the ice era (Extended Data Table 1a). Of these significant occurrence trends, 12 satellite-era and 10 ice-era patterns are robustly significant across all three reanalyses (Fig. 2). The majority of robust satellite-era trends occur in summer and autumn, while robust ice-era trends are more evenly distributed over summer, autumn and winter. These patterns are diverse, and include anticyclonic, cyclonic and ‘dipole’ circulations (Extended Data Figs 4 and 5). Patterns with robust trends in both satellite and ice eras are limited to summer and autumn over western Asia and eastern North America.

While the number of significant trends in pattern persistence varies from 5 to 10 across the individual reanalyses (Extended Data Table 1a),

only three robust pattern persistence trends are identified in each period (Fig. 2). Robust maximum duration trends are more prevalent, including five in the satellite era and six in the ice era. These are predominantly associated with summer anticyclonic patterns, although the maximum duration of central Asia winter troughing events demonstrates a robust ice-era increase (Extended Data Figs 4 and 5). In regions with robust trends in multiple patterns, those patterns are generally complimentary. For example, in summer over eastern North America, robustly increasing satellite-era trends in anticyclonic patterns co-occur with robustly decreasing trends in cyclonic patterns.

We next explore the extent to which trends in mid-atmospheric circulation patterns have influenced the likelihood of temperature extremes. For each period, we compute area-weighted trends in the seasonal occurrence of temperature extremes for all days, and for those days associated with each SOM pattern (for example, Fig. 3a, j–m; see Methods). The three reanalyses generally agree on the direction of all-days trends: consistent with enhanced radiative forcing and global warming, most regions and seasons show positive trends in hot occurrence, and negative trends in cold occurrence (Fig. 2).

Hot extremes are projected to increase due to the dynamic and thermodynamic effects of global warming^{1,23}. Consistent with other assessments^{1,2}, we find substantial increases in extreme heat occurrence over the mid-latitudes (Extended Data Figs 6 and 7). For instance, the regional-mean occurrence of summer hot days over Europe, western Asia and eastern North America has increased 0.10, 0.16 and 0.13 d yr⁻¹ yr⁻¹, respectively, over the satellite era (Fig. 3a and Extended Data Table 2a–c). By definition, one would expect (on average) ~4.5 5th/95th percentile events per 3-month season, meaning that an increase of 0.10 d yr⁻¹ yr⁻¹ accumulated over the course of the satellite era (35 years) yields an additional ~3.5 d yr⁻¹, an ~75% increase.

Heatwaves, similar to those which occurred in western Russia in 2010 and Europe in 2003, develop when persistent anticyclonic patterns, often referred to as ‘atmospheric blocking’, initiate a cascade of self-reinforcing, heat-accumulating physical processes^{24,25}. In addition to the increasing trends in extreme heat occurrence, robust positive trends in the occurrence, persistence and maximum duration of satellite-era summer mid-atmospheric anticyclonic patterns are detected over Europe (Fig. 3c, g), western Asia (Fig. 4a, e), and eastern North America (Extended Data Fig. 4c). Robust positive trends in the occurrence of satellite-era anticyclonic patterns are also detected—along with increasing hot extremes—in autumn over eastern North America (Fig. 4c, g), eastern Asia (Fig. 4d, h) and central North America, and in spring over Europe (Extended Data Fig. 1).

Increases in hot extremes may result from dynamic changes (namely greater occurrence and persistence of anticyclonic patterns) as well as from thermodynamic changes such as land cover change or global warming (reflected in the increased intensity of extreme temperature when anticyclonic patterns occur). Over Europe, the summer occurrence of circulations similar to dipole patterns with ridging over the eastern half of the domain (Fig. 3c) increased 0.45 d yr⁻¹ yr⁻¹ over the satellite era, while the persistence and maximum duration increased 0.05 and 0.19 d event⁻¹ yr⁻¹, respectively (Fig. 3g). The trend in the frequency of hot events coincident with this pattern (0.06 d yr⁻¹ yr⁻¹; Fig. 3k and Extended Data Table 2a) accounts for ~61.5% of the total trend in hot extremes over Europe (0.10 d yr⁻¹ yr⁻¹; Fig. 3a). In addition, the number of hot extremes per pattern occurrence has increased for all four patterns (Fig. 3n–q and Extended Data Table 2a). Under the assumption of pattern stationarity, we perform a quantitative partitioning of the dynamic and thermodynamic contributions to the overall extreme temperature trend, as well as to the extreme temperature trends associated with each pattern¹⁰ (Methods). This partitioning reveals that ~27.3% of the 0.10 d yr⁻¹ yr⁻¹ overall increasing trend in hot extremes is driven by the dynamic influence of increased occurrence of the dipole pattern (along with ~35.2% thermodynamic and ~–1.0% interaction influences). Additionally, of the

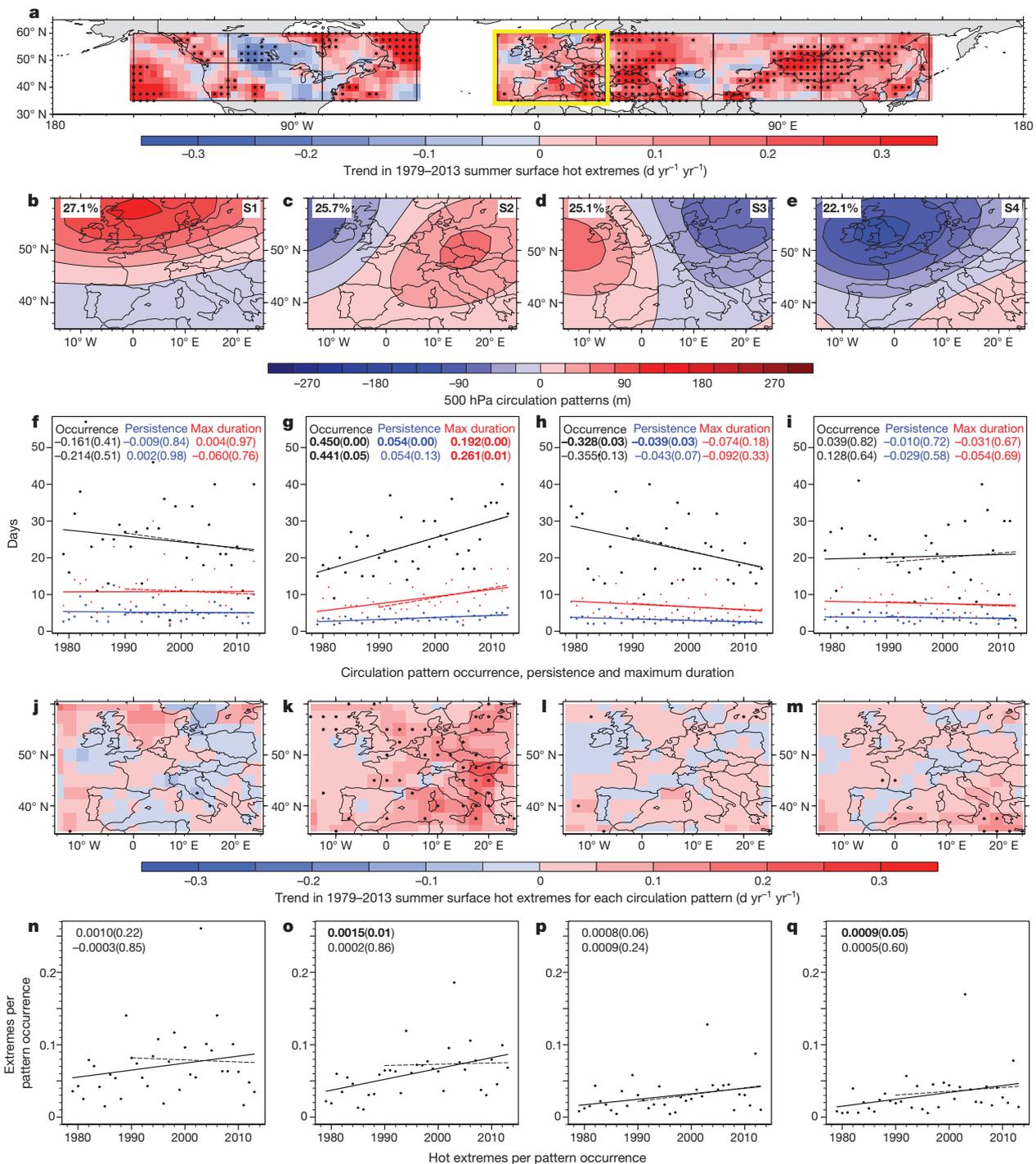


Figure 3 | Trends in circulation patterns and hot extremes over Europe.

a, 1979–2013 trends in summer hot extreme occurrences for all regional domains based on 2-m maximum/minimum temperatures from NCEP-DOE-R2. **b–e**, SOM-derived mid-atmospheric circulation patterns (500 hPa geopotential height anomalies) over Europe. White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. **f–i**, Time series of SOM circulation pattern occurrence (black (d yr^{-1})), persistence (blue (d event^{-1})) and maximum duration (red (d event^{-1})). The slope of the trend line (yr^{-1}) and P values (in parentheses) are colour coded,

$0.06 \text{ d yr}^{-1} \text{ yr}^{-1}$ portion of the hot extremes trend associated with this pattern, $\sim 57.3\%$ is attributable to thermodynamic influences and $\sim 44.3\%$ to increased pattern occurrence (Extended Data Table 2a). Together, these results suggest that the observed increase in extreme summer heat over Europe is attributable to both increasing frequency of blocking circulations and changes in the surface energy balance.

with the values from 1979–2013 (solid trend line) displayed above those from 1990–2013 (dashed trend line). **j–m**, Spatially rendered trends in hot extreme occurrences for days that correspond to each SOM circulation pattern. **n–q**, Time series of the area-weighted mean of hot extremes per pattern occurrence, referred to throughout the text as a measure of the intensity of temperature extremes associated with each pattern. Statistically significant trends (5% significance level; Methods) are shown by stippling in the mapped panels and by bold font in the scatter plots.

Similar results are found in other regions that exhibit robust upward trends in anticyclonic patterns (Fig. 4 and Extended Data Table 2).

Global warming is also generally expected to decrease the frequency of cold extremes¹. In autumn over eastern Asia, the occurrence of satellite-era cold extremes decreased $0.08 \text{ d yr}^{-1} \text{ yr}^{-1}$ (Extended Data Fig. 6g and Extended Data Table 2f), indicating a reduction of

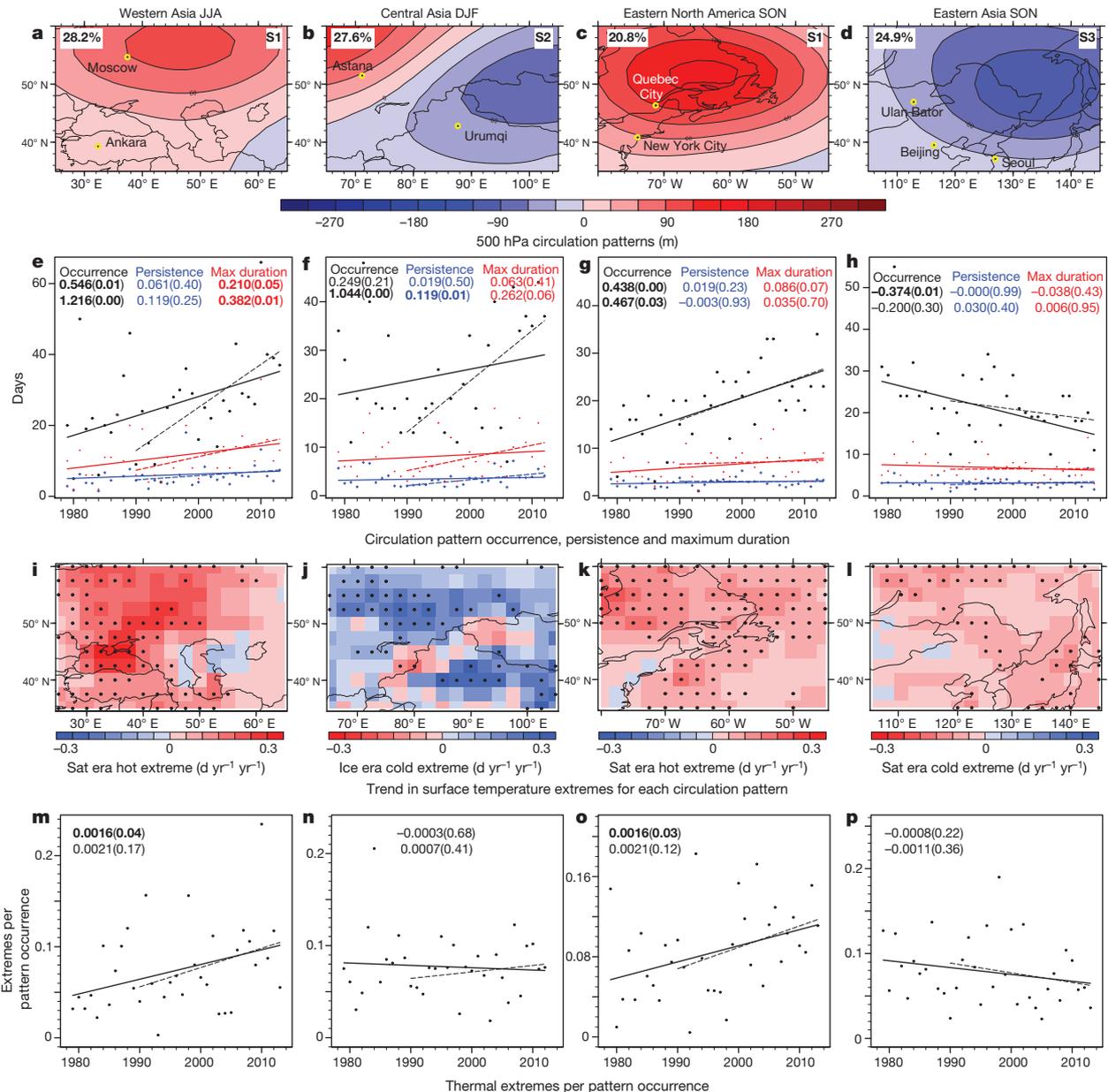


Figure 4 | Circulation pattern and thermal extreme trends for selected regions. Trends in thermal extreme occurrences for selected regions and seasons based on 2-m maximum/minimum temperatures from NCEP-DOE-R2. **a–d**, SOM-derived mid-atmospheric circulation patterns (500 hPa geopotential height anomalies) over western Asia in summer (**a**), central Asia in winter (**b**), eastern North America in autumn (**c**), and eastern Asia in autumn (**d**). White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. In contrast to Fig. 3, just one of the four SOM circulation patterns is displayed for each region. **e–h**, Time series of SOM circulation pattern occurrence (black (d yr⁻¹)), persistence (blue (d event⁻¹)) and maximum duration (red (d event⁻¹)). The slope of the trend line (yr⁻¹)

~60% over the 35-year period. A majority of this decreasing trend (0.05 d yr⁻¹ yr⁻¹; Extended Data Table 2f) is attributable to changes associated with one pattern type: cyclonic circulations capable of advecting cold air equator-ward (Fig. 4d). Less frequent occurrence of cyclonic patterns (Fig. 4h), in conjunction with less intense cold temperature anomalies when cyclonic patterns occur (Fig. 4p), drives ~62.5% of the overall decreasing trend (partitioned ~35.4% dynamic, ~21.9% thermodynamic, ~5.3% interaction). Of the 0.05 d yr⁻¹ yr⁻¹ decrease in extreme cold associated with the trend in cyclonic patterns, partitioning indicates ~56.5% dynamic and ~35.0% thermodynamic influences (Extended Data Table 2f).

and *P* values (in parentheses) are colour coded, with the values from 1979 to 2013 (solid trend line) displayed above those from 1990 to 2013 (dashed trend line). **i–l**, Spatially rendered trends in thermal extreme occurrences for days that correspond to each SOM circulation pattern. **m–p**, Time series of the area-weighted mean of temperature extremes per pattern occurrence, referred to throughout the text as a measure of the intensity of temperature extremes associated with each pattern. Statistically significant trends (5% significance level; Methods) are shown by stippling in the mapped panels and by bold font in the scatter plots. Refer to Extended Data Figs 6 and 7 for satellite-era and ice-era trends in temperature extremes over the regional domains.

In contrast to this expected extreme cold decrease, winter cold extremes over central Asia have increased 0.07 d yr⁻¹ yr⁻¹ over the ice era (Extended Data Fig. 7a). One-hundred-and-forty-nine per cent of this trend (0.10 d yr⁻¹ yr⁻¹; Fig. 4j) occurred when mid-atmospheric circulation was similar to a pattern of troughing in the south and east, and ridging in the northwest (Fig. 4b; ~111.4% dynamic, ~25.6% thermodynamic, ~11.6% interaction). Trend percentages exceeding 100% indicate that other circulation patterns provide negative contributions. Occurrence and persistence of this dipole pattern robustly increased (1.0 d yr⁻¹ yr⁻¹ and 0.12 d event⁻¹ yr⁻¹, Fig. 4f) at the expense of all other circulations (Extended Data Table 2d).

Partitioning indicates that ~75.0% of the extreme cold trend associated with this pattern is due to the dynamic influence of increased pattern occurrence, with ~17.2% linked to thermodynamic influences (Extended Data Table 2d).

Substantial dynamic contributions to the overall trend in cold extremes could be expected given that circulations that support the equator-ward advection of Arctic air will bring anomalously cold temperatures to lower-latitude locales²⁰. Increased occurrence of such patterns has previously been observed, and linked to reduced regional sea-ice and decreased baroclinicity over the Barents and Kara seas^{4,17,26–28}. Positive thermodynamic contributions to the extreme cold trend indicate processes that are in opposition to the direct warming effects of enhanced radiative forcing. For example, positive thermodynamic contributions from three of the four winter patterns over central Asia (Extended Data Table 2d) suggest that these thermodynamic contributions are largely independent of atmospheric circulation, and therefore potentially related to surface processes such as increased snow cover and enhanced diabatic cooling^{4,28}.

The circulation trends detected here cannot as yet be attributed to anthropogenic or natural causes, nor can they be projected to continue into the future. Attribution and projection will require an increased understanding of the causes of the circulation trends, including the ability to identify the signal of an anthropogenically forced trend from the noise of internal decadal-scale climate variability^{16,29}. However, our quantitative partitioning, in conjunction with targeted climate model simulations^{16,29,30}, offers the potential to fingerprint dynamic and thermodynamic climate influences in isolation, which in turn may facilitate attribution of the observed trends, and projection of future trends. We hypothesize that the main assumption of our quantitative partitioning—pattern stationarity—is justified given the expectation that circulation responses to enhanced radiative forcing are likely to reinforce pre-existing modes of natural variability^{15,16}. A related assumption is that the reanalyses act as reasonable proxies for the state of the three-dimensional atmosphere through time. Given uncertainties in the data assimilation and numerical modelling that underpin atmospheric reanalysis, we have restricted our identification criteria to those trends that are statistically significant in all three reanalyses.

Our approach finds robust trends in mid-atmospheric circulation patterns over some regions, and suggests that both dynamic and thermodynamic effects have contributed to observed changes in temperature extremes over the past 35 years. Although thermodynamic influences have largely dominated these changes, dynamic influences have been critical in some regions and seasons. Long-term projections of future dynamic contributions are challenging given the substantial underlying decadal-scale variability, as well as the uncertain impact of anthropogenic forcing on mid-latitude circulation^{15,16}. However, given our finding that many patterns have exhibited increasing (decreasing) intensity of extreme hot (cold) events, and that those trends are coincident with a nearly categorical increase in thermodynamic forcing, the observed trends of increasing hot extremes and decreasing cold extremes could be expected to continue in the coming decades, should greenhouse gases continue to accumulate in the atmosphere.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 November 2014; accepted 27 April 2015.

- Field, C. B. *et al.* (eds) *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (Cambridge Univ. Press, 2012).
- Donat, M. G. *et al.* Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *J. Geophys. Res.* **118**, 2098–2118 (2013).
- Francis, J. A. & Vavrus, S. J. Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophys. Res. Lett.* **39**, L06801 (2012).
- Liu, J., Curry, J. A., Wang, H., Song, M. & Horton, R. M. Impact of declining Arctic sea ice on winter snowfall. *Proc. Natl Acad. Sci. USA* **109**, 4074–4079 (2012).

- Petoukhov, V., Rahmstorf, S., Petri, S. & Schellnhuber, H. J. Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes. *Proc. Natl Acad. Sci. USA* **110**, 5336–5341 (2013).
- Screen, J. A. & Simmonds, I. Caution needed when linking weather extremes to amplified planetary waves. *Proc. Natl Acad. Sci. USA* **110**, E2327 (2013).
- Barnes, E. A., Dunn-Sigouin, E., Masato, G. & Woolings, T. Exploring recent trends in Northern Hemisphere blocking. *Geophys. Res. Lett.* **41**, 638–644 (2014).
- Screen, J. & Simmonds, I. Exploring links between Arctic amplification and mid-latitude weather. *Geophys. Res. Lett.* **40**, 959–964 (2013).
- Kohonen, T. *Self-Organizing Maps* 501 (Springer, 2001).
- Cassano, J. J., Uotila, P., Lynch, A. H. & Cassano, E. N. Predicted changes in synoptic forcing of net precipitation in large Arctic river basins during the 21st century. *J. Geophys. Res.* **112**, G04S49 (2007).
- Johnson, N. C., Feldstein, S. B. & Tremblay, B. The continuum of Northern Hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organizing maps. *J. Clim.* **21**, 6354–6371 (2008).
- Lee, S. & Feldstein, S. B. Detecting ozone- and greenhouse gas-driven wind trends with observational data. *Science* **339**, 563–567 (2013).
- Diffenbaugh, N. S. *et al.* in *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (eds Field, C. B. *et al.*) 137–141 (IPCC, Cambridge Univ. Press, 2014).
- Field, C. B. *et al.* in *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (eds Field, C. B. *et al.*) 1–32 (IPCC, Cambridge Univ. Press, 2014).
- Shepherd, T. G. Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.* **7**, 703–708 (2014).
- Deser, C., Phillips, A. S., Alexander, M. A. & Smoliak, B. V. Projecting North American climate over the next 50 years: uncertainty due to internal variability. *J. Clim.* **27**, 2271–2296 (2014).
- Cohen, J. *et al.* Recent Arctic amplification and extreme mid-latitude weather. *Nat. Geosci.* **7**, 627–637 (2014).
- Palmer, T. Record-breaking winters and global climate change. *Science* **344**, 803–804 (2014).
- Wallace, J. M., Held, I. M., Thompson, D. W. J., Trenberth, K. E. & Walsh, J. E. Global warming and winter weather. *Science* **343**, 729–730 (2014).
- Screen, J. A. Arctic amplification decreases temperature variance in northern mid-to high-latitudes. *Nature Clim. Change* **4**, 577–582 (2014).
- Hartmann, D. L. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) 159–254 (IPCC, Cambridge Univ. Press, 2013).
- Simmonds, I. Comparing and contrasting the behaviour of Arctic and Antarctic sea ice over the 35-year period 1979–2013. *Ann. Glaciol.* **56**, 18–28 (2015).
- Diffenbaugh, N. S. & Ashfaq, M. Intensification of hot extremes in the United States. *Geophys. Res. Lett.* **37**, L15701 (2010).
- Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C. & Vila-Guerau de Arellano, J. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.* **7**, 345–349 (2014).
- Diffenbaugh, N. S., Pal, J. S., Trapp, R. J. & Giorgi, F. Fine-scale processes regulate the response of extreme events to global climate change. *Proc. Natl Acad. Sci. USA* **102**, 15774–15778 (2005).
- Inoue, J., Hori, M. E. & Takaya, K. The role of Barents Sea ice in the wintertime cyclone track and emergence of a warm-Arctic cold-Siberian Anomaly. *J. Clim.* **25**, 2561–2568 (2012).
- Mori, M., Watanabe, M., Shiogama, H., Inoue, J. & Kimoto, M. Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades. *Nat. Geosci.* **7**, 869–873 (2014).
- Cohen, J., Furtado, J., Barlow, J. M., Alexeev, V. & Cherry, J. Arctic warming, increasing fall snow cover and widespread boreal winter cooling. *Environ. Res. Lett.* **7**, 014007 (2012).
- Screen, J. A., Deser, C., Simmonds, I. & Tomas, R. Atmospheric impacts of Arctic sea-ice loss, 1979–2009: Separating forced change from atmospheric internal variability. *Clim. Dyn.* **43**, 333–344 (2014).
- Screen, J. A., Simmonds, I., Deser, C. & Tomas, R. The atmospheric response to three decades of observed Arctic sea ice loss. *J. Clim.* **26**, 1230–1248 (2013).

Acknowledgements Work by D.E.H., D.S., D.L.S. and N.S.D. was supported by NSF CAREER Award 0955283, DOE Integrated Assessment Research Program Grant No. DE-SC005171DE-SC005171, and a G.J. Lieberman Fellowship to D.S. Contributions from N.C.J. were supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections program award NA14OAR4310189. B.R. acknowledges support from the US Air Force Office of Scientific Research (FA9550-13-1-0043), the US National Science Foundation (DMS-0906392, DMS-CMG-1025465, AGS-1003823, DMS-1106642, and DMS-CAREER-1352656), the Defense Advanced Research Projects Agency (DARPA YFA N66001-111-4131), and the UPS Foundation (SMC-DBNKY). We thank B. Santer, J. Cattiaux, D. Touma, and J. S. Mankin for discussions that improved the manuscript. Computational resources for data processing and analysis were provided by the Center for Computational Earth and Environmental Science in the School of Earth, Energy, and Environmental Sciences at Stanford University.

Author Contributions D.E.H. conceived the study, D.E.H., N.C.J., D.S., D.L.S. and N.S.D. designed the analysis and co-wrote the manuscript. D.E.H., N.C.J. and D.S. provided analysis tools. D.E.H. performed the analysis. B.R. provided and described the multiple hypothesis testing and transformation analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.E.H. (danethan@stanford.edu).

METHODS

Categorization of circulation patterns. We use SOM cluster analysis^{9–12} to categorize large-scale circulation patterns over seven Northern Hemisphere domains³¹ using daily 500 hPa geopotential height anomaly fields from the NCAR/NCEP-R1 (ref. 32), NCEP-DOE-R2 (ref. 33) and ECMWF ERA-Interim (ref. 34) reanalyses. Daily anomalies are calculated by subtracting the seasonal cycle (calendar-day mean) from each grid cell. Reanalyses are analysed individually to maintain their physical consistency, and to facilitate their intercomparison. The SOM's unsupervised learning algorithm requires neither a priori knowledge of which types of circulation patterns might be detected, nor the specific geographic regions in which they might occur. Geopotential height anomaly fields from each day are assigned to one of a pre-defined number of nodes, according to pattern similarity. The final SOM patterns are obtained by minimizing the Euclidean distance between iteratively updated nodes and their matching daily geopotential height anomaly fields¹¹. Each SOM pattern can therefore be viewed as a representative composite of relatively similar circulation patterns.

Owing to global-scale warming, trends in geopotential height anomalies record both altered atmospheric circulation patterns and the thermal expansion of the troposphere. To isolate the signal of circulation pattern change, previous clustering analyses have assumed uniform thermal dilation and removed either the domain average³⁵ or domain average linear trend³⁶ from the daily-scale anomalies. In our analysis, we find that 1979–2013 trends in Northern Hemisphere geopotential height anomalies are non-uniform in both magnitude and sign (Fig. 1), and demonstrate substantial seasonal, regional and latitudinal differences (Extended Data Fig. 3e). These findings suggest that for the relatively short period of our analysis, an assumption of uniform thermal dilation is inappropriate. Moreover, the strong spatial heterogeneity indicates the importance of large-scale dynamics in the regional geopotential height trends, and so the removal of local geopotential height trends would conflate dynamic changes with thermal dilation. Therefore, in the main text we present results and conclusions based on raw geopotential height data. However, despite the lack of uniform expansion, we have conducted an analysis that attempts to account for the effects of thermal dilation. SOM analyses are performed on geopotential heights that have been detrended by removing the seasonal mean hemispheric trend (Extended Data Fig. 3e) from each grid cell. Results from this analysis indicate that the magnitude, significance and sign of circulation trends are sensitive to the method of controlling for thermal expansion (Extended Data Fig. 3f–j). Despite this sensitivity, the conclusions presented in the main text are supported, in that both raw and detrended analyses generally suggest trends of similar magnitude, sign and significance.

Based on domain-wide pattern correlations between daily height field anomalies and different SOM node counts¹² (Extended Data Fig. 8), we divide circulation patterns over each domain into four SOM nodes (for example, Fig. 3b–e). To determine a suitable number of nodes, a suite of different node counts were analysed. We found that four nodes were sufficiently great in number to capture a diversity of highly generalized circulation patterns, but sufficiently few to facilitate convenient presentation and, critically, to prevent overly similar SOM patterns¹². To test the sensitivity of our results to the number of nodes, we present 2-, 4-, 8- and 16-node SOMs for the summer season over the European domain (Extended Data Figs 1 and 2). Based on these analyses, it is apparent that a 2-node SOM is insufficient to capture the diversity of circulation patterns that are found in the reanalysis data (Extended Data Fig. 1a), whereas 8- and 16-node SOMs produce nodes with overly similar circulation patterns (Extended Data Figs 1c and 2). Examination of the 4-, 8- and 16-node SOMs largely verifies the conclusions drawn from the 4-node SOM: that the occurrence of patterns with ridging over the eastern half of the European domain has increased over time, while the occurrence of complimentary patterns has decreased. We note that these pattern trends are not identified in the 2-node SOM, confirming that two nodes are too few to capture specific circulation patterns that are critical for extreme temperature occurrence. Similar node-count analyses for other regions/seasons likewise verify the conclusions drawn from the 4-node analyses that are presented in the main text (not shown).

Calculation of robust trends in circulation patterns. For each season in each year, we calculate (1) the total number of days on which each SOM pattern occurs (occurrence (d yr^{-1})); (2) the mean length of consecutive occurrence (persistence (d event^{-1})); and (3) the longest consecutive occurrence (maximum duration (d event^{-1})). A trend in one of these characteristics is considered robust when the trend in that pattern is (1) statistically significant in all three reanalyses, and (2) of the same sign in all three reanalyses. Trends with regression coefficients that surpass the 5% significance (95% confidence) threshold are considered statistically significant. Trends are calculated across satellite-era and ice-era annual time series using the approach of ref. 37, which allows us to account for temporal dependence. Here, the trends are calculated using linear least squares regression, but to account for temporal dependence, the confidence bounds of annual time-series trends with

lag-one autocorrelation greater than the 5% significance level are recalculated by adjusting the number of degrees of freedom used to compute the regression coefficient significance³⁷. As a result, unlike a simple linear regression, this approach does not rely on the independent and identically distributed assumption for the residuals, but instead accounts for temporal dependence using a red noise assumption.

The approach of ref. 37 also assumes that the distribution of the residuals is Gaussian. Using the Anderson–Darling³⁸ test for normality, we find that 91–100% of residual distributions in each metric of each reanalysis do not reject the null hypothesis of Gaussianity when multiple hypothesis testing controlling the family-wise error rate³⁹ (FWER) at the 5% significance level is considered (Extended Data Table 1b). The Gaussianity assumption is therefore largely appropriate. However, due to the identification of non-normality in some distributions, particularly in the persistence and maximum duration metrics, we apply Box–Cox power transformations⁴⁰ to all distributions. Using the Anderson–Darling test, we find that 96–100% of the distributions of the residuals in the transformed setting do not reject the null hypothesis of Gaussianity at the 5% significance level. Furthermore, when multiple hypothesis testing is considered by controlling the FWER at the 5% level, 100% of the individual tests are non-significant (Extended Data Table 1b). In addition, the number of pattern trends identified as significant in the transformed case is largely consistent with the non-transformed regression analysis. However, additional significant trends in the persistence and maximum duration metrics are identified when Box–Cox transformations are used (Extended Data Table 1a). The large overlap of results between the transformed and non-transformed analyses suggests that although individual residual distributions may vary, the Gaussian assumption applied throughout this study is for the most part quite robust, although in some cases not valid. In sum, the non-transformed analysis that fits a linear relationship to circulation pattern metrics allows for a relatively simple classification of two short analysis periods, while simultaneously accounting for temporal dependence in a large number ($>2,000$) of individual time series (7 regions \times 4 nodes \times 4 seasons \times 3 characteristics \times 2 time periods \times 3 reanalyses).

Because SOM nodes are calculated independently for each reanalysis, individual SOM patterns must be matched between the three reanalyses in order to determine whether an individual pattern shows robust results across all three reanalyses. To determine which SOM patterns are the closest match between the three reanalyses, the root mean square error (RMSE) is calculated between the SOM patterns of one reanalysis and those of the other reanalyses. Patterns with the smallest RMSEs are considered matches. Although we undertake a multi-reanalysis robustness evaluation, further work is needed to confirm that other available reanalyses (such as CFSR and MERRA) show the same trends.

Multiple hypothesis testing of linear trends. It is possible that some of the trends identified as significant in any individual reanalysis could occur by chance. In addition to screening for those patterns that are significant in all three reanalyses (our ‘robustness’ criterion), we also employ formal multi-hypothesis testing using several methodologies. The first is the familywise error rate (FWER). This type of error metric controls the probability of falsely rejecting any null hypothesis, and is considered one of the strictest forms of error control³⁹. Since a certain number of false rejections can happen by chance alone, one can account for this formally by using the k -familywise error rate⁴¹ (k -FWER). The k -FWER controls the probability of falsely rejecting k or more null hypotheses, and aims to formalize the concept that some of the hypotheses will be rejected by chance. One option for the value of k is to use the expected number of hypotheses that will be rejected at a given significance level. For instance, in our study, out of 112 total ‘local’ hypotheses, 5 or 6 hypotheses will be significant at the 5% significance level by chance ($112 \times 0.05 = 5.6$). In this case, one can evaluate the probability that 7 or more hypotheses are falsely rejected, since on average about 6 could be rejected as significant by chance. The third metric is the false discovery rate (FDR), which controls the expectation of the ratio given by the number of false rejections divided by the total number of rejections³⁹.

All of the above measures of error control aim to guard against hypotheses being falsely declared as significant in the context of multiple tests. To be thorough, we have implemented all three types of error control at both global significance levels of 5% and 10%. The results of these analyses are summarized in Extended Data Table 2. We note that all three metrics heavily favour the null as they are designed to protect against the possibility of false positives. Despite this, the presence of local tests that reject the null represents a strong confirmation of the significance of those local tests. The fact that a number of local hypotheses still prevail as significant, even after imposing much stricter multiple testing error controls, arises partly from the fact that some of the local P values indicate trends that are so highly significant that they can withstand the stricter multiple testing error control metrics. This rigorous multiple-testing error control yields increased credibility to the scientific conclusions of robust trends in pattern occurrence.

Temperature extremes. Daily-scale hot and cold extreme occurrences are calculated using temperature anomalies at each grid cell. Temperature anomalies are computed by removing the seasonal cycle from daily reanalysis 2-m maximum/minimum temperatures. Similar to previous studies^{1,2}, temperature extremes are calculated based on the statistical distribution of daily temperature anomalies²⁰. Hot/cold extreme thresholds are defined as the 95th/5th percentile value of the 1979–2013 daily 2-m maximum/minimum temperature anomaly distribution (for example, for JJA, the population of daily-maximum temperature anomalies from the months of June, July and August in the years 1979–2013). Hot/cold extreme occurrences are defined as days on which the daily temperature anomalies are greater/less than (or equal to) the hot/cold extreme thresholds. Reanalysis temperature extremes are qualitatively similar to those found in station-based observations². Given this similarity, we use the reanalysis temperatures in order to maintain internal physical consistency between daily 2-m temperatures and daily atmospheric circulation (as represented by the 500 hPa SOM circulation patterns). Trends in temperature extreme occurrence are computed across satellite-era and ice-era annual time series following the methodology of ref. 37.

Quantitative partitioning. To determine the dynamic and thermodynamic contributions to trends in temperature extreme occurrence, we adapt the climate change partitioning methodology of Cassano *et al.*¹⁰. Our adapted methodology partitions the contributions of dynamic and thermodynamic changes to (1) the overall trend in temperature extreme occurrence and (2) the trends associated with individual SOM circulation patterns. Previous applications of the Cassano *et al.*¹⁰ methodology indicate that partitioning is largely insensitive to the number of SOM nodes used in the analysis⁴². All trends in temperature extremes in the below methodology are area-weighted averages. Following Cassano *et al.*¹⁰:

$$E = \sum_{i=1}^K E_i f_i$$

where E is the frequency of extreme temperature occurrence, f_i is the frequency of occurrence of SOM pattern i , E_i is the frequency of extreme temperature occurrence when SOM pattern i occurs, and K is the total number of SOM nodes. We decompose E and f into time mean and deviation from time mean components:

$$E = \sum_{i=1}^K (\bar{E}_i + E'_i)(\bar{f}_i + f'_i)$$

Now we differentiate the above equation with respect to time, noting that the mean values are constants:

$$\frac{dE}{dt} = \sum_{i=1}^K \left(\bar{f}_i \frac{dE'_i}{dt} + \bar{E}_i \frac{df'_i}{dt} + \frac{d}{dt}(E'_i f'_i) \right)$$

The derivative on the left-hand side provides the area-weighted average trend in the seasonal occurrence of temperature extremes for all days. The summation on the right-hand side, from left to right, provides the thermodynamic, dynamic and interaction contributions for days associated with each SOM pattern, i .

The thermodynamic contribution of each circulation pattern's extreme temperature trend assumes that each SOM pattern is stationary in time, and that trends in extremes that result during this pattern are the result of influences unrelated to circulation, such as changes in long-wave radiation from increasing

greenhouse gas concentrations, or changes in surface fluxes of moisture and/or radiation resulting from changes in land cover. The thermodynamic contribution associated with each circulation pattern is determined by taking the product of the trend in the intensity of temperature extremes and the mean occurrence of the circulation pattern. Trends in the intensity of temperature extremes are computed by calculating the trend in area-weighted extreme occurrence per pattern occurrence (for example, Fig. 3n–q).

The dynamic contribution of each circulation pattern's extreme temperature trend assumes that, on average, a circulation pattern is associated with a portion of the total extreme event trend, and that changes in the occurrence frequency of that circulation pattern will modify the occurrence frequency of extreme events. The dynamic contribution associated with each circulation pattern is determined by taking the product of the trend in circulation pattern occurrences and the mean number of extreme events per pattern occurrence.

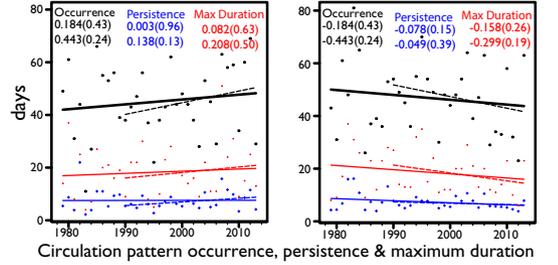
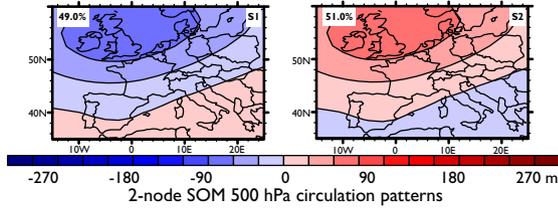
The third component represents the interaction between dynamic and thermodynamic changes, and captures contributions that result from changes in the dynamic component acting on changes in the thermodynamic component, such as the positive/negative feedbacks of surface–atmosphere interactions. The interactive term is determined by computing the trend in the product of circulation pattern occurrence deviations and intensity of temperature extreme deviations.

Code availability. SOM code is available at <http://www.cis.hut.fi/projects/somtoolbox/>. All other analysis code is available upon request from the corresponding author (danethan@stanford.edu).

Reanalysis data sets. NCAR/NCEP-Reanalysis 1 data was downloaded from <http://www.esrl.noaa.gov/psd/>; NCEP-DOE-Reanalysis 2 data was downloaded from <http://www.esrl.noaa.gov/psd/>; ECMWF ERA-Interim data was downloaded from <http://www.ecmwf.int/>.

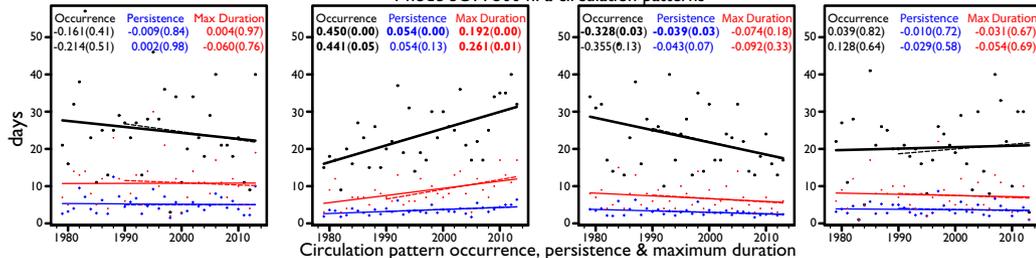
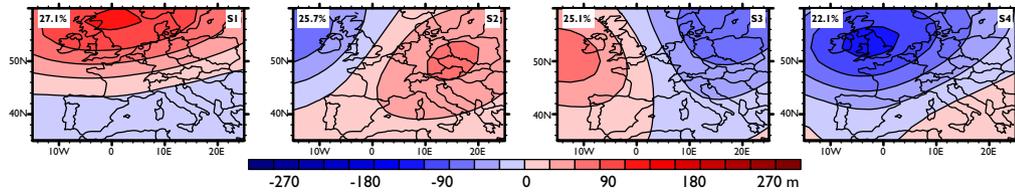
31. Screen, J. A. & Simmonds, I. Amplified mid-latitude planetary waves favour particular regional weather extremes. *Nature Clim. Change* **4**, 704–709 (2014).
32. Kalnay, E. *et al.* The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
33. Kanamitsu, M. *et al.* NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Am. Meteorol. Soc.* **83**, 1631–1643 (2002).
34. Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).
35. Driouech, F., Déqué, M. & Sánchez-Gómez, E. Weather regimes–Moroccan precipitation link in a regional climate change simulation. *Glob. Plan. Ch.* **72**, 1–10 (2010).
36. Cattiaux, J., Douville, H. & Peings, P. European temperatures in CMIP5: origins of present-day biases and future uncertainties. *Clim. Dyn.* **41**, 2889–2907 (2013).
37. Zwiers, F. W. & von Storch, H. Taking serial correlation into account in tests of the mean. *J. Clim.* **8**, 336–351 (1995).
38. Anderson, T. W. & Darling, D. A. Asymptotic theory of certain “goodness-of-fit” criteria based on stochastic processes. *Ann. Math. Stat.* **23**, 193–212 (1952).
39. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc., B* **57**, 289–300 (1995).
40. Box, G. E. P. & Cox, D. R. An analysis of transformations. *J. R. Stat. Soc. B* **26**, 211–243 (1964).
41. Lehmann, E. & Romano, J. P. Generalizations of the familywise error rate. *Ann. Stat.* **33**, 1138–1154 (2005).
42. Skific, N., Francis, J. A. & Cassano, J. J. Attribution of projected changes in atmospheric moisture transport in the Arctic: A self-organizing map perspective. *J. Clim.* **22**, 4135–4153 (2009).

a



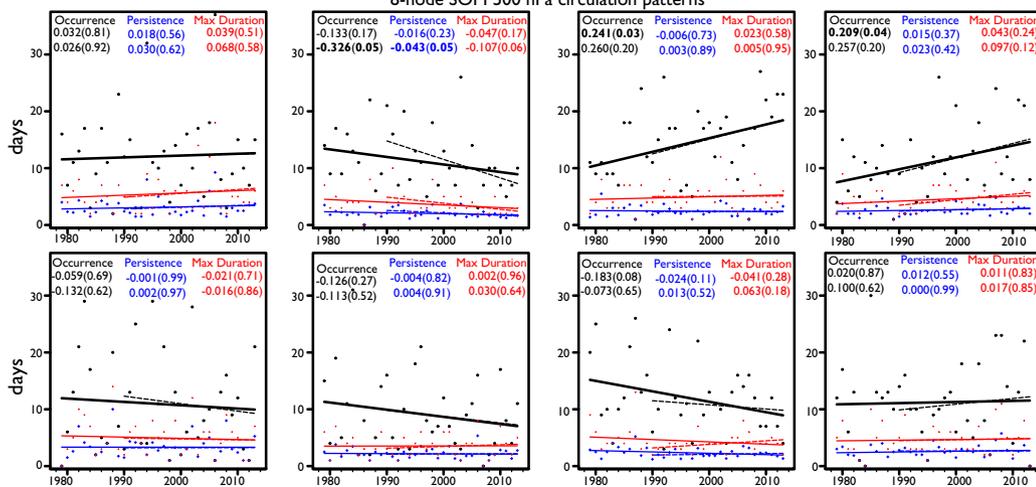
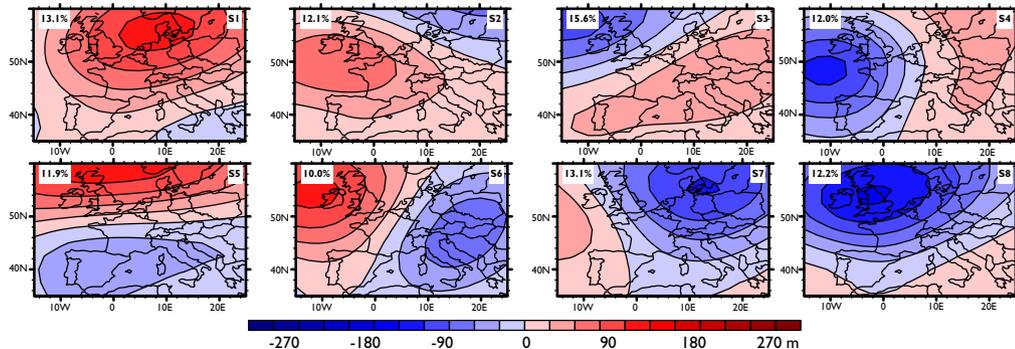
Circulation pattern occurrence, persistence & maximum duration

b



Circulation pattern occurrence, persistence & maximum duration

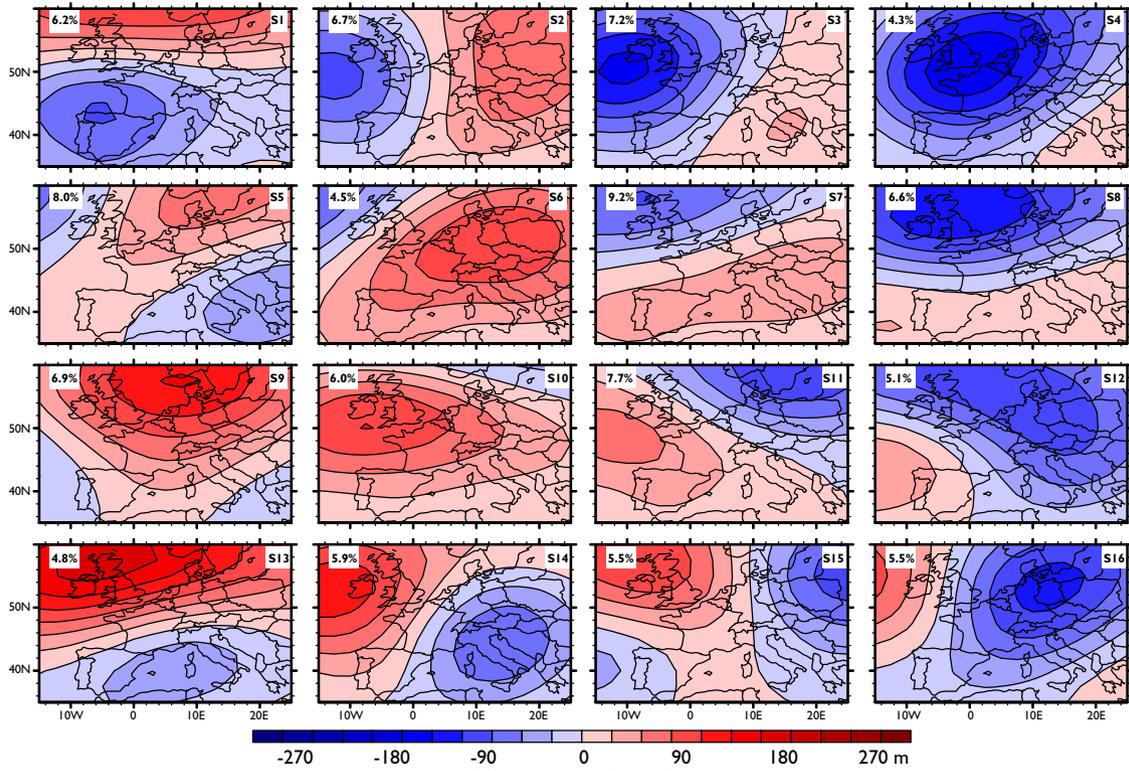
c



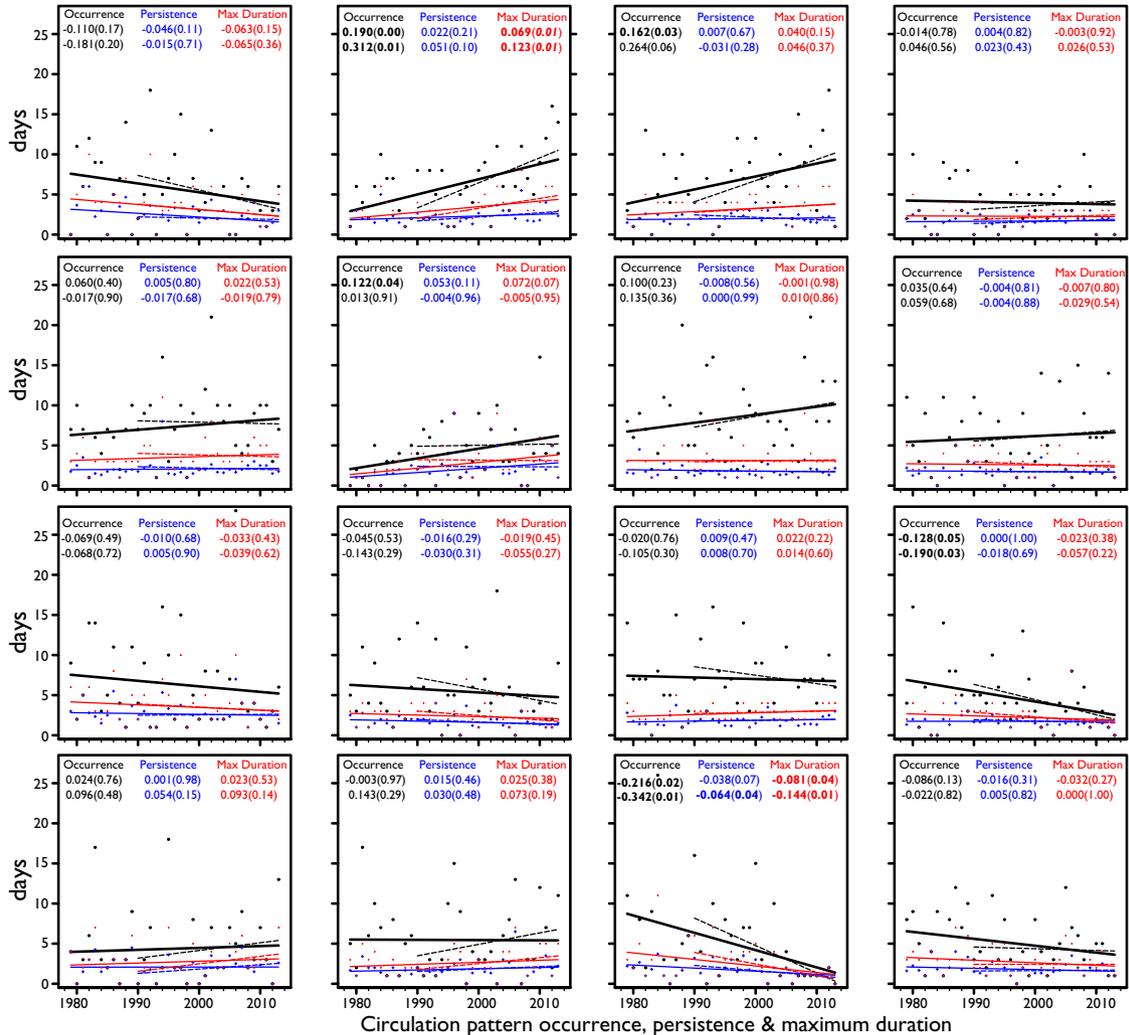
Circulation pattern occurrence, persistence & maximum duration

Extended Data Figure 1 | 2-, 4- and 8-node SOM analyses. SOM-derived mid-atmospheric summer (JJA) circulation patterns (500 hPa geopotential height anomalies) over Europe using 2- (a), 4- (b) and 8-node (c) analyses. White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. Time series of SOM circulation pattern occurrence (black (d yr^{-1})), persistence (blue (d event^{-1})) and maximum duration (red

(d event^{-1})). The slope of the trend line (yr^{-1}) and P values (in parentheses) are colour coded, with the values from 1979 to 2013 (solid trend line) displayed above those from 1990 to 2013 (dashed trend line). Statistically significant trends (5% significance level; Methods) are shown by bold fonts in the scatter plots. Geopotential height fields are sourced from the NCEP-DOE-R2 reanalysis³³.

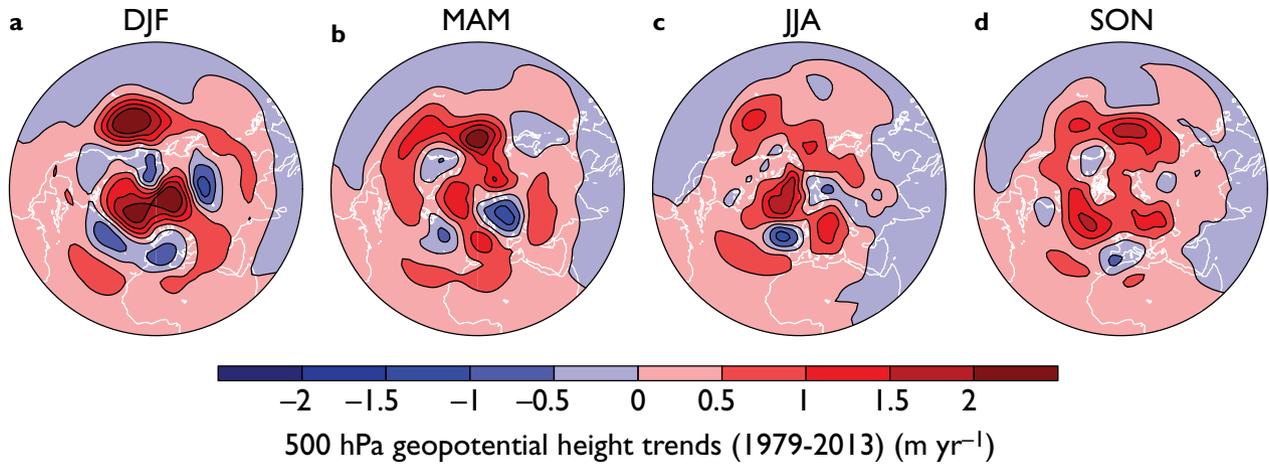


16-node SOM 500 hPa circulation patterns



Extended Data Figure 2 | 16-node SOM analysis. SOM-derived mid-atmospheric summer (JJA) circulation patterns (500 hPa geopotential height anomalies) over Europe derived from a 16-node analysis. White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. Time series of SOM circulation pattern occurrence (black (d yr^{-1})), persistence (blue (d event^{-1})) and maximum duration (red (d event^{-1})).

The slope of the trend line (yr^{-1}) and P values (in parentheses) are colour coded, with the values from 1979 to 2013 (solid trend line) displayed above those from 1990 to 2013 (dashed trend line). Statistically significant trends (5% significance level; Methods) are shown by bold fonts in the scatter plots. Geopotential height fields are sourced from the NCEP-DOE-R2 reanalysis³³.



e

1979-2013 Trends in area-weighted geopotential heights (m yr ⁻¹)				
	DJF	MAM	JJA	SON
N.Hemisphere	0.347	0.304	0.232	0.293
Euro	-0.210	0.708	0.221	0.053
wAsia	0.570	0.361	0.656	0.585
cAsia	-0.271	0.491	0.130	0.151
eAsia	0.005	0.100	0.439	0.551
wNAAm	0.370	0.314	0.438	0.386
cNAAm	0.188	0.409	0.207	0.550
eNAAm	0.370	0.319	0.527	0.818

f

Trend Removed	Euro-JJA-1979-2013											
	A			C-A			A-C			C		
	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR
None	-0.161	-0.009	0.004	0.45	0.054	0.192	-0.328	-0.039	-0.074	0.039	-0.01	-0.031
N.Hemisphere	-0.196	-0.013	-0.002	0.42	0.05	0.173	-0.33	-0.037	-0.067	0.105	-0.004	-0.005

g

Trend Removed	wAsia-JJA-1979-2013											
	A			C-A			A-C			C		
	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR
None	0.546	0.061	0.21	-0.366	-0.04	-0.135	0.057	0.005	-0.02	-0.237	0.031	0.035
N.Hemisphere	0.352	0.034	0.127	-0.301	-0.03	-0.126	0.114	0.004	0	-0.165	0.036	0.046

h

Trend Removed	eNAAm-SON-1979-2013											
	A			A-C			C-A			C		
	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR
None	0.438	0.019	0.086	-0.285	-0.011	-0.063	0.097	0.013	-0.037	-0.25	-0.007	0.003
N.Hemisphere	0.366	0.017	0.075	-0.266	-0.005	-0.055	0.091	0.01	-0.041	-0.19	-0.002	0.003

i

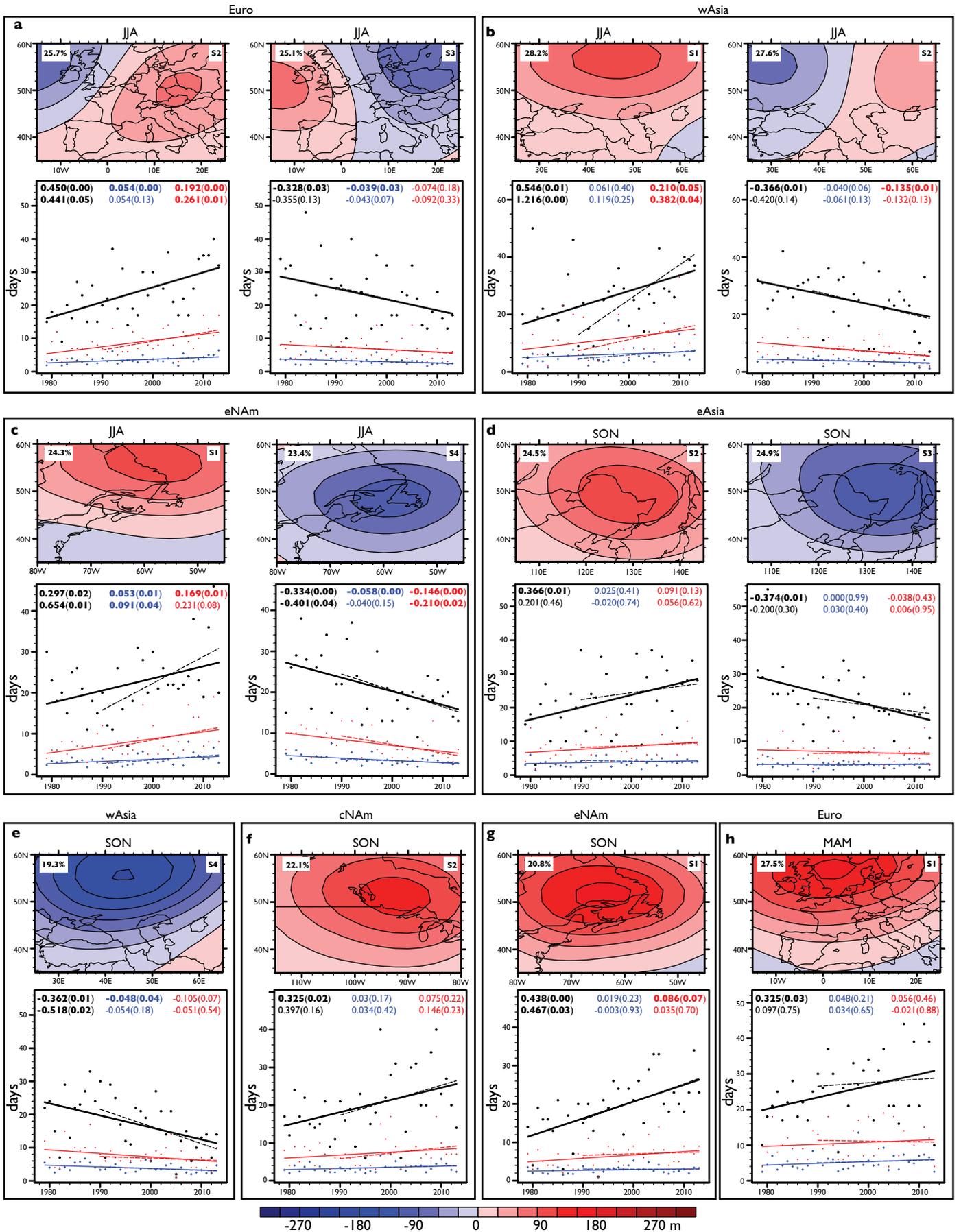
Trend Removed	eAsia-SON-1979-2013											
	A-C			A			C			C-A		
	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR
None	-0.114	-0.01	0.015	0.366	0.025	0.091	-0.374	0	-0.038	0.122	0.02	0.049
N.Hemisphere	-0.101	-0.003	0.015	0.217	0.01	0.036	-0.229	0.034	-0.017	0.113	0.023	0.049

j

Trend Removed	cAsia-DJF-1990-2013											
	A			A-C			C-A			C		
	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR	OCC	PER	DUR
None	-0.127	0.03	0.127	1.044	0.119	0.262	-0.621	0.013	-0.07	-0.296	-0.005	-0.042
N.Hemisphere	-0.169	0.037	0.103	1.105	0.144	0.319	-0.701	0.021	-0.056	-0.235	0.004	0.006

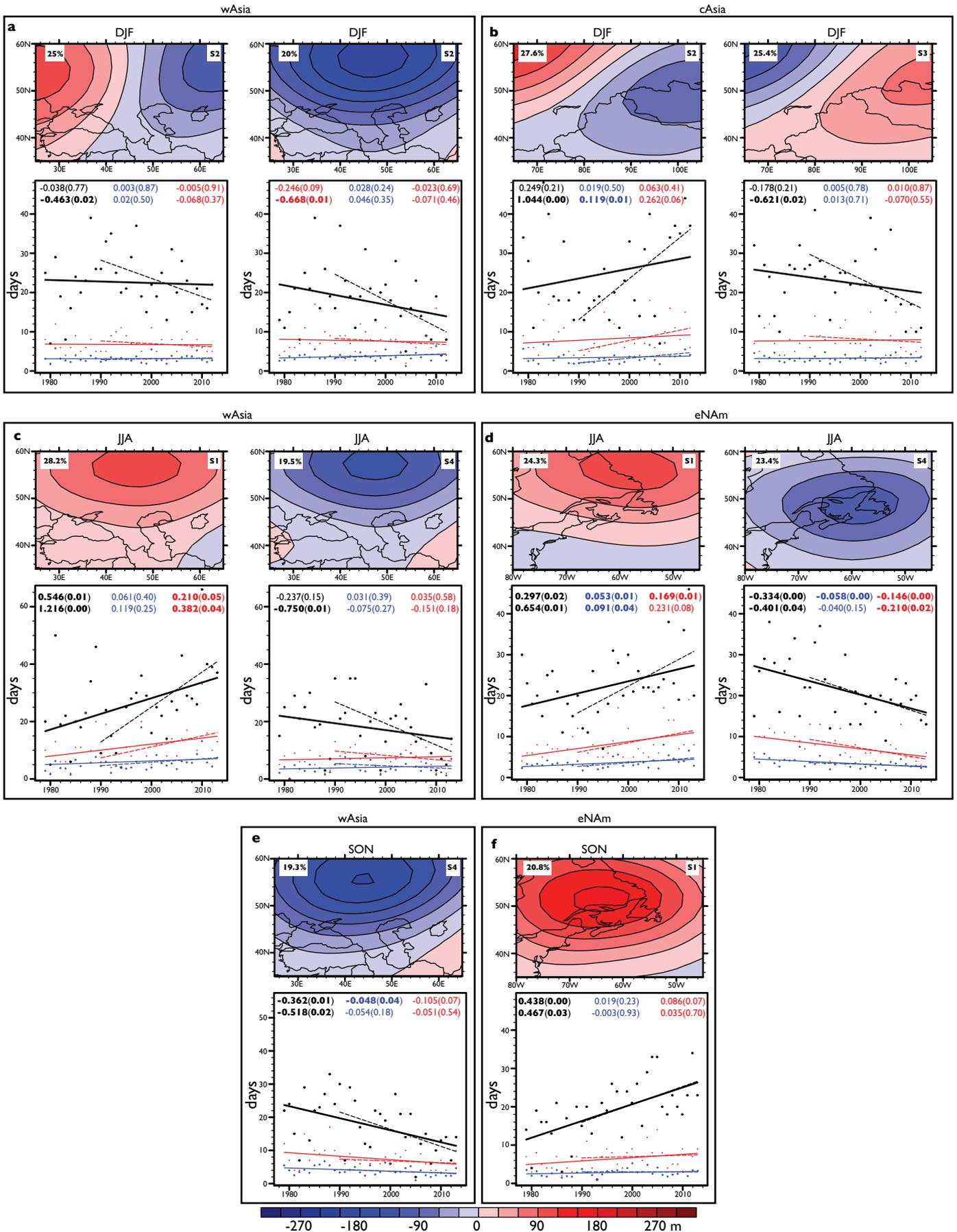
Extended Data Figure 3 | Geopotential height trends and thermal dilation adjustment. **a–d**, Northern Hemisphere polar projections of 1979–2013 seasonal trends in 500 hPa geopotential heights (same as Fig. 1, reproduced here for convenience). **e**, Area-weighted trends in seasonal geopotential heights over the Northern Hemisphere and regional SOM domains. **f–j**, Trends in raw and detrended geopotential height SOM pattern occurrence (OCC), persistence (PER) and maximum duration (DUR) in

units of $\text{d yr}^{-1} \text{yr}^{-1}$ for domains and seasons highlighted in the main text. The magnitudes of the (removed) seasonal Northern Hemisphere trends can be found in **e**. Grid cells highlighted in grey contain trends significant at the 5% level (Methods). SOM circulation patterns are abbreviated as follows: A, anticyclonic; C, cyclonic; and combinations of the two represent dipole patterns and west–east configurations. Geopotential height fields are sourced from the NCEP-DOE-R2 reanalysis³³.



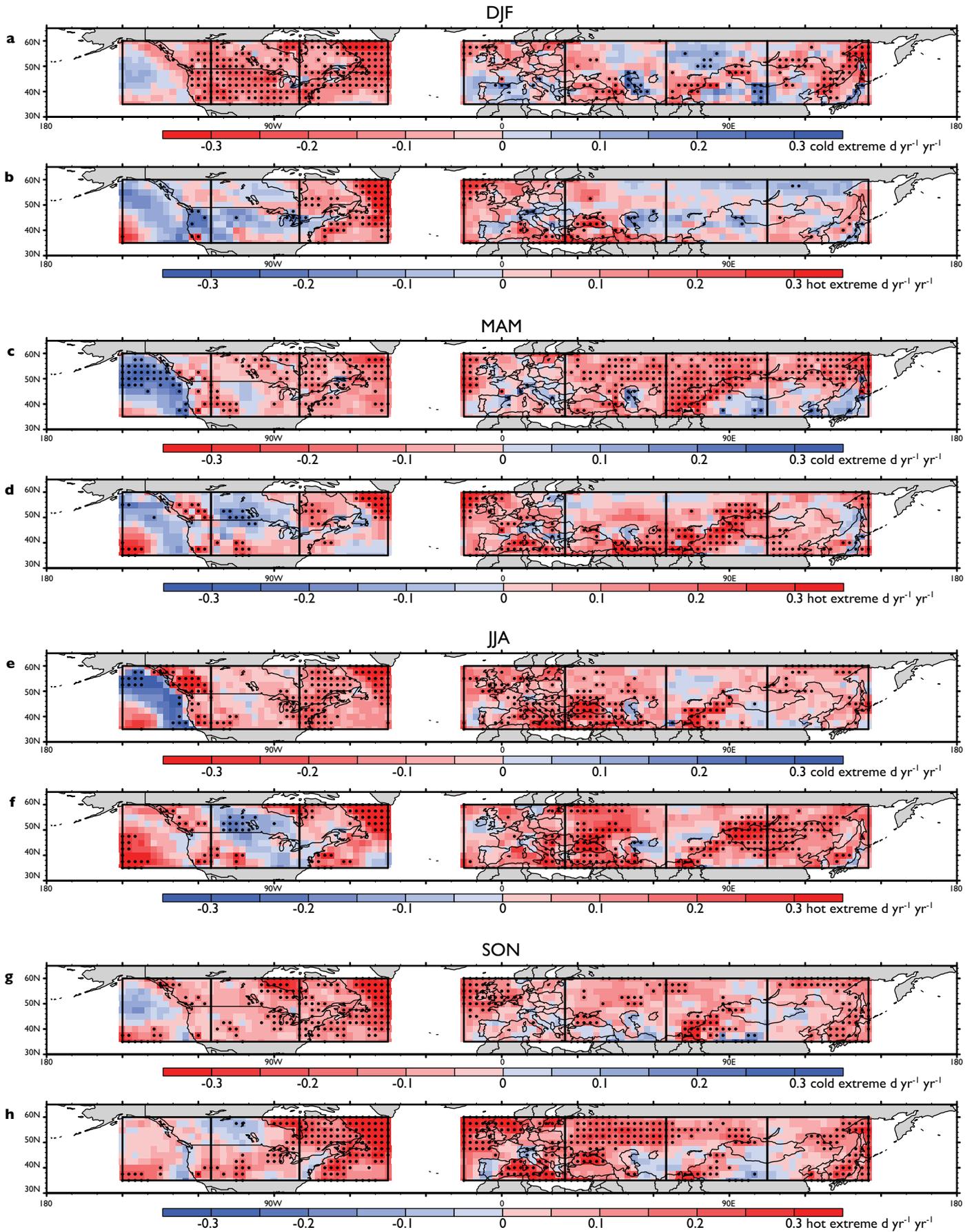
Extended Data Figure 4 | 1979–2013 (satellite era) robust atmospheric circulation pattern trends. Time series of circulation pattern occurrence (black (d yr^{-1})), persistence (blue (d event^{-1})) and maximum duration (red (d event^{-1})) from the NCEP-DOE-R2 reanalysis³³: **a**, summer over Europe; **b**, summer over western Asia; **c**, summer over eastern North America;

d, autumn over eastern Asia; **e**, autumn over western Asia; **f**, autumn over central North America; **g**, autumn over eastern North America; and **h**, spring over Europe. Statistically significant trends ((yr^{-1}) ; 5% significance level; Methods) are identified by bold font in the scatter plots.



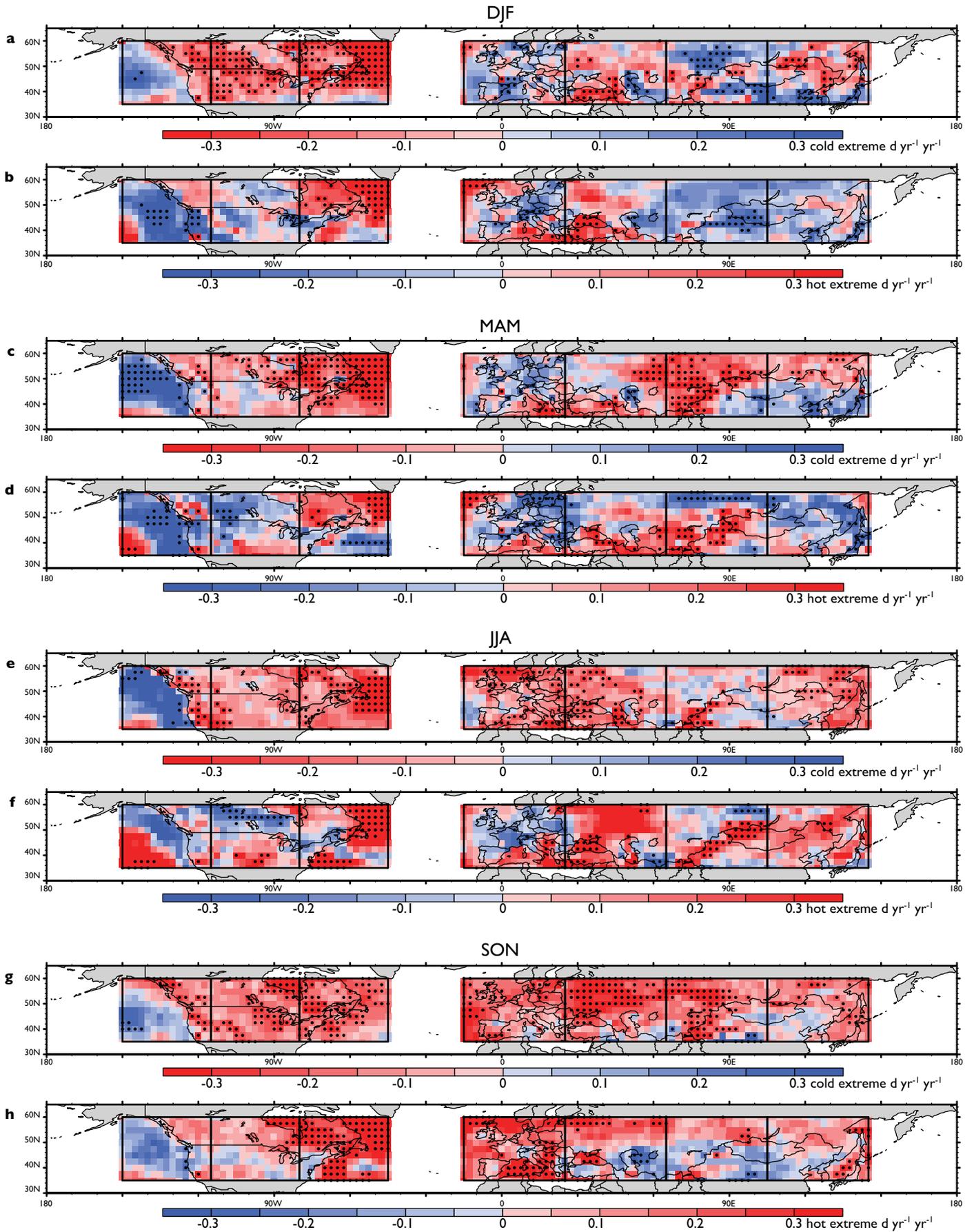
Extended Data Figure 5 | 1990–2013 (ice era) robust atmospheric circulation pattern trends. Time series of circulation pattern occurrence (black (d yr^{-1})), persistence (blue (d event^{-1})) and maximum duration (red (d event^{-1})) from the NCEP-DOE-R2 reanalysis³³: **a**, winter over western Asia;

b, winter over central Asia; **c**, summer over western Asia; **d**, summer over eastern North America; **e**, autumn over western Asia; and **f**, autumn over eastern North America. Statistically significant trends (yr^{-1} ; 5% significance level; Methods) are identified by bold font in the scatter plots.



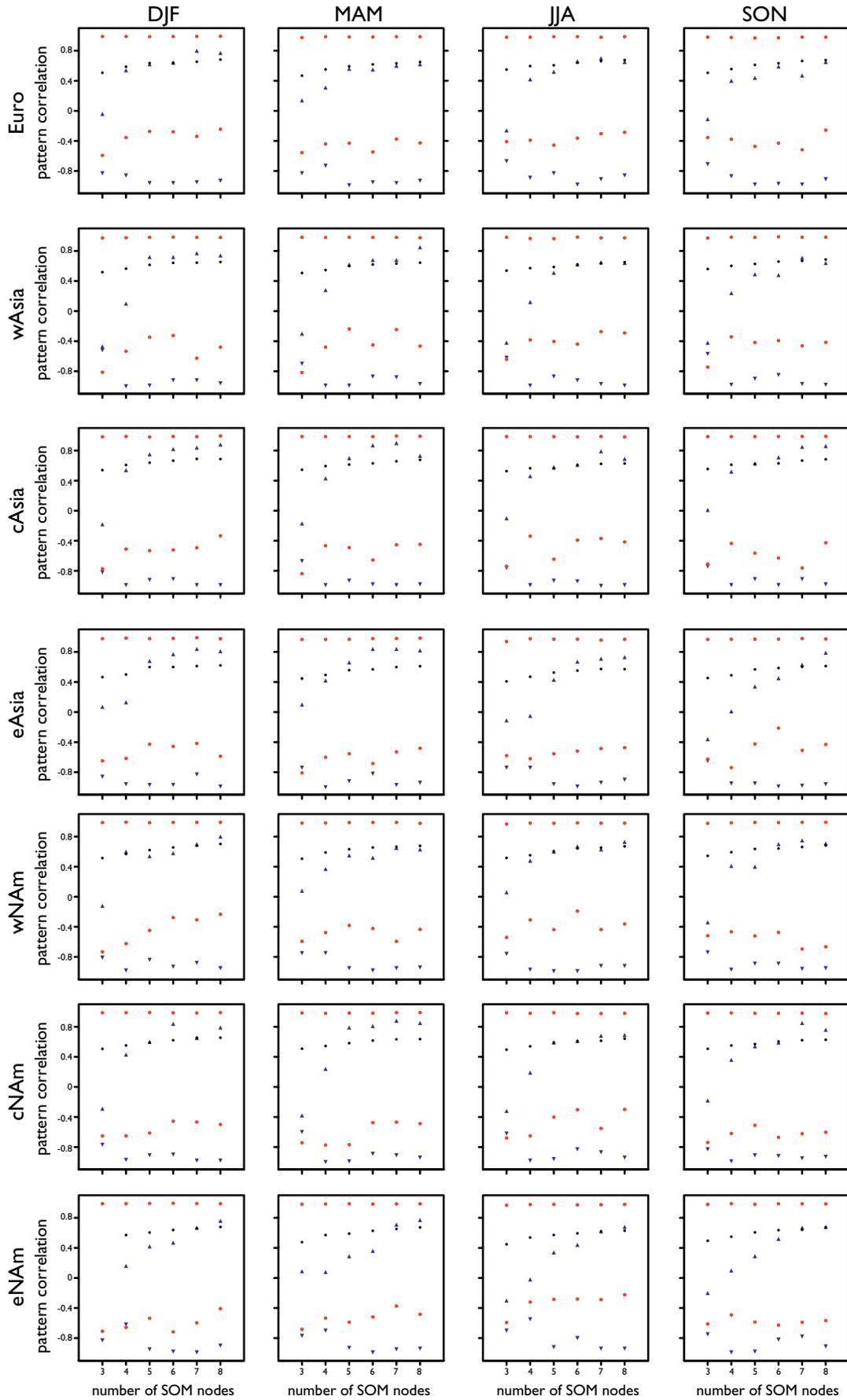
Extended Data Figure 6 | 1979–2013 (satellite era) Northern Hemisphere extreme temperature occurrence trends. Satellite-era extreme temperature trends ($\text{d yr}^{-1} \text{yr}^{-1}$) for winter cold (a) and hot (b) occurrences; spring cold (c) and hot (d) occurrences; summer cold (e) and hot (f) occurrences; and

autumn cold (g) and hot (h) occurrences. Trends are calculated from the NCEP-DOE-R2 reanalysis 2-m daily maximum/minimum temperatures³³. Grid cells with statistically significant trends (5% significance level; Methods) are stippled.



Extended Data Figure 7 | 1990–2013 (ice era) Northern Hemisphere extreme temperature occurrence trends. Ice-era extreme temperature trends ($\text{d yr}^{-1} \text{yr}^{-1}$) for winter cold (a) and hot (b) occurrences; spring cold (c) and hot (d) occurrences; summer cold (e) and hot (f) occurrences; and autumn cold

(g) and hot (h) occurrences. Trends are calculated from the NCEP-DOE-R2 reanalysis 2-m daily maximum/minimum temperatures³³. Grid cells with statistically significant trends (5% significance level; Methods) are stippled.



● Max/Min pattern correlation of daily geopotential height anom. fields with matching SOM node pattern
 ● Mean pattern correlation of all daily geopotential height anom. fields with matching SOM node patterns
 ▲ Maximum SOM pattern to SOM pattern correlation
 ▼ Minimum SOM pattern to SOM pattern correlation

Extended Data Figure 8 | Sensitivity of pattern similarity to number of SOM nodes. To determine an adequate number of SOM nodes, we follow a modified version of the methodology introduced by ref. 12, wherein the mean pattern correlation of all daily geopotential height anomaly fields and their matching SOM node patterns are computed for a suite of different SOM node counts (3, 4, 5, 6, 7 and 8), for all regions and all seasons (black dots). We also compute the maximum/minimum pattern correlation of daily geopotential height anomaly fields with their matching SOM node pattern (red dots) and the maximum/minimum SOM-pattern-to-SOM-pattern correlation (blue

triangles). The goal is to select an adequate number of nodes such that: (1) the mean pattern correlation of all daily geopotential height anomaly fields is relatively large; (2) the minimum pattern correlation of daily geopotential height anomaly fields is relatively large; and (3) the maximum SOM-pattern-to-SOM-pattern correlation is relatively small. Similar to ref. 12, we find that four SOM nodes are generally sufficient to capture the different modes of atmospheric variability, but small enough that SOM patterns depict distinct circulations. Geopotential height anomaly fields are sourced from the NCEP-DOE-R2 reanalysis³³.

Extended Data Table 1 | Significant reanalysis circulation pattern trends and summary of multiple hypothesis testing

a

Statistically Significant Trends ($P < 0.05$)							
	n	satellite-era			ice-era		
		occ	per	dur	occ	per	dur
Linear Regression Analysis							
NCAR/NCEP-R1	112	17	6	7	15	10	9
NCEP-DOE-R2	112	16	5	7	13	8	9
ERA-Interim	112	16	6	10	14	9	9
Box-Cox Transformation Analysis							
NCAR/NCEP-R1	112	15	9	8	16	9	10
NCEP-DOE-R2	112	15	6	7	16	10	10
ERA-Interim	112	16	8	12	14	9	11

b

Anderson-Darling Test for Normality							
	n	satellite-era			ice-era		
		occ	per	dur	occ	per	dur
Original Data							
NCAR/NCEP-R1	112	103	47	56	102	76	68
NCEP-DOE-R2	112	102	57	58	102	74	70
ERA-Interim	112	99	62	49	105	76	65
Original Data with FWER Multiple Hypothesis Testing							
NCAR/NCEP-R1	112	112	102	102	112	103	107
NCEP-DOE-R2	112	112	100	103	112	107	109
ERA-Interim	112	112	100	102	112	104	106
Box-Cox Transformed Data							
NCAR/NCEP-R1	112	110	109	110	109	110	112
NCEP-DOE-R2	112	111	110	110	110	111	112
ERA-Interim	112	110	111	109	111	109	108
Box-Cox Transformed Data with FWER Multiple Hypothesis Testing							
NCAR/NCEP-R1	112	112	112	112	112	112	112
NCEP-DOE-R2	112	112	112	112	112	112	112
ERA-Interim	112	112	112	112	112	112	112

c

Multiple Hypothesis Testing							
Control Type	Level	satellite-era			ice-era		
		occ	per	dur	occ	per	dur
NCAR/NCEP-R1							
FWER	5%	2	0	0	0	1	0
	10%	2	0	1	1	1	0
k-FWER	5%	4	1	2	1	2	0
	10%	7	2	3	4	2	2
FDR	5%	0	0	0	0	0	0
	10%	2	0	0	0	0	0
NCEP-DOE-R2							
FWER	5%	2	0	0	0	1	0
	10%	2	0	1	0	1	0
k-FWER	5%	2	1	2	2	2	0
	10%	5	2	4	5	3	0
FDR	5%	0	0	0	0	0	0
	10%	2	0	0	0	0	0
ERA-Interim							
FWER	5%	0	0	1	0	0	0
	10%	1	0	1	0	0	0
k-FWER	5%	3	2	2	2	0	1
	10%	7	2	2	5	3	3
FDR	5%	0	0	0	0	0	0
	10%	0	0	0	0	0	0

a. We indicate the number of linear regression and Box-Cox transformed statistically significant occurrence (occ), persistence (per) and maximum duration (max dur) trends that surpass the 95% confidence threshold (5% significance level; Methods) for each reanalysis in each analysis period (1979–2013 is 'satellite-era' and 1990–2013 is 'ice-era'). For a given metric (occ, per and max dur) in a given period (satellite-era and ice-era), there are four patterns per region, over four seasons, for 7 domains. There are therefore 112 total trends in each metric in each period. b. Results of Anderson-Darling³⁸ tests of normality indicating the number of original and Box-Cox transformed⁴⁰ distributions of residuals that do not reject the null hypothesis of Gaussianity at the 5% level. For both the original and Box-Cox Anderson-Darling tests, we also apply multiple hypothesis testing to control the familywise error rate³⁹ (FWER) at the 5% significance level. c. To protect against the possibility of false positives in the linear regression trend analyses, for each reanalysis, we use three different multiple hypothesis error control methodologies to assess the number of locally significant hypotheses that surpass the global 5% or 10% significance level under strict error control requirements. FWER³⁹, k-familywise error rate⁴¹ (k-FWER), and false discovery rate³⁹ (FDR) analyses are applied (Methods). We note that all three metrics heavily favour the null. The presence of local tests that reject the null therefore represents a strong confirmation of the significance of those local tests, as those local tests are so highly significant that they can withstand the stricter multiple testing error control criteria.

Extended Data Table 2 | Quantitative partitioning of temperature extreme trends for select SOM analyses

1979-2013 Europe JJA Hot Extreme Partitioning												
	Pat occ trend	Mean pat occ	T95 per pat occ trend	Mean T95 per pat occ	Dyn trend	Thermo trend	Inter trend	Total trend	Percent of trend	Percent dynamic	Percent thermo	Percent inter
SOM1	-0.16	24.94	0.001	0.07	-0.01	0.02	0.00	0.01	10.1	-112.0	237.9	-25.9
SOM2	0.45	23.66	0.002	0.06	0.03	0.04	0.00	0.06	61.5	44.3	57.3	-1.6
SOM3	-0.33	23.09	0.001	0.03	-0.01	0.02	0.00	0.01	8.7	-109.3	200.9	8.4
SOM4	0.04	20.31	0.001	0.03	0.00	0.02	0.00	0.02	19.7	5.9	96.6	-2.5
Overall:					0.01	0.10	0.00	0.10				

1979-2013 Western Asia JJA Hot Extreme Partitioning												
	Pat occ trend	Mean pat occ	T95 per pat occ trend	Mean T95 per pat occ	Dyn trend	Thermo trend	Inter trend	Total trend	Percent of trend	Percent dynamic	Percent thermo	Percent inter
SOM1	0.55	25.94	0.002	0.07	0.04	0.04	0.03	0.11	70.3	36.5	37.8	25.7
SOM2	-0.37	25.40	0.001	0.05	-0.02	0.04	0.00	0.02	11.8	-92.3	198.0	-5.7
SOM3	0.06	22.69	0.001	0.03	0.00	0.02	-0.01	0.02	12.0	9.9	122.3	-32.1
SOM4	-0.24	17.97	0.001	0.03	-0.01	0.02	-0.01	0.01	5.9	-72.2	240.8	-68.6
Overall:					0.02	0.12	0.01	0.16				

1979-2013 Eastern North America JJA Hot Extreme Partitioning												
	Pat occ trend	Mean pat occ	T95 per pat occ trend	Mean T95 per pat occ	Dyn trend	Thermo trend	Inter trend	Total trend	Percent of trend	Percent dynamic	Percent thermo	Percent inter
SOM1	0.30	22.31	0.002	0.07	0.02	0.04	0.01	0.07	51.4	29.9	56.8	13.3
SOM2	-0.02	21.63	0.001	0.05	0.00	0.02	0.00	0.02	13.1	-5.1	114.4	-9.3
SOM3	0.05	26.51	0.001	0.04	0.00	0.04	0.00	0.03	26.7	6.8	106.7	-13.6
SOM4	-0.33	21.54	0.001	0.04	-0.01	0.03	0.00	0.01	8.7	-105.6	234.8	-29.1
Overall:					0.01	0.12	0.00	0.13				

1990-2013 Central Asia DJF Cold Extreme Partitioning												
	Pat occ trend	Mean pat occ	T05 per pat occ trend	Mean T05 per pat occ	Dyn trend	Thermo trend	Inter trend	Total trend	Percent of trend	Percent dynamic	Percent thermo	Percent inter
SOM1	-0.13	21.04	0.001	0.04	0.00	0.01	0.00	0.01	9.0	-79.0	177.8	1.3
SOM2	1.04	24.61	0.001	0.07	0.08	0.02	0.01	0.10	149.6	74.6	18.1	7.2
SOM3	-0.62	22.87	0.000	0.02	-0.01	0.01	-0.01	-0.01	-21.3	98.0	-45.4	47.4
SOM4	-0.30	21.74	-0.001	0.05	-0.02	-0.02	0.01	-0.03	-37.2	59.7	68.2	-27.8
Overall:					0.04	0.02	0.01	0.07				

1979-2013 Eastern North America SON Hot Extreme Partitioning												
	Pat occ trend	Mean pat occ	T95 per pat occ trend	Mean T95 per pat occ	Dyn trend	Thermo trend	Inter trend	Total trend	Percent of trend	Percent dynamic	Percent thermo	Percent inter
SOM1	0.44	18.89	0.002	0.08	0.04	0.03	0.00	0.07	37.0	51.7	42.5	5.8
SOM2	-0.29	28.40	0.001	0.03	-0.01	0.03	0.00	0.02	10.5	-45.3	156.9	-11.6
SOM3	0.10	25.91	0.003	0.06	0.01	0.07	0.01	0.08	43.8	6.6	86.3	7.1
SOM4	-0.25	17.80	0.002	0.03	-0.01	0.03	0.00	0.02	8.7	-44.1	161.9	-17.8
Overall:					0.03	0.16	0.00	0.19				

1979-2013 East Asia SON Cold Extreme Partitioning												
	Pat occ trend	Mean pat occ	T05 per pat occ trend	Mean T05 per pat occ	Dyn trend	Thermo trend	Inter trend	Total trend	Percent of trend	Percent dynamic	Percent thermo	Percent inter
SOM1	-0.11	19.43	-0.001	0.05	-0.01	-0.02	0.01	-0.02	18.4	40.2	152.7	-92.9
SOM2	0.37	22.29	-0.001	0.03	0.01	-0.01	0.00	0.00	0.8	-1637.5	1844.1	-106.5
SOM3	-0.37	22.69	-0.001	0.08	-0.03	-0.02	0.00	-0.05	62.5	56.5	35.0	8.4
SOM4	0.12	26.60	-0.001	0.04	0.00	-0.02	0.00	-0.02	18.3	-31.2	119.7	11.5
Overall:					-0.02	-0.07	0.01	-0.08				

Trends are partitioned for Europe (a), western Asia (b), and eastern North America (c) satellite-era summer hot extremes; central Asia ice-era winter cold extremes (d); eastern North America satellite-era autumn hot extremes (e); and eastern Asia satellite-era autumn cold extremes (f). From the left column moving right: (1) SOM node number; (2) trend in pattern occurrence ('pat-occ' d yr⁻¹ yr⁻¹); (3) mean pattern occurrence for the period ('pat-occ' d yr⁻¹); (4) trend in temperature extremes per pattern occurrence (intensity of extreme occurrence) ('ext-occ' d pat-occ d yr⁻¹ yr⁻¹); (5) mean temperature extreme per pattern occurrence ('ext-occ' d pat-occ d yr⁻¹ yr⁻¹); (6) trend in the dynamic influences ('ext-occ' d yr⁻¹ yr⁻¹); (7) trend in the thermodynamic influences ('ext-occ' d yr⁻¹ yr⁻¹); (8) trend in the interaction of dynamic and thermodynamic influences ('ext-occ' d yr⁻¹ yr⁻¹); (9) the total trend in extreme occurrence for each SOM pattern ('ext-occ' d yr⁻¹ yr⁻¹); (10) the per cent of the total trend in extreme occurrences that occur during each pattern (pattern trend percentages sum to 100%, meaning contributions from individual SOM patterns can be either positive or negative; that is, trend percentages greater than 100% indicate that other circulation patterns provide negative contributions); (11) the per cent of column 10 that is due to dynamic influences; (12) the per cent of column 10 that is due to thermodynamic influences; and (13) the per cent of column 10 that is due to interactive influences (dynamic, thermodynamic and interactive influence percentages for individual SOM patterns sum to 100%, meaning contributions from individual influences can be either positive or negative). The overall trends for the domain are presented below the individual SOM rows. All data are sourced from the NCEP-DOE-R2 reanalysis³⁹.