Title

Decadal predictability of late winter precipitation in western Europe through an ocean–jet stream connection

Journal

Nature Geoscience, 12(8)

ISSN

1752-0894

Authors

Simpson, Isla R
Yeager, Stephen G
McKinnon, Karen A
et al.

Publication Date

2019-08-01

DOI

10.1038/s41561-019-0391-x

Peer reviewed

nature geoscience

ARTICLES

https://doi.org/10.1038/s41561-019-0391-x

Decadal predictability of late winter precipitation in western Europe through an ocean–jet stream connection

Isla R. Simpson[1*], Stephen G. Yeager[1], Karen A. McKinnon[2] and Clara Deser[1]

The characteristics of the North Atlantic jet stream play a key role in the weather and climate of western Europe. Although much of the year-to-year variability in the jet stream arises from internal atmospheric processes that are inherently unpredictable on timescales beyond a few days to weeks, any low-frequency variability or long-term trends that can be considered forced by slowly varying boundary conditions offer the potential for extended range predictability of climatological conditions in western Europe. Here we demonstrate that station-based precipitation observations have displayed pronounced multidecadal variability over the past century in western Europe during the late winter. We then use these precipitation observations as an independent verification of the multidecadal Atlantic jet stream variability found in reanalysis products. Both signals are highly correlated with sea surface temperature variability in the North Atlantic that is well predicted in initialized decadal prediction experiments with a coupled general circulation model. Combining the model-based predictions of the sea surface temperature with the observed relationship between precipitation and sea surface temperature, we show that there is great potential for skilful predictions of the forthcoming decadal average of March precipitation in western Europe, with hindcasts for the UK and Portugal yielding anomaly correlation coefficients of 0.82 and 0.69, respectively.

The Atlantic Ocean exhibits variability on decadal- to multidecadal timescales, often termed the Atlantic multidecadal variability (AMV) or Atlantic multidecadal oscillation[1]. This is typically characterized using basin wide sea surface temperature (SST) anomalies[2] with a variety of methods proposed to isolate the internal variability from externally forced trends (Frankignoul et al.[3] and references therein). Although a component of the SST variability in the North Atlantic is driven directly by the influence of wind variations on the oceanic mixed layer[4], at longer timescales the role of the ocean circulation becomes increasingly important. This is evidenced by the evolution of the spatial pattern of SSTs that accompanies the AMV from a tripolar/horseshoe pattern on short timescales to localized SST anomalies in the North Atlantic subpolar gyre (SPG) on longer timescales[5,6].

Low-frequency SPG SST variability is skilfully predicted on decadal timescales in general circulation model simulations initialized with observation-based ocean and sea ice states[6,7]. If it impacts the atmospheric circulation, it may result in predictable low-frequency variability in climatological conditions in western Europe. During the summertime, North Atlantic SSTs are thought to influence the atmospheric circulation and European temperatures and/or precipitation[8–11], but robust evidence of an influence during the wintertime has proved challenging to find. Although observation-based reanalysis data sets suggest that a positive AMV is accompanied by a negative wintertime average North Atlantic oscillation[12,13], modelling studies have produced mixed results when attempting to verify this relationship[14]. In general, models do exhibit a negative North Atlantic oscillation-like anomaly under positive AMV conditions, but the atmospheric anomalies are typically smaller than those observed[15–20].

However, models may be deficient in their representation of the SST influence on the atmosphere, in which case the predictability that could be harnessed from North Atlantic SST variations in the real world may be far greater than that inferred from models[21]. Indeed, Simpson et al.[6] argued that the strength of the observed connection between the AMV and the jet stream maximizes in the late winter, March in particular, and that models are deficient in this late winter AMV–jet stream connection. Following the analysis of Simpson et al.[6], we here investigate the possibility that, given the observed relationship between the AMV and the jet stream, late winter precipitation variability in western Europe may be far more predictable than models currently imply.

Multidecadal variability in western European precipitation

Defining multidecadal variability as the standard deviation of 20 year running means, Fig. 1a–h reproduces the results of Simpson et al.[6] and demonstrates the pronounced multidecadal variability in reanalysis 700 hPa zonal wind (U700) in March and its connection to the AMV. A detailed discussion of the causal inferences that can be made from Fig. 1e–h is provided in Simpson et al.[6] and is discussed further below. In brief, it was argued that a positive AMV results in a more equatorward and zonal North Atlantic jet stream in March (Fig. 1h), although the mechanisms involved remain to be understood.

Extending this analysis to precipitation observations based on rain gauges (Methods), it was found that multidecadal variability in western European precipitation also occurs in March (Fig. 1i–l). Only the March variability exceeds expectations from the sampling of year-to-year variability (Fig. 1l).

Reconstructing jet stream variability using precipitation

Precipitation observations can be exploited as an independent verification of the multidecadal jet stream variability. Consider the U700NA index (area averaged U700 anomalies in the North

[1]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA. [2]Institute of the Environment and Sustainability, Department of Statistics, University of California, Los Angeles, CA, USA. *e-mail: islas@ucar.edu

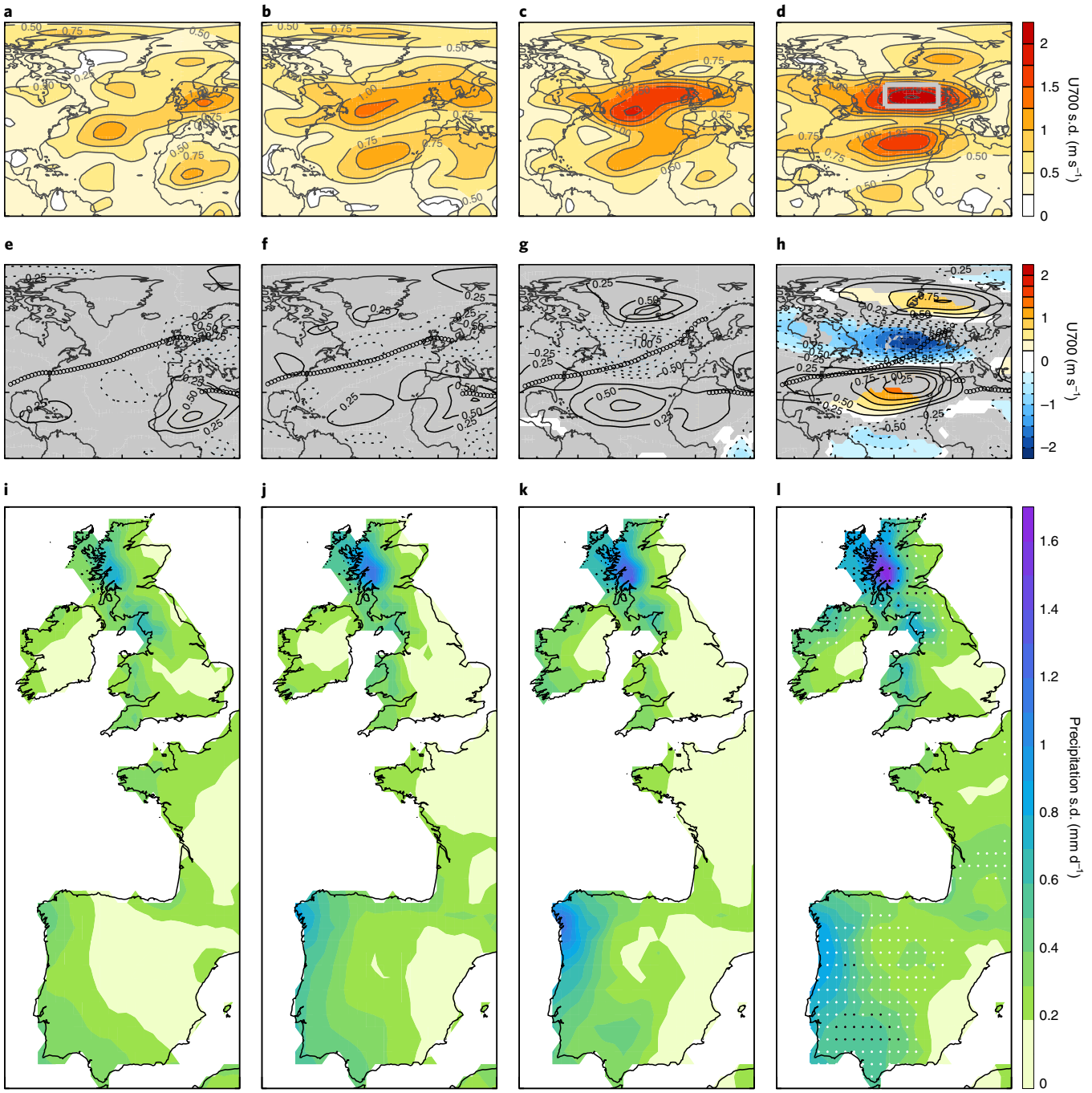

Fig. 1 | Variability in U700 and precipitation from 1901–2016. a–d, Standard deviation of 20 yr running means of ERA-20C+ERA-Interim U700 (Methods) for December (**a**), January (**b**), February (**c**) and March (**d**). The grey box in **d** depicts the averaging region used for the U700NA index (40–10° W, 50–60° N). **e–h**, Regression of U700 onto the AMV for December (**e**), January (**f**), February (**g**) and March (**h**). Grey shading, no significant difference from zero at the 95% level by a resampling methodology (Methods); black circles, the climatological jet latitude at each longitude. **i–l**, Standard deviation of 20 yr running means of Climatic Research Unit (CRU) precipitation for December (**i**), January (**j**), February (**k**) and March (**l**). The white and black dots represent where the standard deviation of 20 yr running means exceeds expectations from interannual sampling at the 95% and 99% levels, respectively (Methods). Panels **a–h** are adapted from Simpson et al.[6], American Meteorological Society.

Atlantic (Fig. 1d)). U700NA has exhibited low-frequency fluctuations in March over the past century (Fig. 2d,e, black lines), with a minimum in the 1940s and 1950s associated with a more zonal jet stream (Fig. 3h), which rose to a maximum in the 1980s and 1990s associated with a more tilted jet stream (Fig. 3i). Given the limited observational constraint in reanalyses back to 1900 and the possible introduction of inhomogeneities as the assimilated observations evolved, one may question quantitative estimates of low-frequency variability from such products[22,23]. However, precipitation observations provide an independent, quantitative verification of the multidecadal jet stream variability.

We reconstructed the U700NA variability using the interannual relationship between March U700NA and March precipitation over the satellite-era record, during which reanalyses are more

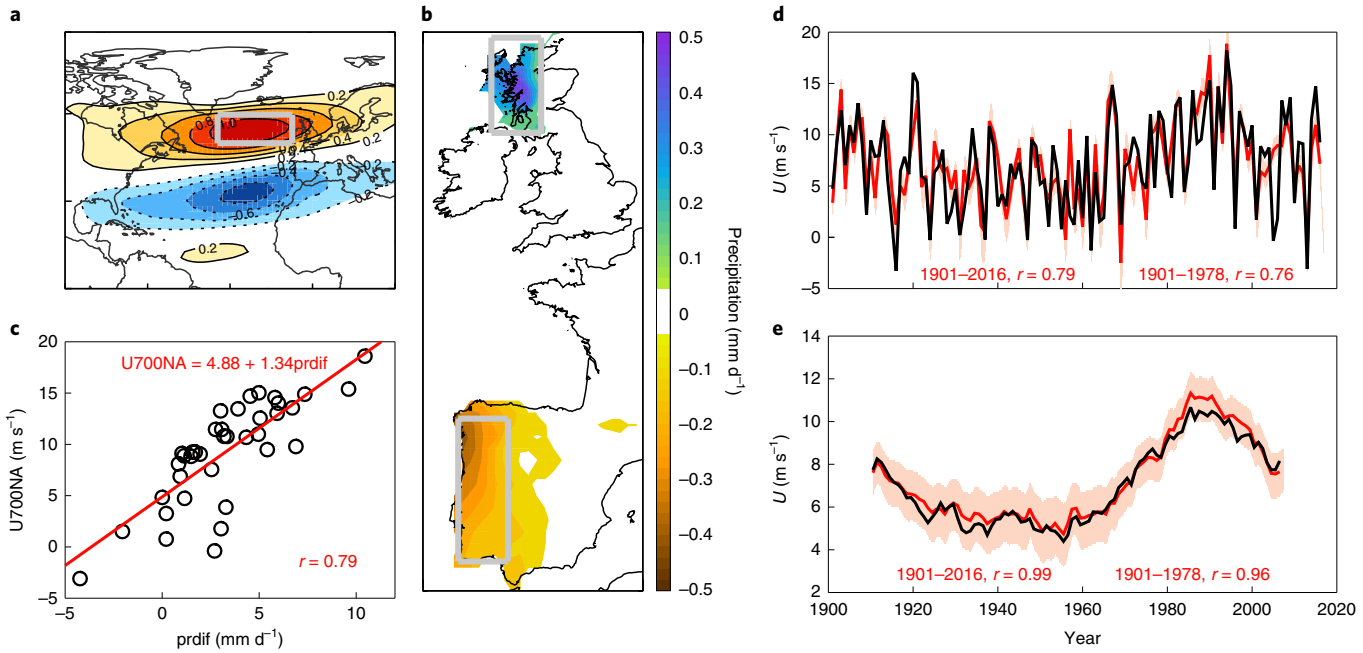

Fig. 2 | Reconstructing variations in U700NA from precipitation. **a**, 1979–2016 ERA-Interim interannual regression of March U700 onto U700NA. Grey box, averaging region used for U700NA (40–10° W, 50–60° N). **b**, 1979–2016 interannual regression of March precipitation onto ERA-Interim U700NA. Grey boxes, Scotland (7–4° W, 55–59° N) (top) and Portugal (9–6° W, 37–43° N) (bottom), indicate regions used to construct prdif (Scotland–Portugal). **c**, 1979–2016 March U700NA versus March prdif with the regression and correlation (*r*) given in red. **d**, Yearly March U700NA. Black line, ERA-20C+ERA-Interim; red line, reconstruction; pink shading, 95% confidence interval on the reconstruction (Methods); red text gives the correlation between the actual and reconstructed time series. **e**, As for **d**, but for 20 yr running means.

observationally constrained. The interannual regression of March U700NA onto the U700NA index for ERA-Interim reanalysis from 1979 to 2016 (Fig. 2a) demonstrates that an increased U700NA is associated with a strengthened U700 to the west of the UK and reduced U700 to the west of the Iberian Peninsula. This is accompanied by increased precipitation over Scotland and reduced precipitation over western Iberia (Fig. 2b). Taking the precipitation difference (prdif) between the area averages over western Scotland and western Iberia (grey boxes in Fig. 2b), we used the 1979–2016 interannual relationship between prdif and U700NA (Fig. 2c) to reconstruct U700NA back to 1901 (Methods). Interannually, prdif and U700NA are correlated at 0.79, so U700NA variability is associated with ~62% of interannual variance in the March prdif. The reconstruction of U700NA based on the 1979–2016 prdif–U700NA relationship (Fig. 2d, red line) performs remarkably well, with the correlation between the reconstructed and reanalysis interannual time series of around 0.79 being retained for the full record. Even more striking is the correspondence at low frequencies (Fig. 2e). A correlation of 0.99 indicates that almost all the low-frequency variance in prdif is associated with variations in U700NA. The success of the reconstruction indicates that the relationship between U700NA and prdif does not change substantially between high and low frequencies. Overall, the rain-gauge-based precipitation data in western Europe provide an independent, quantitative verification of the low-frequency variability in the March jet stream.

The relationship between precipitation and SST variability

At low frequencies, a positive AMV is associated with reduced westerly U700 west of the UK and increased westerly U700 west of Iberia in March (Fig. 3a,b), accompanied by reduced precipitation over the UK and enhanced precipitation over Iberia (Fig. 3c,d). Note that the U700 anomalies associated with the AMV have a slightly different structure to those that accompany the interannual U700NA variability: less of a southwest–northeast tilt results in precipitation anomalies that extend further south over the UK (compare Fig. 2a with Fig. 3b).

We now focus attention on the grey boxed regions in Fig. 3c,d, referred to as the UK and Portugal hereafter, chosen because they exhibit both a high correlation and substantial precipitation anomalies in association with the AMV. Monthly regressions of precipitation in these regions onto the AMV (Fig. 3e and Methods) show a significant relationship between precipitation and the AMV in March, as also found for the jet stream by Simpson et al.[6].

The time series in Fig. 3f,g further demonstrate the low-frequency relationship between precipitation and the AMV. The difference in precipitation between 20-year periods of extreme positive and negative AMV is as much as 2 mm d^{-1} over the UK and Portugal. To put this low-frequency variability in perspective, the climatological March precipitation for two 31-year periods is shown in Fig. 3j,k with the accompanying jet stream climatologies shown in Fig. 3h,i. The period 1935–1965 was characterized by a positive phase of the AMV and a rather zonal March jet stream (Fig. 3h), which presumably guided the Atlantic storm systems more towards Iberia and less towards the UK, which led to the western UK and western Iberia having similar March precipitation climatologies (Fig. 3j). This is in stark contrast to 1980–2010. On average, 1980–2010 is characterized by a negative AMV and, although it is not the 31-year period with the most negative AMV, we show it here because it covers much of our satellite-era observational record. During 1980–2010, the jet stream was strongly tilted (Fig. 3i) and the UK was characterized by a far greater climatological precipitation than western Iberia. (Fig. 3k). The March drying trend in Portugal between these two periods has already been recognized (for example, Trigo and DaCamara[24] and references therein).

Predicting decadal averaged precipitation

Could this relationship between low-frequency SSTs and precipitation be put to practical use in predicting precipitation variations?

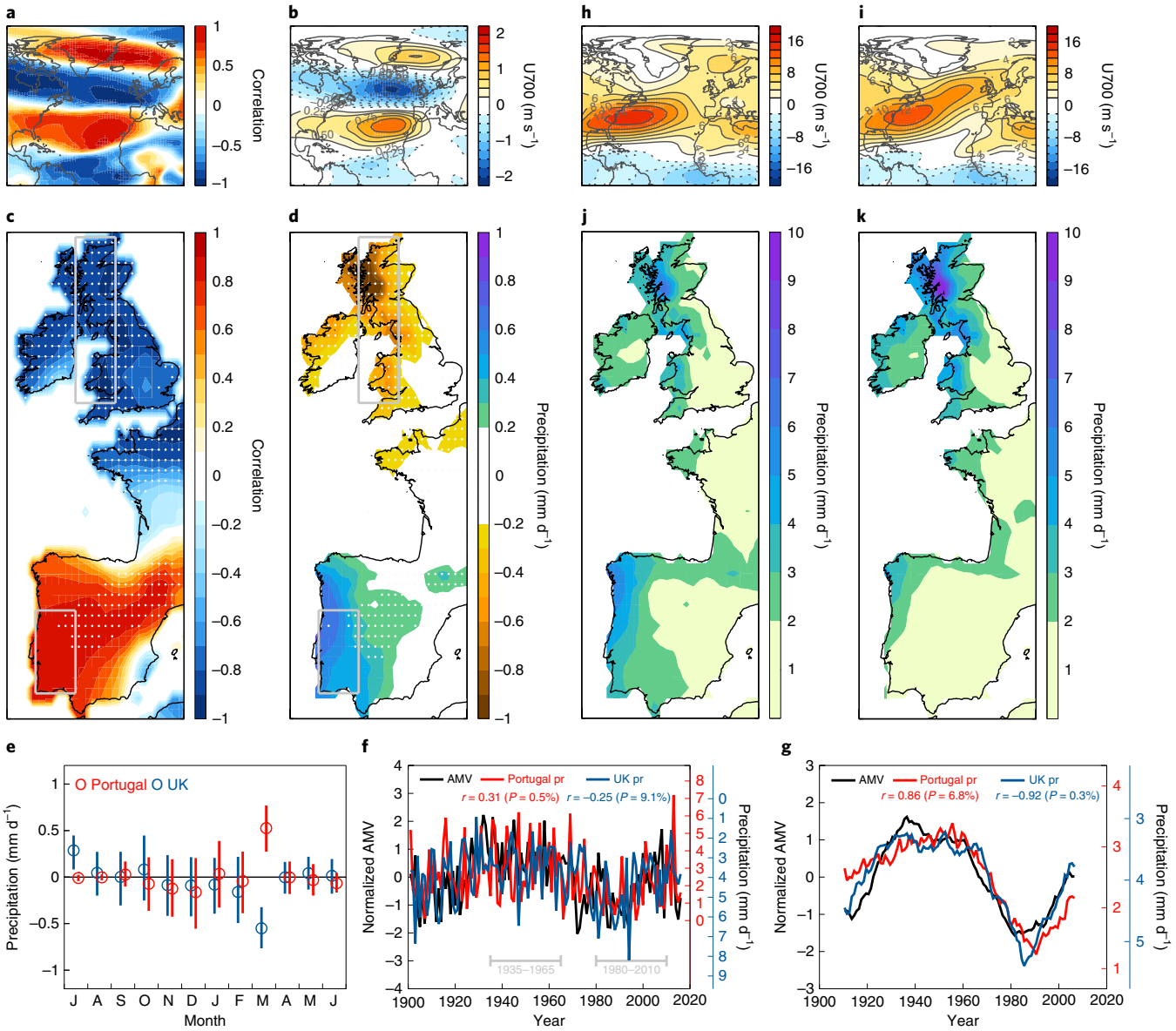

Fig. 3 | Linking March precipitation to March AMV, via the jet stream. a,**b**, 20 yr running mean correlation between U700 and AMV (**a**) and regression of U700 onto AMV (**b**). **c**,**d**, As **a** and **b**, respectively, but for precipitation. Grey boxes, UK (6–3° W, 51–59° N) (top) and Portugal (9–6° W, 37–41° N) (bottom); white stippling, indicates significant correlation and/or regression coefficients at 95% confidence (Methods). **e**, Monthly regression of 21 yr running mean UK and Portugal precipitation onto AMV with 95% confidence intervals (Methods). **f**,**g**, Yearly (**f**) and 20 yr running mean (**g**) AMV, Portugal and UK precipitation (pr). Portugal, red axis; UK, blue axis (note the axis inversion). **h**,**i**, U700 for 1935–1965 (**h**) and 1980–2010 (**i**). **j**,**k**, Precipitation for 1935–1965 (**j**) and 1980–2010 (**k**). Panels **h** and **i** reproduced from Simpson et al.[6], American Meteorological Society.

Initialized decadal prediction experiments with the Community Earth System Model (CESM), referred to as the Decadal Prediction Large Ensemble (DPLE) (Methods), can successfully predict the SST variability for the subsequent decade over much of the North Atlantic (Fig. 4a,d). The predictability in the SPG arises from the ocean initialization and the predictability in the lower latitudes arises primarily from external forcings[7] (Supplementary Fig. 5). Decadal averaged precipitation and the North Atlantic jet stream exhibit a similar connection to the AMV as found with the 20 year running means discussed above (compare Fig. 3 and Supplementary Fig. 6) and a closer inspection of the SSTs that are most strongly correlated with precipitation in western Europe reveals a dipole structure, with high correlations in the SPG and oppositely signed correlations to the southwest (Fig. 4b,c). The AMV index used in the above analysis is associated with a similar pattern, which is why it is also highly correlated with precipitation. Even though the DPLE does not produce the jet stream or precipitation variability observed to accompany the SST variability (see the precipitation predictions below and Simpson et al.[6]), the fact that it can predict the SSTs of relevance raises the possibility that the SST prediction along with the observed SST–precipitation relationship could be used for precipitation prediction.

Hindcasts based on the DPLE ensemble mean SST predictions and the observed SST–precipitation relationship (Methods) show good agreement with the observed precipitation (compare the red lines and black points in Fig. 4e,h). The low-frequency rise and fall in precipitation from the 1960s to the 1990s and subsequent decline and increase is captured for the UK and Portugal, respectively. In

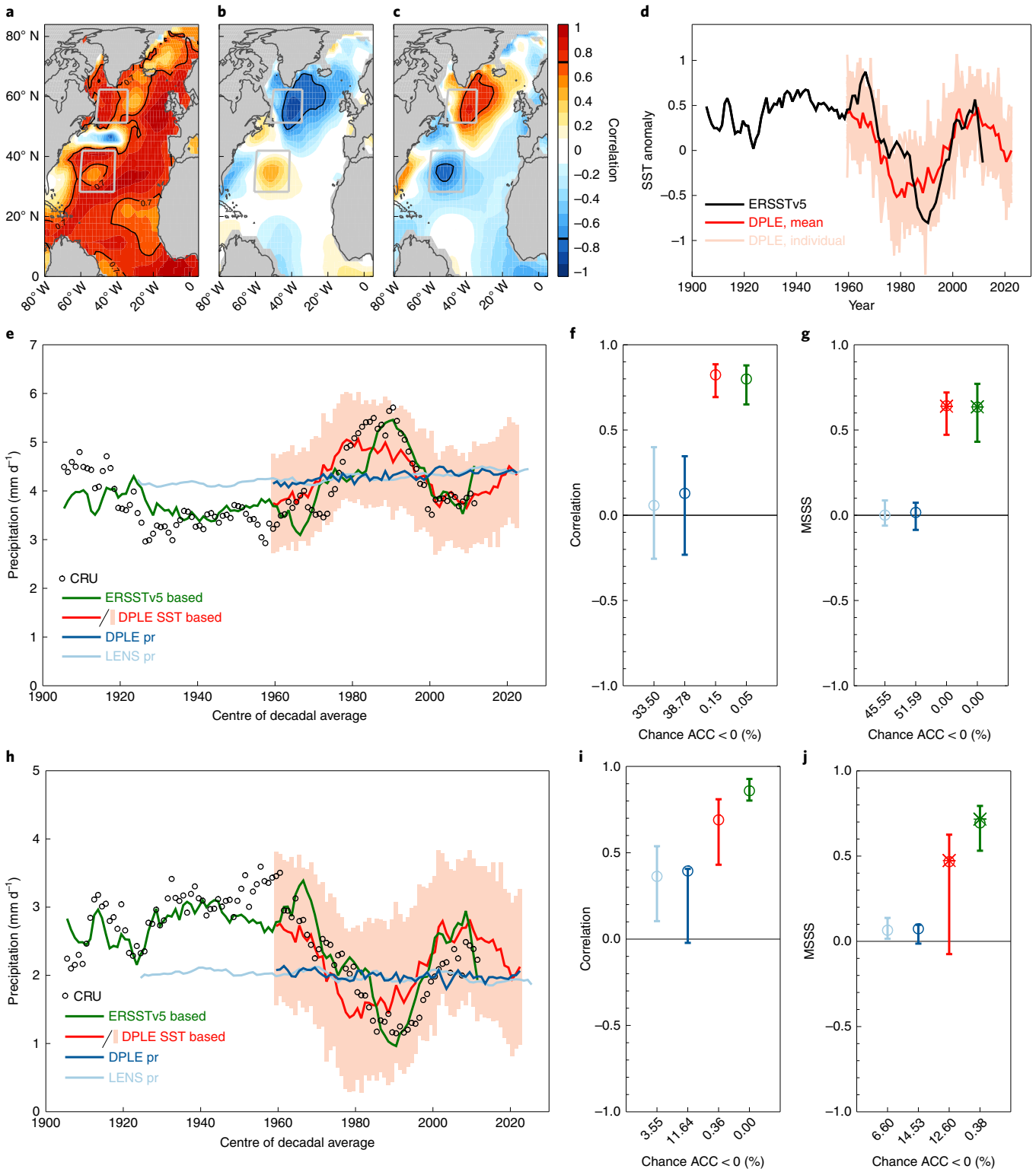

Fig. 4 | Precipitation predictions for the UK and Portugal (Fig. 3d, grey boxes). All the panels use March decadal averages. **a**, Correlation between DPLE and the observed SST. **b**,**c**, Correlation between the observed UK (**b**) and Portugal (**c**) precipitation and SST. **d**, Anomalies in the difference between the northern (55–35° W, 48–59° N) and southern (62–42° W, 26–39° N) SST (boxes in **a**–**c**), relative to the 1964–2016 average. **e**, UK observed and forecast decadal averaged precipitation. The pink shading shows the 95% confidence interval on DPLE SST-based forecasts. **f**, UK precipitation hindcast ACC, 10–90th percentile range and chance (%) that ACC < 0 (Methods). Key as in **e**. **g**, As **f**, but for MSSS against climatology. Asterisk, MSSS using only the period prior to when the DPLE begins to fit the observed precipitation–SST regression. **h**–**j**, as **e**–**g**, but for Portugal.

contrast, the actual precipitation predictions by CESM fail (blue), as expected given the apparent inability of models to capture the relationship between North Atlantic SSTs and the jet stream[6].

The DPLE SST-based hindcasts can only be assessed back to 1955, but the 'prediction' based on observed SSTs gives an indication of the prediction that would be made if SSTs were

predicted perfectly in advance from 1901 (green lines in Fig. 4e,h). As well as capturing the fluctuations since the 1960s this 'prediction' based on observed SSTs also captures the gradual decline/increase from the beginning of the 20th century to the 1940s in the UK/Portugal.

For the UK, the anomaly correlation coefficient (ACC) (Methods) of 0.82 for the DPLE SST-based prediction is far greater than that for the direct DPLE and Large Ensemble (LENS) predictions (Fig. 4f). The mean squared skill score (MSSS) (Methods) of 0.64 demonstrates that the DPLE SST-based prediction is considerably more skilful than climatology (Fig. 4g). For Portugal, the ACC of 0.69 exceeds that of the direct model predictions (Fig. 4i) and, although the MSSS of 0.47 suggests considerably more skill than climatology, the short record available for verification leads to large uncertainties on this metric (Fig. 4j). Skill may be increased further if the SST predictions could be improved (green in Fig. 4) and a similar degree of skill is obtained when using SSTs to predict U700NA (Supplementary Fig. 7). We also compared the DPLE SST-based prediction with those from five other models in Supplementary Fig. 15, which are discussed in Supplementary Section 6. In brief, DPLE exhibits a higher SST skill than the other models and, therefore, exhibits a higher skill in SST-based precipitation predictions, although a number of the models do show promise.

Although our predictions are based on local SSTs in the North Atlantic, the pattern of SST variability that accompanies the AMV is global (Fig. 1 of Simpson et al.[6]) and the ultimate driver need not necessarily be local to the North Atlantic. Even if this is the case, provided that the same global pattern of SSTs continues to accompany the AMV, forecasts made based on the AMV or associated North Atlantic SST anomalies should continue to be valid. We do, however, think that a local forcing of the jet stream by North Atlantic SSTs is far more likely and, although the mechanisms involved remain an open question, it is reasonable to hypothesize that the jet stream responds directly to the altered meridional temperature gradient in the North Atlantic basin[18].

As discussed in more detail in Simpson et al.[6], that the DPLE can predict the low-frequency SPG SST variability at a ten-year lead time without predicting the jet stream variability (Fig. 13 of Simpson et al.[6]) is our best line of evidence that the observed SST–jet stream connection results from an influence of the SSTs on the jet stream. It tells us that the observed connection does not simply represent an influence of the jet stream on the ocean mixed layer. Combined with the fact that external forcings alone cannot reproduce the observed SPG SST variability within CESM (Supplementary Fig. 5), we can deduce that the ocean circulation must be involved and that whatever driving forces gave rise to the SST variability had already left their mark ten years previously, within the ocean/sea ice initialization states. One could attempt to argue that the SST–jet stream connection need not be causal and that some external factor is driving both separately. However, then an explanation is needed as to why the initial ocean state felt the influence of that driving factor ten years before the jet stream did. A more likely explanation is that the SST variability influences the jet stream and, for some reason, our models are deficient in representing this connection.

The majority of the skill arises from the SPG SST predictability (Supplementary Fig. 13), which is due to the ocean initialization (Supplementary Fig. 5)[7]. The skill is, however, enhanced by the inclusion of the longer-term externally forced low latitude trend in the SSTs (compare Fig. 4 with Supplementary Fig. 13). Continued accurate projections of the lower latitude externally forced trend would, therefore, allow this extra increase in skill to continue. In the SPG, Simpson et al.[6] argue that the variability is largely internal in origin, but in terms of predictability it should not matter if SPG variability is internally generated or externally forced, provided the initialized predictions continue to exhibit skill. It is possible that as the atmosphere and SSTs change further under anthropogenic forcing, the SST–jet stream connection, or the optimum SST regions for prediction, may change. However, it is currently impossible to assess this without a general circulation model that successfully reproduces the connections.

An obvious question is why is this signal restricted to March? As discussed in Simpson et al.[6] and shown here in Fig. 1e–h, there are indications that the March signal in U700 is actually one that grows over the course of the winter from weak anomalies in the central North Atlantic in January that strengthen and shift slightly poleward to reach their maximum amplitude in a prime position to influence the precipitation over the UK and Portugal in March. By March, the central North Atlantic anomalies are also accompanied by anomalies in the tropics and polar regions. If the signal does, indeed, grow over the season, then March need not be special—it is just late enough in the winter that the SST forced signal has time to reach a maximum amplitude and move to a position that influences the UK and Portugal. Whether this is actually the case can only be determined unambiguously once a mechanistic understanding is obtained.

In summary, evidence is provided here for pronounced multidecadal precipitation variability in western Europe during March associated with jet stream variability that is highly correlated with North Atlantic SSTs. If the SST–precipitation relationship proves to be robust, it offers the potential for decadal predictability of the low-frequency variability in western European precipitation, as demonstrated here. Although this predictability appears to be limited to the month of March, it may still prove useful for industries that are impacted by March precipitation, most notably the Portuguese wine industry[25]. We end by providing a prediction for the precipitation over the coming decade, which may be verified, or otherwise, as the next ten years unfold. Given the DPLE SST-based predictions and their uncertainties shown in Fig. 4e,h, we expect that precipitation over the coming decade will be somewhere in the middle of the range of decadal averages observed over the past century. For the UK/Portugal, there is only a 0.5%/4.7% chance that it will be as wet/dry as the decade from 1986 to 1995 and only a 3.1%/4.8% chance that it will be as dry/wet as the decade from 1936 to 1945.

Although a longer record for verification and an improved understanding of the physical mechanisms that relate SSTs to the jet stream would help to improve the confidence that a user may have in these predictions, this analysis represents a first step towards exploiting both the observed relationship between SSTs and precipitation and the skilful model predictions of SST to provide skilful decadal predictions for western European precipitation. It is hoped that this could be improved on in the future as our physical understanding increases and model improvements are made.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41561-019-0391-x.

Received: 18 January 2019; Accepted: 21 May 2019;
Published online: 08 July 2019

References

1. Kushnir, Y. Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Clim.* **7**, 141–157 (1994).
2. Trenberth, K. E. & Shea, D. J. Atlantic hurricanes and natural variability in 2005. *Geophys, Res, Lett.* **33**, L12704 (2006).
3. Frankignoul, C., Gastineau, G. & Kwon, Y.-O. Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic multidecadal oscillation and the Pacific decadal oscillation. *J. Clim.* **30**, 9871–9895 (2017).
4. Clement, A. et al. The Atlantic multidecadal oscillation without a role for ocean circulation. *Science* **350**, 320–324 (2015).

5. Delworth, T. L. et al. The central role of ocean dynamics in connecting the North Atlantic oscillation to the extratropical component of the Atlantic multidecadal oscillation. *J. Clim.* **30**, 3789–3805 (2017).
6. Simpson, I. R., Deser, C., McKinnon, K. A. & Barnes, E. A. Modelled and observed multi-decadal variability in the North Atlantic jet stream and its connection to sea surface temperatures. *J. Clim.* **31**, 8313–8338 (2018).
7. Yeager, S. G. et al. Predicting near-term changes in the Earth system: a large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bull. Am. Meteorol. Soc.* **99**, 1867–1886 (2018).
8. Sutton, R. T. & Dong, B. Atlantic Ocean influence on a shift in European climate in the 1990s. *Nat. Geosci.* **5**, 788–792 (2012).
9. Dong, B., Sutton, R. T., Woollings, T. & Hodges, K. Variability of the North Atlantic summer storm track: mechanisms and impacts on European climate. *Environ. Res. Lett.* **8**, 034037 (2013).
10. Dunstone, N. et al. Skilful seasonal predictions of summer European rainfall. *Geophys. Res. Lett.* **45**, 3254–3426 (2018).
11. Osso, A., Sutton, R., Shaffrey, L. & Dong, B. Observational evidence of European summer weather patterns predictable from spring. *Proc. Natl Acad. Sci. USA* **115**, 59–63 (2018).
12. Ting, M., Kushnir, Y. & Li, C. North Atlantic multidecadal SST oscillation: external forcing versus internal variability. *J. Mar. Syst.* **133**, 27–38 (2014).
13. Gastineau, G. & Frankignoul, C. Influence of the North Atlantic SST variability on the atmospheric circulation during the twentieth century. *J. Clim.* **28**, 1396–1416 (2015).
14. Hodson, D. L. R. et al. Climate impacts of recent multidecadal changes in Atlantic Ocean sea surface temperature: a multimodel comparison. *Clim. Dyn.* **34**, 1041–1058 (2010).
15. Kushnir, Y. & Held, I. M. Equilibrium atmospheric response to North Atlantic SST anomalies. *J. Clim.* **9**, 1208–1220 (1996).
16. Davini, P., von Hardenberg, J. & Corti, S. Tropical origin for the impacts of the Atlantic multidecadal variability on the Euro-Atlantic climate. *Env. Res. Lett.* **10**, 094010 (2015).
17. Omrani, N.-E., Keenlyside, N. S., Bader, J. & Manzini, E. Stratosphere key for wintertime atmospheric response to warm Atlantic decadal conditions. *Clim. Dyn.* **42**, 649–663 (2014).
18. Peings, Y. & Magnusdottir, G. Forcing of the wintertime atmospheric circulation by the multidecadal fluctuations of the North Atlantic Ocean. *Environ. Res. Lett.* **9**, 034018 (2014).
19. Peings, Y. & Magnusdottir, G. Wintertime atmospheric response to Atlantic multidecadal variability: effect of stratospheric representation and ocean-atmosphere coupling. *Clim. Dyn.* **47**, 1029–1047 (2016).
20. Ruprich-Robert, Y. et al. Assessing the climate impacts of the observed Atlantic multidecadal variability using the GFDL CM2.1 and NCAR CESM1 global coupled models. *J. Clim.* **30**, 2785–2810 (2017).
21. Eade, R. et al. Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.* **41**, 5620–5628 (2014).
22. Chang, E. K. M. & Yau, A. M. W. Northern Hemisphere winter storm track trends since 1959 derived from multiple reanalysis datasets. *Clim. Dyn.* **47**, 1435–1454 (2016).
23. Dell'Aquila, A. et al. Benchmarking Northern Hemisphere midlatitude atmospheric synoptic variability in centennial reanalysis and numerical simulations. *Geophys. Res. Lett.* **43**, 5442–5449 (2016).
24. Trigo, R. M. & DaCamara, C. C. Circulation weather types and their influence on the precipitation regime in Portugal. *Int. J. Climatol.* **20**, 1559–1581 (2000).
25. Santos, J. A., Malheiro, A. C., Karremann, M. K. & Pinto, J. G. Statistical modelling of grapevine yield in the Port Wine region under present and future climate conditions. *Int. J. Biometeorol.* **55**, 119–131 (2011).

Acknowledgements

The CESM project is supported primarily by the National Science Foundation (NSF). This material is based on work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement no. 1852977. S.G.Y. acknowledges the support of the NSF Collaborative Research EaSM2 Grant OCE-1243015. Computing and data storage resources, which include the Cheyenne supercomputer (https://doi.org/10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. NOAA_ERSST_v5 data were provided by the NOAA/OAR/ESRL PSD at www.esrl.noaa.gov/psd/. The data analysis was performed with IDL (Excelis Visual Information Solutions).

Author contributions

I.R.S. conceived the study and conducted the analysis. All the authors contributed to decisions regarding the methodology and contributed to the interpretation of the results. S.G.Y. led the DPLE model experiments used extensively in the analysis. I.R.S. wrote the manuscript with input from all the authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41561-019-0391-x.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to I.R.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Observation-based data sets. The observation-based analysis considers the period that spans 1901 to 2016 as this is the period of overlap of the precipitation and zonal wind data sets used. Zonal wind on the 700 hPa level (U700) from ERA-20C[26] from 1901 to 2010 and ERA-Interim[27] from 1979 to 2016 were used to describe the jet stream variability. Only observations of the surface pressure and marine surface winds were used to constrain ERA-20C, whereas a much more extensive range of observations was used to constrain ERA-Interim. When the full 1901–2016 record was considered, ERA-20C from 1901 to 2010 was combined with ERA-Interim from 2011 to 2016, referred to as ERA-20C+ERA-Interim. These reanalyses are in good agreement with other reanalyses in terms of their jet stream variability (Simpson et al.[6] gives a more extensive discussion). For consistency with Simpson et al.[6], U700 was interpolated onto a 2°×2° longitude–latitude grid and isotropically smoothed in the spectral domain, retaining only scales larger than the total wavenumber 42 according to equation (9) in Sardeshmukh and Hoskins[28] with $n_o = 42$ and $r = 1$.

For precipitation, we primarily made use of the CRU time series (CRU TS) version 4.01 precipitation data set[29], referred to here as the CRU. This gridded monthly precipitation data set is derived from station-based observations and spans from 1901 to 2016 with a 0.5°×0.5° horizontal resolution. Similar conclusions are found with the Global Precipitation Climatology Center (GPCC) version 2018 precipitation data set at 0.5° resolution[30]. Equivalent figures to Figs. 1–4 are shown and discussed for GPCC in Supplementary Figs. 1–4.

SST variability is characterized using National Oceanic and Atmospheric Administration (NOAA)'s Extended Reconstruction SST version 5 (ERSSTv5)[31,32] and the index of the AMV considered in this study is that of Trenberth and Shea (2006)[2]. This involves taking the average monthly SST over the area 80° W to 0° and from 0° to 60° N and then subtracting the area-averaged SST from 60° S to 60° N over all the ocean basins in an attempt to remove the externally forced signal. Frankignoul et al.[3] assessed the ability of various methods to isolate the internal variability from the externally forced signal in large ensembles and found this method to work well in the North Atlantic. Furthermore, as discussed in Simpson et al.[6], similar conclusions as to the connection between the AMV and the North Atlantic jet stream were found using the Optimal Linear Inverse Model method favoured by Frankignoul et al[3] and the conclusions are also not dependent on the SST data set considered. Throughout the analysis, the AMV index is normalized to have a zero mean and s.d. = 1 after any low-pass filtering has been performed.

Model-based data sets. To provide decadal predictions for the climatological precipitation in Fig. 4, we made use of the SST and precipitation predictions from the CESM DPLE[7]. This is a 40-member ensemble of hindcasts, initialized on 1 November each year from 1954 to 2017 and run for 10 yr. They are run under the Coupled Model Intercomparison Project phase 5 'historical' forcings prior to 2006 and RCP8.5 forcings thereafter (ref. [33] and references therein) and are initialized with ocean and sea ice states determined from a reanalysis-forced simulation with the ocean and sea ice models. The atmospheric initial conditions are those from an equivalent 40-member ensemble of uninitialized CESM simulations, referred to as the LENS[33]. We also made use of the LENS precipitation predictions in Fig. 4.

Comparing multidecadal variability to expectations from interannual sampling. In Fig. 1i–l, the s.d. of the 20 yr running mean precipitation is compared to expectations from the sampling of year-to-year variability. This was achieved by randomly resampling individual years with replacement to obtain a resampled time series of equivalent length to the observational record, in which all autocorrelation was lost. The 20 yr running mean of this resampled time series was obtained and its s.d. calculated. This was repeated 1,000 times. The significance was determined based on where the observed 20 yr running mean s.d. sat in terms of the percentiles of this resampled distribution. Spatial autocorrelation was accounted for by using the false discovery rate (FDR) method[34] with control levels $\alpha_{FDR} = 0.1$ and 0.02 for significance at the 95% and 99% levels, respectively (equation (3) of Wilks[34]).

Reconstruction of the U700 variability from precipitation. The analysis presented in Fig. 2 shows the reconstruction of the North Atlantic U700NA index (area-averaged U700 from 40° W to 10° W and 50° N to 60° N) based on the prdif between western Scotland (7–4° W, 55–58° N) and Portugal (9–6° W, 37–43° N). This was achieved by first considering the interannual regression of U700NA onto prdif using ERA-Interim U700 and CRU precipitation from 1979 to 2016, that is:

$$\mathrm{U700NA}(iy) = a + b \times \mathrm{prdif}(iy) + \varepsilon(iy) \quad (1)$$

where iy refers to the year, a and b are the regression coefficients and ε refers to residuals. The regression coefficients a and b were determined from 1979–2016 interannual variability (quoted in Fig. 2c). These regression coefficients were used along with observations of the prdif to reconstruct U700NA back to 1901 assuming:

$$\mathrm{U700NA}(iy) = a + b \times \mathrm{prdif}(iy) \quad (2)$$

(red line in Fig. 2d), which can then be compared with the ERA-20C time series (black line in Fig. 2d). The 20 yr running mean of ERA-20C and the reconstruction are also shown in Fig. 2e. A 95% confidence interval on the reconstructed U700NA time series was obtained by resampling, with replacement, the residuals in equation (1), constructing a bootstrapped time series of U700NA according to equation (1) but with $\varepsilon(iy)$ replaced by these resampled residuals and then recalculating the regression coefficients a and b. A total of 1,000 estimates of a and b were determined to produce 1,000 estimates of the reconstructed U700NA. The pink shading in Fig. 2d,e shows the 2.5th to 97.5th percentile range of these bootstrapped time series.

Correlations and regressions with respect to the AMV. The correlation between a field X and the AMV is calculated as the Pearson correlation coefficient between X and the AMV, as shown in Fig. 3a,c for the 20 yr running mean U700 and precipitation, respectively, and quoted for the time series in Fig. 3f,g. The low-frequency regression of a field X onto the AMV is given by:

$$X(i) = a + b \times \mathrm{AMV}(i) + \varepsilon(i) \quad (3)$$

where i refers to the ith running mean, a and b are the regression coefficients and $\varepsilon(i)$ are the residuals. The regression coefficient b is shown in Fig. 3b,d for the 20 yr running mean U700 and precipitation, respectively.

Given that the focus is on the low-frequency variability within a relatively short observational record, care must be taken to assess the likelihood that correlations or regression coefficients may have arisen by chance. To assess the significance of correlations or regression coefficients between two time series TS1 and TS2 of length N, as in Fig. 3a–d, the resampling method of Delworth et al.[5] was employed. This involves reshuffling both TS1 and TS2 by choosing random numbers i and j between 1 and N. The segment for TS1(TS2) from year = $i(j)$ to year = N is then pieced together with the segment from year = 1 to year = $i(j) - 1$. The correlation, or regression coefficient, of these shuffled time series was assessed and this was repeated 10,000 times with different values of i and j. This was found to yield roughly 1,600 independent correlation/regression coefficient estimates and a P value was obtained for the correlation/regression coefficients based on the percentile at which the value occurred within the distribution of the 10,000 randomly resampled values. This methodology was used for the significance tests provided in Fig. 1e–h, Fig. 3c,d and the P values quoted in Fig. 3f,g. For the mapped significance, grid points that were significant at the 95% level were determined by accounting for the spatial autocorrelation using the FDR method, with the control level $\alpha_{FDR} = 0.1$ (ref. [34]).

To assess the magnitude of the uncertainty in a regression coefficient determined from low-frequency variability (as shown in Fig. 3e), a resampling procedure was also used, whereby the time series was divided into high- and low-frequency components. The low-frequency component is the running mean considered and the high frequency is the remainder. For Fig. 3e, we used 21 yr running means, as opposed to 20 yr running means, so that the low-frequency component was centred on a given year, which led to a straightforward calculation of the high-frequency remainder. To determine the uncertainty, the high-frequency remainder was reshuffled by resampling individual years with replacement, preserving the relationship between the two variables considered. The running mean of this reshuffled high-frequency time series was then obtained and combined with the original low-frequency time series to obtain new low-frequency time series. The regression coefficient for these new time series was then calculated and this procedure repeated 1,000 times. The uncertainty range on the regression coefficient is the 2.5th to 97.5th percentile range of these resampled regression coefficients.

Decadal prediction methodology. To predict the precipitation based on SST we used, as a predictor, the SST difference between the area averages of the northern and southern grey boxes in Fig. 4a–c. The northern box is from 55–35° W, 48–59° N and the southern box is from 62–42° W, 26–39° N and the difference between these is referred to as SST_{n-s}. This predictor was found to be one that optimized the skill in the forecast and sensitivity to this choice is discussed in the Supplementary Information. Hindcasts/forecasts were performed for anomalies from the baseline climatology of 1964–2016, denoted $(.)'$, with anomalies for the DPLE computed relative to a lead-time-dependent model climatology that spanned the same years (Supplementary Information gives more details). All the forecasts were then converted into actual precipitation values by adding back the baseline climatology of the observed precipitation.

Predictions were performed for decadal averages, referred to with an overbar $\overline{(.)}$ and the decadal averaged $\overline{\mathrm{SST}'}_{n-s}$ from the DPLE can be compared with the observations in Fig. 4d. Taking the UK as an example, the regression of the decadal averaged CRU precipitation ($\overline{\mathrm{pr}'}_{\mathrm{UK}}$) onto the ERSSTv5 $\overline{\mathrm{SST}'}_{n-s}$ was performed for 1901–2016 to obtain the regression coefficients a and b, that is:

$$\overline{\mathrm{pr}'}_{\mathrm{UK}}(i) = a + b \times \overline{\mathrm{SST}'}_{n-s}(i) + \varepsilon(i) \quad (4)$$

where i refers to the ith decadal mean. These regression coefficients were then used along with the ensemble mean predictions of the decadal averaged $\overline{\mathrm{SST}'}_{n-s}$ from the DPLE ($\overline{\mathrm{SST}'}_{n-s,\mathrm{DPLE}}$) to perform hindcasts (hc) (and forecasts) of the decadal averaged $\overline{\mathrm{pr}'}_{\mathrm{UK}}$ according to:

$$\overline{\mathrm{pr}'}_{\mathrm{UK,hc}}(i) = a + b \times \overline{\mathrm{SST}'}_{n-s,\mathrm{DPLE}}(i) \quad (5)$$

These hindcasts/forecasts are shown in red in Fig. 4e,h and the detailed methodology of the calculation of the uncertainty (pink shading in Fig. 4e,h) is provided in the Supplementary Information.

The 'prediction' based on observed SSTs was performed by replacing $\overline{\mathrm{SST}'}_{n-s,\mathrm{DPLE}}$ in equation (5) with the observed decadal average SST anomalies (green lines in Fig. 4e,h). In addition, the predictions from the CESM DPLE and LENS are simply the decadal averaged March precipitation from the simulations (dark and light blue lines in Fig. 4e,h).

Two measures of hindcast skill were assessed, as recommended by Goddard et al.[35], for each of the prediction methods over the period of common overlap between all data sets, which consists of the decadal average that starts in 1955 to the decadal average that starts in 2007. The ACC of a hindcast (hc) with respect to observations (obs) is given by:

$$\mathrm{ACC(hc)} = \frac{\mathrm{cov(hc, obs)}}{\sigma_{\mathrm{hc}}\sigma_{\mathrm{obs}}} \quad (6)$$

Here, hc and obs refer to the anomalies from the average over all hindcasts being assessed, cov refers to the covariance between hc and obs, and σ refers to the s.d. of the time series.

The MSSS of the hindcast relative to climatology is given by:

$$\mathrm{MSSS(hc)} = 1 - \frac{\mathrm{MSE}_{\mathrm{hc}}}{\mathrm{MSE}_{\mathrm{obs}}} \quad (7)$$

where MSE refers to the mean squared error, which for decadal hindcasts $j = 1$, n is given by:

$$\mathrm{MSE}_{\mathrm{hc}} = \frac{1}{n}\sum_{j=1}^{j=n}(\mathrm{hc}(j) - \mathrm{obs}(j))^2 \quad (8)$$

that is, the mean squared error of the hindcast relative to observations. $\mathrm{MSE}_{\mathrm{obs}}$ refers to the mean squared error of the observations relative to climatology:

$$\mathrm{MSE}_{\mathrm{obs}} = \frac{1}{n}\sum_{j=1}^{j=n}(\mathrm{obs}(j))^2 \quad (9)$$

that is, the mean squared anomaly of observations relative to the average of the observed decadal means.

Uncertainty estimates on the ACC and the MSSS were obtained following the methods outlined in Goddard et al.[35] and Yeager et al.[7], which make use of a block bootstrapping methodology that involves resampling in time, and across DPLE or LENS ensemble members in the case of model-based predictions, as discussed in detail in the Supplementary Information. The possibility of forced trends in precipitation, unrelated to SST_{n-s} variability, is not considered in the DPLE SST-based predictions, although the high skill of the hindcasts suggest that any such forced trends were small. This may change, however, if any forced signal that is unrelated to SST_{n-s} becomes larger in the future.

In Fig. 4g,h, additional skill estimates are provided (asterisks) in which regression (4) was performed only with the years prior to the beginning of the DPLE, that is, the period of verification was not included in the estimates of the regression coefficients a and b, which results in a true out-of-sample test. The MSSS is comparable to that of the forecasts using the full record for regression (4).

A discussion of the sensitivity to the above methodological choices, such as the SST index considered, the region used for the precipitation average and the lead time considered, is provided in the Supplementary Information. In short, only minor changes to the skill, or the uncertainty estimates, were found when varying these aspects within reasonable bounds and the skill does not solely arise from the first few years of the decadal predictions. In addition, although the DPLE does, in some sense, make use of 'future' information prior to 2006 in the form of the imposed historical forcings, the majority of the skill in the precipitation hindcasts arises from the subpolar North Atlantic SST prediction and the SST prediction skill in that region arises from the ocean initialization, not the external forcings (Supplementary Fig. 5). An additional ten-member ensemble of simulations analogous to the DPLE except that volcanic forcing is omitted was also performed and this demonstrates skill that is comparable to that of the DPLE itself (Supplementary Fig. 5d). Therefore, volcanic eruptions, which are entirely unpredictable, are not the source of the skill that arises from external forcings and it is possible that the externally forced skill could continue in future forecasts provided the scenarios of future forcings are adequate. This is discussed in more detail in Supplementary Section 3.

Data availability

All data sets used in this study are publicly available. The CESM large ensemble and DPLEs are available through NCAR's Climate Data Gateway at www.earthsystemgrid.org/. ECMWF reanalyses are available from https://apps.ecmwf.int/datasets/. The CRU TS version 4.01 precipitation data set is available at https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.01/, GPCC precipitation is available at https://doi.org/10.5676/DWD_GPCC/FD_M_V2018_050 and ERSSTv5 SSTs are available at https://doi.org/10.7289/V5T72FNM.

Code availability

All analysis codes will be made available from I.R.S. (islas@ucar.edu) on request.

References

26. Poli, P. et al. ERA-20C: an atmospheric reanalysis of the twentieth century. *J. Clim.* **29**, 4083–4097 (2016).
27. 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).
28. Sardeshmukh, P. D. & Hoskins, B. J. Spatial smoothing on the sphere. *Mon. Weath. Rev.* **112**, 2524–2529 (1984).
29. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
30. Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A. & Ziese, M. *GPCC Full Data Monthly Product Version 2018 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data* (Global Precipitation Climatology Centre, 2018); https://doi.org/10.5676/DWD_GPCC/FD_M_V2018_050
31. Huang, B. et al. Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J. Clim.* **30**, 8179–8205 (2017).
32. Huang, B. et al. *NOAA Extended Reconstructed Sea Surface Temperature (ERSST)* Version 5 (NOAA, 2017); https://doi.org/10.7289/V5T72FNM
33. Kay, J. E. et al. The Community Earth System Model (CESM) Large Ensemble Project: a community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.* **96**, 1333–1349 (2015).
34. Wilks, D. S. 'The stippling shows statistically significant grid points', how research results are routinely overstated and overinterpreted and what to do about it. *Bull. Am. Meteorol. Soc.* **97**, 2263–2273 (2016).
35. Goddard, L. et al. A verification framework for interannual-to-decadal predictions experiments. *Clim. Dyn.* **40**, 245–272 (2013).