

## ARTICLE OPEN



# Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions

Steven C. Hardiman<sup>1</sup>✉, Nick J. Dunstone<sup>1</sup>, Adam A. Scaife<sup>1,2</sup>, Doug M. Smith<sup>1</sup>, Ruth Comer<sup>1</sup>, Yu Nie<sup>3</sup> and Hong-Li Ren<sup>4</sup>

The signal-to-noise paradox that climate models are better at predicting the real world than their own ensemble forecast members highlights a serious and currently unresolved model error, adversely affecting climate predictions and introducing uncertainty into climate projections. By computing the magnitude of feedback between transient eddies and large-scale flow anomalies in multiple seasonal forecast systems, this study shows that current systems underestimate this positive eddy feedback, and that this deficiency is strongly linked to weak signal-to-noise ratios in ensemble mean predictions. Improved eddy feedback is further shown to be linked to greater teleconnection strength between the El Niño Southern Oscillation and the Arctic Oscillation and to stronger predictable signals. We also present a technique to estimate the potential gain in skill that may come from eliminating eddy feedback deficiency, showing that skill could double in some extratropical regions, significantly improving predictions of the Arctic Oscillation.

npj *Climate and Atmospheric Science* (2022)5:57; <https://doi.org/10.1038/s41612-022-00280-4>

## INTRODUCTION

The signal-to-noise paradox<sup>1,2</sup> highlights a serious deficiency in climate models and has recently generated a lot of interest in the climate prediction community. In a perfect seasonal forecast system, each forecast ensemble member should behave equivalently to the observations, in the sense that the skill of the model in predicting the observations (or the correlation of the model ensemble mean with the observations,  $r_{mo}$ ) should be statistically indistinguishable from the skill of the model in predicting its own forecast ensemble members (or the average correlation of any given ensemble member with the mean of all other members,  $r_{mm}$ ). In other words the Ratio of Predictable Components ( $RPC = r_{mo}/r_{mm}$ ) should equal 1. The signal-to-noise paradox occurs when RPC is greater than 1, as is found to be the case in almost all current seasonal forecast systems<sup>3</sup>. Forecasts in regions where  $RPC > 1$  are ‘under-confident’, and this leads to a number of issues, including the apparently paradoxical result that the model is better at predicting the real world than its own ensemble forecast members. The standard metrics for forecast skill will then underestimate the skill that is potentially available in a forecast system<sup>2</sup>. Additionally, a very large ensemble is needed to extract the predictable signal, and the variance of the ensemble mean must be inflated to match the observed predictable signal<sup>1</sup>. Small signal-to-noise ratios also lead to uncertainty in decadal predictions, and it requires very large ensembles to remove this uncertainty. For example, the North Atlantic Oscillation (NAO) predictable signals are around 2–3 times too small in seasonal forecasts, and could be as much as ten times too small in decadal forecasts<sup>4,5</sup>. Since the error variance decreases linearly with ensemble size, such weak signals require 4–9 (seasonal) and 100 (decadal) times more ensemble members than would a perfect model, representing a significant computational cost<sup>4,6,7</sup>. Therefore there is a current need for large ensembles, and a significant effort in the seasonal

and decadal forecast communities to try to explain and resolve the signal-to-noise paradox.

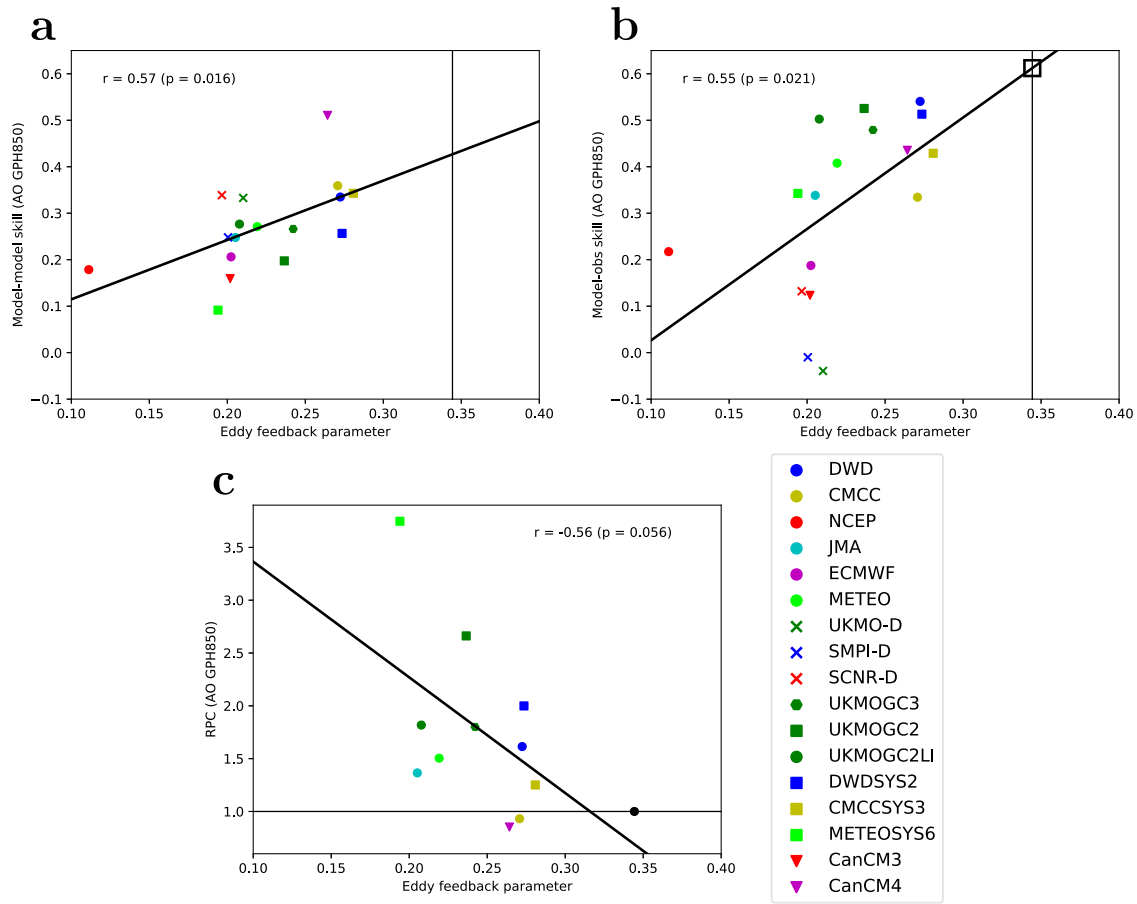
Currently, theories as to why the paradox exists include: a deficiency of atmospheric eddy feedback in climate models due to insufficient spatial resolution<sup>8</sup>, a deficiency in simulated ocean eddies alongside weak ocean–atmosphere coupling<sup>9–11</sup>, a bias in surface drag leading to too much baroclinic instability<sup>12</sup>, and inaccurate regime persistence<sup>13,14</sup>. Some of these may be interrelated<sup>15</sup>, but in this study, we focus on the first of these ideas, a deficiency in atmospheric eddy feedback.

Eddy feedback is the process whereby interaction with small-scale transient eddies amplifies large-scale quasi-stationary climate anomalies in the mid-latitudes. It is an essential part of the process whereby remote influences impact the tropospheric jets<sup>16–27</sup>. Eddy feedback is therefore crucial for simulating the correct strength of the change to the tropospheric jets caused by remote influences in climate models, from monthly through to centennial time scales. However, eddy feedback has been shown to be deficient in current climate models<sup>8,24,28</sup>. Although there have been several hypotheses relating to the interaction of wave propagation and breaking with the latitudinal and vertical structure of the jets in models<sup>29–31</sup>, none have yet suggested a means of resolving this model deficiency.

By computing the magnitude of eddy feedback in a range of seasonal forecast systems we demonstrate that these systems are deficient in eddy feedback, and we show an important link between this deficiency and the signal-to-noise error in those systems. We consider how reducing the eddy feedback deficiency in forecast systems has the potential to improve their skill. A mechanism whereby this may occur, via the influence of the El Niño Southern Oscillation (ENSO) on the Arctic Oscillation<sup>32,33</sup> is discussed. Finally, the potential for more accurate eddy feedback to improve regional forecast skill is considered.

<sup>1</sup>Met Office, FitzRoy Road, Exeter, Devon, UK. <sup>2</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, Devon, UK. <sup>3</sup>Laboratory for Climate Studies, National Climate Center, China Meteorological Administration, Beijing, PR China. <sup>4</sup>Chinese Academy of Meteorological Sciences, Beijing, PR China.

✉email: [steven.hardiman@metoffice.gov.uk](mailto:steven.hardiman@metoffice.gov.uk)



**Fig. 1 Eddy feedback relation to skill and RPC.** Eddy feedback parameter (EFP) correlated with model–model skill ( $r_{mm}$ ), model–observed skill ( $r_{mo}$ ) and RPC (the ratio of predictable components, where  $RPC = r_{mo}/r_{mm}$ ), as calculated from the AO index using geopotential height at 850 hPa (AO GPH850). EFP is defined as the correlation squared between the December–January–February horizontal EP-flux divergence and the zonal mean zonal wind, at each latitude at 500 hPa, area weighted 25°N–72°N. The value of EFP computed from ERA5 reanalysis data is shown with a black vertical line in **a** and **b** and as a black circle in **c**. Thick black lines show linear least squares regression lines. In **b** the point where the regression line crosses the observed value of EFP is marked with a hollow black box as discussed in relation to Fig. 4. Models used in **c** are those that have significance at the 90% level in **b**, equating to an  $r_{mo}$  value greater than 0.277. In Table 1, the same subset of models is used at all levels. See ‘Methods’ section for explanation of different symbols.

## RESULTS

### Signal-to-noise

In order to investigate a link between eddy feedback and signal-to-noise ratio, data from seventeen seasonal forecast systems is used, alongside the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis version 5, ERA5<sup>34,35</sup> dataset. We define an eddy feedback parameter (EFP) for ERA5 and for each forecast system<sup>28</sup>, quantifying the link between atmospheric wave driving and the zonal mean wind. Predictable components ( $r_{mm}$  and  $r_{mo}$ ) are formed from an Arctic Oscillation (AO) index. Details of the reanalysis data and all forecast systems, equations for the eddy feedback parameter, and definitions of the AO and the predictable components are all given in the ‘Methods’ section.

We find that the forecast system predictable components, both the model–observed skill ( $r_{mo}$ ) and the model–model skill ( $r_{mm}$ ), are significantly positively correlated with eddy feedback (Fig. 1a and b). Systems are all deficient in eddy feedback, thus Fig. 1 demonstrates that systems with greater and more realistic values of EFP also exhibit increased skill. The values of  $r_{mm}$  are around 0.25, lower than the values of  $r_{mo}$  which are around 0.4, demonstrating the signal-to-noise error. The RPC becomes ill-defined as model skill ( $r_{mo}$ ) tends to zero<sup>2</sup>, thus we focus on systems with significant skill when considering the RPC. Figure 1c shows a significant negative correlation between eddy feedback

and RPC in these skilful systems. Again, since RPC is greater than 1 in the vast majority of systems (the signal-to-noise paradox), a greater and more realistic value of EFP leads to a value of RPC closer to 1 and, therefore, a reduced signal-to-noise error. Indeed, the regression line in Fig. 1c crosses  $RPC = 1$  very close to the value of EFP calculated from observational reanalysis. This figure supports the hypothesis that increased eddy feedback, potentially through increased horizontal resolution, would improve weak signal-to-noise ratios in forecast systems<sup>8</sup>.

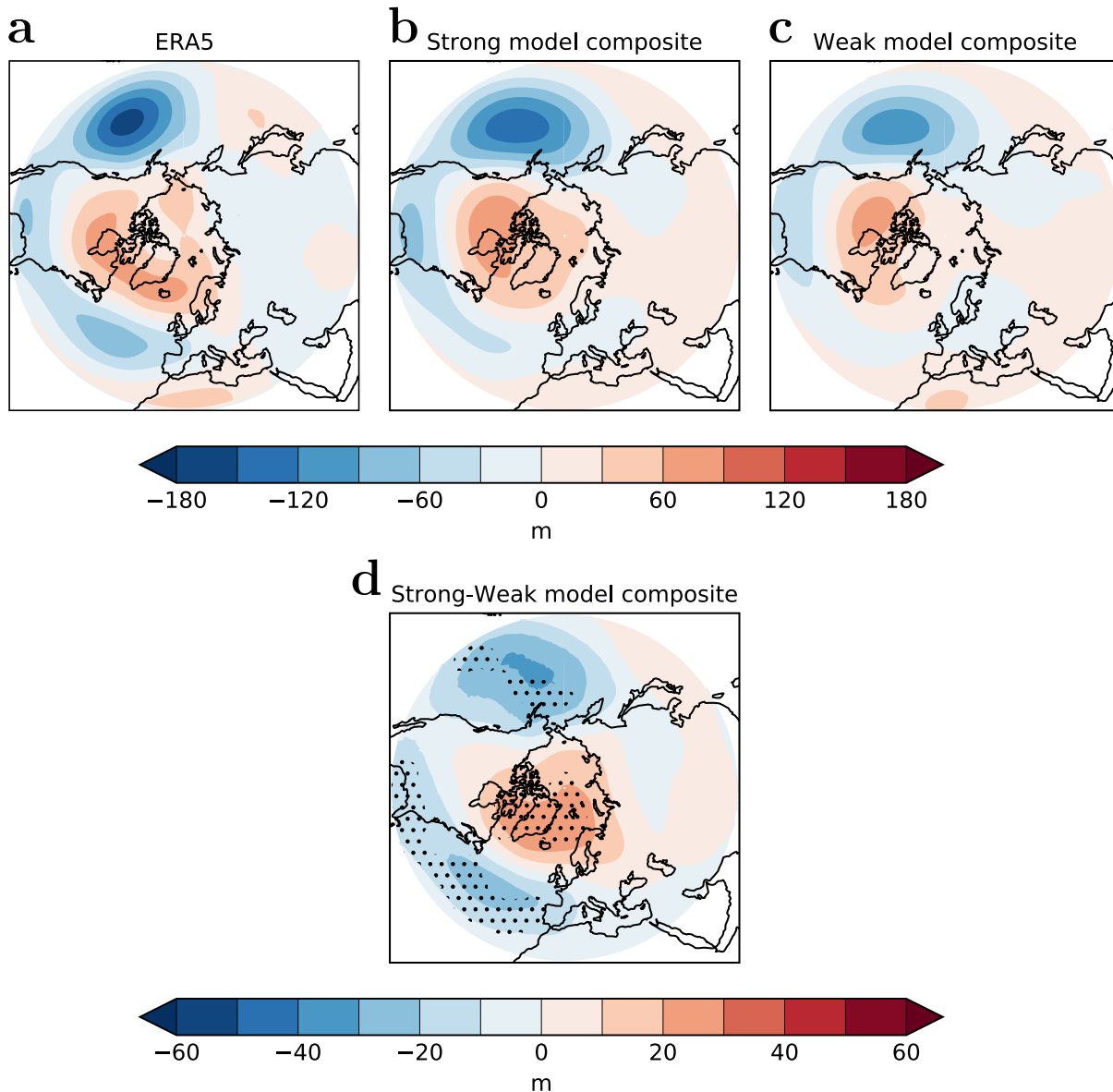
Figure 1 shows results calculated at 850 hPa, but all results shown are largely independent of height (see Table 1 where results are shown using geopotential height (GPH) at 500 hPa and 850 hPa, and using mean sea-level pressure (MSLP)). It is stressed that, throughout this study, it should be remembered that correlations between two quantities do not demonstrate causality, but rather show where there is a significant link or relationship between those quantities. Nevertheless, Fig. 1 shows improvements in simulated eddy feedback to be linked to improved signal-to-noise ratios, and also linked to increased system skill.

### ENSO teleconnection

The impact of ENSO on the AO is well documented, with teleconnection pathways going via the troposphere and stratosphere. The importance of eddy feedback for the response via the

**Table 1.** Correlation ( $r$ ) between eddy feedback parameter (EFP) and skill/RPC.

	AO (GPH500)	AO (GPH850)	AO (MSLP)	NAO (MSLP)
$r$ ( $r_{mm}$ vs EFP)	0.48 ( $p = 0.050$ )	0.57 ( $p = 0.016$ )	0.53 ( $p = 0.028$ )	0.53 ( $p = 0.028$ )
$r$ ( $r_{mo}$ vs EFP)	0.53 ( $p = 0.028$ )	0.55 ( $p = 0.021$ )	0.42 ( $p = 0.092$ )	0.21 ( $p = 0.418$ )
$r$ (RPC vs EFP)	-0.52 ( $p = 0.080$ )	-0.56 ( $p = 0.056$ )	-0.58 ( $p = 0.046$ )	-0.54 ( $p = 0.067$ )

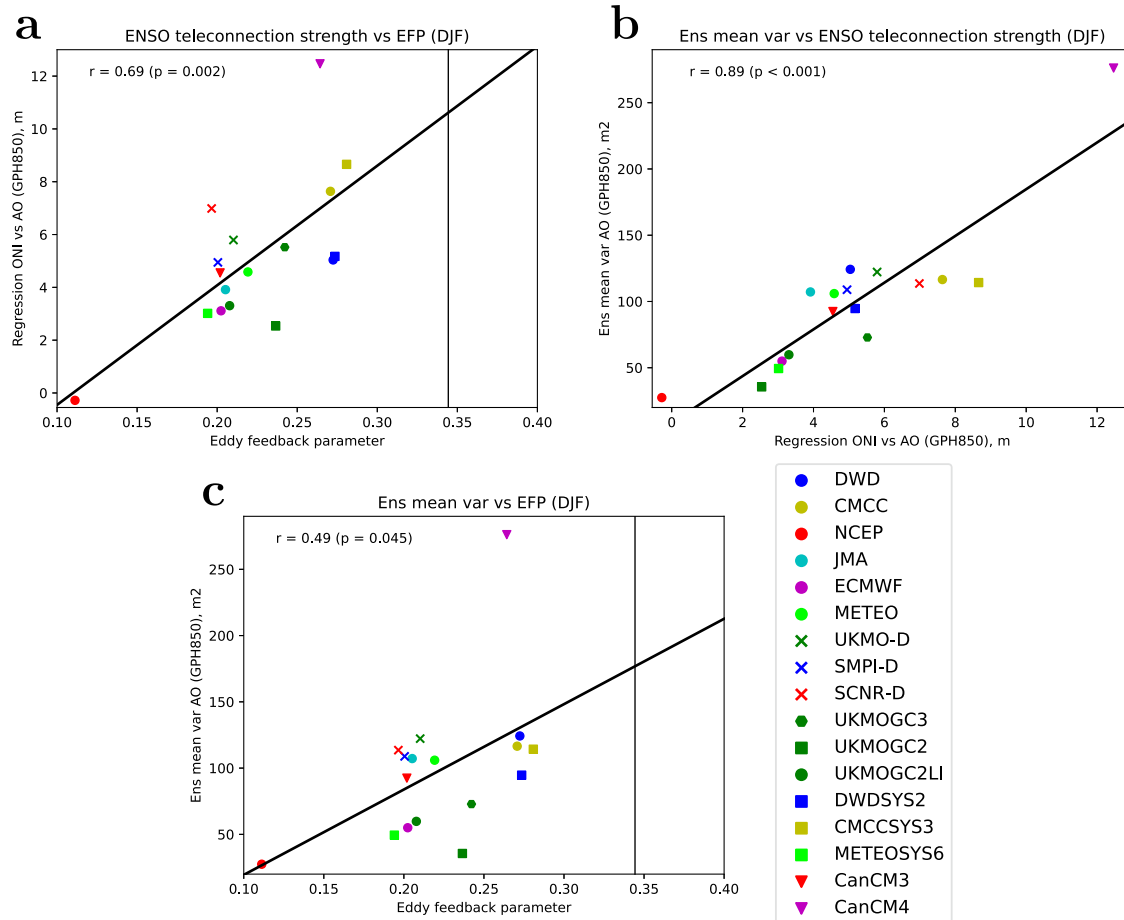


**Fig. 2 ENSO teleconnection.** ENSO anomalies of the 500 hPa GPH field for **a** ERA5, composites over models with **b** strong and **c** weak eddy feedback<sup>24</sup>, and **d** the difference between strong and weak composites (**b–c**). ENSO anomalies are formed by taking GPH anomalies in El Niño years (Niño 3.4 index > 1.5K) minus those in La Niña years (Niño 3.4 index < 1.5K). Model composites based on eddy feedback are defined in the Methods section. Stippling in **d** shows where the difference in model composites is statistically significant at the 95% level.

tropospheric pathway has long been known<sup>36</sup>. In this study, we focus on the stratospheric teleconnection pathway, for which the mechanism is as follows. El Niño leads to an intensification and eastward shift of the climatological Aleutian cyclone in the North Pacific, causing increased planetary wave flux to enter the stratosphere<sup>37</sup>, where it leads to a weaker than average winter

stratospheric polar vortex<sup>38</sup>. This weak vortex then projects onto a negative AO<sup>25</sup>. The opposite occurs for La Niña. One way of diagnosing this teleconnection is to consider GPH anomalies, composited onto El Niño and La Niña years.

Figure 2 shows maps of the difference in GPH 500 hPa anomalies between El Niño and La Niña years, for the reanalysis



**Fig. 3 EFP, teleconnection strength, and predictable signals.** Correlation of ensemble mean **a** eddy feedback parameter (EFP) and ENSO teleconnection strength, **b** ENSO teleconnection strength and the predicted AO signal (measured by the ensemble mean variance), and **c** EFP and the predicted AO signal. The EFP and ENSO teleconnection strength are defined in the Methods section. The value of the EFP computed from ERA5 reanalysis data is shown with a black vertical line in **a** and **c**.

(Fig. 2a), and for composites across systems with strong eddy feedback (Fig. 2b) and weak eddy feedback (Fig. 2c)—see ‘Methods’ section for details. As expected, the signal in the reanalysis is that of a negative AO in response to El Niño, particularly strong in the north Atlantic, projecting onto the NAO, and in the north Pacific (strengthening of the climatological Aleutian cyclone<sup>37</sup>). The same signal is captured in the forecast systems, with both composites showing a deficiency in the magnitude of the signal, but the strong eddy feedback composite being closer to the reanalysis than the weak eddy feedback composite, consistent with that found in earlier-generation climate models<sup>24</sup>. In addition, the difference between the strong and weak composites is a clear AO signal that looks very similar to that in the reanalysis, and is statistically significant (Fig. 2d).

This is highly suggestive of a link between eddy feedback and the strength of the ENSO teleconnection to the AO captured in a forecast system, namely that the strength of the teleconnection increases and becomes more realistic as the strength of the eddy feedback increases and becomes more realistic. To quantify this, the ENSO teleconnection ‘strength’ is defined as the regression coefficient between the normalised Niño 3.4 index and the AO index in a forecast system (see Methods section for details). A strongly positive correlation of 0.69 is found between the ENSO teleconnection strength and the eddy feedback parameter across all the forecast systems<sup>33</sup> (Fig. 3a).

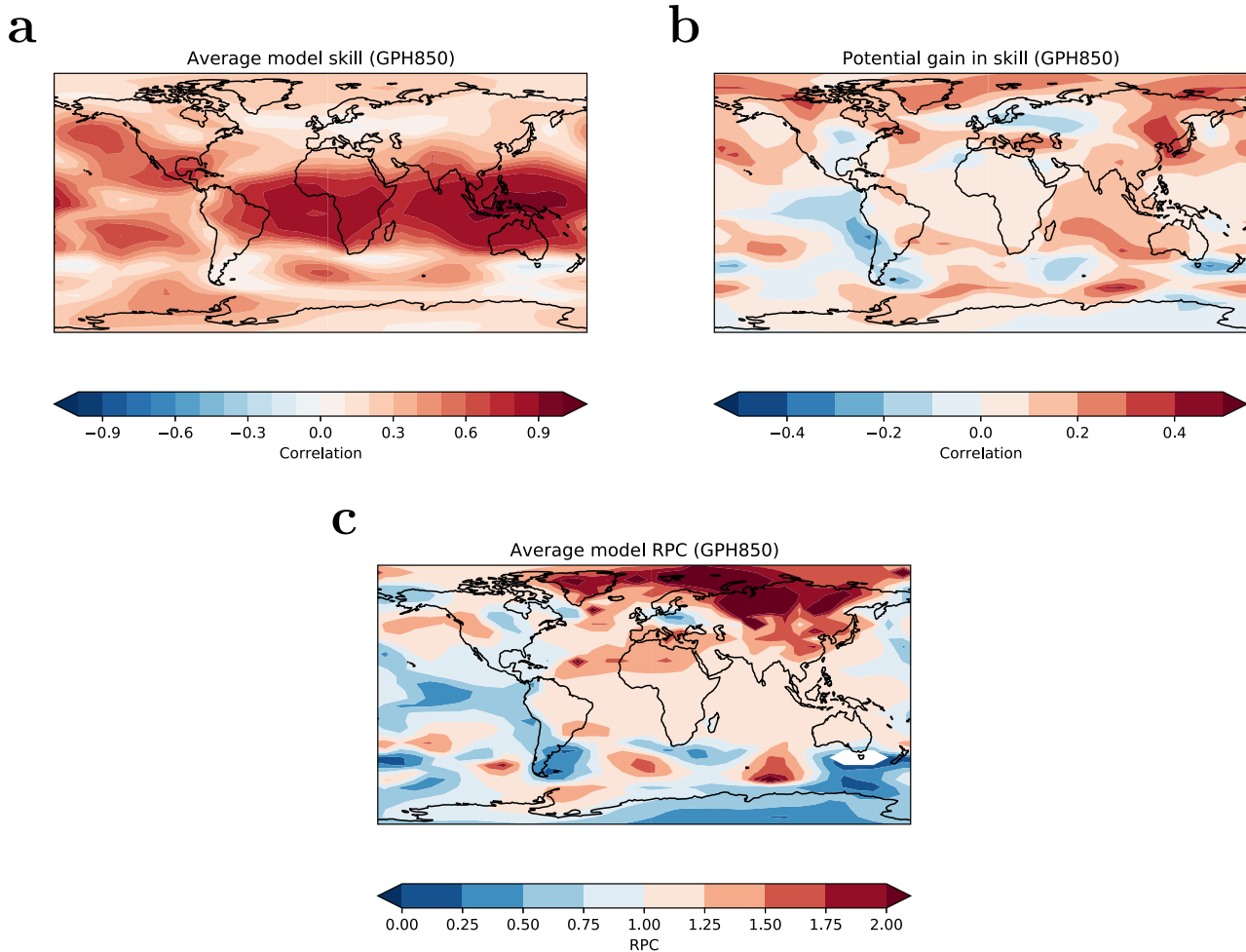
To further investigate the implications of this link, we find a very strong positive correlation between the ENSO teleconnection

strength and the predictable signal, measured by the ensemble mean variance, in the forecast systems of 0.89 (Fig. 3b). Further, we find a significantly positive correlation between eddy feedback and the predictable signal in the forecast systems of 0.49 (Fig. 3c). In summary, our results suggest that increased eddy feedback is related to a stronger ENSO teleconnection which, in turn, leads to stronger simulated predictable signals.

### Potential gain in regional skill

We have shown that increased/more realistic eddy feedback is associated with increased ENSO teleconnection strength, improved signal-to-noise ratio, and increased AO forecast skill. We now consider the potential gain in regional prediction skill if the eddy feedback in forecast systems was at realistic levels.

Taking the AO as an example, we define the current average model skill as the average, across all model systems, of  $r_{mo}$  (i.e. the average of all  $r_{mo}$  values in Fig. 1b). It is also possible to calculate the potential skill that could be achieved by increasing the eddy feedback in forecast systems to equal that in the reanalysis. This is defined as the skill value where the regression line across all systems crosses the ERA5 value of eddy feedback parameter, shown by a hollow black box in Fig. 1b. The potential gain in skill is then simply the potential skill minus the current average model skill. By using  $r_{mo}$  scatter plots similar to Fig. 1b, but defined over smaller regions, this approach is applied more generally to consider the potential gain in regional prediction skill (further details are given in the Methods section).



**Fig. 4 Model skill.** Maps of **a** Average model skill, **b** potential gain in skill by improving EFP to observed values (see Methods section for details) and **c** Ratio of Predictable Components averaged across skilful models (those used in Fig. 1c). In **b**, red colours denote regions where skill could potentially be improved by a more accurate simulation of the eddy feedback parameter. In **c**, red colours denote regions where the signal-to-noise paradox is evident, and white indicates regions where model skill is negative for all models.

The current average model skill is found to be high in the tropics and around the region of the Aleutian low in the extratropics<sup>3</sup> (Fig. 4a). It is generally lower in the extratropics, but significant in the NAO regions (the Azores and Iceland<sup>6,39</sup>). The potential gain in skill that arises from correct eddy feedback is shown in Fig. 4b. In the extratropics this is comparable in magnitude to the current average skill, particularly in the NAO regions and east Asia, suggesting the potential to double the skill of seasonal forecasts in these regions. As might be expected, the regions where there is potentially the greatest skill to be gained by improved eddy feedback correspond closely to those regions where the signal-to-noise error is largest in the skilful forecast systems (Fig. 4c)<sup>1,3</sup>, consistent with improved eddy feedbacks driving larger predictable signals. There are also some regions (for example central North America and continental northern Europe) where improvement in eddy feedback would not lead to any increase in model skill, and in these regions, the model is currently over-confident ( $RPC < 1$ ).

## DISCUSSION

In this study, we have investigated one potential cause of the signal-to-noise paradox—that of a deficiency in eddy feedback in current seasonal forecast systems. We have considered the impacts of improving the accuracy of the eddy feedback in these systems.

We find that increased eddy feedback is strongly linked with a reduced signal-to-noise error, with linear regression suggesting that the error would be completely removed ( $RPC = 1$ ) if the eddy feedback in forecast systems was equal to that in the reanalysis.

Consistent with improving the RPC, we find that increased eddy feedback increases model–model predictability ( $r_{mm}$ ) and model–observed skill ( $r_{mo}$ ). In particular, increased and more accurate eddy feedback is found to increase the strength of the ENSO teleconnection to the AO in forecast systems which, in turn, leads to stronger predictable signals. The potential gain in skill from corrected eddy feedback is considered regionally. It is found that correlation skill scores could roughly double in many extratropical regions, particularly in the region of the NAO, if eddy feedback were of a realistic magnitude. Thus, improving eddy feedback in forecast systems may yield significant improvement in the skill of seasonal forecasts in the extratropics.

The regions where there is potentially the greatest skill to be gained by improved eddy feedback correspond closely to those regions where the signal-to-noise error is largest. This emphasises the point that improving eddy feedback in forecast systems may be important for both improving the signal-to-noise error in these systems, and increasing their skill.

It is important to note that our analysis does not prove cause and effect - we cannot say that imposing the observed eddy feedback in a forecast system, or including a parameterisation of the missing eddy feedback, would cure the signal-to-noise

**Table 2.** Forecast system ensemble members.

Forecast system	Number of ensemble members
DWD	30
CMCC	40
NCEP	24
JMA	10
ECMWF	25
METEO	25
UKMO-D	9
SMPI-D	9
SCNR-D	9
UKMOGC3	28
UKMOGC2	56
UKMOGC2LI	56
DWDSYS2	30
CMCCSYS3	40
METEOSYS6	25
CanCM3	10
CanCM4	10

paradox. For example, if larger predictable signals were achieved by another means, this would impact the simulated jet stream and therefore potentially be realised as increased eddy feedback. Nevertheless, our results are still strongly suggestive of eddy feedback deficiency contributing to the signal-to-noise error.

The eddy feedback parameter used in this study is a seasonal mean, zonal mean, quantity as recently used to study the atmospheric circulation response to climate change<sup>28</sup>. To investigate, in addition, the transient component of the synoptic eddy forcing, and the effects of eddy feedback on the full three-dimensional flow, one might use the eddy-induced growth rate<sup>40,41</sup>. This may help in better understanding the physical mechanisms underlying eddy feedback, and is a topic for future research.

Since it is unlikely that model resolution will increase in the near future to anything like that required to accurately and fully resolve eddy feedback<sup>8</sup>, our results motivate an increased effort to both understand the physical mechanisms underlying eddy feedback and to conceive how to implement imposed eddy feedback or a parameterisation of missing eddy feedback in forecast systems.

## METHODS

### Data

The ECMWF Reanalysis version 5, ERA5<sup>34,35</sup>, is used as the ‘observations’ or ‘truth’ in this study. Seasonal forecast systems are from a variety of sources. The most up-to-date systems from the Copernicus Climate Data Store (C3S) for which all required data is available are the ECMWF SEAS5 system (ECMWF<sup>42</sup>), Météo-France System 7 (METEO<sup>43</sup>), the Centro Euro-Mediterraneo sui Cambiamenti Climatici SPS3.5 (CMCC<sup>44</sup>), the Deutscher Wetterdienst GCF5.2.1 (DWD<sup>45</sup>), the National Centers for Environmental Prediction CFS version 2 (NCEP<sup>46</sup>), the Japan Meteorological Agency CPS2 (JMA<sup>47</sup>), and the land-initialised versions of the UK Met Office Global Seasonal Forecast system version 5 (UKMOGC2LI<sup>48</sup>). These are shown as circles on all scatter plots in this study. Older seasonal forecast system versions also obtained from C3S are DWD GCF5.2.0 (DWDSYS2<sup>45</sup>), CMCC SPS3 (CMCCSYS3<sup>49</sup>), METEO System 6 (METEOSYS6<sup>50</sup>), and climatological-land versions of UKMO GloSea5 (UKMOGC2<sup>48</sup>), denoted by squares on the scatter plots. There are three systems from the DEMETER ensembles<sup>51</sup> for which the required data is available, denoted UKMO-D, SMPI-D, and SCNR-D. These are denoted by crosses on the scatter plots. There are two systems from the North American Multi-Model Ensemble (NMME) for which data was obtainable, namely the CCCma climate model version 3 (CanCM3<sup>52</sup>) and version 4

(CanCM4<sup>53</sup>), denoted by triangles on the scatter plots. Finally, the new Met Office Global Seasonal Forecast system version 6, also available on C3S (UKMOGC3<sup>48</sup>), is added as a hexagon on the scatter plots. The number of ensemble members available from each system is listed in Table 2.

For all of these systems, forecasts start at the beginning of November. Daily zonal wind ( $u$ ) and meridional wind ( $v$ ) at 500 hPa, monthly mean GPH at 500 hPa and 850 hPa, and monthly MSLP is used for all winter months (December–January–February), a forecast lead time of 2–4 months. The years 1993–2016 are available for C3S systems (23 winters), 1970–2001 for DEMETER systems (31 winters), and 1981–2012 for NMME systems (31 winters). For each system, the corresponding period is used from ERA5.

### Eddy feedback parameter

Eddy feedback, the feedback between small-scale transient eddies and large-scale quasi-stationary climate anomalies, can be quantified as the link between resolved wave driving and the zonal mean wind in climate models. Here, the eddy feedback parameter is defined as follows. Using daily  $u$  and  $v$  data at 500 hPa, the zonal acceleration due to the quasi-geostrophic component of the horizontal EP-flux divergence<sup>54</sup> is computed:

$$\frac{\nabla \cdot \mathbf{F}_H}{\rho a \cos \phi} = \frac{-1}{a \cos^2 \phi} \frac{d(\overline{u'v'} \cos^2 \phi)}{d\phi} \quad (1)$$

where  $\rho$  is density,  $\phi$  is latitude,  $a$  is the radius of the Earth, overbar represents a zonal mean, and  $'$  represents the residual after removing the zonal mean. The December–January–February mean of this zonal acceleration is then formed from the daily values, for each year. Next, the December–January–February mean of the zonal mean zonal wind,  $\bar{u}$  is computed for each year. Then the correlation at each latitude, across years, between the zonal acceleration and  $\bar{u}$  is calculated. The area-weighted average, 25°N–72°N, of this correlation squared is the eddy feedback parameter.

This definition is identical to that used previously<sup>28</sup>, except simplified in two ways due to the limited amount of data available for some systems. Firstly, we only use data at 500 hPa, as opposed to a vertical integral between 600 hPa and 200 hPa. Data at 500 hPa is available for all systems, and the signal has been found to be coherent across the different vertical levels<sup>28</sup>. Secondly, only the quasi-geostrophic component of the horizontal EP-flux divergence is included. However, the other components of the horizontal EP-flux divergence are small. Thus, these two simplifications make the calculations possible across many more systems, and should not impact the results.

### Predictable components

The ratio of predictable components is defined as

$$RPC = \frac{r_{mo}}{r_{mm}} \quad (2)$$

where  $r_{mo}$ , the skill of the model in predicting the observations, is defined for each system as the correlation across years between ERA5 quantities and those from the system ensemble mean<sup>1,2</sup>. The definition of  $r_{mm}$ , the model–model skill, has been previously given as<sup>1</sup>

$$r_{mm} = \sqrt{\frac{\sigma_{ens\ mean}^2}{\sigma_{total}^2}} \quad (3)$$

where  $\sigma_{ens\ mean}^2$  is the variance of the ensemble mean and  $\sigma_{total}^2$  is the total variance of individual ensemble members. It has also been defined as the average of the correlation of a single ensemble member with the mean of all other ensemble members<sup>2</sup>.

Here we use an improved calculation of  $r_{mm}$  that is less sensitive to ensemble size. Equation (3) will be an over-estimate for  $r_{mm}$  when the number of ensemble members used is small<sup>1</sup>. Further, the average ensemble member correlation<sup>2</sup> is found to be an under-estimate when the number of ensemble members used is small. Therefore,  $r_{mm}$  is here defined as follows.

Starting from equations for the total variance and the ensemble mean variance<sup>55</sup>:

$$\sigma_{\text{total}}^2 = \sigma_{\text{signal}}^2 + \sigma_{\text{noise}}^2 \quad (4)$$

$$\sigma_{\text{ens mean}}^2 = \sigma_{\text{signal}}^2 + \frac{\sigma_{\text{noise}}^2}{N} \quad (5)$$

combine and rearrange to obtain

$$\sigma_{\text{signal}}^2 = \frac{N\sigma_{\text{ens mean}}^2 - \sigma_{\text{total}}^2}{N - 1} \quad (6)$$

Then  $r_{\text{mm}} = \sigma_{\text{signal}}/\sigma_{\text{total}}$  such that

$$r_{\text{mm}} = \sqrt{\frac{N(\sigma_{\text{ens mean}}^2/\sigma_{\text{total}}^2) - 1}{N - 1}} \quad (7)$$

The total variance is computed, across years, by using the individual ensemble members and bootstrapping using 1,000,000 random samples with replacement. The ensemble mean variance is computed with a single calculation. The equation for  $r_{\text{mm}}$  used here takes a value in between the estimates for equation (3) and the average ensemble member correlation, and is found to be stable regardless of the number of ensemble members (not shown).

## AO and NAO indices

The AO index is defined as the zonal mean area-weighted 30°N–60°N mean minus the zonal mean area-weighted 60°N–90°N mean<sup>48</sup>, and computed using GPH at 500 hPa and 850 hPa, and MSLP. The NAO index is defined as an area-weighted mean centred around the Azores (28.5–20°W, 36–40°N) minus an area-weighted mean centred around Iceland (25–16.5°W, 63.5–70°N)<sup>56</sup>, and computed using MSLP.

## ENSO teleconnection

The El Niño Southern Oscillation (ENSO<sup>57</sup>) index used in this study is the December–January–February mean of sea surface temperature anomalies in the Niño 3.4 region (5°N–5°S, 120–170°W) obtained from [https://origin.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ONI\\_v5.php](https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php). The regression coefficient, across years, of this normalised Niño 3.4 index with the AO index (calculated using GPH 850 hPa) is used to define the strength of the ENSO teleconnection. For each system, the ensemble mean AO index is regressed against the Niño 3.4 index to give the value of the teleconnection strength (in metres) shown in Fig. 3a and b.

For the maps in Fig. 2, we first divide the systems into three roughly equal-sized groups: those with high values of the eddy feedback parameter (DWD, CMCC, DWDSYS2, CMCCSYS3, and CanCM4)—a ‘Strong’ model composite, those with low values of the eddy feedback parameter (NCEP, SMPI-D, SCNR-D, METEOSYS6, and CanCM3)—a ‘Weak’ model composite, and the remaining seven form a ‘neutral’ composite (although this composite is not used). Next, we define years for which there were El Niño events and La Niña events. In order to have roughly equal numbers of events, we here define the threshold such that an event is said to have occurred if the magnitude of the Niño 3.4 index is greater than 1.5K. Using ERA5, and each system individually, we form El Niño minus La Niña differences of the GPH 500 hPa field. In the case of the models, these are then averaged to form the composite maps shown in Fig. 2b and c. In the case of the observations, El Niño minus La Niña differences are formed using ERA5 for each of the three model periods (1993–2016, 1970–2001, and 1981–2012) with the average of these differences shown in Fig. 2a. The result is found to be insensitive to the exact period chosen.

## Potential skill and RPC maps

In order to compute the skill maps in Fig. 4, the following method is used. Using the GPH 850 hPa field, a scatter plot, equivalent to that shown in Fig. 1b, is computed for each area-weighted 10° longitude × 5° latitude region of the globe (a total of 36 × 36 = 1296 plots). For each plot, the average model skill (shown in Fig. 4a) is the straight average, across all model systems, of  $r_{\text{mo}}$ . The potential skill is the value of  $r_{\text{mo}}$  where the linear least-squares best fit line crosses the ERA5 value of eddy feedback parameter (equivalent to that shown by a hollow black box in Fig. 1b). This makes the assumption that the relationship between eddy feedback and skill remains linear at least up to the observed value of eddy feedback. The potential gain in skill (shown in Fig. 4b) is then the potential skill minus the average model skill, in each region.

RPC maps are computed by dividing average model–observed skill ( $r_{\text{mo}}$ , shown in Fig. 4a) by average model–model skill ( $r_{\text{mm}}$ , computed in the same way) for each region for each model. The RPC values in each region are then averaged across all skilful models (those included in Fig. 1c) to produce the RPC map shown in Fig. 4c.

## DATA AVAILABILITY

ERA5 data and data for all C3S systems are available from the Copernicus Climate Data Store, <https://cds.climate.copernicus.eu/>. Data for the DEMETER systems can be obtained from <https://www.ecmwf.int/en/forecasts/dataset/development-european-multimodel-ensemble-system-seasonal-interannual-prediction>, and data for the NMME systems can be obtained from <https://www.cpc.ncep.noaa.gov/products/NMME/>.

## CODE AVAILABILITY

For Python code used to produce all graphs in this study, please see <https://zenodo.org/record/6472589>

Received: 11 March 2022; Accepted: 22 June 2022;

Published online: 11 July 2022

## REFERENCES

- Eade, R. et al. Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.* **41**, 5620–5628 (2014).
- Scaife, A. A. & Smith, D. A signal-to-noise paradox in climate science. *npj Clim. Atmos. Sci.* **1**, 28 (2018).
- Baker, L. H., Shaffrey, L. C., Sutton, R. T., Weisheimer, A. & Scaife, A. A. An inter-comparison of skill and overconfidence/underconfidence of the wintertime North Atlantic Oscillation in multimodel seasonal forecasts. *Geophys. Res. Lett.* **45**, 7808–7817 (2018).
- Smith, D. M. et al. Robust skill of decadal climate predictions. *npj Clim. Atmos. Sci.* **2**, 13 (2019).
- Klavans, J. M., Cane, M. A., Clement, A. C. & Murphy, L. N. NAO predictability from external forcing in the late 20th century. *npj Clim. Atmos. Sci.* **4**, 22 (2021).
- Scaife, A. A. et al. Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.* **41**, 2514–2519 (2014).
- Smith, D. M. et al. North Atlantic climate far more predictable than models imply. *Nature* **583**, 796–800 (2020).
- Scaife, A. A. et al. Does increased atmospheric resolution improve seasonal climate predictions? *Atmos. Sci. Lett.* **20**, e922 (2019).
- Ossó, A., Sutton, R., Shaffrey, L. & Dong, B. Development, amplification, and decay of Atlantic/European summer weather patterns linked to spring north Atlantic sea surface temperatures. *J. Climate* **33**, 5939–5951 (2020).
- Zhang, W., Kirtman, B., Siqueira, L., Clement, A. & Xia, J. Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model. *Clim. Dyn.* **56**, 2895–2913 (2021).
- Haarsma, R. J. et al. Sensitivity of winter North Atlantic-European climate to resolved atmosphere and ocean dynamics. *Sci. Rep.* **9**, 13358 (2019).
- Wicker, W., Greatbatch, R. J. & Claus, M. Sensitivity of a simple atmospheric model to changing surface friction with implications for seasonal prediction. *Q. J. R. Meteorol. Soc.* **148**, 199–213 (2021).
- Strommen, K. & Palmer, T. N. Signal and noise in regime systems: a hypothesis on the predictability of the North Atlantic Oscillation. *Q. J. R. Meteorol. Soc.* **145**, 147–163 (2019).
- Strommen, K. Jet latitude regimes and the predictability of the North Atlantic Oscillation. *Q. J. R. Meteorol. Soc.* **146**, 2368–2391 (2019).
- Nie, Y., Zhang, Y., Chen, G. & Yang, X.-Q. Delineating the barotropic and baroclinic mechanisms in the midlatitude eddy-driven jet response to lower-tropospheric thermal forcing. *J. Atmos. Sci.* **73**, 429–448 (2016).
- Lau, N.-C. Variability of the observed midlatitude storm tracks in relation to low-frequency changes in the circulation pattern. *J. Atmos. Sci.* **45**, 2718–2743 (1988).
- Robinson, W. A. The dynamics of low-frequency variability in a simple model of the global atmosphere. *J. Atmos. Sci.* **48**, 429–441 (1991).
- Held, I. M. & Philipps, P. J. Sensitivity of the eddy momentum flux to meridional resolution in atmospheric GCMs. *J. Clim.* **6**, 499–507 (1993).
- Feldstein, S. & Lee, S. Mechanisms of zonal index variability in an aquaplanet GCM. *J. Atmos. Sci.* **53**, 3541–3556 (1996).
- Feldstein, S. B. The dynamics of NAO teleconnection pattern growth and decay. *Q. J. R. Meteorol. Soc.* **129**, 901–924 (2003).

21. Lorenz, D. J. & Hartmann, D. L. Eddy-zonal flow feedback in the northern hemisphere winter. *J. Clim.* **16**, 1212–1227 (2003).
22. Kug, J.-S. & Jin, F.-F. Left-hand rule for synoptic eddy feedback on low-frequency flow. *Geophys. Res. Lett.* **36**, L05709 (2009).
23. Barnes, E. & Hartmann, D. Dynamical Feedbacks and the Persistence of the NAO. *J. Atmos. Sci.* **67**, 851–865 (2010).
24. Kang, I.-S., Kug, J.-S., Lim, M.-J. & Choi, D.-H. Impact of transient eddies on extratropical seasonal-mean predictability in DEMETER models. *Clim. Dyn.* **37**, 509–519 (2011).
25. Kidston, J. et al. Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nat. Geosci.* **8**, 433–440 (2015).
26. Hitchcock, P. & Simpson, I. R. Quantifying eddy feedbacks and forcings in the tropospheric response to stratospheric sudden warmings. *J. Atmos. Sci.* **73**, 3641–3657 (2016).
27. Ren, H.-L., Zhou, F., Nie, Y. & Zhao, S. Dynamic synoptic eddy feedbacks contributing to maintenance and propagation of intraseasonal NAO. *Geophys. Res. Lett.* **49**, e2021GL096508 (2022).
28. Smith, D. et al. Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nat. Commun.* **13**, 727 (2022).
29. Robinson, W. A. On the self-maintenance of midlatitude jets. *J. Atmos. Sci.* **63**, 2109–2122 (2006).
30. Gerber, E. P. & Vallis, G. K. Eddy-zonal flow interactions and the persistence of the zonal index. *J. Atmos. Sci.* **64**, 3296–3311 (2007).
31. Gray, S. L., Dunning, C. M., Methven, J., Masato, G. & Chagnon, J. M. Systematic model forecast error in Rossby wave structure. *Geophys. Res. Lett.* **41**, 2979–2987 (2014).
32. Held, I. M., Lyons, S. W. & Nigam, S. Transients and the extratropical response to El Niño. *J. Atmos. Sci.* **46**, 163–174 (1989).
33. L'Heureux, M. L. et al. Strong relations between ENSO and the Arctic Oscillation in the North American Multimodel Ensemble. *Geophys. Res. Lett.* **44**(11), 654–11,662 (2017).
34. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
35. Bell, B. et al. The ERA5 global reanalysis: preliminary extension to 1950. *Q. J. R. Meteorol. Soc.* **147**, 4186–4227 (2021).
36. Hoerling, M. P. & Ting, M. Organization of extratropical transients during El Niño. *J. Climate* **7**, 745–766 (1994).
37. Ineson, S. & Scaife, A. A. The role of the stratosphere in the European climate response to El Niño. *Nat. Geosci.* **2**, 32–36 (2009).
38. Butler, A. H. & LM.Polvani, E. Niño, La Niña, and stratospheric sudden warmings: a reevaluation in light of the observational record. *Geophys. Res. Lett.* **38**, L13807 (2011).
39. Thornton, H. E. et al. Skilful seasonal prediction of winter gas demand. *Environ. Res. Lett.* **14**, 024009 (2019).
40. Ren, H.-L., Jin, F.-F., Kug, J.-S. & Gao, L. Transformed eddy-PV flux and positive synoptic eddy feedback onto low-frequency flow. *Clim. Dyn.* **36**, 2357–2370 (2011).
41. Ren, H.-L., Jin, F.-F. & Kug, J.-S. Eddy-induced growth rate of low-frequency variability and its mid-to late winter suppression in the northern hemisphere. *J. Atmos. Sci.* **71**, 2281–2298 (2014).
42. Johnson, S. J. et al. SEAS5: the new ECMWF seasonal forecast system. *Geosci. Model Dev.* **12**, 1087–1117 (2019).
43. Batté, L., L. Dorel, C. Ardilouze, and J.-Fo. Guérémy. Documentation of the METEO-FRANCE seasonal forecasting system 7. Meteo France. <http://www.umr-cnrm.fr/IMG/pdf/system7-technical.pdf> (2019)
44. Gualdi, S. et al. The new CMCC Operational Seasonal Prediction System. In: Centro Euro-Mediterraneo sui Cambiamenti Climatici (TN0288) (2020).
45. Fröhlich, K. et al. The German climate forecast system: GCF5. *J. Adv. Model Earth Syst.* **13**, e2020MS002101 (2021).
46. Saha, S. et al. The NCEP climate forecast system version 2. *J. Clim.* **27**, 2185–2208 (2014).
47. Takaya, Y. et al. Japan Meteorological Agency/Meteorological Research Institute-Coupled Prediction System version 2 (JMA/MRI-CPS2): atmosphere-land-ocean-sea ice coupled prediction system for operational seasonal forecasting. *Clim. Dyn.* **50**, 751–765 (2018).
48. MacLachlan, C. et al. Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Q. J. R. Meteorol. Soc.* **141**, 1072–1084 (2015).
49. Sanna, A. et al. CMCC-SPS3: the CMCC seasonal prediction system 3. In: Centro Euro-Mediterraneo sui Cambiamenti Climatici (RP0285) (2017).
50. Dorel, L., C. Ardilouze, M. Déqué, L. Batté, and J.-Fo. Guérémy. Documentation of the Meteo-france pre-operational seasonal forecasting system. Meteo France. <http://www.umr-cnrm.fr/IMG/pdf/system6-technical.pdf> (2017).
51. Palmer, T. N. et al. Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). *BAMS* **85**, 853–872 (2004).
52. Merryfield, W. J. et al. The Canadian Seasonal to Interannual Prediction System (CanSIPS): an overview of its design and operational implementation. Tech. Note, Environment Canada, 51 pp (2011).
53. Merryfield, W. J. et al. The Canadian seasonal to interannual prediction system. Part I: Models and initialization. *Mon. Wea. Rev.* **141**, 2910–2945 (2013).
54. Andrews, D. G., J. R. Holton, and C. B. Leovy. Middle Atmosphere Dynamics. p. 489 (Academic Press, Inc.1987).
55. Siebert, S. et al. A Bayesian framework for verification and recalibration of ensemble forecasts: how uncertain is NAO predictability? *J. Clim.* **29**, 995–1012 (2016).
56. Dunstone, N. J. et al. Skilful predictions of the winter North Atlantic oscillation one year ahead. *Nat. Geosci.* **9**, 809–814 (2016).
57. Trenberth, K. E. The definition of El Niño. *BAMS* **78**, 2771–2778 (1997).

## ACKNOWLEDGEMENTS

The work of S.C.H., N.J.D., A.A.S., Y.N., and H.-L.R. was supported by the UK-China Research and Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund. H.-L.R. is also sponsored by the National Natural Science Foundation of China (41775066). The work of R.C. was supported by the UK Public Weather Service research programme. The work of D.M.S. was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra, and by the European Commission Horizon 2020 EUCP project (GA 776613).

## AUTHOR CONTRIBUTIONS

S.C.H. performed the analysis and wrote the manuscript, S.C.H., N.J.D., A.A.S., and D.M.S. contributed to the development of the method and the interpretation of the results, R.C. developed the method for ensemble member independent  $r_{mmv}$ , A.A.S. suggested the method to infer the gain in skill when eddy feedback is increased to observed levels, and Y.N. and H.-L.R. provided useful comments on the manuscript, including points to emphasise, and references to the existing literature.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Steven C. Hardiman.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2022