

Meeting Report: The Signal-to-Noise Paradox in Climate Forecasts:

Revisiting our Understanding and Identifying Future Priorities

Antje Weisheimer^{1,2*}, Laura H. Baker³, Jochen Bröcker^{4,5}, Chaim I. Garfinkel⁶, Steven C. Hardiman⁷, Dan L.R. Hodson³, Tim N. Palmer⁸, Jon I. Robson³, Adam A. Scaife⁷, James A. Screen⁹, Theodore G. Shepherd⁴, Doug M. Smith⁷ and Rowan T. Sutton³

¹ National Centre for Atmospheric Science (NCAS), University of Oxford, Physics Department, AOPP, UK

² European Centre for Medium-Range Weather Prediction (ECMWF), Reading, UK

³ National Centre for Atmospheric Science (NCAS), University of Reading, UK

⁴ Department of Meteorology, University of Reading, UK

⁵ Department of Mathematics and Statistics, University of Reading, UK

⁶ Earth Science Institute, Hebrew University, Jerusalem, Israel

⁷ United Kingdom Met Office, Exeter, UK

⁸ University of Oxford, Physics Department, AOPP, UK

⁹ Department of Mathematics and Statistics, University of Exeter, UK

* Corresponding author, email: Antje.Weisheimer@physics.ox.ac.uk

Info box:

Workshop on the Signal-to-Noise Paradox in Climate Forecasts

What: Nearly 40 participants from universities (Oxford, Reading, Exeter, Leeds, Lincoln), the National Centre for Atmospheric Science (NCAS) and operational forecasting centers (Met Office, ECMWF) in the UK as well as from the Hebrew University in Jerusalem and NCAR in Boulder gathered - primarily in person - to discuss our current understanding of the so-called signal-to-noise paradox in climate forecasts and develop ideas to resolve it.

Where: University of Oxford, UK

1

Early Online Release: This preliminary version has been accepted for publication in *Bulletin of the American Meteorological Society*, may be fully cited, and has been assigned DOI 10.1175/BAMS-D-24-0019.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

© 2024 The Author(s). Published by the American Meteorological Society. This is an Author Accepted Manuscript distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License



When: 25-26 September 2023

Background

The signal-to-noise paradox (SNP) in model-based climate forecasting refers to counterintuitive situations where time series of ensemble-mean forecasts correlate better with observations of the real world than with individual members of the model forecast ensemble. It implies that the predictability of the real world exceeds the predictability within the model world.

The expected correlation of observations with a forecast ensemble mean is related to the signal-to-noise ratio of the forecasting system in a monotonic but nonlinear way (*Kumar, 2009*). Here, “signal” refers to the temporal variability of the ensemble mean, and “noise” to the variability of the ensemble members about the ensemble mean. The SNP occurs when the correlation between the ensemble-mean forecast and observations is greater than expected given the signal-to-noise ratio of the forecasting system.

The SNP is often quantified with the Ratio of Predictable Components (RPC) between the real and model world. The predictable component of the observations is estimated from the correlation coefficient between the ensemble-mean signal and the observations, and the (squared) predictable component of the model is estimated from the fraction of the signal variance to the total model variance. This latter fraction is identical to the (squared) correlation coefficient between the ensemble-mean signal and the individual ensemble members. If the RPC is significantly larger than 1, then the observations are more predictable than the model ensemble realizations which constitutes the SNP.

Evidence supporting the existence of the SNP in seasonal and decadal climate predictions was first presented by *Scaife et al. (2014)* and *Eade et al. (2014)*. They described instances where the predictable components of the winter atmospheric circulation over the North Atlantic were sometimes lower in models than in observations. Despite numerous studies exploring different facets of the SNP since the comprehensive review by *Scaife & Smith (2018)*, a conclusive solution to the problem has not yet been reached. The Oxford workshop provided a dedicated platform not only to present our current understanding to an expert audience

but, more importantly, to critically discuss gaps and problems in our state of knowledge. The primary objective of the workshop was to achieve a better and more complete comprehension of the paradox, as well as to identify suggestions towards its resolution. During the workshop it became clear that our community, including the authors of this report, hold a range of perspectives on the problem, and our meeting report reflects this diversity of ideas and evidence presented at the meeting.

Shaping the problem

The statistical interpretation

The correlation between the ensemble mean and observations can be visualized in a simple trigonometric framework with the three sides of the triangle corresponding to the variances of the ensemble mean, the observation and the forecast error (i.e. the mean squared difference between the ensemble mean and the observation, *Bröcker et al., 2023*). In this way, the behavior of the RPC, and possible emergence of the SNP, can be studied as a function of fairly primitive statistics. From the several conclusions emerging, the sensitive dependence of the RPC on these statistics, particularly if the model signal is small, is one of the most noticeable. Consequently, in this situation it is easy to misdiagnose a forecast as suffering from the SNP, or to miss its presence.

Formally, a small signal-to-noise ratio can be either due to a deficit in the signal, to excessive noise, or to a combination of both. In the statistical interpretation, the SNP tends to manifest itself in a too weak predicted ensemble-mean signal compared with the observed signal (*Siebert et al., 2016; Falkena et al., 2022; Williams et al., 2023*).

A non-linear perspective in terms of circulation regimes

A complementary interpretation of the SNP using dominant and persistent circulation regimes has been put forward by *Strommen & Palmer (2019)* and *Zhang & Kirtman (2019)*. A nonlinear circulation regime can be likened to a potential well in state space. Here, the signal can be understood as the regime centroid and the noise as the width of the regime. In general, models predict regime centroids and the ensemble-mean signal of fluctuating between these preferred regimes reasonably well. However, the simulated regimes are not as statistically significant as in observational data, implying that the potential wells associated with these regimes are too shallow and broad compared with the deeper and sharper regimes of the real

world. Such blurred regimes would be characterized by weak persistence within the regime centroids. The resulting signal-to-noise ratios from realistic signals of regime centroids but too strong noise components of shallow regimes will thus be small.

Detection of the paradox

Following the earlier findings, the number of studies commenting on the signal and noise characteristics has been rapidly growing. While we cannot mention all of them here, new evidence for, but also against the existence of the SNP in climate predictions was presented at the workshop.

As part of a skill comparison of the winter North Atlantic Oscillation (NAO) across several seasonal forecast models, *Baker et al. (2018)* demonstrated that most but not all individual systems had $RPC > 1$. Consistently with *Stockdale et al. (2015)*, *Charlton-Perez et al. (2019)* found that the low signal-to-noise ratio is predominantly a feature of the lower and middle troposphere linked to weak signal amplitudes in early winter and is not present in the stratosphere. *Smith et al. (2020)* in a very large ensemble of decadal predictions of the NAO showed high correlation skill but very low amplitudes of the ensemble-mean signals, resulting in a high RPC. While the SNP was initially identified for the winter season, new studies have shown evidence that it also occurs for summer precipitation over Northern Europe and the Tibetan Plateau (*Dunstone et al., 2018; Yeager et al., 2018; Hu & Zhou, 2021*). More recently, *Dunstone et al. (2023)* diagnosed the SNP in predictions of the summer NAO with spurious weak dynamical signals in the model.

It is important to contrast and learn from situations where the SNP is not apparent. In seasonal hindcasts for recent decades, *Falkena et al. (2022)* found that whilst the system exhibited a signal deficit in the winter NAO index of about a factor of 2, the zonal flow regimes - both the positive NAO and a negative counterpart of Scandinavian blocking - showed no signal deficit at all. The deficit for the NAO index thus resulted primarily from the negative NAO (blocked) regime, where the predictable signal in the model did not align with the observations.

Applying the concept of predictable modes of climate variability to seasonal forecasts, *Hodson et al. (2023)* found that the SNP does not affect all predictable modes over the North Atlantic region equally. The amplitude of the NAO-like mode, which is influenced by Indian

Ocean temperatures, is significantly underestimated. In contrast, the Pacific North America (PNA)-like mode, which is strongly related to tropical Pacific variability, shows less of an underestimation.

Seasonal forecasts of the Southern Annular Mode (SAM) during late spring were analyzed in *Seviour et al. (2014)* and *Byrne et al. (2019)*. Both studies did not find a significant SNP, but the ensemble-mean correlation with observations was twice that expected, raising the possibility that a southern hemisphere SNP may be present in that system.

Whether the SNP already emerges in sub-seasonal predictions is yet unclear. Preliminary results presented at the workshop showed a possible weak problem for the NAO at timescales of approx. 30 days but its disappearance at longer lead times is indication of a lack of robustness.

The combination of multiple single-model forecast ensembles into a multi-model ensemble has often been used as a pragmatic “ensemble of opportunity” to overcome problems in single-model systems. Its success in providing more skillful and reliable forecasts relies on model error compensation and reduced overconfidence (*Palmer et al., 2004*). However, diagnosing a SNP is problematic in a multi-model ensemble because the ensemble members are not designed to represent model uncertainty in a coherent theoretical way and will thus primarily reflect the diversity of the individual modelling system deficiencies with respect to signals and noise. Notwithstanding, multi-model ensembles are widely used for climate predictions and weak predictable signals need to be calibrated to obtain the most accurate and reliable forecasts (*Smith et al., 2022*).

Hypotheses for mechanisms contributing to the SNP

Non-stationarity and intermittency characteristics

Before we discuss physical and dynamical mechanisms that may lie behind the SNP, it is instructive to reflect on the non-stationarity and intermittent character of the paradox. In the context of coupled and uncoupled seasonal hindcasts covering a 110-year period, it was shown that the predictable components undergo low-frequency variations, linked to significant multi-decadal skill variations (*Weisheimer et al., 2019, 2020*). While the North Atlantic circulation in winter showed regions with $RPC > 1$ towards the end of the 20th century, the early and particularly the middle decades of the century did not. Considering the

substantial uncertainty in estimating the RPC for the winter NAO index, robust periods of values significantly above 1 were not detected. Any mechanism to explain the SNP for the recent past will thus also have to explain the absence of the paradox for earlier historical decades.

A closer look at the contributions from individual years (or seasons) to the covariances that form the basis for the correlation and RPC analyses revealed that the interannual variability is very large, leading to pronounced intermittency in the contributions of individual years to the overall estimate of the SNP (*Weisheimer et al., 2019*). The signal-to-noise problems of the winter NAO are dominated by a few years with larger amplitudes of the observed and modelled anomalies, while most of the years contribute very little to the SNP estimates. For the winter NAO, four of the most contributing 5 years were characterized by negative NAO flow patterns over the North Atlantic and one year was the strongest positive NAO winter during the 20th century. These findings imply that concentrating efforts on specific key extreme years or seasons to understand the paradox should be possible.

Weak teleconnections and signal deficits

Teleconnections between the tropics and the extratropics are the backbone mechanism for predictable signals in the tropics to influence remote regions in the extratropics. However, the strength of these teleconnections appears to be systematically underestimated in sub-seasonal and seasonal forecast models (*Garfinkel et al., 2022; Hardiman et al., 2022; Di Capua et al., 2023; Molteni & Brookshaw, 2023; Roberts et al., 2023; Williams et al., 2023*). While some of this bias may be due to sea surface temperature (SST) biases (*Beverley et al., 2023*), the signal deficit because of too weak teleconnection strength is also evident with prescribed observed SST or before the SST biases develop (*Chen et al., 2020; Garfinkel et al., 2022*).

The stratosphere provides a pathway for tropical signal propagation to the extratropical troposphere and the Quasi-Biennial Oscillation (QBO) has been identified as another likely source of the SNP for the NAO (*Garfinkel et al., 2018*). With the QBO-NAO teleconnections being too weak in seasonal forecasts, diagnostically amplifying the teleconnection increased the ensemble-mean NAO signal and resolved the signal-to-noise problem (*O'Reilly et al., 2019*).

In the Euro-Atlantic sector during winter, the predictable seasonal signal seems largely associated with El Niño Southern Oscillation (ENSO) events, hence with just a small number of years. The ENSO signal in circulation regimes can be enhanced using Bayesian regularization and is evident even in 10-member ensembles (*Falkena et al., 2023*). As noted earlier, the model signal matches that of the observations for the zonal flow regimes, but not for the blocked flow regimes. Regressing the observations against the ensemble-mean forecast provides a simple and robust measure of whether there is a signal deficit, which does not rely on being able to estimate the noise in the observations (*Falkena et al., 2022*).

Transient eddy feedback

Eddy feedback describes the interaction of small-scale transient eddies with large-scale climate anomalies in the mid-latitudes. It is an essential part of the process that governs the intensity and persistence of North Atlantic weather patterns (*Zhao et al., 2023*). It has been hypothesized that for climate models to simulate the correct amplitude of the tropospheric response to remote influences, they need to correctly simulate eddy feedback. One example of this is the jet stream response to Arctic sea-ice loss (*Screen et al., 2022*). Eddy feedback is too weak in almost all current climate models over the North Atlantic (*Smith et al., 2022*). It has recently been argued that increased and more accurate eddy feedback is linked with stronger predictable signals, increased skill, and a reduced signal-to-noise error in seasonal forecast systems of the northern extratropics (*Hardiman et al., 2022*).

There are mounting indications though that this bias is less pronounced in the Southern Hemisphere (e.g., *Screen et al., 2022*). The reasons for such a hemispheric asymmetry in the eddy feedback strength are yet to be understood, with the role of topography and stationary waves being plausible candidates.

Air-sea coupling

Inadequate representation of air-sea interactions in current climate models has also been proposed as a potential cause or contributor to the SNP. *Ossso et al. (2020)* presented evidence that two-way coupling between the North Atlantic Ocean and atmosphere in summer was not well simulated in a coupled climate model, which implied weaker predictable signals in the model. *Zhang et al. (2021)* reported that decadal predictability is underestimated in CMIP5 models and suggested that ocean-atmosphere coupling in ocean-eddy rich regions

(such as the Gulf Stream) could be an important factor. *Patrizio et al. (2023)* recently described further evidence supporting this hypothesis for boreal winter in the North Atlantic.

Model resolution as a plausible fundamental underlying problem

To a large extent, the model errors in the representation of the dynamical and physical mechanisms likely contributing to the SNP can be seen as direct consequences of coarse model resolution, both in the atmosphere and in the ocean. The strength of the eddy-feedback is sensitive to resolution and lower resolution models have been shown to significantly underestimate the strength of this feedback (*Scaife et al., 2019*). New results recently illustrated that increasing the horizontal model resolution substantially improved the performance including skill and the signal-to-noise characteristics of decadal predictions (*Yeager et al., 2023*). It is currently less clear which specific aspects of increased resolution contributed to this overall improvement, and a direct comparison with other systems is lacking.

It is plausible that problems with circulation regime structures, local air-sea interactions and eddy feedback could be solved with enough resolution. However, the model resolutions that may be required to achieve significant increases in model fidelity are not currently computationally affordable for global seasonal and decadal prediction systems.

Take-away conclusions

The discussions between the participants during and after the workshop revealed that our community is lacking a commonly agreed approach to the puzzle that the SNP represents. Nonetheless, the following key points of consensus have been identified:

- *The SNP is neither a simple-to-understand nor a simple-to-fix problem.* Understanding and resolving the paradox is a challenge whose cause defies easy comprehension and straightforward solutions. Although there has been progress in unravelling the intricacies of the paradox, it is evident that the conundrum remains unsolved, even after a decade-long debate within our community. A variety of approaches is needed.
- *The commonly-employed RPC diagnostic requires careful interpretation.* The RPC is highly sensitive to the relative magnitudes of the variances of the observations, the ensemble-mean forecast and the error of the ensemble mean. When coupled with weak correlations and short observed records, this results in high sampling

uncertainties that can significantly impact the RPC estimates and compromise their robustness. Alternative diagnostics should therefore be explored. For example, the regression coefficient of the observations against the ensemble-mean forecast provides a simple diagnostic of the signal deficit in models, i.e., whether the predictable signal is stronger than the predicted signal.

- *The SNP exhibits intermittent behavior, often dominated by a few decisive forecast years.* Correlation-based diagnostics including skill, RPC and regression provide the means to identify these key contributing years which should form the basis for in-depth case studies on the underlying physical and dynamical processes involved. The interannual intermittency is active on top of low-frequency multi-decadal variations of the signal-to-noise characteristics.
- *Multi-model ensembles are inadequate to reliably diagnose the SNP.* Concepts from single-model ensembles do not readily apply to multi-model ensembles of opportunity and, if applied without caution, may lead to erroneous conclusions. Signal-to-noise errors can occur in multi-model ensembles because of different model responses rather than anomalously small predictability seen in single models. Therefore, careful interpretations of signal-to-noise problems in multi-model ensembles are required.
- *Further diagnostics are needed to investigate the influence of eddy feedback on the paradox.* Enhanced diagnostics based on e.g., potential vorticity and E-vectors, may be essential for a more thorough understanding of the processes involved in the feedback mechanisms.
- *The potential emergence of the paradox on shorter timescales opens up new avenues to scrutinize the underlying physical hypotheses.* Should evidence of signal-to-noise errors on multi-weekly timescales prove robust, it could offer opportunities for more comprehensive testing as sub-seasonal forecasts are frequently integrated into high-resolution operational forecasting activities.
- *The signal-to-noise paradox is likely not attributable to a single model deficiency.* Instead, we assume it arises from the collective impact of various deficient physical and dynamical mechanisms. For example, eddy feedback on the large-scale circulation, air-sea interactions, circulation regime structures and tropical-

extratropical teleconnections are interconnected factors which all contribute to the complexity of the phenomenon. Too coarse model resolution, including topography, has been argued to be a broader underlying issue with many of these.

Based on these conclusions, we outline the following future research directions towards resolving the SNP.

1. Models with improved and more realistic representations of a range of physical and dynamical processes will be required. Further increasing the model resolution will help achieve these.
2. The intermittent nature of the paradox provides a basis for focused case studies to test the described hypotheses (“windows of opportunities”).
3. The regime dependence of the SNP over the North Atlantic requires testing and quantifying in multiple forecasting systems to establish more robustly which modes of variability or circulation regimes are worst affected.
4. If the SNP arises from processes that amplify predictable signals in the real world but are poorly represented in current models, it will need to be argued as to why these processes should preferentially amplify only the signals and not also the noise (i.e. internal variability). This puzzle is particularly apparent when considering hypotheses such as the role of eddy feedback, which might well be expected to influence noise as well as signals.
5. Detecting the SNP requires large ensembles, posing a practical constraint due to high computational costs. The quickly evolving field of Machine Learning and, more broadly, Artificial Intelligence is likely to provide impetus for exploring alternative methods of generating large ensembles.

Distinguishing the signal from the noise requires both scientific knowledge and self-knowledge: the serenity to accept things we cannot predict, the courage to predict the things we can, and the wisdom to know the difference. – Nate Silver (2012)

Acknowledgements

We thank all participants of the workshop for their contributions and engagement. Support from NCAS is kindly acknowledged.

References

- Baker, L.H., L.C. Shaffrey, R.T. Sutton, A. Weisheimer and A.A. Scaife (2018). An intercomparison of skill and overconfidence/underconfidence of the wintertime North Atlantic Oscillation in multimodel seasonal forecasts. *Geophys. Res. Lett.*, **45**, 7808-7817, <https://doi.org/10.1029/2018GL078838>
- Beverley, J. D., M. Newman, and A. Hoell (2023). Rapid Development of Systematic ENSO-Related Seasonal Forecast Errors. *Geophys. Res. Lett.*, **50**, e2022GL102249.
- Bröcker, J., A.J. Charlton-Perez and Weisheimer, A. (2023). A statistical perspective on the signal-to-noise paradox. *Q. J. R. Meteorol. Soc.*, **149**, 911-923 doi.org/10.1002/qj.4440
- Byrne, N.J., Shepherd, T.G. and Polichtchouk, I. (2019) Subseasonal-to-seasonal predictability of the southern hemisphere eddy-driven jet during austral spring and early summer. *J. Geophys. Res.*, **124**, 6841–6855, <https://doi.org/10.1029/2018JD030173>.
- Charlton-Perez, A.J., Bröcker, J., Stockdale, T.N. and Johnson, S. (2019) When and where do ECMWF seasonal forecast systems exhibit anomalously low signal-to-noise ratio? *Q. J. R. Meteorol. Soc.*, **145**, 3466–3478. <https://doi.org/10.1002/qj.3631>.
- Chen, R., I. Simpson, C. Deser, and B. Wang (2020). Model biases in the simulation of the springtime North Pacific ENSO teleconnection. *J. Clim.*, **33**, 9985-10002.
- Di Capua, G., D. Coumou, B. van den Hurk, Weisheimer, A., A. Turner, and R. Donner (2023). Validation of boreal summer tropical-extratropical causal links in seasonal forecasts. *Weather Clim. Dynam.*, **4**, 701-723, doi.org/10.5194/wcd-4-701-2023
- Dunstone, N. and 12 co-authors (2018). Skilful seasonal predictions of Summer European rainfall. *Geophys. Res. Lett.*, **45**, 3246-3254, <https://doi.org/10.1002/2017GL076337>
- Dunstone, N. and 11 co-authors (2023). Skilful predictions of the summer North Atlantic Oscillation. *Commun Earth Environ* **4**, 409, <https://doi.org/10.1038/s43247-023-01063-2>

- Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L. and Robinson, N. (2014). Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.*, **41**, 5620–5628, <https://doi.org/10.1002/2014GL061146>
- Falkena, S.K.J., J. de Wiljes, Weisheimer, A., and T.G. Shepherd (2022). Detection of Interannual Ensemble Forecast Signals over the North Atlantic and Europe using Atmospheric Circulation Regimes. *Q. J. R. Meteorol. Soc.*, **148**, 434-453, <https://doi.org/10.1002/qj.4213>
- Falkena, S.K.J., J. de Wiljes, Weisheimer, A., and T.G. Shepherd (2023). A Bayesian approach to atmospheric circulation regime assignment. *J. Clim.*, **36**, 8619-8636, doi.org/10.1175/JCLI-D-22-0419.1
- Garfinkel, C.I., C. Schwartz, D. Domeisen, S.-W. Son, A. Butler and I. White (2018). Extratropical atmospheric predictability from the quasi-biennial oscillation in subseasonal forecast models. *J. Geophys. Res.*, **123**, 7855-7866, <https://doi.org/10.1029/2018JD028724>
- Garfinkel, C. I., W. Chen, Y. Li, C. Schwartz, P. Yadav, and D. Domeisen (2022). The winter North Pacific teleconnection in response to ENSO and the MJO in operational subseasonal forecasting models is too weak. *J. Clim.*, **35**, 8013-8030. <https://journals.ametsoc.org/view/journals/clim/35/24/JCLI-D-22-0179.1.xml>
- Hardiman, S.C., Dunstone, N.J., Scaife, A.A., Smith, D.M., Comer, R., Nie, Y. and Ren, H.-L. (2022). Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions. *npj Climate and Atmospheric Science*, **5**, 57. <https://doi.org/10.1038/s41612-022-00280-4>.
- Hodson, D., R. Sutton and A. Scaife (2023). Signal-to-noise and predictable modes of variability in winter seasonal forecasts. *Q. J. R. Meteorol. Soc.*, **149**, 2598-2616, <https://doi.org/10.1002/qj.4522>
- Hu, S. and T. Zhou (2021). Skillful prediction of summer rainfall in the Tibetan Plateau on multiyear time scales. *Sci. Adv.*, **7**, eabf9395, [DOI: 10.1126/sciadv.abf9395](https://doi.org/10.1126/sciadv.abf9395).

- Kumar, A. (2009). Finite Samples and Uncertainty Estimates for Skill Measures for Seasonal Prediction. *Mon. Wea. Rev.*, **137**, 2622-2631, <https://doi.org/10.1175/2009MWR2814.1>
- Molteni, F. and A. Brookshaw (2023). Early- and late-winter ENSO teleconnections to the Euro-Atlantic region in state-of-the-art seasonal forecasting systems. *Clim. Dyn.*, **61**, 2673–2692, <https://doi.org/10.1007/s00382-023-06698-7>
- O'Reilly, C., A. Weisheimer, T. Woollings, L. Gray and D. MacLeod (2019). The importance of stratospheric initial conditions for winter North Atlantic Oscillation predictability and implications for the signal-to-noise paradox. *Q. J. R. Meteorol. Soc.*, **145**, 131-146, <https://doi.org/10.1002/qj.3413>
- Ossó, A., Sutton, L. Shaffrey and B. Dong (2020). Development, Amplification, and Decay of Atlantic/European Summer Weather Patterns Linked to Spring North Atlantic Sea Surface Temperatures. *J. Clim.*, **33**, 5939-5951, <https://doi.org/10.1175/JCLI-D-19-0613.1>
- Palmer, T.N. and 24 co-authors (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853-872, <https://doi.org/10.1175/BAMS-85-6-853>
- Patrizio, C.R, P.J. Athanasiadis, C. Frankignoul, D. Iovino, S. Masina, L. Famooss Paolini and S. Gualdi (2023). Improved Extratropical North Atlantic Atmosphere–Ocean Variability with Increasing Ocean Model Resolution. *J. Clim.*, **36**, 8403-8424, <https://doi.org/10.1175/JCLI-D-23-0230.1>
- Roberts, C., M. Balmaseda, L. Ferranti and F. Vitart (2023). Euro-Atlantic Weather Regimes and Their Modulation by Tropospheric and Stratospheric Teleconnection Pathways in ECMWF Reforecasts. *Mon. Wea. Rev.*, **151**, 2779-2799, <https://doi.org/10.1175/MWR-D-22-0346.1>
- Scaife and 21 co-authors (2014). Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.*, **41**, 2514-2519, <https://doi.org/10.1002/2014GL059637>

- Scaife, A. and Smith, D. (2018). A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, **1**, 28, <https://doi.org/10.1038/s41612-018-0038-4>.
- Scaife, A. and 12 co-authors (2019). Does increased atmospheric resolution improve seasonal climate predictions? *Atmos. Sci. Lett.*, **20**, e922, <https://doi.org/10.1002/asl.922>
- Screen, J. A., Eade, R., Smith, D. M., Thomson, S., & Yu, H. (2022). Net equatorward shift of the jet streams when the contribution from sea-ice loss is constrained by observed eddy feedback. *Geophys. Res. Lett.*, **49**, <https://doi.org/10.1029/2022GL100523>
- Seviour, W., S. Hardiman, L. Gray, N. Butchart, C. MacLachlan and A. Scaife (2014). Skillful seasonal prediction of the Southern annular mode and Antarctic ozone. *J. Clim.*, **27**, 7462–7474, <https://doi.org/10.1175/JCLI-D-14-00264.1>
- Siegert, S., Stephenson, D.B., Sansom, P.G., Scaife, A.A., Eade, R. and Arribas, A. (2016). A Bayesian framework for verification and recalibration of ensemble forecasts: How uncertain is NAO predictability? *J. Clim.*, **29**, 995–1012. <https://doi.org/10.1175/JCLI-D-15-0196.1>
- Silver, Nate (2012). *The Signal and the Noise: Why So Many Predictions Fail – but Some Don't*. Penguin Group, 534 pages, ISBN 978-1-59-420411-1.
- Smith, D. and 38 co-authors (2020). North Atlantic climate far more predictable than models imply. *Nature*, **583**, 796-800, <https://doi.org/10.1038/s41586-020-2525-0>
- Smith, D. and 31 co-authors (2022). Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nat. Commun.* **13**, 727, <https://doi.org/10.1038/s41467-022-28283-y>
- Stockdale, T.N., F. Molteni and L. Ferranti (2015). Atmospheric initial conditions and the predictability of the Arctic Oscillation. *Geophys. Res. Lett.*, **42**, 1173–1179, <https://doi.org/10.1002/2014GL062681>.
- Strommen, K. and Palmer, T.N. (2019) Signal and noise in regime systems: A hypothesis on the predictability of the North Atlantic Oscillation. *Q. J. R. Meteorol. Soc.*, **145**, 147–163, <https://doi.org/10.1002/qj.3414>.

- Weisheimer, A., D. Decremmer, D. MacLeod, C. O'Reilly, T. Stockdale, S. Johnson and T.N. Palmer (2019). How confident are predictability estimates of the winter North Atlantic Oscillation? *Q. J. R. Meteorol. Soc.*, **145**, 140-159, [doi:10.1002/qj.3446](https://doi.org/10.1002/qj.3446)
- Weisheimer, A., D. Befort, D. MacLeod, T.N. Palmer, C. O'Reilly and K. Strommen (2020). Seasonal forecasts of the 20th Century *Bull. Amer. Meteor. Soc.*, **101**, E1413-E1426, [doi:10.1175/BAMS-D-19-0019.1](https://doi.org/10.1175/BAMS-D-19-0019.1)
- Williams, N., A. Scaife and J. Screen (2023). Underpredicted ENSO teleconnections in seasonal forecasts. *Geophys. Res. Lett.*, **50**, <https://doi.org/10.1029/2022GL101689>
- Yeager, S. and 20 co-authors (2018). Predicting Near-Term Changes in the Earth System: A Large Ensemble of Initialized Decadal Prediction Simulations Using the Community Earth System Model. *Bull. Amer. Meteor. Soc.*, **99**, 1867-1886, <https://doi.org/10.1175/BAMS-D-17-0098.1>
- Yeager, S. and 8 co-authors (2023). Reduced Southern Ocean warming enhances global skill and signal-to-noise in an eddy-resolving decadal prediction system. *npj Clim Atmos Sci*, **6**, 107, <https://doi.org/10.1038/s41612-023-00434-y>
- Zhang, W. and B. Kirtman (2019). Understanding the Signal-To-Noise Paradox with a Simple Markov Model. *Geophys. Res. Lett.*, **46**, 13308-13317, <https://doi.org/10.1029/2019GL085159>
- Zhang, W., B. Kirtman, L. Siqueira, A. Clement and J. Xia (2021). Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model. *Clim. Dyn.*, 56, 2895-2913, <https://doi.org/10.1007/s00382-020-05621-8>
- Zhao, S., H.-L. Ren, F. Zhou, A. Scaife and Y. Nie (2023). Phase asymmetry in synoptic eddy feedbacks on the negatively-skewed winter NAO. *Atm. Res.*, **288**, 106725, <https://doi.org/10.1016/j.atmosres.2023.106725>