SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Assessing the impact of stochastic physics (SPPT) on sub-decadal time-scales
Computer Project Account:	spgbdjb
Start Year - End Year :	Jan. 2019 – Dec. 2019
Principal Investigator(s)	Daniel J Befort, Antje Weisheimer (ECMWF & University of Oxford)
Affiliation/Address:	University of Oxford Atmospheric, Oceanic and Planetary Physics Clarendon Laboratory Department of Physics University of Oxford Oxford, OX1 3PU
Other Researchers (Name/Affiliation):	Christopher H. O’Reilly (University of Oxford)

The following should cover the entire project duration.

Summary of project objectives

The aim of this project is to assess the impact of stochastic physics (SPPT) on multi-year time-scales (up to three winters ahead). Experiments with and w/o SPPT with a forecast length of 28 months initialized in November from 1981 to 2014 were conducted using the coupled long-range system. The main focus of the analysis is on ENSO as a coupled process with predictability on seasonal and perhaps multi-year time scales. Some results showing the impact in the North Pacific are also presented.

Summary of problems encountered

None

Experience with the Special Project framework

The yearly timing of the submission deadlines meant that my application went in as a late special project as my current position only just started after the June deadline. Unfortunately, only 50% of the ~10 million SBUs asked for have been granted. This caused some problems as this amount hasn't been sufficient to perform the planned experiments which couldn't easily be shortened substantially. A collaboration with the Long-Range Team of the Earth System Predictability Section in the Research Department allowed us to use RD units to finish off the simulations. In general, 10 million SBUs is not a lot of resources for running seasonal scale hindcasts. Thus, it would be beneficial if, in the future, the fixed annual application deadline could be reconsidered for a perhaps more frequent application cycle. In my case there would have been nearly 18 months between when I would have liked to apply to the start date of funding if I had used the standard application in the following year instead of the late application. Alternatively, if the upper limit for late applications could be raised, that would be very useful too.

It is not fully clear to us in how far publishing the progress and final reports with results compromises scientific publications in peer-reviewed journals as these often require that the results were not made available elsewhere.

Summary of results

Motivation for this project

During the past decade, large efforts have been made to develop decadal prediction systems. However, while stochastic parameterisations have been extensively utilised for short to seasonal forecasts, initialised decadal climate predictions are still very much performed without accounting for these intrinsic model uncertainties. Given the encouraging results obtained on seasonal timescales, this special project aimed to extend the analysis of model uncertainty beyond seasonal prediction to the multi-annual time scales by running targeted model simulations using ECMWF's coupled model system.

Model simulations

To assess the impact of stochastic physics (SPPT) on multi-year time scales, two hindcast sets using ECMWF's coupled model CY46R1 have been conducted. The setup of both experiments is identical: i) initialized each 1st November from 1981 to 2014 ii) 10 ensemble members iii) 28-month forecasts, iv) atmospheric horizontal resolution TCo199 (approx. 50km); 1-degree ocean resolution. The only difference between the runs is that the stochastic physics scheme SPPT has been switched off

(ECMWF IFS w/o SPPT) in one experiment, whereas it is included in the other experiment (ECMWF IFS w/ SPPT).

The initialisation of these simulations followed the standard protocol of long-range coupled experiments. ERA-5 was used as initial conditions in the atmosphere and a standard 1 degree ocean experiment based on relaxation of ORA5 was used to initialise NEMO and LIM.

Methods

The reliability of the different predictions is assessed using the spread-skill relationship, which is based on the ratio (SoE) between the average ensemble spread and the RMSE of the ensemble mean. Generally speaking, a perfect reliable forecasting system is one where the verification is indistinguishable from the ensemble. It can be shown analytically that for a perfect ensemble the time-mean ensemble spread (standard deviation around the ensemble mean) should equal the time-mean RMSE of the ensemble mean forecast (Palmer et al., 2006). The relationship implies that for perfectly reliable predictions the measure spread-over-error SoE = 1, within the sampling uncertainty. Values of SoE > 1 indicate overdispersive (underconfident), whereas SoE < 1 values indicate underdispersive (overconfident) ensemble predictions.

Previous studies on the impact of stochastic physics on seasonal time-scales proved that the tropical Pacific is the region where the stochastic schemes have the largest and significant impact. As the ENSO regions are also an important source of global teleconnections, we focus the analysis in this study on the sea surface temperatures over the NINO3.4 region (170-120W; 5S-5N) of the central tropical Pacific. We also show global statistics for summer and winter seasons which are included in the 28 months. Besides analysing the effect of SPPT in the tropics, we investigate in how far it also affects the midlatitudes. This is done by assessing the skill of the extratropical large-scale atmospheric circulation by analysing predictions of sea-level pressure anomalies.

Simulations used in this project have been corrected for lead-time dependent biases. Here, anomalies are calculated using all initialization dates between 1981 until 2014.

Results

Skill and reliability of global sea surface temperatures

Skill is estimated by calculating the ensemble mean anomaly correlation coefficient (ACC) per grid point for all boreal summer (JJA) and winter (DJF) seasons during the first 28 months (Fig. 1). The skill decreases with forecast time for both ensembles, with mostly significant positive values occurring during the first winter season (forecast time 2-4 month). Only a few areas show significant skill in the 3rd winter (forecast time 26-28 months).

Using SPPT leads to higher ACC values over the tropical ocean basins, which is in line with previous studies on the impact of SPPT on seasonal time-scales (Weisheimer et al., 2011; 2014). Only marginal differences between both ECMWF IFS simulations with regards to ACC skill is found for the 2nd summer and the 3rd winter. Next, reliability of both ensembles is assessed using the spread-over-error metric (SoE). Some striking differences between the ECMWF experiments w/ and w/o SPPT are found (Fig. 2), e.g. over the tropical Pacific basin in the 1st winter, which is in agreement with previous studies (Weisheimer et al., 2011). Here, the ECMWF IFS w/o SPPT is significantly overconfident, whereas the usage of SPPT improves the reliability drastically. Overconfidence over the tropical Pacific Ocean is apparent in ECMWF IFS w/o SPPT for all seasons. This is different if using SPPT, which leads to a reliable ensemble prediction system for the 1st and 2nd winter as well as for the 2nd summer, whereas in the 3rd winter the SPPT experiment shows large overconfident areas over the tropical Pacific as well. A completely different pattern with regards to reliability is found for the 1st summer, for which including stochastic physics leads to a significantly underconfident ensemble over parts of the Pacific.

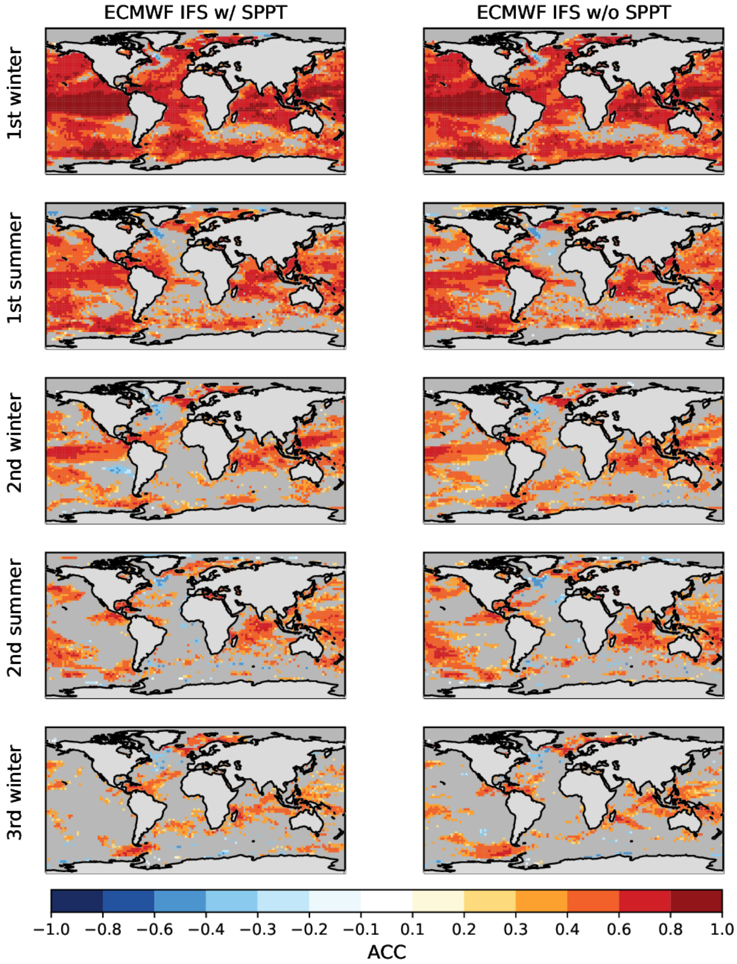

Figure 1 ACC scores for SSTs and different forecast times (1st to 3rd winter/ 1st & 2nd summer). Summer averages are calculated for June to August, whereas December to February is used for boreal winter season. All years between 1981 and 2014 are used.

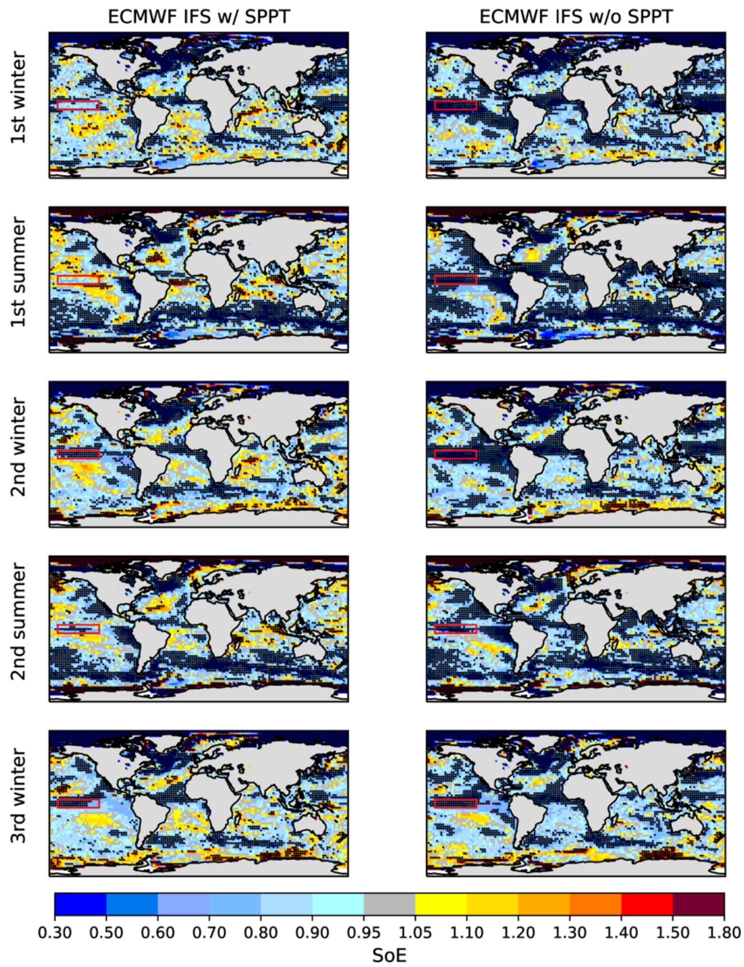

Figure 2 Same as Figure 1 but for spread over error (SoE). Dots for areas significantly different to 1 with 95% confidence (200 samples)

Skill and reliability of ENSO indices

ACC skill for 3-monthly smoothed SSTs over the NINO3.4 central tropical Pacific region for each ECMWF IFS ensemble is shown in Figure 3a. ACC is high for both ECMWF IFS simulations for the first 16-17 months. After the second year spring barrier to ENSO predictability (Cane, 1991; Webster & Yang, 1992) in March-April-May (MAM), the ACC becomes similarly low between both ensembles.

Figure 3b illustrates root mean square error (RMSE, solid lines) and spread (dashed lines) over forecast time in each ensemble, whereas Figure 3c shows reliability measured by the ratio of RMSE and spread (SoE). The w/ SPPT ensemble exhibits higher reliability than the ECMWF IFS w/o SPPT ensemble. This enhanced reliability is to a large extent related to increased spread within the w/ SPPT ensemble but also due to a reduced error (Fig. 3b). During the first half year of the forecasts ECMWF IFS w/ SPPT has an excessively large spread which is related to SPPT modification introduced into the recent model cycle. ECMWF's operational seasonal forecasts from SEAS5 based on an older model cycle do not show such an overdispersion (Stockdale et al, 2018; Fig 3c cyan line). The overconfidence of equatorial Pacific Ocean SSTs in IFS ensembles w/o SPPT on seasonal time-scales has been found in previous studies, but here it is shown for the first time that the overconfidence continues beyond annual time-scales.

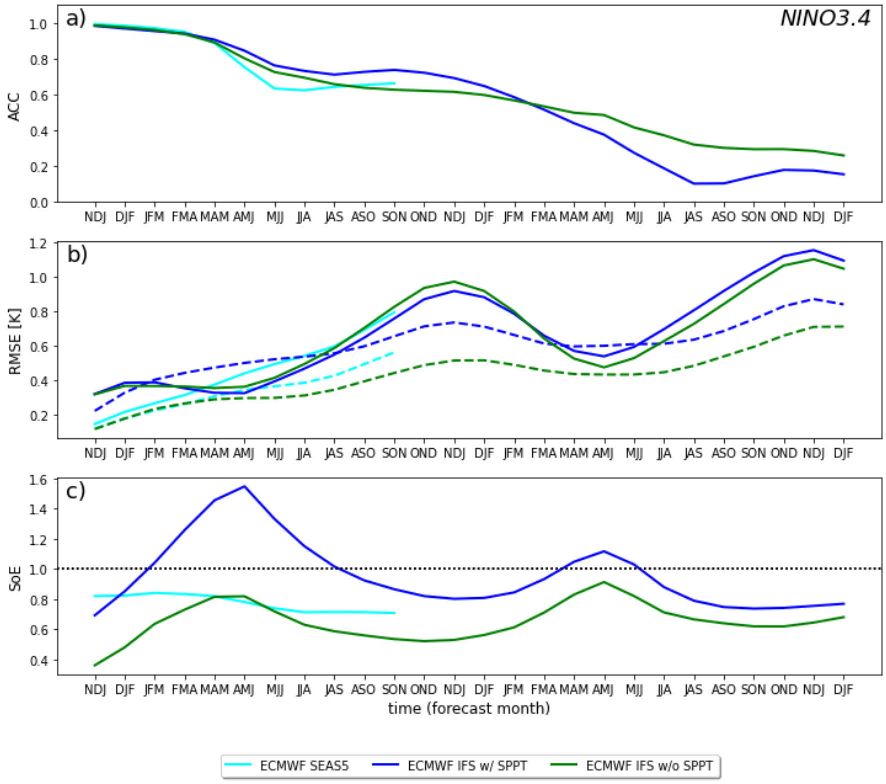

Figure 3 Anomalous correlation coefficients for NINO3.4 using ERA5 as reference, b) same as a) but for RMSE (solid lines) and spread (dashed lines), and c) same as a) but for SoE.

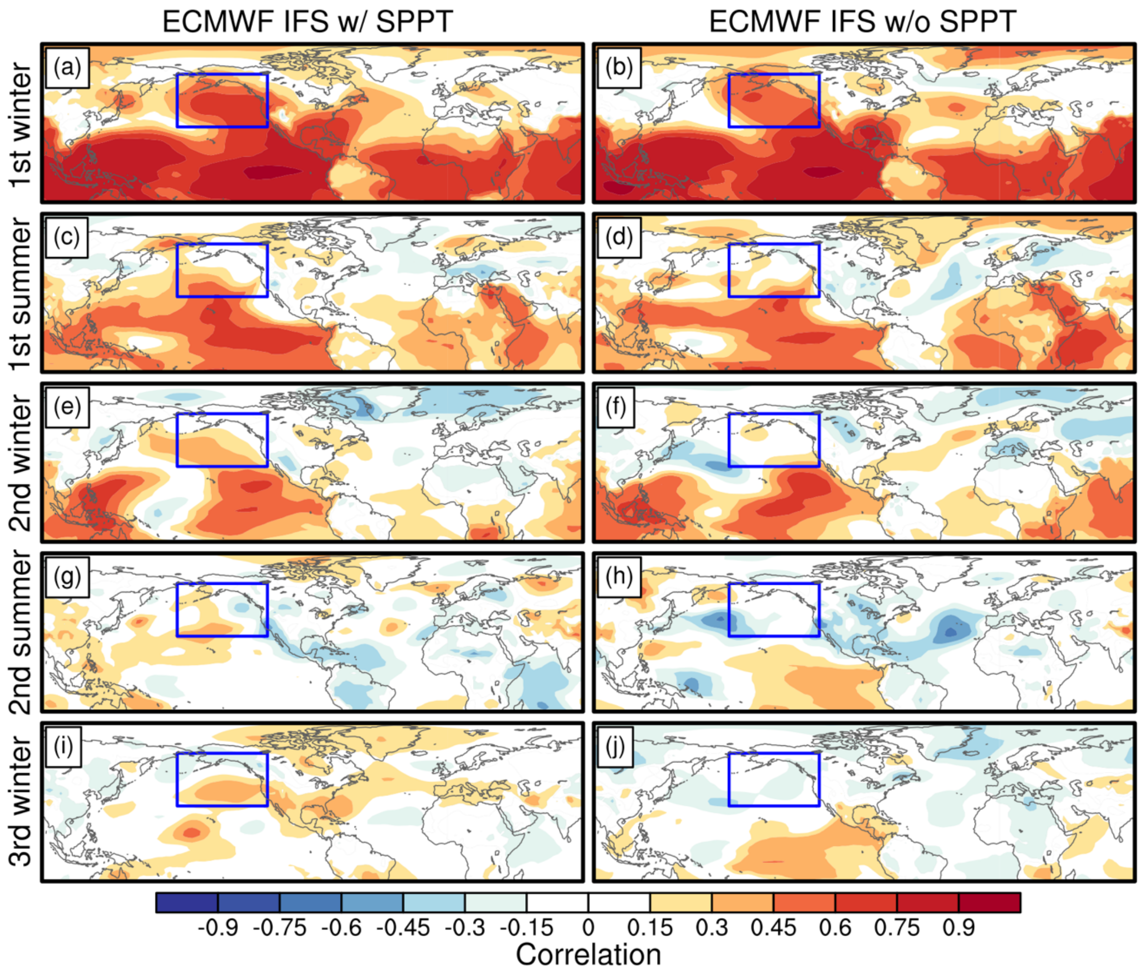

Figure 4 ACC scores for sea-level pressure at different forecast times (1st to 3rd winter/ 1st & 2nd summer). Summer averages are calculated for June to August, whereas December to February is used for boreal winter season. All years between 1981 and 2014 are used.

Analysis of atmospheric circulation skill in the NH extratropics

To examine the influence of SPPT outside of the tropics in the near-term predictions we now analyse the predictions of sea level pressure (SLP) in the NH extratropics. Figure 4 illustrates correlation skill for both ECMWF IFS simulations with and without SPPT. High levels of skill are evident over the tropics during the 1st winter, which is typical of seasonal forecasting systems (e.g. Smith et al., 2012). In the extratropics, significant levels of skill are generally limited to the extratropical North Pacific where ENSO teleconnections are strong (Trenberth et al., 1998), whereas elsewhere the signal due to the relatively small ensemble size becomes weak. Skill drops off in both the tropics and extratropics by the 1st summer but in the 2nd winter, there remains significant skill in the tropics in both the ECMWF IFS ensembles. However, the ECMWF IFS w/ SPPT ensemble is the only system with substantial levels of skill in the extratropical North Pacific.

Into the 2nd summer skill continues to drop and by the 3rd winter there is very little skill left, even in the tropics, which is consistent with the SST analysis above. One curious feature in the ECMWF IFS w/ SPPT ensemble during the 3rd winter is the region of relatively high skill in the extratropical North Pacific. Typically, skill in this region is thought to originate from teleconnection with the tropical Pacific and is associated with skill in the tropical SLP (as seen in the 1st and 2nd winter, for example). SLP skill in the tropics is absent in the 3rd winter in the ECMWF IFS w/ SPPT ensemble, so it is not clear whether the extratropical skill in the 3rd winter is robust. Addressing this, will require further analysis and is left for future work.

Next, the evolution of the skill of the extratropical North Pacific SLP anomalies over lead-time is examined by analysing a North Pacific index region (as shown by the blue boxes in Figure 4). The evolution of the correlation skill of the North Pacific SLP index is shown in Figure 5 (with significant skill indicated by the solid coloured dots). Across the extended first winter season, both the ECMWF IFS ensembles exhibit similarly high levels of skill. Into the spring season, skill in the ECMWF IFS w/ SPPT ensemble drops off slightly more quickly than the w/o SPPT ensemble and by the 1st summer season none of the ensembles exhibit significant skill. Beyond the first summer season, the ECMWF IFS w/o SPPT ensemble shows no significant correlation skill for the North Pacific SLP index, whereas, in the ECMWF IFS w/ SPPT ensemble, significant skill returns in the extended winter during the second year of the forecast. SLP anomalies in this region are strongly influenced by tropical Pacific SST anomalies through an atmospheric teleconnection, so the increased skill in the second winter season is entirely consistent with the levels of skill seen in the ECMWF IFS w/ SPPT ensemble for the NINO 3.4 SST index (i.e. Figure 3a).

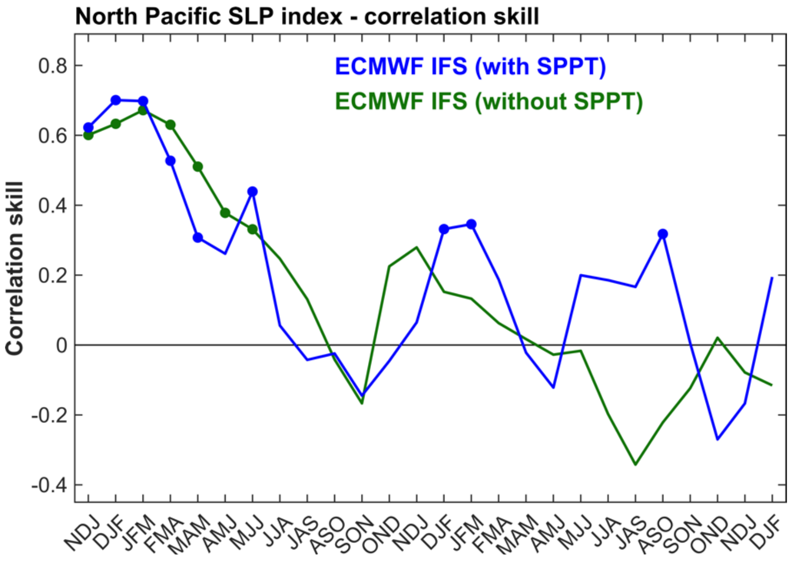

Figure 5 Anomalous correlation coefficients for North Pacific SLP index, using ERA5 as reference. Filled circles indicate where the correlations are significant at the 95% level.

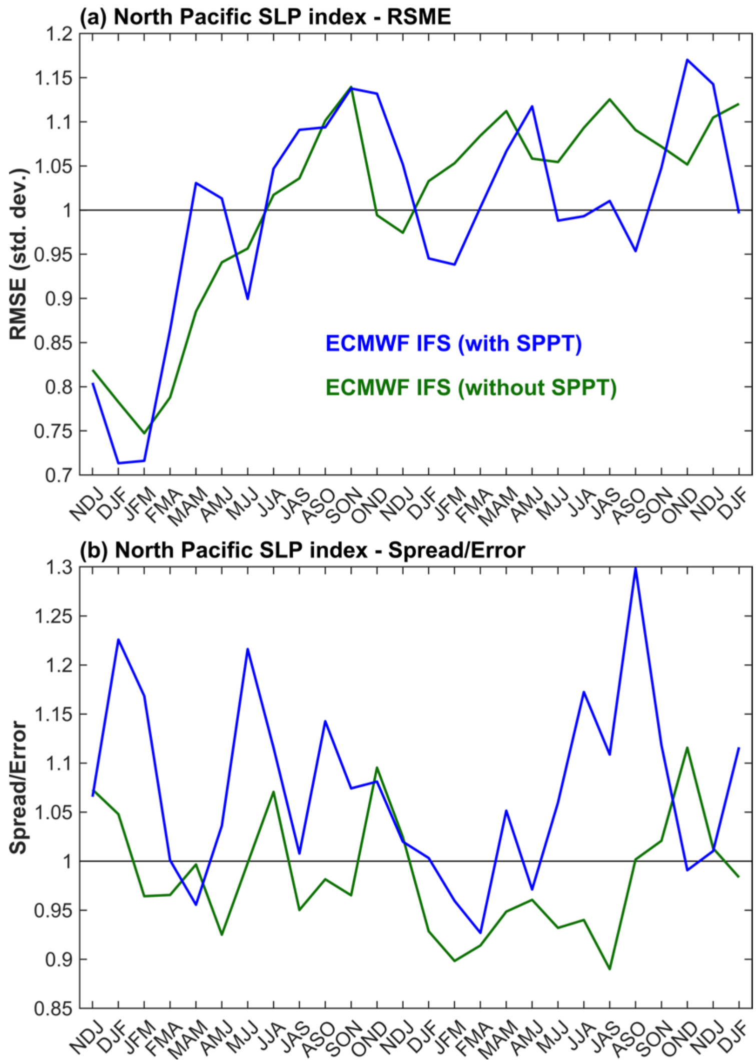

Figure 6 (a) RMSE and (b) Spread/Error ratio for the North Pacific SLP index, using ERA5 as reference.

We also analysed the RMSE and SoE metrics of the North Pacific SLP index (Fig. 6). The RMSE evolution for the ensembles is largely consistent with the correlation skill evolution (i.e. Figure 5),. The RMSE of the North Pacific SLP index is essentially saturated by the first summer season and beyond in the ECMWF IFS w/o SPPT ensemble. In the ECMWF IFS w/ SPPT ensemble, however, there is notably lower RMSE during the extended second winter season, consistent with the higher correlation skill. The SoE ratio is close to one during this second season, indicating that the increased skill seen in the second winter season is also reliable. Beyond the second winter, there are some instances of lower skill in the ECMWF IFS w/ SPPT ensemble, compared with the w/o SPPT ensemble. However, the predictions have a SoE ratio of 1.2 or higher, indicating lower reliability in the ensemble, which may betray a lack of robustness of this apparent skill as previously suggested.

Conclusions

Targeted model simulations with ECMWF's coupled model have been carried out within this project to investigate the impact of stochastic physics beyond seasonal time scales. Two 10-member ensembles initialized each November between 1981 and 2014 with a total forecast time of 28 months have been carried out. These 2 hindcast sets are identical except that one uses stochastic physic perturbations (w/ SPPT), whereas the other does not (w/o SPPT). Results show that SPPT positively impacts skill and reliability especially for SSTs over the tropical oceans up to about 18 months. The w/o SPPT ensemble is heavily overconfident over the NINO3.4 region, whereas enhanced reliability is found for the w/ SPPT ensemble up to the 2nd winter (DJF; forecast time: 14-16 months) primarily due to increases in ensemble spread. The benefits of SPPT are not limited to the tropics, as despite the relatively small ensemble size, SPPT seems to improve the skill of the large-scale atmospheric circulation over the extratropical North Pacific in the second winter of the forecasts. Moreover, these forecasts are found to be reliable, in a statistical sense, which increases confidence in the utility of predictions made with SPPT.

List of publications/reports from the project with complete references

These simulations are part of a study comparing different methodologies used to represent model uncertainty. Main focus is put on skill and reliability of current single model decadal prediction ensembles, ensembles w/ and w/o SPPT (simulations carried out within this special project) and a multi-model ensemble. These results are currently prepared for publication.

Future plans

The simulation carried out in this special project are further used as control experiments for the special project *spgbdjb "Disentangling the local and remote effect of SPPT on seasonal timescales"* (PI: Daniel J. Befort). This new project aims to assess the remote and local effect of stochastic physics on the circulation over the extratropics.

References

Cane, M. A. (1991). Forecasting El Niño with a geographical model. *Teleconnections Connecting World-Wide Climate Anomalies,* M. Glantz, R. Katz, and N. Nicholls, Eds., Cambridge University Press, 345–369.

Palmer, TN, R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher and L.A. Smith (2006). Ensemble prediction: A pedagogical perspective. *ECMWF Newsletter*, **106**, 10-17.

Smith, D. M., Scaife, A. A., & Kirtman, B. P. (2012). What is the current state of scientific knowledge with regard to seasonal and decadal forecasting?. *Environmental Research Letters*, **7(1),** 015602.

Stockdale, T.N., M. Alonso-Balmaseda, S. Johnson, L. Ferranti, F. Molteni, L. Magnusson, S. Tietsche, F. Vitart, D. Decremer, A. Weisheimer, C.D. Roberts, G. Balsamo, S. Keeley, K. Mogensen, H. Zuo, M. Mayer and B.M. Monge-Sanz (2018). SEAS5 and the future evolution of the long-range forecast system. *ECMWF Tech. Memo*, **835**, doi:10.21957/z3e92di7y

Trenberth, K.E., Branstator, G.W., Karoly, D., Kumar, A., Lau, N. C., & Ropelewski, C. (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *Journal of Geophysical Research: Oceans*, 103(C7), 14291-14324.

Webster, P. J., and S. Yang (1992). Monsoon and ENSO: Selectively interactive systems. *Quart. J. Roy. Meteor. Soc.,* **118,** 877–926.

Weisheimer A, Palmer TN and Doblas-Reyes FJ. (2011). Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Lett.* **38**: L16703. https://doi.org/10.1029/2011GL048123.

Weisheimer A, Corti S, Palmer TN, Vitart F. (2014). Addressing model error through atmospheric stochastic physical parametrizations: impact on the coupled ECMWF seasonal forecasting system. *Philosophical Transactions of the Royal Society A* **372**: 20130290, https://doi.org/10.1098/rsta.2013.0290.