SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year	2025
Project Title:	Investigating multi-decadal climate variations in seasonal forecasts: The role of aerosol and greenhouse gas forcings
Computer Project Account:	spgbwrig
Principal Investigator:	Mr. Matthew Wright Dr. Antje Weisheimer (ECMWF; University of Oxford) Prof. Tim Woollings (University of Oxford)
Affiliation:	University of Oxford
Name of ECMWF scientist(s) collaborating to the project (if applicable)	Dr. Antje Weisheimer (ECMWF and University of Oxford) Dr. Tim Stockdale (ECMWF) Dr. Retish Senan (ECMWF)
Start date of the project:	January 2023
Expected end date:	December 2025

Computer resources allocated/used for the current year and the previous one
(if applicable)
Please answer for all project resources

		Previous year		Current year	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	23,600,000	23,600,000	29,600,000	256,000
Data storage capacity	(Gbytes)	189,130	~40,000	294,410	~40,000

Summary of project objectives (10 lines max)

This project aims to build on the work of Weisheimer et al. (2017, 2022, 2023), to investigate multidecadal variations in seasonal hindcast skill throughout the twentieth century. The experiments run in this project will test hindcast skill's sensitivity to various sources of external forcing, initially focusing on aerosol loads and, later, greenhouse gas emissions.

We will work closely with the CONFESS project team at ECMWF to run hindcasts with perturbed aerosol mass mixing ratios. The hindcasts will be analysed, focusing on comparing the low skill mid-century period to the high-skill recent period, and quantifying the impact of aerosol forcing on seasonal forecast dynamic and thermodynamic responses.

Summary of problems encountered (10 lines max)

It took a few months to modify the IFS source code in such a way that we could run seasonal hindcasts initialised from ERA-20C, with time-varying aerosol forcings. This meant that the first set of simulations could not be run until the second half of 2023. The CONFESS aerosol dataset we are basing our work on only goes back to 1970. Therefore, we have had to modify and develop our own aerosol dataset (adapted from CMIP6 forcings). This added a substantial amount of work and methodological development in the first year. Analysing the results of the first set of experiments has led to an unexpected focus on the circulation off Africa, and mostly the impacts of natural aerosols – this is not a problem, but a slight change of focus. Since the natural aerosols have such a dominating impact on the forced response, we are perturbing only anthropogenic aerosols for future experiments.

Summary of plans for the continuation of the project (10 lines max)

The analyses of the initial experiments (DJF runs in 1925-1950 and 1985-2010, with doubled, halved and 'best guess' aerosol mass mixing ratios) are now almost complete, and are being written up into a form ready for publication.

These results are informing the next set of experiments, will involve removing all anthropogenic aerosols from the forcing data and running hindcasts with only anthropogenic aerosol forcing in certain regions: Europe, North America, South Asia and the Middle East, and East Asia. This will enable us to isolate the impact of anthropogenic aerosols from natural aerosols and give an understanding of the processes leading to aerosols' impact on weather and climate on the seasonal timescale. These runs will be compared to each other, and to existing control simulations.

List of publications/reports from the project with complete references

N/A yet – we are still analysing results from the first set of experiments. Once these analyses are completed, it is possible that a publication will be submitted.

Summary of results

Additional experiments were run to complement those presented in the last project report, which analysed an initial set of hindcasts initialised on 1 November in 1925-1950 and 1985-2010. These additional experiments have been analysed to provide more insight into the impact of aerosol forcing on seasonal forecasting skill, and on the mean state and variability in the seasonal forecasting model.

Additional DJF experiments for 1951-1984

We conducted 3 sets of hindcast experiments (doubled aerosol forcing, halved aerosol forcing, and best guess aerosol forcing) to 'fill the gaps' between previous experiments (see previous report for details). These confirmed the results of the previous experiments: that the seasonal forecast skill is insensitive to aerosol forcing for large-scale climate indices (see Figure 1). They also show a similar response in the mean state to the 1925-1950 and 1985-2010 experiments, which seems stationary in time. This corresponds to significant cooling in tropical Atlantic Ocean SSTs, and in west and central African surface temperatures, and an associated circulation response over Africa (see column 1, Figure 2, and previous report).

For SST indices, the skill is almost identical in all three aerosol experiments, to each other and to previous hindcast experiments. All datasets exhibit multidecadal variability in hindcast skill, with a minimum in skill in the mid-century period (approximately 1940-1970). This is similar in the Pacific North American (PNA) index. In the North Atlantic Oscillation (NAO) index, there are some differences in skill, which warrants further investigation; this will be addressed with the next set of experiments, which will isolate the impacts of European and North American aerosols on seasonal forecasting skill, variability and mean state.

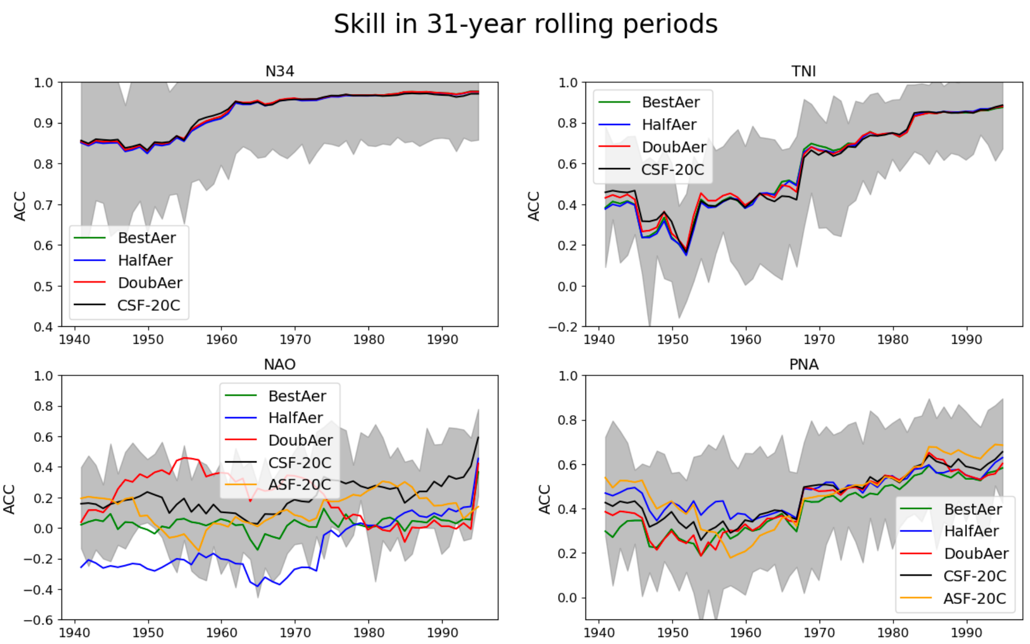

Figure 1: The 31-year rolling anomaly correlation coefficient skill metric for DJF mean climate indices, against ERA-20C: (a) the Nino 3.4 index; (b) the Trans-Nino index; (c) the North Atlantic Oscillation Index; (d) the Pacific North American index, smoothed over 31 years. BestAer (green) shows hindcasts with best guess aerosol forcing; DoubAer (red) shows double this forcing; HalfAer (blue) half this forcing. CSF-20C (black) and ASF-20C (yellow) are also shown for comparison. The grey shading shows indicative errors for BestAer, calculated using bootstrapping with replacement.

DJF experiments with perturbed ocean initial conditions for 1940-1949 and 1990-1999

The initial DoubAer (doubled aerosol mass mixing ratios) and HalfAer (halved aerosol mass mixing ratios) experiments effectively gave the model a step-change in aerosol forcing at the initialisation date, as the CERA-20C used for initialisation assumes real world aerosol forcing. This means that the results were potentially influenced by initialisation shock, and it is hard to tell whether the diagnosed changes to the mean state are due to this sudden change or represent a true equilibrium response to doubled/halved aerosol forcing. To address this, we ran a further set of hindcast experiments for two decades (initialised on 1 November every year in 1940-1949 and 1990-1999). These decades were chosen to provide a decade in the mid-century, low-skill period (1940s), and in the late century, high-skill period (1990s).

Two further sets of experiments were run: DoubAer PertIC, and HalfAer PertIC. The aerosol forcing in the atmosphere is identical to DoubAer and HalfAer (double and half the best guess aerosol forcing, respectively). For DoubAer PertIC, ocean initial conditions are calculated by subtracting the average difference (across all members and years) in February mean ocean heat content between DoubAer and BestAer, from the CERA-20C initial conditions (with adjustments to salinity using the ocean equation of state). For HalfAer PertIC, a similar process is repeated using the difference between HalfAer and BestAer. This enables us to see the longer-term impacts of aerosol forcing, by perturbing ocean initial conditions to represent prior impacts of aerosols. Otherwise, the perturbed initial condition experiments were identical to the initial set of experiments, with 11 ensemble members per start date.

Figure 2 shows some of the results from these perturbed initial condition experiments, compared to those from the original experiments. The first column shows the impact of increased aerosol forcing in the initial experiments, with cooled SSTs in the tropical Atlantic, a warming patch in the tropical Atlantic where upwelling is reduced, and cooled surface air temperatures over central and western Africa, northern India and eastern China. The second column shows the same comparison (doubled minus halved aerosol forcing), but for the two experiments with perturbed initial conditions. There is clearly more cooling over the tropical Atlantic, including a cooling over the warming hole and more cooling of the tropical Pacific SSTs. The surface air temperatures show a stronger cooling signal over the tropical Pacific but are otherwise similar.

The difference in response (column 1 minus column 2) is shown in the third column: positive values indicate more cooling in the perturbed initial condition experiments. There are small differences in the air temperature responses, but the perturbed initial conditions lead to much stronger SST cooling in the tropical Pacific, tropical and subtropical Atlantic, and in the southern Indian Ocean. This shows that perturbing the initial conditions – which arguably more realistically represents the integrated impact of perturbed aerosol forcing – lead to a similar response in SSTs and 2m temperatures, with SST changes having a stronger magnitude. This shows that the surface air temperatures can mostly adjust within the first four months of integration (this is a ‘fast’ response), but the ocean temperatures and SSTs are still responding to the initial aerosol forcing after 4 and even 8 months (a ‘slow’ response). There are further differences in dynamical variables, which will not be discussed here, but are being analysed in preparation for publication.

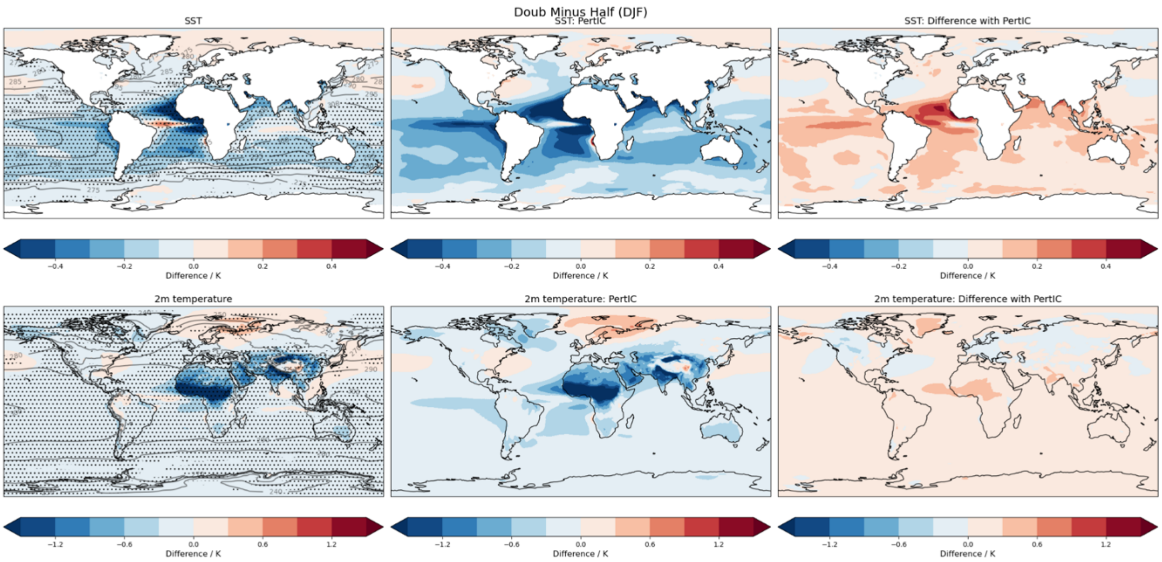

Figure 2: Column 1: filled contours show DoubAer minus HalfAer (DJF means) from initial experiments (averaged over all years and members), with BestAer climatology in line contours, and hatching representing regions with significant differences at the 2.5% level using a KS test. Column 2: DoubAer PertIC minus HalfAer PertIC (averaged over all years and members, noting that there are fewer. Column 3: column 1 minus column 2.
Row 1 shows sea surface temperatures (SSTs); row 2 shows 2m temperatures.

<u>JJA experiments for 1950-2010</u>

We also produced equivalent DoubAer, BestAer and HalfAer experiments, but initialised on 1 May each year in 1950-2010 with 17 ensemble members, to compare the response to aerosol forcing in summer and winter. These experiments were completed late in 2024, and results are still being analysed. Initial results are outlined here, with more analysis to follow.

As shown in Figure 3, the skill of JJA-mean SST indices is still insensitive to aerosol forcing. However, there is more variation in the skill of the NAO and PNA indices between the aerosol forcing experiments. For the NAO, there is no simple relationship between aerosol forcing and JJA-mean NAO skill, and this requires further investigation. For the PNA, there is substantial decadal variability in the skill in all experiments, including CSF-20C, ASF-20C and BestAer. However, it appears that doubling and halving the aerosol forcing negatively impacts skill in this period, with HalfAer having the worst skill and DoubAer the second worst (indeed, both skills are negative, indicating predictions of the opposite sign to what happened in reality). Investigating this further (not shown), it appears that the minimum in skill is primarily due to the seasonal forecasting model predicting an El Nino in 1980, leading to a positive PNA index, when in fact there were neutral conditions in the Pacific and a strong negative PNA index. Further investigation is required to fully understand the differences in skill between BestAer, DoubAer and HalfAer.

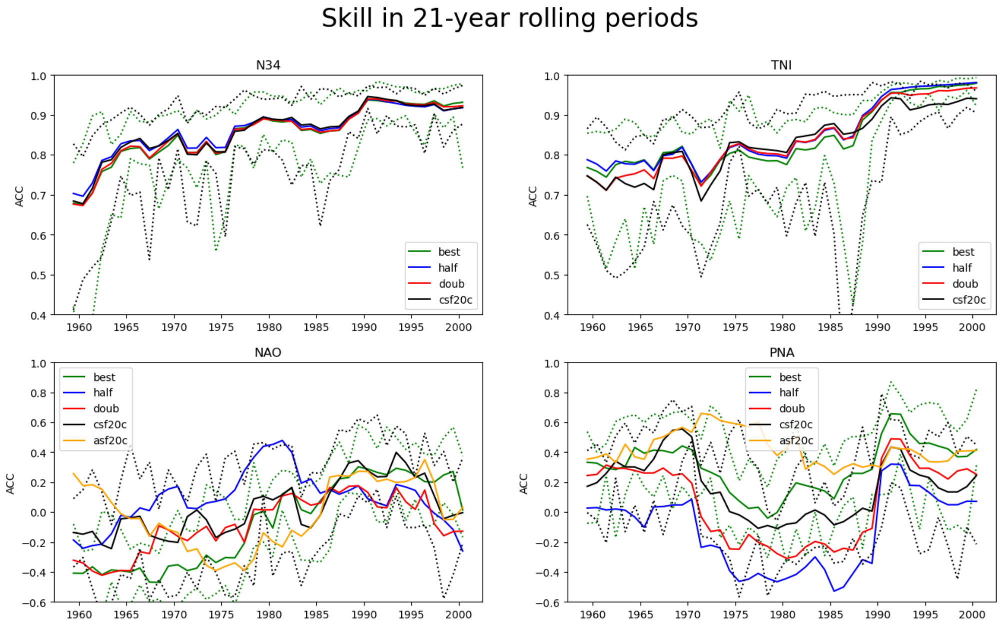

Figure 3: The 21-year rolling anomaly correlation coefficient skill metric for JJA mean climate indices, against ERA-20C: (a) the Nino 3.4 index; (b) the Trans-Nino index; (c) the North Atlantic Oscillation Index; (d) the Pacific North American index, smoothed over 31 years. BestAer (green) shows hindcasts with best guess aerosol forcing; DoubAer (red) shows double this forcing; HalfAer (blue) half this forcing. CSF-20C (black) and ASF-20C (yellow) are also shown for comparison. The grey shading shows indicative errors for BestAer, calculated using bootstrapping with replacement.

Figure 4 shows the mean state response to aerosol forcing by displaying the JJA-mean difference (averaged across all years and members) between DoubAer and HalfAer. Some key features stand out:

- Cooling of SSTs in the northern subtropical Atlantic (displaced further north than in the DJF runs)
- Warming of SSTs in the equatorial Atlantic (again probably due to a change in upwelling location)
- Cooling of SSTs in the Mediterranean and North Atlantic
- More patchy cooling of air temperatures than in DJF experiments, with cooling seen over parts of Africa, southern Europe, India and China
- Large, significant changes to sea level pressure and geopotential height, including a wave-like pattern over the North Atlantic; this pattern projects weakly onto a negative summer NAO and is accompanied by a southerly shift of the jet stream
- Strong and significant changes to zonal winds: at the surface, a northwards shift of easterlies; and at upper levels, polewards shifts of subtropical jets
- As in the DJF experiments, the inter-tropical convergence zone appears to shift southwards over the Atlantic, as shown by a southwards shift of precipitation and changes to outgoing longwave radiation. This is attributed to a stronger inter-hemispheric temperature gradient when aerosol forcing is increased.

It appears, as in the DJF experiments, that much of the signal is originating in the tropical Atlantic, forced by mineral dust aerosols from the Sahara Desert. There is a large response to aerosol forcing, and further work is needed to fully understand the mechanisms for the circulation changes. Further investigations will include study of global and regional monsoon intensity, which has been shown to be impacted by aerosol forcing, and analysis of the summer PNA and NAO mean state and variability to unpick the differences in skill.

A final set of experiments is planned, as outlined above, to attribute the impacts of anthropogenic aerosol forcing from Europe, North America, South Asia and East Asia. These will be completed shortly and compared to the existing set of experiments.

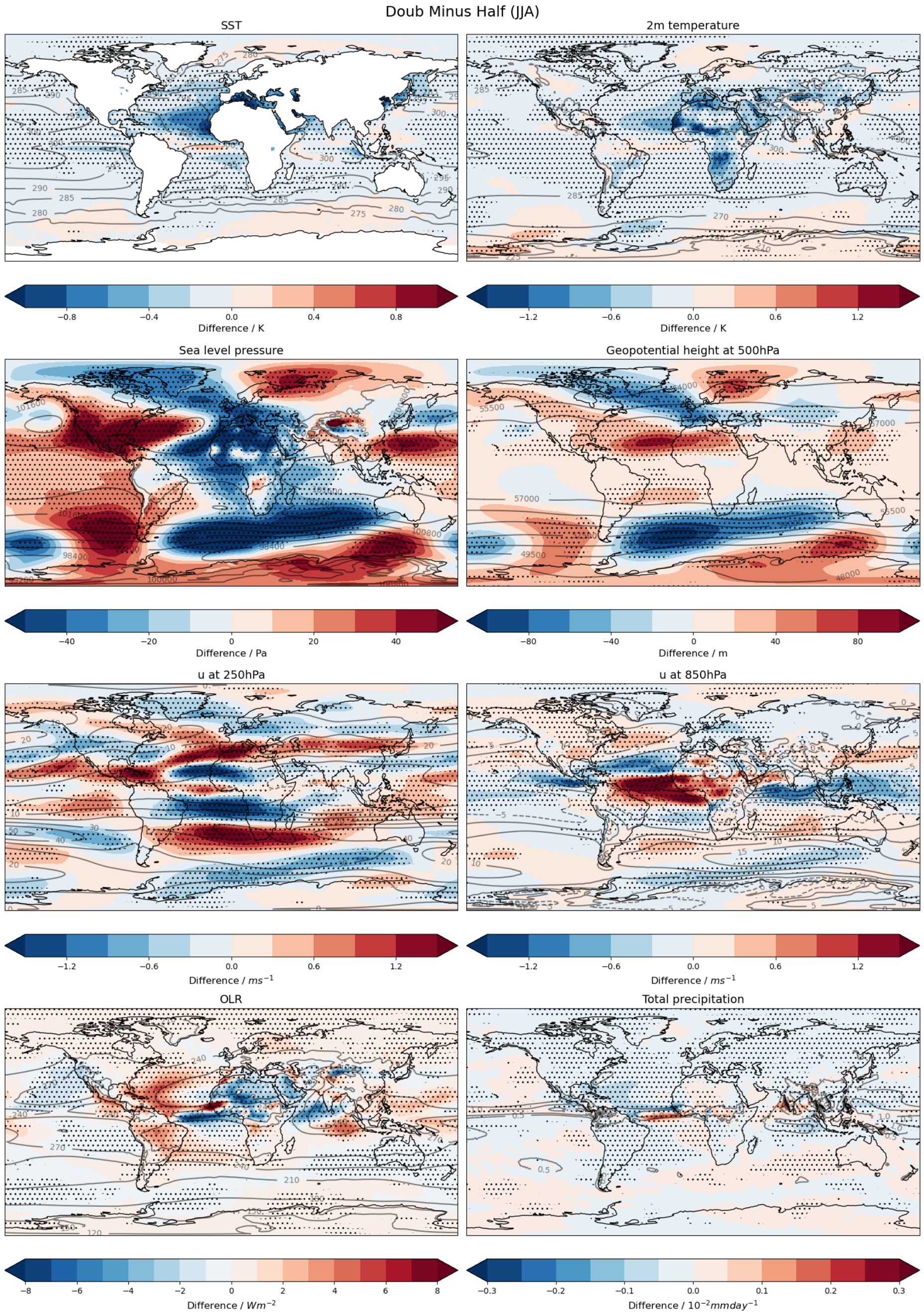

Figure 4: Maps showing JJA-mean DoubAer minus HalfAer (averaged over all members and years) to demonstrate the mean state response to increased aerosol forcing). Eight variables are shown, labelled above the map showing them, and with a corresponding colour bar below. Contours show BestAer climatology, and hatching indicates significant differences between DoubAer and HalfAer at the 5% level.