

Revisiting the Historical Wintertime Drying of the Mediterranean in the LESFMIP Simulations

David Avisar^{1,2} and Chaim I. Garfinkel¹

¹Fredy & Nadine Herrmann Institute of Earth Sciences, The Hebrew University of Jerusalem, Israel, ²The Department of Applied Mathematics, Israel Institute for Biological Reserach, Ness-Ziona, 7410001, Israel

Correspondence: David Avisar (david.avisar@gmail.com)

Abstract. Simulations from the Large Ensemble Single Forcing Model Intercomparison Project are used to isolate the impact of greenhouse gases (GHGs) and anthropogenic aerosols for historical (1850-2014) wintertime drying in the Mediterranean region, ~~and to clarify the importance of different dynamical mechanisms for the intermodel spread.~~ Increasing GHGs have already led to a clear ridging signal across the Mediterranean and a precipitation reduction of up to 15%. Anthropogenic aerosols, on the other hand, led to Mediterranean troughing in most models and hence cancelled out much of the GHG induced signal. The net effect when all forcings are present is a weak drying, and this weak drying and subtle ridging is consistent with recent observational evidence that the anthropogenically forced wintertime drying in the Mediterranean region has not yet robustly emerged. There is pronounced intermodel ~~and intramodel~~ spread in both the sea level pressure and precipitation responses to both GHGs and aerosols, however, and the relation between this spread and the spread in 9 different climatic metrics is explored to help clarify dynamical mechanisms and the causes of intermodel spread. A stronger tendency towards Mediterranean ridging is found in models and ensemble members with a more pronounced North Atlantic warming hole, a stronger stratospheric polar vortex, and to a lesser degree with a larger poleward shift of the eddy-driven jet. While these three sensitivities are as expected, others are not. Namely, a larger increase of global mean temperature is associated with troughing over the Mediterranean, opposite to naive expectations. Moreover, the single-forcing experiments indicate that a warmer land relative to the ocean (over the Mediterranean) is associated with troughing, rather than the previously proposed ridging. Other sensitivities are weak: the spread in the historical response cannot be explained by spread in shifts of the Hadley cell edge or the zonal-mean subtropical jet. ~~Many of these sensitivities differ, however, when considering each model's large ensemble as compared to treating all available members as a single large ensemble. This result highlights the importance of multi-model ensembles over single-model ensembles in revealing relations between dynamical and climate sensitivities.~~ Overall, the results of this work highlight that aerosols ~~had a detectable influence on~~ have influenced Mediterranean climate in the historical climate, ~~implying that removal of these aerosols will have an impact in future decades~~ and helped mitigate the greenhouse gas induced drying.

1 Introduction

Anthropogenic activity has led to a continual increase in greenhouse gases (GHGs) concentrations in the atmosphere over the past century. The impact of this increase on surface climate has been documented in successive IPCC reports and is expected to strongly affect human societies over the upcoming decades. In addition to warming, increased GHGs are expected to cause a drying of subtropical dry regions and a wetting of tropical and mid-to-high-latitude wet regions, as the atmospheric moisture holding capacity increases (Held and Soden, 2006) and moisture transport intensifies (Seager et al., 2010). This projected drying of the subtropics is particularly pronounced in the Mediterranean basin, and indeed the Mediterranean region has been called a ‘hot spot’ of climate change (Giorgi, 2006; Cos et al., 2022). Moreover, there are indications for a prolonged Mediterranean drying even under climate change mitigation scenarios (Delworth et al., 2022). Nevertheless, an updated comprehensive observational study has concluded that precipitation across the Mediterranean is mainly stationary during the last ~150 years, with high spatial and temporal variability, and does not support an anthropogenic-induced drying (Vicente-Serrano et al., 2025). Seager et al. (2024) also concluded that the expected drying has not yet emerged in most of the Mediterranean.

Although the future projections of the temperature increase and precipitation decrease in the Mediterranean region are statistically robust in nearly all models, they suffer from a large variability in their magnitude between the models (Cos et al., 2022; Garfinkel et al., 2020; Seager et al., 2024). Some of this inter-model variability is associated with the structural differences in each model, however some of it is associated with internal variability that is inherent in the chaotic climate system (Garfinkel et al., 2020; Zappa et al., 2015). Only large ensembles can be useful in isolating these two possibilities, or in detecting whether the Mediterranean drying signal has emerged from the noise caused by internal variability. Furthermore, the dynamical mechanisms that underlie the Mediterranean drying, as well as the drying across the subtropics in general, are not yet fully understood and are still under active research (Seager et al., 2019; He and Soden, 2017; Tuel and Eltahir, 2020; Keller et al., 2025), even as the uncertainty in future hydroclimate in these regions is driven almost exclusively by uncertainty in dynamical changes (Elbaum et al., 2022). Recently, the intensity of subtropical drying was found to be sensitive to poorly constrained parameterizations (Garfinkel et al., 2024). Finally, in other regions, climate drivers other than GHGs have been shown to play a large, and in some cases dominant, role in historical and future trends (e.g., Kang et al., 2024). Such findings motivate the examination of the role of these other drivers, and specifically of aerosols, on the large-scale atmospheric circulation and on multi-annual and decadal changes across the Mediterranean as well.

To address the gaps mentioned above, a we use model output from the Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP) (Smith et al., 2022) ~~has been proposed and launched, as,~~ LESFMIP is a component of the implementation plan of the World Climate Research Programs’s lighthouse activity on Explaining and Predicting Earth System Change (Findell et al., 2023). ~~The LESFMIP~~ It is composed of a coordinated series of historical model experiments in which large ensembles of global climate simulations are applied under a single-forcing configuration. The external forcings that can influence climate on multi-annual and decadal scales that are considered within the LESFMIP are: GHGs, anthropogenic aerosols, solar irradiance, volcanic aerosols, ozone, and land use. Moreover, the LESFMIP is designed as large ensemble experiments in order to

distinguish forced signals from the internal variability of the atmosphere. This combination of large ensembles and single forcings allows for the isolation of the impacts of each external driver on the changing climate.

In this work we present a first analysis of the historical LESFMIP simulations, based on data from large ensembles of 10 different climate models. The data used for this study and the methodology we adopt are introduced in Section 2. Motivated by the interest in the Mediterranean region wintertime – December, January, February (DJF) – historical drying, in Section 3 we address the impact that GHGs and anthropogenic aerosols have on precipitation and sea-level pressure in the LESFMIP models: in the multi-model mean, single-model means, and when considering each individual ensemble member of each model. In particular, we demonstrate that anthropogenic aerosols have a sizable impact across the Mediterranean region and even beyond, and furthermore, we explore the possibility that several climate phenomena (Mediterranean land-sea temperature contrast, the North Atlantic (NA) local minimum in surface warming, global mean near-surface temperature, stratospheric polar vortex, polar temperature amplification, tropical belt width, and near-surface jet) might be related to and have a role in the ~~inter- and intra-model~~ inter-model spread in the single forcing experiments of changing GHGs and anthropogenic aerosols. As will be shown below, these are the two forcings with the strongest impact on climate change in this region on multi-decadal timescales. We discuss the results and conclude in Section 4.

2 Data and Methodology

~~In this study we~~ We use the monthly data of the historical LESFMIP experiments ~~for~~ forced by time-varying greenhouse gases (hist-GHG), anthropogenic aerosols (hist-aer), solar irradiance (hist-sol), volcanic aerosols (hist-volc), and ~~Ozone~~ ozone (hist-totalO3), ~~with a focus on the first two of these. The historical~~. In addition, we use the ‘historical’ experiment that includes all of the forcings (hereafter referred to as ‘all-forcings’). The LESFMIP simulations span the period 1850-2020 (though we also use some of the DAMIP simulations from CMIP6 which end in 2014). The monthly data for the fields pr, psl, tas, and ua (precipitation flux, air pressure at sea-level, near-surface air temperature, and 3d eastward wind, respectively) was downloaded from the ESGF for the available ensemble members of 9 climate models. The models’ names and the number of available ensemble members for each single forcing experiment are specified in Table 1. For a 10th model, CESM2, the data was downloaded from www.earthsystemgrid.org. This model has ~~an~~ a hist-GHG experiment analogous to that of the LESFMIP. CESM2 did not perform the hist-aer experiment, however the influence of anthropogenic aerosols can be deduced by adding matching ensemble members of its hist-AAER and hist-BMB experiments (historical anthropogenic aerosols and historical biomass burning aerosols, respectively) followed by subtraction of a pre-industrial control run. Note that for CESM2 the meteorological fields ~~follow different naming and positions of the vertical levels, relative to~~ are archived at somewhat different pressure levels than in the LESFMIP models, though very similar. These differences are noted below when relevant. In all, ~ 300 ensemble members are available for hist-GHG ~~and~~, close to 240 for hist-aer, and almost 400 for the all-forcings. In the data analysis we focus on the northern-hemisphere winter season averages (DJF) and on the centennial-scale response that is defined as the difference between the average during 1990-2014 and that during 1851-1920, for the relevant season. We choose

long epochs to reduce as much as possible the noise and extract the forced signals. Figures for other periods are included in the supplemental material, to test the sensitivity to this choice, and show similar results.

Table 1. Number of model ensemble members for the various single-forcing experiments. ACC=ACCESS-ESM1-5, Can=CanESM5, CMCC=CMCC-CM2-SR5, GISS=GISS-E2-1-G, Had=HadGEM3-GC31-LL, IPSL=IPSL-CM6A-LR, MIROC=MIROC6, MPI=MPI-ESM1-2-LR, and Nor=NorESM2-LM

	ACC	Can	CMCC	GISS	Had	IPSL	MIROC	MPI	Nor	CESM2
hist-GHG	10	50	10	45	55	10	50	30	23	15
hist-aer	10	30	10	45	55	10	10	30	23	15
hist-volc	10	50	10or9	40	50	-	10	30	20	-
hist-solar	9	50	-	40	50	-	10	30	20	-
hist-totalO3	-	10	-	5	50	-	10	30	20	-
<u>historical</u>	<u>40</u>	<u>65</u>	<u>11</u>	<u>46</u>	<u>55</u>	<u>33</u>	<u>50</u>	<u>51</u>	<u>44</u>	<u>~</u>

90 The ~~downloaded monthly all-forcings~~ simulated pressure at sea-level (PSL) is compared to the monthly HadSLPv2 sea-level pressure gridded observations (Allan and Ansell, 2006). These observations have 5 degrees spatial resolution and cover the years 1850-2014.

The LESFMIP data is analyzed in the following ways. First, for each single-forcing experiment~~and for each season, and for the all-forcings run, we calculate the seasonal~~ (winter: December, January, February (DJF), spring: March, April, May
95 (MAM), summer: June, July, August (JJA), and fall: September, October, November (SON)) ~~, we calculate the~~ multi-model means ~~(i.e., the average of the individual models ensemble means) of the historical responses of the responses (1990-2014 vs. 1851-1920)~~ for the precipitation and sea-level pressure fields. The ~~centennial-scale response is defined, hereafter, as the difference between the seasonal average during 1990-2014 and that during 1851-1920. The~~ multi-model mean responses are calculated to asses the relative ~~significance contribution~~ of each of the single forcings for the response over the NA, Europe
100 and the Mediterranean. Second, for each model separately, we calculate the wintertime (DJF) ensemble mean responses for the eastward component of the wind at 700 hPa (691 hPa for the CESM2 model; hereafter U700), sea-level pressure, near-surface temperature (temperature at 992 hPa for CESM2), and precipitation (PRECC+PRECL fields for CESM2) for the hist-GHG~~and~~, hist-aer, and the all-forcings experiments, focusing on the above-mentioned regions. As shown below, there is substantial intermodel spread in the forced responses in these regions, and this motivates the last step of ~~our~~ the analysis.
105 Namely, we regress the wintertime response of the Mediterranean region ~~pressure at sea-level (denoted PSL hereafter) PSL~~ with the wintertime responses of 9 large-scale climatic measures/metrics that previous literature has linked to Mediterranean drying (Garfinkel et al., 2020). ~~This regression is performed in two ways: The regression is calculated~~ across all available ensemble members from all models, ~~and also for all available ensemble members for each model separately. In addition, we calculate the corresponding correlation coefficients, and the correlation coefficients based on the models' ensemble means.~~ Based on
110 these regressions, we analyze the possibility that the considered climatic measures contribute or are linked to the ~~intra-and~~ inter-model spread in the ~~simulated historical~~ Mediterranean PSL and precipitation change in the LESFMIP simulations. The definitions of the Mediterranean region PSL and the climatic measures considered are as follows.

- 115 1. **Mediterranean region mean PSL and precipitation:** Unless noted otherwise, the Mediterranean region is defined hereafter within 0°-30°E and 32°N-48°N; see red rectangle in Figures 4 and 1. ~~The mean PSL or precipitation are merely their spatial average within this Mediterranean region borders~~ 1 and 4. For the PSL we calculate an area-weighted (by the cosine of latitude) mean across the region, and for the precipitation a simple spatial mean.
- 120 2. **Mediterranean region land-ocean near surface temperature contrast:** Tuel and Eltahir (2020) argue that slower-paced warming over the Mediterranean relative to adjacent land regions will lead to a local ridge centered over the Mediterranean. Following Tuel and Eltahir (2020), we compute the land-sea temperature contrast in the Mediterranean region defined within 5°W-38°E and 28°N-48°N. Land ~~and ocean grid cells are determined based on the land-fraction field of the CESM2 model (followed by a ‘linear’ interpolation to the spatial grid of the other models).~~ Land (ocean) grid cells are chosen ~~such that~~ where the land fraction is ≥ 0.8 (≤ 0.2). The temperature difference is obtained by subtracting the ocean mean temperature from the land mean temperature (area-weighted means). As will be apparent below, land generally warms more than ocean in hist-GHG but cools more than ocean in hist-aer. Therefore, this index is positive for
- 125 hist-GHG and negative for hist-aer.
3. **North Atlantic warming-hole surface temperature:** The warming hole influences the Atlantic storm track and has been mechanistically linked to the ridge over Europe (Gervais et al., 2019; Keller et al., 2025). We consider ocean grid cells in the NA region where the multi-model mean temperature response indicates a warming hole, within 50°W-10°W, 45°N-65°N (Vacca et al., 2025). Mean temperature is calculated following area-weighting by the cosine of latitude.
- 130 4. **Global mean ~~near-surface~~ near-surface temperature** ~~÷~~ is calculated as an area-weighted average. A larger increase in global mean temperature will be associated with a stronger thermodynamic response to climate change. However, in many regions climate change arises not via a thermodynamic mechanism but rather through a change in the circulation that is not proportional to the global mean temperature rise (Grise and Polvani, 2016; Elbaum et al., 2022; Garfinkel et al., 2024)
- 135 5. **Stratospheric polar vortex:** is expected to ~~have influence on~~ influence precipitation in the Mediterranean region (~~Dai et al., 2025~~), ~~and is calculated as an average of~~, with a weak vortex enhancing precipitation in the western and central Mediterranean (Dai et al., 2025). The polar vortex strength is quantified as the zonal-mean eastward component wind at 10 hPa averaged between latitudes 60°N and 75°N. For the CESM2 model, we use the 14 hPa level of the U field.
- 140 6. **Polar amplification of near surface temperature:** Polar amplification is expected to lead to an equatorward jet shift, which in turn could lead to enhanced precipitation in the Mediterranean region (Garfinkel et al., 2020). Following Garfinkel et al. (2020), the polar amplification is defined as the ratio of the temperature response averaged between latitudes 60°N to 87.5°N (denoted ‘pole’) and that between latitudes 30°S to 30°N (denoted ‘trop’), hence

$$\frac{\left(T_{\text{pole}}^{1990\text{to}2014} - T_{\text{pole}}^{1851\text{to}1920}\right)}{\left(T_{\text{trop}}^{1990\text{to}2014} - T_{\text{trop}}^{1851\text{to}1920}\right)} . \quad (1)$$

Note that as long as both the ‘pole’ and ‘trop’ temperature responses are either positive or negative (as will turn out to be the case for hist-GHG and hist-aer, respectively), the polar amplification ratio will be positive. Here, as well, temperature means are calculated as area-weighted averages.

7. An equatorward shift of the eddy-driven jet or a contraction of the tropical circulation is expected to lead to more precipitation in the Mediterranean (Zappa et al., 2015; Schmidt and Grise, 2017a). Three different metrics of the edge of the tropical belt in the northern hemisphere are used. They are all calculated using the PyTropD python package that is based on Adam et al. (2018):

- (a) **Latitude of the Hadley-cell edge:** is calculated by ~~The~~the pyTropD ‘PSI’ function (with the ‘Psi_500’ method), which computes the zero-crossing latitude of the meridional mass stream function (ψ) at 500 hPa poleward of the tropical stream function extrema. This function is based on the zonal-mean meridional wind.
- (b) **Latitude of the subtropical jet:** is calculated by the pyTropD ‘STJ’ function (with the ‘adjusted_max’ method) as the latitude of the maximum of the zonal-mean zonal wind averaged between the 100 and 400 hPa levels minus that at 850 hPa.
- (c) **Latitude of the midlatitude Eddy-driven jet:** is calculated by the ‘EDJ’ function in pyTropD as the latitude of the maximum of the midlatitude zonal-mean zonal wind at 850 hPa.

8. **Averaged zonal-wind at 700 hPa in the Northern North Atlantic:** the ~~averaged zonal-wind is calculated within individual models’ ensemble means show high intermodel spread in the response of U700 in the North Atlantic region bounded by 50°W-5°E, 50°N-65°N, where the models ensemble means show high intermodel spread in the response of the eastward-wind component at 700hPa (green rectangle in Figures 5 and 6), motivating us to examine the possibility of downstream impacts from U700 in this region on the Mediterranean.~~

~~In addition to the above analyses, we calculate the bias of PSL in each model relative to the monthly HadSLP sea-level pressure gridded observations (Allan and Ansell, 2006). These observations have 5 degrees spatial resolution and cover the years 1850-2014. The biases were calculated based on the DJF means and after the model results were interpolated to the HadSLP grid.~~

3 Results

3.1 ~~Multi-model mean~~ Comparison of the historical responses response with observations and evaluation of the contribution of each single-forcings

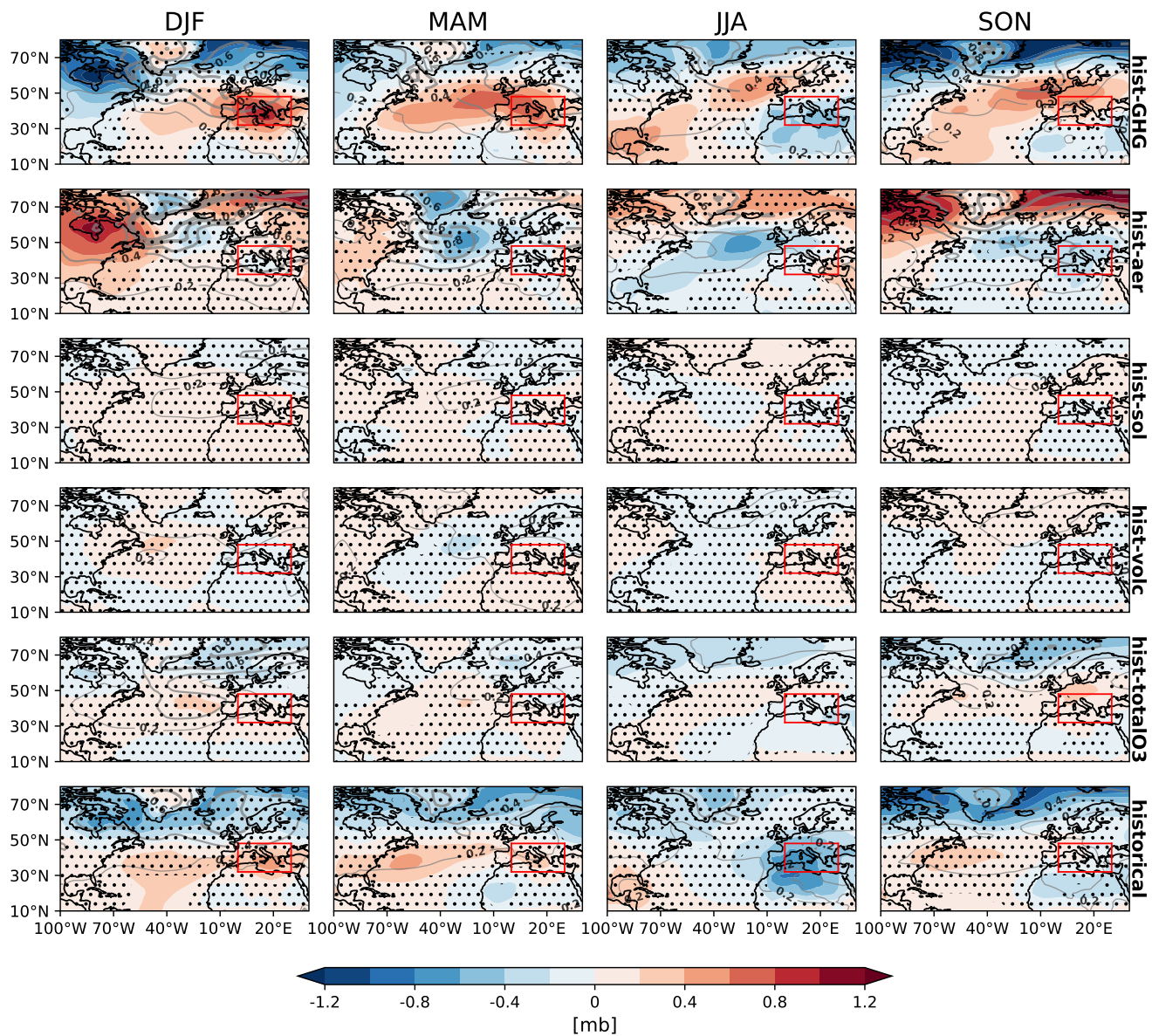
~~-0.7cm Multi-model means of the seasonal (column-wise) precipitation responses (1990-2014 vs. 1851-1920) in the single-forcing experiments (row-wise). The red rectangle marks the Mediterranean region as defined in the main text. Regions without stippling indicates where at least 80% of the models agree on the sign of the response. Gray contours indicate the intermodel spread in the forced response.~~

175 -0.7cmMulti-model means of the seasonal (column-wise) PSL responses (1990-2014 vs. 1851-1920) in the single-forcing experiments (row-wise). The red rectangle marks the Mediterranean region as defined in the main text. Regions without stippling indicates where at least 80% of the models agree on the sign of the response.

We begin with the multi-model mean ~~precipitation and PSL responses. These~~ PSL response for each single-forcing experiment and the all-forcings experiment; these are shown in ~~Figures 4 and 1, respectively, for the different single-forcings experiments~~ in each season Figure 1. Grid-points where at least 80% of the models agree on the response sign are not stippled. Gray contours represent the standard deviation of the intermodel responses (scaled by in-line values and by the line-thickness). ~~Among the single-forcing experiments, A robust anomalous ridge over the Mediterranean develops in hist-GHG and hist-aer indicate the largest historical response signals for both the precipitation and PSL fields (a detectable and robust responsesignal in the precipitation field is also indicated under hist-totalO3 during summer). Under hist-GHG, a robust negative precipitation~~ response is indicated during all seasons across the Mediterranean and across the NA Ocean between the Iberian peninsula and the Caribbean. A robust positive PSL response is found across the Mediterranean under hist-GHG mainly during winter, but also during spring and fall. Under hist-aer, the to a lesser degree in the all-forcings in DJF. More broadly, a positive North Atlantic Oscillation (NAO) response is also seen in both experiments in DJF. The DJF NAO response is also evident in MAM and SON, however the foci over the Mediterranean is less pronounced. The effect of GHGs and all-forcings in JJA is qualitatively different, and will be the focus of a future paper. The effect of aerosols over Europe is nearly as strong as that of GHGs, however there is less agreement across models as to the sign of the response, except in JJA, which again is outside the scope of this paper which focuses on winter. These effects are similar if we average over different centennial-scale periods (Supplemental Figure S1).

Figure 2 shows the observed PSL response between the same periods. A meridional dipole response is evident in DJF and ~~MAM generally resembling the multi-model mean precipitation-all-forcings and GHG-induced responses in the LESFMIP models (the pattern in the observations in DJF is similar if the latter epoch starts at 1980 or 1970 instead of 1990; not shown). However, the signals differ in magnitude (note the different colorbars for Figure 1 and 2), which is not surprising since the observed anomalies include both the forced response and also internal variability. In order to assess whether the modeled response is discrepant with the observed PSL response, in Figure 3 we contrast, separately for two regions, the distribution of~~ the PSL response as obtained in the all-forcings LESFMIP experiment for all of the models' ensemble members (color-coded bars) with the observed response (vertical thick black line). For the Mediterranean region (Figure 3a, solid red box on Figure 2) the observed response lies well within the LESFMIP historical distribution, and PSL responses across the entire field are mainly opposite in sign relative to that under hist-GHG, but are less robust (and thus stippled), and sometimes weaker (by their absolute value) across the Mediterranean and the NA. The 36% (143/396) of the ensemble members exceed the observed PSL response. If we instead focus on the region to the west of France in which the observed response is stronger (Figure 3b, dashed red box on Figure 2), the LESFMIP multi-model standard deviation of the PSL historical response, which corresponds to the mean shows a weak response, however 8% (32/396) of the ensemble members nonetheless exceed the observed PSL response. Hence, there is no evidence for a model vs. observations discrepancy over this period.

Multi-Model Mean of PSL response, 1990-2014 vs. 1851-1920



-0.7cm

Figure 1. Multi-model means of the seasonal (column-wise) PSL responses (1990-2014 vs. 1851-1920) in the single-forcing and all-forcings experiments (row-wise). The red rectangle marks the Mediterranean region as defined in the main text. Regions without stippling indicate where at least 80% of the models agree on the sign of the response. Gray contours quantify the multi-model standard deviation and indicate the intermodel spread. Similar figures but for different averaging periods are shown in Supplemental Figure S1.

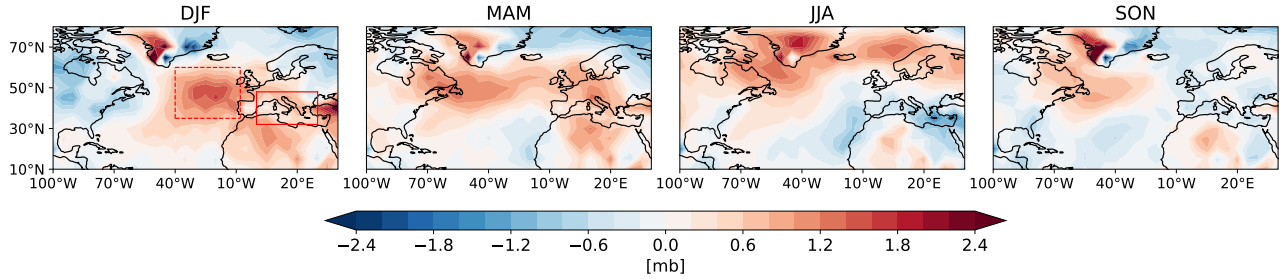


Figure 2. Seasonal (column-wise) PSL responses (1990-2014 vs. 1851-1920) in the HadSLPv2 observations.

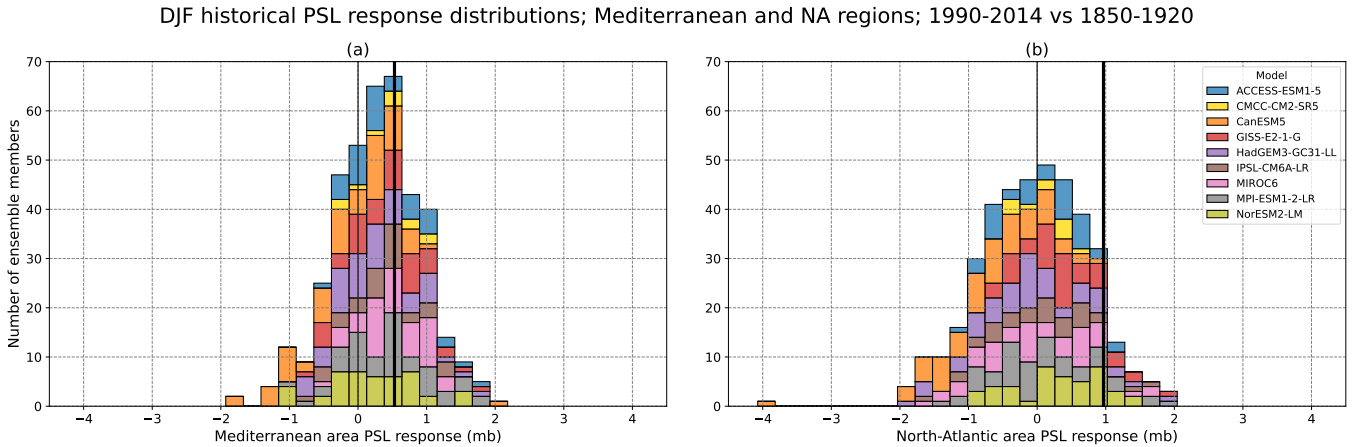


Figure 3. DJF PSL response distribution in the historical experiment by the models' (color-coded) ensemble members in the Mediterranean (left; solid red box on Figure 2) and in the North Atlantic (right; dashed red box on Figure 2). The vertical thick black line refers to the observed response based on the HadSLPv2.

210

The multi-model mean precipitation response is shown in Figure 4. Consistent with the hist-GHG and all-forcings PSL meridional dipole in Figure 1, precipitation decreases in the Mediterranean region and increases in Northern Europe in response to GHGs in all seasons, and to a lesser degree in all-forcings in DJF, MAM, and SON. The precipitation signals in the European sector extend westward towards North America, and include increased precipitation in Canada and reduced precipitation in the Caribbean. Aerosols, on the other hand, have an opposite signed response with increased precipitation over North Africa and

reduced precipitation over Northern Europe, but there is generally less intermodel agreement than for GHGs. These patterns are similar if we average over different centennial-scale periods (Supplemental Figure S2). The responses in hist-GHG and hist-aer are larger than in any of the other single forcing experiments for both precipitation and PSL (though note a robust response in JJA in hist-totalO3 during summer for precipitation that mimics the GHG induced response). In most regions aerosols and GHGs have opposite signed responses for both PSL and precipitation, and hence the historical all-forcings changes are relatively muted and are less robust across models than for GHG alone. The spread across the different model-means, models' means is quantified with the multi-model standard deviation gray contours. This spread is higher across the NA, Europe and the Mediterranean, relative to that across other regions. This is in particular notable during winter in both hist-GHG and hist-aer. Intermodel spread in winter precipitation is particularly large near the coastline of North Africa than in other regions for both hist-GHG and hist-aer.

Comparing the simulated precipitation to observations is challenging due to the unreliability and limited nature of observed precipitation data in the early parts of the historical simulation. Nevertheless, more recent precipitation data is more reliable and extensive, and we therefore show in Supplemental Figure S3 trends in precipitation for each all-forcings LESFMIP model ensemble mean in the periods considered by Vicente-Serrano et al. (2025). In agreement with Vicente-Serrano et al. (2025), recent trends in precipitation (i.e., 1981-2014) indicate wettening in most of the Mediterranean (and not only there) with the majority of the models indicating a positive precipitation trend. Drying trends in most of Mediterranean in the historical simulations only emerge for start-dates of 1931 or earlier, however many individual models simulate wettening trends over Southern Europe (especially Southern France, Italy, and the Balkans) even for a start date in 1871.

Based on the above comparison of the LESFMIP PSL response and the observed one, it may be suggested that the ridging (and drying) signal across the Mediterranean has not yet emerged out of the system's internal variability, in agreement with Vicente-Serrano et al. (2025) and Seager et al. (2024), because aerosols have canceled out a substantial portion of the GHGs-induced drying. That is, GHGs and aerosols have had a strong but opposite impact on historical wintertime climate in the Mediterranean. Since the responses to aerosols and GHGs are more significant than the responses to any of the other three forcings on these centennial timescales, we focus only on aerosols and GHGs for the rest of this paper. We also exclusively focus on DJF, the rainy season for most of the Mediterranean region, in the rest of this paper.

3.2 Single-model Ensemble ensemble means historical responses

Section 3.1 identified a pronounced forced response to both aerosols and GHGs in winter over Europe and the Mediterranean, but with large intermodel spread in both PSL and precipitation. We now focus on this spread across models, and specifically plot in Figures 5 and 6, for each model separately, the DJF ensemble mean responses in the hist-GHG and hist-aer experiments, respectively, of the eastward wind component at 700 hPa, pressure at sea-level, near-surface temperature, and precipitation (denoted U700, PSL, T2m, and Prcip., respectively, in the figures). The U700 and T2m fields are shown due to their relevance to the indices discussed in section 2. The comparable figure for all-forcings is in Supplemental Figure S4.

The response of U700 to GHGs differs strongly in the North Atlantic differs qualitatively across the models. About half show an intensification of U700 across the UK and southern Scandinavia, and the qualitative nature of the response over the

Multi-Model Mean of Prcip. response, 1990-2014 vs. 1851-1920

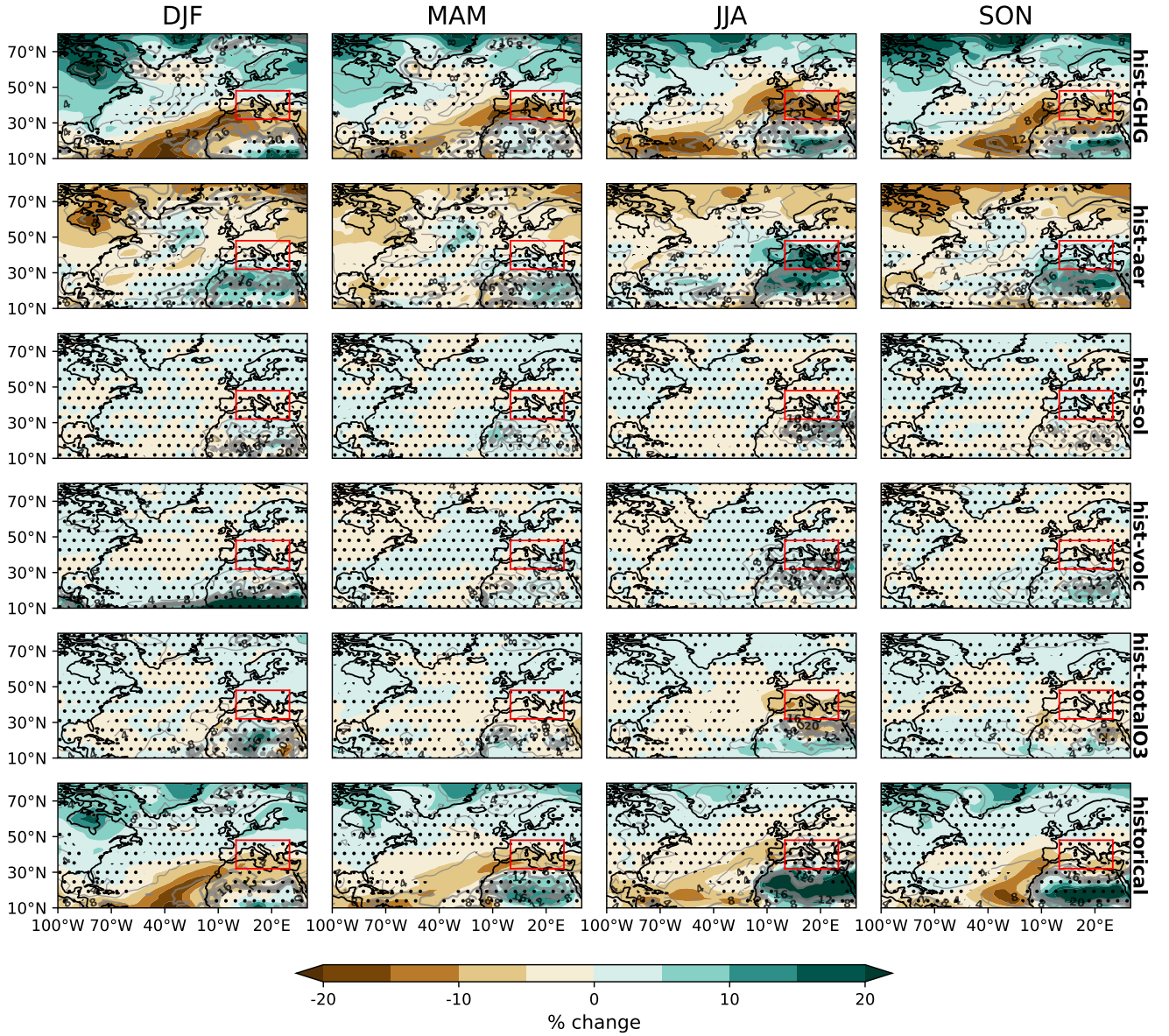


Figure 4. Multi-model means of the seasonal (column-wise) precipitation responses (1990-2014 vs. 1851-1920) in the single-forcing experiments (row-wise). The red rectangle marks the Mediterranean region as defined in the main text. Regions without stippling indicates where at least 80% of the models agree on the sign of the response. Gray contours quantify the multi-model standard deviation and indicate the intermodel spread. Similar figures but for a different averaging period are shown in Supplemental Figure S2.

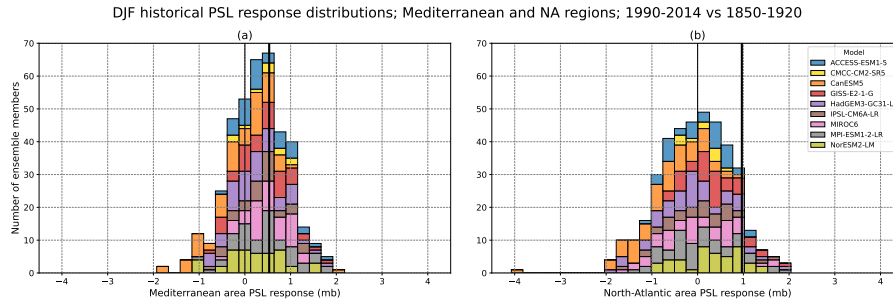
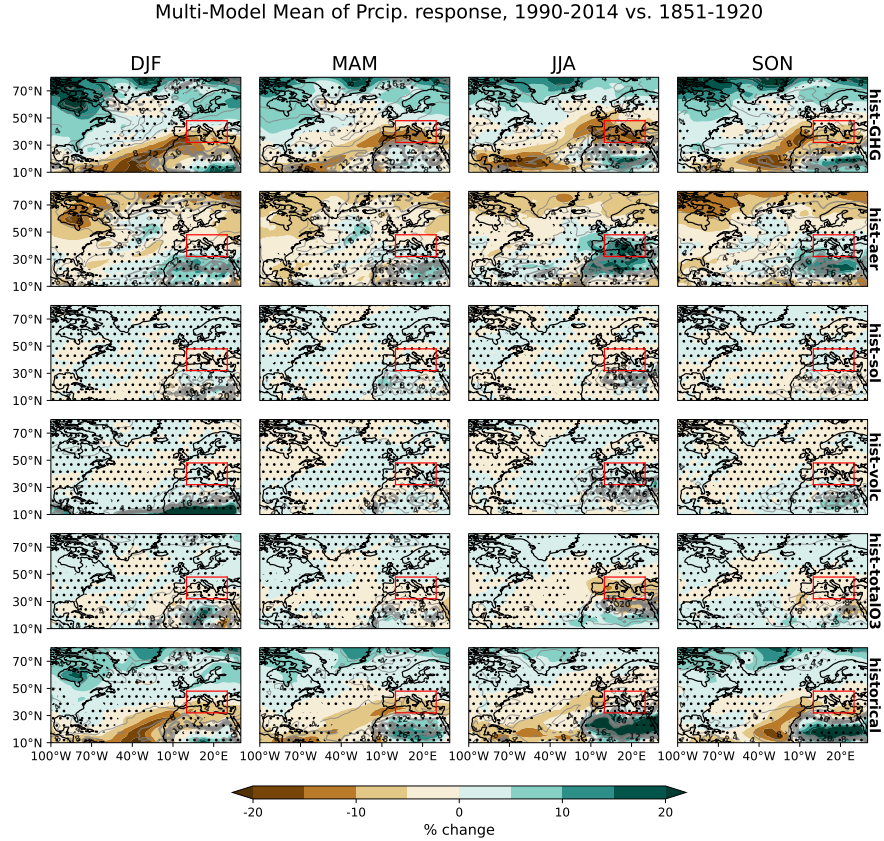


Figure 5. Ensemble means of the DJF ~~historical~~-responses for the U700, PSL, T2m, and Precip. (column-wise) fields, obtained for each model (row-wise) in the hist-GHG experiment. The green, red, and blue rectangles mark region of high intermodel variability of U700, the Mediterranean region, and region of the NA warming-hole, respectively. Similar figures but for a different averaging period are shown in Supplemental Figure S5.

~~North Atlantic varies substantially among models~~but others shown an opposite response. The ensemble means of all of the models, except that of CanESM5, show an increase in PSL across the Mediterranean region, but with variable intensity. About half of the models show a positive PSL response also across Europe and the NA. All of the models show a positive near-surface temperature response, except over the NA south of Greenland, hereafter referred to as the ‘warming hole’ region (Rahmstorf et al., 2015). Four of the models show a cooling in this region of the NA (CESM2, CMCC, MIROC, and MPI), and several more show a local minimum in warming (GISS, NorESM). It is worth noting that the ensemble mean of the CanESM5 model shows the most significant warming response, in particular across the Atlantic, Africa, and Eurasia—; this model is known for its particularly high climate sensitivity (Zelinka et al., 2020). Over most of the region presented in Figures 5 and 6, the ~~ensemble-mean~~-responses in hist-aer are opposite in sign relative to these in hist-GHG. The main exception is seen for HadGEM3-GC31-LL which shows a positive PSL response across almost the entire field shown, that is even larger than that in hist-GHG. Similar patterns are seen if we average over different centennial-scale periods (Supplemental Figures S5 and S6).

In Table 2 we specify for each model ~~ensemble-mean-the averaged-the ensemble-mean~~ precipitation response (in percent; values are rounded to the closest integer) ~~within-averaged over~~ the Mediterranean region. In parenthesis are the averaged precipitation responses within that region but southward of 40°N. Bold values mark that at least 80% of the ensemble members agree on the sign of the ~~averaged-area-averaged~~ precipitation response. Italicized values mark that at least half of the ensemble members agree on the sign (but less than 80%).

Under the GHG forcing, the ensemble means of all models show a reduction in precipitation within the Mediterranean region ~~over the historical period~~between the historical periods, which vary in the range ~~-1 to -10%~~-1 to -10% between the models. According to Figure 5, the precipitation reduction also ~~vary-varies~~ between models in its spatial extent within the Mediterranean, as well as beyond it. The drying is more pronounced south of 40°N (within the Mediterranean region)



-0.7em

Figure 6. Ensemble means of the DJF ~~historical~~-responses for the U700, PSL, T2m, and Prcip. (column-wise) fields, obtained for each model (row-wise) in the hist-aer experiment. The green, red, and blue rectangles mark region of ~~high~~ high intermodel variability of U700, the Mediterranean region, and region of the NA warming-hole, respectively. Similar figures but for a different averaging period are shown in Supplemental Figure S6.

Table 2. Averaged precipitation responses (in %) across the Mediterranean region (extracted from the fields presented in Figures 5 and 6). The values in parenthesis refer to the Mediterranean region southward of 40°N. Italicized and bold values are explained in the main text.

Exp./Model	ACC	Can	CMCC	GISS	Had	IPSL	MIROC	MPI	Nor	CESM2
hist-GHG	-5 (-8)	-3 (-15)	-10 (-12)	-3 (-7)	-3 (-7)	-7 (-12)	-4 (-8)	-7 (-13)	-6 (-11)	-6 (-9)
hist-aer	<i>1 (2)</i>	-8 (-5)	-4 (-1)	1 (2)	-4 (-1)	+0 (3)	-0 (1)	-1 (0)	5 (11)	
hist-(AAER,BMB)										2 (3)

and reaches ~~-15%~~-15%. Under hist-aer, some models indicate drying while others indicate wetting or negligible change; the response vary in the range -8% to 5%. Southward ~~to~~of 40°N (within the Mediterranean), the response tends to be more
270 positive.

Last, the PSL biases for the models ensemble means ~~, for both the hist-GHG and hist-aer experiments,~~for the historical
~~experiment~~experiment are presented in ~~the supplemental material~~supplemental Figure S7. The biases range between ~~-10~~-10 and 10 mb
across the NA, Europe and Mediterranean ~~area~~ and are very similar, in each model, ~~between the two single forcing experiments~~
~~(hist-GHG and hist-aer)~~ and across the three temporal periods. This indicates that the models responses are substantially
275 smaller then their bias. Most of the models show a meridional dipole bias pattern: negative across northern Europe and positive
southward and across the NA. A somewhat different pattern is observed for GISS-E2-1-G, HadGEM3-GC31-LL, MIROC6,
and MPI-ESM1-2-LR. Finally, there does not appear to be a clear relationship between each model’s bias and its PSL response.
For example, although the PSL biases in the top 4 rows of Figure ~~S1~~S7 are very similar, their PSL responses (in particular
for CanESM5) differ substantially (Figure 5). Our various attempts to use mean-state biases to "predict" the forced response
280 in a given model were not successful (details excluded for brevity), which suggests that differences across the models in
processes unrelated to the large-scale time-mean dynamical state are more important factors in explaining intermodel spread
in the response. This motivates the next section, which tries to unravel what factors might, instead, explain the spread in the
responses.

3.3 Towards Understanding the Intermodel Spread

285 In this section we examine what large-scale climate or circulation response might explain the spread in the regional wintertime
response of the PSL and precipitation to identical forcing over the historical period. We do this by analyzing the multi-model
correlations of the Mediterranean region PSL responses and the responses of a group of climatic or circulation measures. The
left (right) columns in the following figures refer to results for the hist-GHG (hist-aer) experiment.

The upper panels of Figure 7 show scatter plots of the DJF PSL ~~historical~~ response versus the precipitation response. Each
290 data point (color-coded dots) corresponds to an ensemble member of a specific model. The corresponding linear regression
equation~~and~~, the correlation coefficient, and its p-value, calculated based on all >240 individual ensemble members, are
indicated in red text outside the red brackets. Stars mark the ensemble-mean responses of each model. ~~(In Section 4 we~~
~~refer to the correlation coefficients that were calculated based on the responses of each model separately),~~and the correlation
and p-value when using these ensemble-mean responses only are indicated in the square brackets. Figure 7 demonstrates the

Figure 7. Mediterranean PSL response vs. responses of Mediterranean precipitation (top panels), Mediterranean land-ocean near-surface temperature difference (middle panels), and near-surface temperature within the NA warming-hole region (blue rectangle in Figures 5 and 6; bottom panels) obtained for each ensemble member of the LESFMIP models (color coded dots) for the hist-GHG and hist-aer experiments (left and right panels, respectively). Stars refer to the individual models' ensemble-means. The correlation coefficient and its p-value when using the individual ensemble-members (-means) are indicated outside (inside) the square brackets.

295 expected strong relationship between PSL and precipitation under both GHG and aer forcings, with a correlation coefficient of about -0.85 ~~under both GHG and aer forcings. It also indicates a~~ and a reduction of precipitation by 7.5mm/day for every 1hPa increase in PSL. The fit is similar if we use the ensemble-mean response in each model only. However, there is large inter- and intra-model spread in the Mediterranean precipitation and PSL responses to the single forcings. The spread in the PSL response varies in the range ~~$\sim(-2.5) \sim (-2.5)$~~ to 3 mb under the GHG forcing and ~~$\sim(-2) \sim (-2)$~~ to 3.5 mb under aer. Still, $\sim 78\%$ of the

300 hist-GHG simulations show a positive PSL response and $\sim 74\%$ show a reduction of precipitation. Similar spread is indicated in the hist-aer experiment with $\sim 52\%$ of the simulations showing a negative PSL response and $\sim 32\%$ a precipitation increase. The dominant role of ~~dynamical uncertainty~~ uncertainty in circulation responses for uncertainty in precipitation changes in this region is consistent with previous work (Zappa et al., 2015; Elbaum et al., 2022). The rest of this section is devoted to understanding this intermodel and across-ensemble spread.

305 The middle row in Figure 7 ~~show~~ shows the PSL response vs. that of the land-ocean near-surface temperature difference (see Section 2 for definition). In both the hist-GHG and hist-aer experiments, a more pronounced relative land warming is associated with troughing over the Mediterranean. This is in contrast with the results of Tuel and Eltahir (2020) who found that a relative Mediterranean cooling (that is, a relative land warming) is associated with ridging and drying. It is worth noting that the CanESM5 model contributes substantially to this relationship: if excluded, the correlation coefficients reduce to ~~-0.31 and~~

310 ~~-0.40 -0.34 and -0.35~~ for hist-GHG and hist-aer, respectively, but these are nonetheless opposite to the relationship predicted by Tuel and Eltahir (2020). ~~In conclusion~~ Thus, under both single forcings experiments there seems to be a relationship between the PSL response and the response of the land-ocean temperature contrast, but it appears to be opposite to that suggested by previous work. We return to this in the discussion.

Next, we consider the relationship between the ~~response of the~~ responses of Mediterranean PSL and ~~that of~~ the NA Ocean

315 temperature in the 'warming hole' region defined in Section 2; the corresponding scatter plots are shown in the bottom row of Figure 7. For both single forcing experiments, a stronger NA warming hole is associated with a stronger Mediterranean ridge, which is consistent with results of previous studies (Delworth et al., 2022; Keller et al., 2025). The correlation between the two responses is higher under the hist-aer experiment than under hist-GHG. In this case, excluding the CanESM5 model reduces the correlation coefficients to ~~-0.42 and -0.71~~ -0.43 and -0.71 for hist-GHG and hist-aer, respectively, but this is nonetheless

320 highly statistically significant (p-values of 10^{-12} and 10^{-33} for hist-GHG and hist-aer, respectively).

In Figure 8 we show the correlations between the PSL response and the responses of three large-scale climatic measures: ~~mean-global-mean~~ near-surface ~~global~~ temperature (top), stratospheric polar vortex (middle), and polar amplification of near-surface temperature (bottom). In both hist-GHG and hist-aer, a larger global mean warming is associated with troughing over the Mediterranean. This result is opposite to what one might naively have assumed if the PSL response was thermodynamically driven, and indicates that climate sensitivity (i.e., global-mean surface temperature response) is distinct from dynamical sensitivity (atmospheric circulation response; Grise and Polvani, 2016). ~~Notably, and in particular for hist-GHG, this~~ This negative correlation is ~~due mainly to differences in even larger if we focus on~~ the ensemble mean response for each model ~~and is not manifested by-, however, in~~ most of the ~~individual models ensembles; the spread within most of the models shows essentially no relationship between the PSL and global mean temperature responses~~ cases, the correlation is near-zero if we focus on the spread across each individual model's ensemble (not shown). According to the middle row of Figure 8, a stronger stratospheric polar vortex is associated with ridging over the Mediterranean. This trend is consistent with findings that associate weak stratospheric polar vortex with negative NAO periods and a corresponding increase in Mediterranean precipitation (Domeisen, 2019; Zhang et al., 2023, 2024; Baldwin et al., 2021). According to Figure 8, this relation is even stronger based on the ensemble-means than based on the individual ensemble members, and is stronger in hist-aer than in hist-GHG.

Next, we turn to the polar amplification of near-surface temperature (bottom row of Figure 8). Before discussing its relationship with the PSL response, we first clarify the sign of this metric for the two single forcing experiments. Under GHG (aer) forcing, both the pole and tropics warm (cool), but with a larger temperature change in the pole than the tropics. Hence, since the metric is defined as the ratio of the temperature changes, its sign is generally positive in both ~~historical~~ experiments. However, the slope of the relationship between the PSL response and polar amplification differ in sign between the hist-GHG and hist-aer. This sign difference originates from opposite dynamical responses in the two experiments. For hist-aer, as the pole cools more than the tropics (i.e., as the polar amplification index gets more positive; not shown here), a poleward shift of the jet might be expected and, as a result, a stronger ridge over the Mediterranean. Similar reasoning for the more pronounced polar warming in hist-GHG would explain the troughing trend: polar amplification leads to an equatorward jet shift and a trough over the Mediterranean. Hence, it is to be expected that the signs of the slope in the bottom row should be opposite, as indeed is evident in the bottom row of Figure 8. Still, there is a substantial difference between the magnitude of the correlation coefficients for the two experiments, with a more profound relation for the hist-aer. Moreover, in this case too, the correlation between polar amplification and Mediterranean PSL is much more pronounced in the multi-model data and is much less evident in each individual model's large ensemble.

In Figure 9 we analyze the relationship between the PSL response and that of the northern edge of the tropical belt using three metrics: zero-crossing latitude of the meridional stream function at 500 hPa, latitude of the subtropical jet, and latitude of the eddy-driven jet (first three rows from the top, respectively). ~~In the bottom row, we show the correlation between the PSL response with that of the averaged U700 in the Northern NA where there is a high intermodel spread (region marked by green rectangle in Figures 5 and 6).~~ For the hist-GHG (hist-aer) experiment, about 92% (72%) of the ensemble members show a poleward (equatorward) shift of the tropical belt edge according to the zero-crossing of the stream function metric. However,

Figure 8. Mediterranean PSL response vs. responses of global near-surface temperature (top panels), stratospheric polar vortex (middle panels), and polar temperature amplification (bottom panels) obtained for each ensemble member of the LESFMIP models (color coded dots) for the hist-GHG and hist-aer experiments (left and right panels, respectively). Stars refer to the ensemble-mean of each model. The correlation coefficient and its p-value when using the individual ensemble-members (-means) are indicated outside (inside) the square brackets.

the correlation between this zonally averaged shift and the Mediterranean PSL response does not exceed $|0.20|$. Even more striking is that the relationship in hist-GHG is opposite to naive expectations: members with a stronger poleward shift of the zero-crossing of ψ at 500 hPa, which corresponds to a stronger Hadley cell expansion, actually tend to have more of a trough in the Mediterranean (Schmidt and Grise, 2017b). However, when considering each model ensemble separately (not shown),
 360 eight of the ten individual ~~model-ensembles-models~~ indicate an opposite relation (with correlations between 0.26 and 0.67) which is more in line with expectations (a negative correlation ~~within a given model's large ensemble~~ is only found for two models, CESM2 and CMCC). For hist-aer, three of the ten models show a negative correlation.

According to the subtropical jet metric, ~~~53~~52% of the hist-GHG and hist-aer ensemble members show a poleward and equatorward shift, respectively. There is a lack of relation between the PSL response and the latitude of the subtropical jet
 365 response, which is also supported by the responses of the individual models. In contrast to the ~~previous two~~ two latter metrics, there is a strong relationship between the PSL response and that of the latitude of the eddy-driven jet (third row of Figure 9), both in hist-GHG and hist-aer: members with a stronger poleward jet shift in the zonal mean tend to have a stronger ridge. ~~This relation is also supported by most~~ All of the individual models' ensembles support a positive correlation (although a single model in each of the hist-GHG and hist-aer experiments has a correlation that is smaller than 0.25).

370 In the bottom row of Figure 9, we show the correlation between the PSL response with that of the averaged U700 in the Northern NA region where there is a high intermodel spread (region marked by green rectangle in Figures 5 and 6). The relationship between the Mediterranean ridge and the near-surface jet is even stronger ~~if we consider the jet in the North Atlantic region (bottom row of Figure 9)~~ than that indicated for the zonal mean eddy-driven jet latitude response. Specifically, a stronger U700 between Iceland and the British Isles, which corresponds to a poleward shift of the jet, is highly correlated with
 375 Mediterranean ridging, ~~as expected,~~ both in hist-GHG and hist-aer. This indicates that intermodel spread in the Mediterranean and in the Atlantic are linked.

4 Discussion and Conclusions

Climate models consistently predict a reduction in precipitation during winter in the Mediterranean region in response to climate change. While this reduction has not yet emerged in observations in most of the basin, ~~see e.g. Figure 13 of Seager et al. (2024)~~

Figure 9. Mediterranean PSL response vs. responses of the subtropical latitude of ψ zero-crossing (top row), latitude of the subtropical jet (second row), and of the latitude of the midlatitude eddy-driven jet in the zonal mean (third row) and in the North Atlantic (bottom row) obtained for each ensemble member of the LESFMIP models (color coded dots) for the hist-GHG and hist-aer experiments (left and right panels, respectively). Stars refer to the ensemble-mean of each model. The correlation coefficient and its p-value when using the individual ensemble-members (-means) are indicated outside (inside) the square brackets.

380 (Seager et al., 2024; Vicente-Serrano et al., 2025), it is expected to have serious negative impacts on society and is expected to emerge in the next decade or two (Shaw et al., 2024). ~~In practice, however, models suffer from large variability in the magnitude of the precipitation decrease (Seager et al., 2024), motivating the need to understand the model spread, the mechanisms that drive the drying, and to develop early warning capabilities.~~ It is not clear to what extent the lack of emergence is due to a weak GHG induced response in this region, to a low signal-to-noise ratio, or due to cancellation between the GHG induced drying
 385 and other climate forcings. Furthermore, there is substantial intermodel spread in projected drying which is only partially understood (Elbaum et al., 2022; Garfinkel et al., 2020).

In this study we use the Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP) simulations from 10 models and address part of the above needs. The single forcing configuration of these simulations allows for distinguishing the separate contributions of the different forcings to changes in climate, in parallel with a comparison of the all-forcing
 390 historical simulation to observations. Because there is substantial intermodel ~~and intramodel~~ spread in the sea-level pressure (PSL) and precipitation responses, we also test whether this spread can be explained by the spread in the historical change in several climatic measures: Mediterranean land-sea temperature contrast, near-surface temperatures over the North Atlantic (NA), global mean temperature, stratospheric polar vortex, polar amplification, tropical belt width, and averaged U700 in the northern North Atlantic.

395 We first inspect the multi-model mean centennial changes (1990-2014 vs. 1850-1920) of ~~the~~ precipitation and PSL ~~among the various in each of the five~~ single forcing experiments: GHGs, anthropogenic aerosols, solar irradiance, volcanic aerosols, and ozone (hist-GHG, hist-aer, hist-sol, hist-volc, and hist-totalO3, respectively). Across North America, North Atlantic, Europe, and the Mediterranean region, the ~~response signals responses~~ under hist-GHG and hist-aer are comparable in magnitude. Relative to these, the magnitude of the response ~~signals under to~~ the three other ~~experiments are mostly forcings (ozone,~~
 400 volcanoes, solar) is much smaller. Under hist-GHG, the precipitation response across the Mediterranean is negative and robust during all seasons and that of PSL is positive (i.e., ridging) and robust, except during summer. The decline in precipitation across the Mediterranean approaches 15% in some models in hist-GHG. ~~Generally~~ (Figures 1 and 4). However, the multi-model mean responses under hist-aer are opposite in sign and comparable in magnitude ~~relative to these under the GHG forcing, but are less robust across the Mediterranean region.~~ Because of this cancellation by aerosols of part of the GHG induced drying, the

405 net changes in the historical simulation are weaker than in response to GHG only, and there is no indication of a model vs. observations discrepancy in historical changes (Figure 3).

During the wintertime (and also spring), there is substantial intermodel spread in the PSL response across the NA, Europe, and the Mediterranean, both under hist-GHG and hist-aer. Furthermore, around half of the models ensemble means indicate a warming-hole in the NA Ocean during DJF under hist-GHG, and there is also a large intermodel spread in e.g., polar
410 amplification, the stratospheric polar vortex, land vs. ocean temperature gradients in the Mediterranean region, and globally averaged temperature increase (Figures 5 and 6). Section 3.3 quantifies whether these intermodel spreads are related to intermodel spread in the Mediterranean PSL response. ~~The corresponding relationships are summarized in Table ??, for the hist-GHG and hist-aer experiments. The table indicates both the multi-model correlation coefficients obtained for the wintertime PSL responses vs. the responses of the group of large-scale measures (presented in Section 3.3), and additionally,~~
415 ~~under the ‘relative’ columns, it indicates whether the large ensemble for each individual model shows a similar sensitivity to the multi-model. Namely, the ‘relative’ column specifies the ratio between the averaged individual models response correlation coefficients and the multi-model coefficient (specified in the columns to the left of ‘relative’). The closer the ratio to +1, the more the correlation coefficients of the individual models resemble (or represent) the multi-model coefficient. Based on the multi-model correlations, the wintertime PSL-historical~~ wintertime PSL response is moderately to highly correlated with the
420 responses of the Mediterranean land-ocean temperature contrast, the NA near-surface temperature within the warming-hole region, the stratospheric polar vortex, the latitude of the eddy-driven jet, and averaged U700 across region of high variability in the NA. In particular, we find no evidence for the mechanism proposed by Tuel and Eltahir (2020): a warmer land relative to the ocean (over the Mediterranean) is associated with troughing rather than ridging. Furthermore, a stronger NA warming hole, a stronger stratospheric polar vortex, poleward shift of the eddy-driven jet and U700 are correlated with a Mediterranean ridging.
425 A high correlation is also obtained between the PSL response and that of the polar amplification, but only under hist-aer. In addition, the LESFMIP simulations indicate that a larger increase of global mean temperature is associated with troughing over the Mediterranean, which indicates that dynamical sensitivity in this region is unrelated to climate sensitivity (Grise and Polvani, 2016). The importance of the NA warming hole, and the lack of importance of land-sea warming in the Mediterranean basin, is in agreement with the mechanistic experiments of Keller et al. (2025). Last, the multi-model spread in the historical
430 PSL response cannot be explained on the basis of response of either a shift of the Hadley cell edge or shift of the zonal-mean subtropical jet.

~~Interestingly~~ Many of these relationships are similar if we focus on the ensemble spread within each individual model. Specifically, the stratospheric polar vortex, the edge of the eddy-driven jet, and averaged U700 responses in each individual member within a given model is correlated with the PSL response within that model in a manner similar to the grand
435 multi-model ensemble or the models’ ensemble-means. But this is not true for all effects. Namely, for the climatic metrics responses that are based on temperature, i.e., the Mediterranean land-ocean temperature contrast, NA warming-hole temperature, and in particular the global-mean surface temperature, the individual ~~models do not represent/support the multi-model correlations with the PSL response (note the relatively low values under the ‘relative’ column). On the other hand, for the stratospheric polar vortex, the edge of the eddy-driven jet, and averaged U700 the individual models show correlations that are closer to~~

440 ~~these obtained for model ensemble members show much smaller correlations than that of the multi-model .In contrast, for~~
~~the PSL vs. Hadley cell responses, the individual models show opposite correlations then that of the multi-model, which~~
~~means that the individual models show the expected relation between the two responses. The with the PSL response. This~~
~~implies that internal variability driven by dynamical processes within a given model ensemble can overwhelm these large-scale~~
~~thermodynamic forcings. These differences between the multi-model correlations and the individual models correlations implies~~
445 ~~that multi-model ensembles might be essential not only to cover a large enough spread for capturing realistic sensitivities values,~~
~~but also for capturing realistically- intra-ensemble correlation imply that single-model large ensembles may not be sufficient~~
~~for capturing~~ relations between, e.g., dynamical and climate sensitivities.

~~Correlation coefficients for the wintertime PSL historical response vs. the responses of the considered climatic matrices for~~
~~the hist-GHG and hist-aer LESFMIP experiments. The ‘relative’ columns specify the ratio between the averaged correlation~~
450 ~~coefficients obtained for the individual models and the multi-model coefficient (see main text). Climatic metric hist-GHG~~
~~relative~~ Our focus on understanding intermodel spread in historical climate has implications for attempts to build storylines of
~~future climate change that explicitly consider intermodel spread in underlying processes (Shepherd et al., 2018). In Europe and~~
~~the Mediterranean, these storyline approaches couch future projections in terms of how each model simulates, e.g., changes in~~
~~the polar vortex or changes in polar amplification while normalizing per degree of global warming (e.g. Zappa and Shepherd, 2017)~~
455 ~~We find no evidence that a stronger global mean warming leads to more pronounced drying in this region, which implies~~
~~that dynamical effects overwhelm thermodynamic mechanisms (in agreement with Seager et al., 2019; Elbaum et al., 2022).~~
~~Furthermore, our results imply that the rate of warming of SSTs in the North Atlantic warming hole region has a stronger impact~~
~~than either the polar vortex or Arctic amplification on Mediterranean drying, and should be included when building storylines of~~
~~European and Mediterranean climate. Finally, the regression coefficients derived from hist-aer relative Mediterranean land-ocean~~
460 ~~temperature contrast -0.61 -0.37 -0.56 -0.19 NA warming-hole t2m -0.66 -0.51 -0.78 -0.42 Global t2m -0.50 -0.15 -0.61 -0.28~~
~~Stratospheric polar vortex -0.39 -0.98 0.68 -0.41 Polar amplification -0.22 -0.05 -0.67 -0.21 Latitude of Hadley-cell edge -0.20~~
~~-1.47 -0.19 -0.92 Latitude of sub tropical jet 0.01 -13.6 -0.13 -1.02 Latitude of eddy-driven jet 0.43 -1.25 -0.70 -0.74 High variability~~
~~region of U7000.69 -0.85 -0.77 -0.65~~

~~Last, we note that for and hist-GHG are not, in general, quantitatively similar. That is, the correlations and regression~~
465 ~~coefficients of the majority of the large-scale metrics their correlations with the with the Mediterranean PSL response are~~
~~higher under the anthropogenic aerosols forcing than under the GHGs forcing. Explaining the reason for that This could arise if~~
~~European sector aerosols have an impact on the local atmospheric circulation. Confirming this requires a more detailed analysis~~
~~of the historical simulations (Elbaum et al., 2022). and hist-aer simulations, including analyzing the potential for reversals of~~
~~trends since the 1980s when European aerosols peaked. Ongoing work is aimed at clarifying this effect.~~

470 ~~In a future analysis, we shall exploit the large ensembles employed within the LESFMIP to partition the simulations~~
~~uncertainty into different contributions (Lehner et al., 2020). Such partitioning is expected to allow for differentiating the~~
~~forced response of the different models. In addition, combined with observations, large ensembles are expected to enable~~
~~further reduction of uncertainties in future climate projections based on emergent constraints that are developed from the~~
~~large ensemble of historical simulations (Hall et al., 2019). Future analysis will focus on drivers of summertime changes in this~~

475 region. Future work that compares the sum of the responses in the various single forcing simulations to the historical simulation is also ongoing, as additivity is a prerequisite for a quantitative attribution statement. Finally, ongoing work is aimed at model vs. observations discrepancies since 1951.

. The LESFMIP data that is used in this work can be downloaded from the ESGF portal <https://esgf-ui.ceda.ac.uk/cog/search/cmip6-ceda/>.
The HadSLP observations can be downloaded from <https://www.metoffice.gov.uk/hadobs/hadslp2/>.

480 . DA: data analysis and manuscript writing. CIG: methodology conceptualization, manuscript writing, providing feedback, and procuring funding.

. The authors declare no competing interests.

. DA and CIG acknowledge the support of the Israel Science Foundation (grant agreement 1727/21). CIG is supported by the US-Israel Binational Science Foundation (BSF) grant 2020316.

- Adam, O., Grise, K. M., Staten, P., Simpson, I. R., Davis, S. M., Davis, N. A., Waugh, D. W., Birner, T., and Ming, A.: The TropD software package (v1): standardized methods for calculating tropical-width diagnostics, *Geosci. Model Dev.*, 11, 4339–4357, <https://doi.org/https://doi.org/10.5194/gmd-11-4339-2018>, 2018.
- Allan, R. and Ansell, T.: A New Globally Complete Monthly Historical Gridded Mean Sea Level Pressure Dataset (HadSLP2): 1850–2004, *J. Clim.*, 19, 5816–5842, <https://doi.org/https://doi.org/10.1175/JCLI3937.1>, 2006.
- Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., Domeisen, D. I. V., Garfinkel, C. I., Garny, H., Gerber, E. P., Hegglin, M. I., Langematz, U., and Pedatella, N. M.: Sudden stratospheric warmings, *Reviews of Geophysics*, 59, e2020RG000708, <https://doi.org/https://doi.org/10.1029/2020RG000708>, 2021.
- Cos, J., Doblas-Reyes, F., Jury, M., Marcos, R., Bretonnière, P.-A., and Samsó, M.: The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections, *Earth Syst. Dyn.*, 13, 321–340, <https://doi.org/https://doi.org/10.5194/esd-13-321-2022>, 2022.
- Dai, Y., Hitchcock, P., Butler, A. H., Garfinkel, C. I., and Seviour, W. J. M.: Assessing stratospheric contributions to subseasonal predictions of precipitation after the 2018 SSW from SNAPSI, EGU sphere, <https://doi.org/https://doi.org/10.5194/egusphere-2025-484>, 2025, 2025.
- Delworth, T. L., Cooke, W. F., V. Naik, D. P., and Zhang, L.: A weakened AMOC may prolong greenhouse gas-induced Mediterranean drying even with significant and rapid climate change mitigation, *Proc. Natl. Acad. Sci. U.S.A.*, 119, e2116655119, <https://doi.org/https://doi.org/10.1073/pnas.2116655119>, 2022.
- Domeisen, D. I. V.: Estimating the Frequency of Sudden Stratospheric Warming Events From Surface Observations of the North Atlantic Oscillation, *Journal of Geophysical Research: Atmospheres*, 124, 3180–3194, <https://doi.org/https://doi.org/10.1029/2018JD030077>, 2019.
- Elbaum, E., Garfinkel, C. I., Adam, O., Morin, E., Rostkier-Edelstein, D., and Dayan, U.: Uncertainty in projected changes in precipitation minus evaporation: Dominant role of dynamic circulation changes and weak role for thermodynamic changes, *Geophysical Research Letters*, 49, e2022GL097725, <https://doi.org/https://doi.org/10.1029/2022GL097725>, 2022.
- Findell, K. L., Sutton, R., Caltabiano, N., Brookshaw, A., Heimbach, P., Kimoto, M., Osprey, S., Smith, D., Risbey, J. S., Wang, Z., Cheng, L., Diaz, L. B., Donat, M. G., M. Ek, J.-Y. L., Minobe, S., Rusticucci, M., Vitart, F., and Wang, L.: Explaining and Predicting Earth System Change A World Climate Research Programme Call to Action, *Bull. Am. Meteorol. Soc.*, 104, E325–E339, <https://doi.org/https://doi.org/10.1175/BAMS-D-21-0280.1>, 2023.
- Garfinkel, C. I., Adam, O., Morin, E., Enzel, Y., Elbaum, E., Bartov, M., Rostkier-Edelstein, D., and Dayan, U.: The Role of Zonally Averaged Climate Change in Contributing to Intermodel Spread in CMIP5 Predicted Local Precipitation Changes, *J. Clim.*, 33, 1141–1154, <https://doi.org/https://doi.org/10.1175/JCLI-D-19-0232.1>, 2020.
- Garfinkel, C. I., Keller, B., Lachmy, O., White, I., Gerber, E. P., Jucker, M., and Adam, O.: Impact of parameterized convection on the storm track and near-surface jet response to global warming: Implications for mechanisms of the future poleward shift, *J. Clim.*, 37, 2541–2564, <https://doi.org/https://doi.org/10.1175/JCLI-D-23-0105.1>, 2024.
- Gervais, M., Shaman, J., and Kushnir, Y.: Impacts of the North Atlantic warming hole in future climate projections: Mean atmospheric circulation and the North Atlantic jet, *J. Climate*, 32, 2673–2689, 2019.
- Giorgi, F.: Climate change hot-spots, *Geophys. Res. Lett.*, 33, L08707, <https://doi.org/https://doi.org/10.1029/2006GL025734>, 2006.
- Grise, K. M. and Polvani, L. M.: Is climate sensitivity related to dynamical sensitivity?, *J. Geophys. Res. Atmos.*, 121, 5159–5176, <https://doi.org/https://doi.org/10.1002/2015JD024687>, 2016.

- Hall, A., Cox, P., Huntingford, C., and Klein, S.: Progressing emergent constraints on future climate change, *Nat. Clim. Chang.*, 9, 269–278, <https://doi.org/https://doi.org/10.1038/s41558-019-0436-6>, 2019.
- He, J. and Soden, B. J.: A re-examination of the projected subtropical precipitation decline, *Nature Clim Change*, 7, 53–57, <https://doi.org/https://doi.org/10.1038/nclimate3157>, 2017.
- Held, I. M. and Soden, B. J.: Robust Responses of the Hydrological Cycle to Global Warming, *J. Clim.*, 19, 5686–5699, <https://doi.org/https://doi.org/10.1175/JCLI3990.1>, 2006.
- Kang, J. M., Shaw, T. A., and Sun, L.: Anthropogenic Aerosols Have Significantly Weakened the Regional Summertime Circulation in the Northern Hemisphere During the Satellite Era, *AGU advances*, 5, e2024AV001318, <https://doi.org/https://doi.org/10.1029/2024AV001318>, 2024.
- Keller, B., Garfinkel, C. I., and Gerber, E. P.: Disentangling Projected Stationary Wave Changes: Implications for Future Drying of the Mediterranean Region, *J. Clim.*, <https://doi.org/https://doi.org/10.1175/JCLI-D-24-0659.1>, 2025.
- Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., and Hawkins, E.: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6, *Earth Syst. Dynam.*, 11, 491–508, <https://doi.org/https://doi.org/10.5194/esd-11-491-2020>, 2020.
- Rahmstorf, S., Box, J., and Feulner, G. e. a.: Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation, *Nature Clim Change*, 5, 475–480, <https://doi.org/https://doi.org/10.1038/nclimate2554>, 2015.
- Schmidt, D. F. and Grise, K. M.: The response of local precipitation and sea level pressure to Hadley cell expansion, *Geophysical Research Letters*, 44, 10–573, 2017a.
- Schmidt, D. F. and Grise, K. M.: The response of local precipitation and sea level pressure to Hadley cell expansion, *Geophysical Research Letters*, 44, 10 573–10 582, <https://doi.org/https://doi.org/10.1002/2017GL075380>, 2017b.
- Seager, R., Naik, N., and Vecchi, G. A.: Thermodynamic and Dynamic Mechanisms for Large-Scale Changes in the Hydrological Cycle in Response to Global Warming, *J. Clim.*, 23, 4651–4668, <https://doi.org/https://doi.org/10.1175/2010JCLI3655.1>, 2010.
- Seager, R., Osborn, T. J., Simpson, Y. K. I. R., Nakamura, J., and Liu, H.: Climate Variability and Change of Mediterranean-Type Climates, *J. Clim.*, 32, <https://doi.org/https://doi.org/10.1175/JCLI-D-18-0472.1>, 2019.
- Seager, R., Yutian, W., Annalisa, C., Simpson, I. R., Osborn, T. J., Kushnir, Y., Lukovic, J., Liu, H., and Nakamura, J.: Recent and near-term future changes in impacts-relevant seasonal hydroclimate in the world’s Mediterranean climate regions, *International Journal of Climatology*, 44, 3792–3820, <https://doi.org/https://doi.org/10.1002/joc.8551>, 2024.
- Shaw, T. A., Arblaster, J. M., Birner, T., Butler, A. H., Domeisen, D. I. V., Garfinkel, C. I., H. Garny, K. M. G., and Karpechko, A. Y.: Emerging Climate Change Signals in Atmospheric Circulation, *AGU advances*, 5, e2024AV001297, <https://doi.org/https://doi.org/10.1029/2024AV001297>, 2024.
- Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., et al.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, *Climatic change*, 151, 555–571, 2018.
- Smith, D. M., Gillett, N. P., Simpson, I. R., Athanasiadis, P. J., Baehr, J., Bethke, I., Bilge, T. A., Bonnet, R., Boucher, O., Findell, K. L., Gastineau, G., Gualdi, S., Hermanson, L., Leung, L. R., Mignot, J., Müller, W. A., Osprey, S., Otterå, O. H., Persad, G. G., Scaife, A., Schmidt, G. A., Shiogama, H., Sutton, R. T., Swingedouw, D., Yang, S., Zhou, T., and Ziehn, T.: Attribution of multi-annual to decadal changes in the climate system: The Large Ensemble Single Forcing Model Intercomparison Project (LESFMP), *Frontiers in Climate*, 4, 1, <https://doi.org/https://doi.org/10.3389/fclim.2022.955414>, 2022.

- 560 Tuel, A. and Eltahir, E.: Why Is the Mediterranean a Climate Change Hot Spot?, *J. Clim.*, 33, 5829–5843, <https://doi.org/https://doi.org/10.1175/JCLI-D-19-0910.1>, 2020.
- Vacca, A. V., Bellomo, K., and Fabiano, F.: On the role of AMOC weakening in shaping wintertime Euro-Atlantic atmospheric circulation, *Clim Dyn*, 63, 273, <https://doi.org/https://doi.org/10.1007/s00382-025-07747-z>, 2025.
- Vicente-Serrano, S., Trambly, Y., and Reig, F. e. a.: High temporal variability not trend dominates Mediterranean precipitation, *Natur*, 639, 658–666, <https://doi.org/https://doi.org/10.1038/s41586-024-08576-6>, 2025.
- 565 Zappa, G. and Shepherd, T. G.: Storylines of Atmospheric Circulation Change for European Regional Climate Impact Assessment, *Journal of Climate*, 30, 6561 – 6577, <https://doi.org/10.1175/JCLI-D-16-0807.1>, 2017.
- Zappa, G., Hoskin, B. J., and Shepherd, T. G.: The dependence of wintertime Mediterranean precipitation on the atmospheric circulation response to climate change, *Environ. Res. Lett.*, 10, 104012, <https://doi.org/https://doi.org/10.1088/1748-9326/10/10/104012>, 2015.
- 570 Zelinka, M. D., Myers, T. A., D. T. McCoy, S. P.-C., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, *Geophys. Res. Lett.*, 47, e2019GL085782, <https://doi.org/https://doi.org/10.1029/2019GL085782>, 2020.
- Zhang, C., Zhang, J., Xia, X., and Li, D.: Impact of Arctic Stratospheric Polar Vortex on Mediterranean Precipitation, *J. Clim.*, 37, 4403–4419, <https://doi.org/https://doi.org/10.1175/JCLI-D-23-0469.1>, 2024.
- Zhang, J., Zhang, C., Zhao, S., Liu, Y., Du, S., Wang, W., Huang, J., and Xu, M.: Impacts of the Arctic stratospheric polar vortex changes on the frontogenesis over the northern middle latitudes during winter, *Atmos. Res.*, 289, 106751, <https://doi.org/https://doi.org/10.1016/j.atmosres.2023.106751>, 2023.
- 575