Climate Dynamics (2025) 63:312
https://doi.org/10.1007/s00382-025-07783-9

ORIGINAL ARTICLE

The potential impacts of improved MJO prediction on the prediction of MJO teleconnections in the UFS global fully coupled model

Jiabao Wang[1] · Daniela I. V. Domeisen[2,3] · Chaim I. Garfinkel[4] · Andrea M. Jenney[5] · Hyemi Kim[6,8] · Zheng Wu[7] · Cheng Zheng[8] · Cristiana Stan[9]

Received: 13 August 2024 / Accepted: 30 June 2025 / Published online: 11 August 2025

Abstract
The value of Madden-Julian oscillation (MJO) prediction for extratropical subseasonal forecasts hinges on the assumption that reliable MJO simulation translates to reliable simulation of its teleconnections. This study examines the prediction of the MJO and its teleconnections in two recently developed NOAA Unified Forecast System (UFS) coupled model prototypes: Prototype 7 (UFS7) and Prototype 8 (UFS8). The MJO is skillfully predicted at a lead time of 27 days in UFS8, which is a considerable improvement (~one-week skill increase) compared to UFS7. The potential effect of this improvement on MJO teleconnections via both tropospheric and stratospheric pathways is examined. UFS8 captures the pattern and amplitude of the geopotential height response in the North Pacific reasonably well and its evolution following active MJO events. The dipole response in the storm tracks over the North Pacific after active MJO events is also better captured in UFS8. In addition, the upward wave propagation and subsequent weakening of the polar vortex are better simulated in UFS8, with comparable strength to that in the reanalysis. Despite the notable improvements listed above, some biases remain: too-fast MJO propagation, an underestimation of geopotential height variability in the North Atlantic and Europe, an underestimation of the precipitation response, failure to capture the temperature evolution, and weaker MJO impacts on the NAO. This study suggests the potential of increasing the MJO teleconnection prediction skill, although not in all variables, by improving MJO predictions in dynamical models with more coupled components and upgraded model physics.

Keywords MJO teleconnections · Subseasonal-to-seasonal forecast · UFS fully coupled model · Tropospheric and stratospheric pathways

1 Introduction

Attempts to reduce disaster risks have relied on weather forecasts and seasonal prediction. However, long-term seasonal forecasts may not be able to predict the exact timing of high-impact events while information gained from weather forecasts may be too late for satisfactory preparation. Subseasonal-to-seasonal (S2S) predictions (i.e., lead times ranging from 2 weeks to 2 months) can be more temporally specific than seasonal forecasts while leaving sufficient time for policymakers, water resource managers, and stakeholders to better mitigate potential impacts. However,

✉ Jiabao Wang
jiw093@ucsd.edu

[1] Center for Western Weather and Water Extremes, Scripps Institute of Oceanography, University of California, San Diego, USA

[2] University of Lausanne, Lausanne, Switzerland

[3] Institute for Atmospheric and Climate Science, ETH Zürich, Zurich, Switzerland

[4] The Fredy and Nadine Hermann Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel

[5] College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, USA

[6] Department of Science Education, Ewha Womans University, Seoul, Republic of Korea

[7] Department of Atmospheric and Environmental Sciences, University at Albany, Albany, USA

[8] School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA

[9] Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, VA, USA

S2S timescales have been considered a "predictability desert" with much lower prediction skill than the weather and seasonal timescales (Vitart et al. 2014; Robertson et al. 2020). Tremendous international efforts are underway to explore and understand sources of S2S predictability, assess S2S prediction skill, and transfer predictability into reliable predictions. The Madden–Julian oscillation (MJO; Xie et al. 1963; Madden and Julian 1971, 1972), which is a unique type of organized convection-circulation coupled system varying on subseasonal timescales in the tropics, is recognized as an important source of subseasonal predictability for midlatitude weather phenomena. The remote impacts of the MJO on midlatitude weather and climate are known as MJO teleconnections. Various types of weather events such as blocking (Henderson et al. 2016), precipitation extremes (Wang et al. 2023), atmospheric rivers (Mundhenk et al. 2018), and extratropical cyclones (Zheng et al. 2018) as well as climate modes such as the North Atlantic Oscillation (NAO; Lin et al. 2009) and the Pacific North American (PNA; Riddle et al. 2013) pattern are found to be closely impacted by tropical variations associated with the MJO.

The influence of the MJO on the extratropics occurs via two teleconnection pathways: (1) the tropospheric pathway through a propagating Rossby wave train in the upper troposphere which, by interacting with the extratropical jets, modulates the circulation in mid-latitudes (e.g., Seo and Son 2012; Seo and Lee 2017) and weather and climate over North America (e.g., Zhou et al. 2012); and (2) the stratospheric pathway through tropospheric-stratospheric coupling which manifests as either vertically propagating Rossby waves associated with the MJO heating or planetary waves produced by tropospheric MJO teleconnections that propagate into the polar stratosphere. As the MJO can influence the conditions of the stratospheric polar vortex with enhanced upward propagating wave activity to the stratosphere (Garfinkel et al. 2012), studies have shown that the MJO can have a stronger and longer-lasting impact on surface weather over the Euro-Atlantic region through this "stratospheric pathway" (Schwartz and Garfinkel 2017; Jiang et al. 2017; Green and Furtado 2019).

The theoretical MJO predictability can be up to 6–7 weeks (Kim et al. 2018; Straus et al. 2023) in a perfect model scenario. Although the operational MJO prediction skill, which varies between 10 and 38 days in the global dynamical models archived in the Subseasonal-to-Seasonal (S2S; Vitart et al. 2017) and Subseasonal Experiment (SubX; Pegion et al. 2019) projects is still much lower than this theoretical predictability, substantial improvement has been made in the past few decades. Increased model resolution, upgraded model physics, and ocean coupling have contributed to the reduction of MJO amplitude and phase errors, leading to a continuous improvement in prediction skill with a roughly 1-day-per-year increase in the forecast lead time from 2002 to 2011 (Vitart et al. 2014; Kim et al. 2018). While research and model development have primarily focused on improving the prediction skill of the MJO, less attention has been paid to evaluating the prediction capability of MJO teleconnections and understanding sources of improvements to predicted MJO teleconnections.

Three recent studies discussed the prediction skill of global MJO teleconnections in the various generations of models that participated in the S2S project (Vitart et al. 2017; Stan et al. 2022; Kim et al. 2023). Vitart et al. (2017) showed that at least four of the models (ECMWF, NCEP, JMA, and BoM) are able to capture the general pattern of 500 hPa geopotential height anomalies after active MJO phases but with a stronger amplitude in the North Pacific and a weaker amplitude in the North Atlantic than observed at a 10-day lead time. Stan et al. (2022) provided a comprehensive overview of prediction skill assessment on various aspects of MJO teleconnections based on process-level diagnostics including general extratropical circulations, storm tracks, temperature, precipitation, tropical cyclones, and meridional heat flux in a newer generation of S2S models. They find that the teleconnection magnitude tends to be stronger than observed in the first two weeks and weaker in the following weeks. In addition, the lead time to skillfully predict general patterns of the PNA is between 8 to 14 days. They also find that S2S models struggle to represent the MJO teleconnection's stratospheric pathway. Kim et al. (2023) compared the prediction skill of geopotential height over the PNA region and the North Atlantic, as well as that of surface temperature over North America, between forecasts initialized with and without an active MJO. They find that the extratropical skill is significantly increased during Week 3–4 lead times when there is an active MJO east of the Maritime Continent. Other studies have focused on the specific impact of the MJO in a particular S2S forecast model (Vitart and Molteni 2010; Lin et al. 2010; Schwartz and Garfinkel 2020; Garfinkel et al. 2022; Vitart and Balmaseda 2024).

The ability of an S2S forecast model to reliably predict MJO teleconnections may depend on model resolution, coupling between the atmosphere and other components, model physics, and numerous other factors. One common caveat of the current S2S hindcast models is that they generally lack some active model components and/or have relatively coarse horizontal and vertical resolutions. Recent studies have shown that in addition to ocean coupling which is the norm for most S2S models, coupling between the atmosphere and other components such as sea ice and waves can lead to improved prediction skill on short- to medium-range weather timescales (e.g., Brassington et al. 2015; Smith et al. 2018; Gentile et al. 2021). Inspired by these studies, the

National Centers for Environmental Prediction (NCEP) developed a fully coupled (atmosphere/land/ocean/sea-ice/wave/aerosols) Unified Forecast System (UFS) to be implemented as the next-generation S2S forecasting system (e.g., Global Ensemble Forecast System v13, GEFSv13) at the National Oceanographic and Atmospheric Administration (NOAA). The development towards a fully coupled model includes incremental upgrades in model physics, resolution, and model components from Prototype 5 to Prototype 8. Recent studies have investigated the changes in prediction skill from Prototype 5 to Prototype 6 (Zheng et al. 2024; Garfinkel et al. 2025). Yet, no assessment has been done on the prediction skill of the MJO and its teleconnections in the latest prototypes. This study aims to fill this gap by providing a comprehensive overview of the prediction skill of MJO teleconnections via both a tropospheric and a stratospheric pathway using the metrics discussed in recent studies (Stan et al. 2022; Zheng et al. 2024; Garfinkel et al. 2025).

Prototype 8 distinguishes itself from previous prototypes by including prognostic aerosols based on NASA's 2nd-generation GOddard Chemistry Aerosol Radiation & Transport (GOCART) model (Chin et al. 2003). As we will demonstrate, there is a notable one-week improvement in the prediction skill of the MJO between UFS Prototypes 7 and 8. Hence, this analysis, which compares MJO teleconnections between the two prototypes, presents a unique opportunity to explore the relationship between advancements in MJO prediction and their potential impacts on MJO teleconnections within a near-operational framework.

The remainder of the paper is organized as follows. The forecast models, verification data, and MJO teleconnection skill metrics used in this study are described in Sect. 2. The metrics are integrated into a Python package with a User-friendly Interface that can be easily applied to both operational and research studies (Stan et al. 2025a, b). The prediction skill of the MJO and its teleconnection pathways are discussed in Sect. 3. Section 4 provides a summary and discussion of the results.

2 Data and models

2.1 Forecast models

In this study, reforecasts from the two latest versions of the UFS model, Prototype 7 and 8 (hereafter referred to as UFS7 and UFS8) are analyzed. Both prototypes consist of atmosphere/land/ocean/sea-ice/wave coupling while UFS8 has additional coupling with aerosols. The standard features between the two prototypes include the Geophysical Fluid Dynamics Laboratory (GFDL)'s Finite Volume Cubed-Sphere (FV3) dynamical core (Putman and Lin 2007; Harris and Lin 2013) with a horizontal resolution of C384 (~0.25°), 127 vertical layers with a model top at 80 km, the Common Community Physics Package (CCPP)-based atmospheric physics package with Noah-MP land model (Niu et al. 2011), the GFDL MOM6 ocean model (Adcroft et al. 2019), the Los Alamos CICE6 sea ice model with mushy thermodynamics, and the WAVEWATCH III wave model, which provides feedback to the atmosphere and ocean.

Differences between UFS7 and UFS8 are generally in the model physics package: UFS8 uses the GFSv17 physics with Thompson microphysics (Thompson and Eidhammer 2014), while UFS7 uses the GFSv16 Physics with GFDL cloud microphysics (e.g., Lin et al. 1983; Chen and Lin 2013). The Thompson scheme optimizes cloud cover and radiative fluxes and utilizes semi-Lagrangian sedimentation for rain and graupel, which effectively reduces the high biases in 3-h accumulated precipitation of less than 2.5 inches. Stefanova et al. (2022) provides a more detailed description and results for UFS Prototypes 1–7. Upgrades in model physics in UFS7 compared to its previous versions include (1) Unified Gravity Wave Physics (UGWP) which consists of a set of drag parameterizations such as the Gravity wave drag parameterization (uGWD.v1 replacing uGWD.v0) to better represent the drag forces due to subgrid-scale orographic and non-stationary sources; (2) updates in the convection, planetary boundary layer (PBL) and surface layer parameterizations which lead to improvement in MJO intensity and propagation (Han et al. 2021); (3) Near-Surface Sea Temperature Model (NSSTM) that calculates a temperature profile below the ocean surface; (4) MERRA2 aerosols (Randles et al. 2017) replacing Optical Properties of Aerosols and Clouds (OPAC); (5) stochastic parameterization of tropical convection using Cellular Automata; (6) updated fractional grid compositing and lower threshold for fractional lakes; and (7) Noah-MP replacing Noah LSM which potentially improves the realism of the land surface, hydrology, and atmosphere interactions. Additional upgrades in model physics in UFS8 include (1) updates in the PBL with a modified version of the sa-TKE-EDMF of Han and Bretherton (2019) to reduce the cold bias in the tropical troposphere found in GFSv16 and positive definite mass flux scheme to remove negative tracers in the PBL and cumulus convection parameterizations; (2) updates in convection with stochastic convective organization using Cellular Automata (Bengtsson et al. 2021, 2022) to improve MJO and updated sa-SAS cumulus convection developed by Kwon and Hong (2017) to improve Convective Available Potential Energy (CAPE); (3) updates in the Noah-MP parameterization for snow, land–atmosphere coupling, roughness length, and sub-grid tiling to correct UFS7 biases; (4) a shift back to the uGWD.v0 Gravity wave drag parameterization to reduce winter high latitude warming;

(5) optimizing the scaling of the Cellular Automata scheme for different resolutions; and (6) turbulent orographic gravity drag being turned off. UFS8 also includes additional one-way aerosol coupling with NASA's GOCART model (Chin et al. 2003).

The reforecast period for both UFS7 and UFS8 is from April 2011 to March 2018 (a total of 7 years with 168 forecasts). The model is initialized on the 1st and 15th of each month with a 35-day forecast period without ensemble generation. The caveats of the UFS prototypes are discussed in detail by Zheng et al. (2024). Here, we want to emphasize caution in interpreting the results given the limited sample size in this relatively short reforecast period (2011–2018) that increases the noise of variability. The noise resides not only in the nature of baroclinic instability and non-linearities in mid-latitudes but also in the varying nature of the MJO itself. MJO episodes vary considerably in period and phase speed, so composite diagnostics assuming a constant phase speed may not provide the whole picture. Additionally, there may be potential impacts from the slow-varying climate modes on the results, such as the El Niño-Southern Oscillation (ENSO). Nonetheless, this study demonstrates an improvement in MJO teleconnection prediction when the fully coupled framework is enabled with updated physics, leading to enhanced MJO prediction. The robustness of the results needs to be tested with a larger sample size.

2.2 Verification data

The reanalysis and observational data used in this study for the comparison with the reforecast include (1) temperature, geopotential height, and horizontal winds from the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-I; Dee et al. 2011); (2) daily precipitation from the Integrated Multi-satellitE Retrievals for GPM (IMERG; Huffman et al. 2015); and (3) interpolated Outgoing Longwave Radiation (OLR) data derived from the NOAA Advanced Very High-Resolution Radiometer (Liebmann and Smith 1996) product. ERA-I is selected to show over an updated reanalysis product for consistency with the recent studies that analyzed the S2S models and UFS Prototypes 5 and 6 (Stan et al. 2022; Zheng et al. 2024; Garfinkel et al. 2025), but the results are highly consistent using the ECMWF ERA5 reanalysis dataset (Hersbach et al. 2020) given the same MJO cases in the two datasets. To allow a consistent comparison, both the reforecast and verification data are interpolated to $1.5° \times 1.5°$ horizontal resolution. The anomalies are also derived similarly: for reforecast data, the climatology is taken for each initialization date (e.g., Jan 01, Jan 15) as a function of forecast lead times (day 1 to day 35) which is calculated separately for UFS7 and UFS8, and then the anomalies are derived as the difference between the forecast field and its climatology. Limited by the reforecast data size, climatology may be skewed by a few significant events. To mitigate the impact on model evaluation, the climatology of the verification data is calculated using the same dates as in the observational record, allowing for a consistent comparison. The verification data will be referred to as reanalysis hereafter.

2.3 MJO teleconnection skill metrics

2.3.1 Real-time multivariate MJO (RMM) index

To quantify the propagation, location (i.e., phase), and amplitude of the MJO, the RMM index developed by Wheeler and Hendon (2004) is calculated as the first two principal components of the combined empirical orthogonal functions of 15° N–15° S averaged OLR and zonal wind anomalies at 850 hPa (U850) and 200 hPa (U200). The model RMM index is then derived by projecting the OLR and wind fields onto the observed eigenvectors following the method of Gottschalk et al. (2010). In this study, the analysis focuses on forecasts initialized from November through March when the MJO is active (RMM amplitude > 1) at initialization to enhance the signal-to-noise ratio in the results. The sample size from MJO phase 1 to phase 8 is 2, 8, 9, 4, 2, 6, 6, 9, respectively. In some tropospheric pathway analyses (e.g., PNA circulations and storm track activity), adjacent MJO phases are combined to enlarge the sample size. Specific MJO phases are depicted in certain analyses, for example, MJO phases 2–3 and 6–7 for PNA circulations and MJO phases 3 and 7 for temperature analysis. These phases are found to have the most significant signals in relation to that particular MJO impact and have been the focus of most previous studies. Re-examining those phases allows for consistent and direct comparisons with the performance of other UFS prototypes and S2S hindcast systems (e.g., Stan et al. 2022; Zheng et al. 2024; Garfinkel et al. 2025).

2.3.2 MJO prediction skill

RMM indices are commonly used in the literature to examine the development of MJO biases in dynamical forecasting systems (e.g., Kim et al. 2019; Xiang et al. 2022). Although a caveat is that the fractional contribution of U200 to RMM indices is considerably higher than the contribution of OLR and U850 anomalies, leading to an overly optimistic estimate of the ability of a model to predict the MJO convection (Kim et al. 2019), this index efficiently reflects the skill and biases in MJO, especially its wind components (Straub 2013). The wind components are of great importance in generating MJO teleconnections, as they interact with the

mean westerly jet, as suggested by Rossby wave theory (e.g., Sardeshmukh and Hoskins 1988; Seo and Lee 2017).

To quantify the MJO prediction skill, two metrics are calculated between the reanalysis and predicted RMM indices following Kim et al. (2018): anomaly correlation coefficient (ACC) and root-mean-squared error (RMSE).

$$ACC(\tau) = \frac{\sum_{t=1}^{t=N}\left[a_1(t)\,b_1(t,\tau) + a_2(t)\,b_2(t,\tau)\right]}{\sqrt{\sum_{t=1}^{t=N}\left[a_1(t)^2 + a_2(t)^2\right]}\sqrt{\sum_{t=1}^{t=N}\left[b_1(t,\tau)^2 + b_2(t,\tau)^2\right]}}$$

$$RMSE(\tau) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[|a_1(t) - b_1(t,\tau)|^2 + |a_2(t) - b_2(t,\tau)|^2\right]}$$

where a_1 and a_2 are RMM1 and RMM2 in reanalysis, b_1 and b_2 are RMM1 and RMM2 in forecast data, t is for initialization time with a lead time of τ days, and N is the total number of predictions.

2.3.3 Sensitivities to the remote influence of periodic events (STRIPES) index

The STRIPES index (Jenney et al. 2019) is employed in this study to quantify the strength and consistency of the regional 500 hPa geopotential height (Z500) and precipitation responses to the propagating MJO. The advantage of this index is that, instead of showing the composite anomaly of a field corresponding to a specific MJO longitudinal location and lead/lag time as many studies do, the STRIPES index compiles the composite anomalies across the entire life cycle of the MJO and across a range of lead times into a single, positive-definite number. A larger value of the STRIPES index indicates a stronger dipole (i.e., displaying both positive and negative anomalies) co-variability of the local variable with the MJO, and hence implies a stronger MJO teleconnection response. A more detailed description of the STRIPES index including the definition, calculation, and interpretation is provided by Jenney et al. (2019). We note that the STRIPES index assumes that the MJO's phase speed varies between spending 5–8 days per phase.

2.3.4 Extratropical cyclone activity (storm tracks)

The extratropical cyclone activity (also referred to as storm tracks) is quantified by calculating the filtered eddy kinetic energy at 850 hPa (EKE850) following Yau and Chang (2020):

$$EKE850(t) = \frac{1}{2}\{[U850(t+24h) - U850(t)]^2 + [V850(t+24h) - V850(t)]^2\}$$

where U850 and V850 represent zonal wind and meridional wind at 850 hPa, respectively. This 24 h filter is equivalent to a bandpass filter that highlights synoptic timescale variability with periods of 1.2–6 days. Previous studies have shown that this filtering method can effectively identify extratropical cyclone activities which, in the Northern Hemisphere, are climatologically most active in the North Pacific and western North Atlantic (e.g., Zheng et al. 2019).

2.3.5 Significance test

The significance of the results is calculated with a bootstrapping test. The data is resampled by randomly selecting synthetic MJO events that have the same number as the true composites. The bootstrapping uses 1000 samples to obtain a sufficiently large sample size to determine the significance. The results are significantly different from climatology if the 0.1 (in the stratospheric pathway) or 0.05 (in the tropospheric pathway) confidence interval is greater or smaller than the climatology, which indicates significant positive or negative anomalies of the variables, respectively. The significance shown in this study suggests a significant difference from the observation's or model's own climatology rather than the difference between the two prototypes.

3 Results

3.1 Tropospheric pathways

We first examine how the prototypes reproduce the overall structure of the MJO and its propagation. Longitude-time composites of reanalysis OLR and U850 anomalies averaged over the tropics for active MJO events initialized during phases 2 and 3 are shown in Fig. 1a, along with results for UFS7 and UFS8 (Fig. 1b–c). In reanalysis, the MJO shows a clear eastward propagation signal from the Indian Ocean to the central Pacific. The MJO's convective center is co-located with low-level convergent winds. Based on the moisture mode theory, the associated easterly winds contribute to moisture recharge to the east of MJO convection and favor its eastward propagation (e.g., Kim et al. 2017; Ahn et al. 2020). Figure 1b–c show that the MJO eastward propagation is generally better captured in UFS8 than in UFS7 with the amplitude of convection and wind anomalies closer to reanalysis and more realistic easterly winds to the east of the active convection center. Note that the improvement in the MJO is primarily in the circulation fields. On the other hand, the MJO tends to propagate at a faster speed in UFS8 compared to UFS7, which may potentially lead to a degradation of NAO predictability as a fast-propagating MJO typically results in a weak positive NAO phase (Yadav and Straus 2017; Yadav et al. 2019, 2024). The opposite

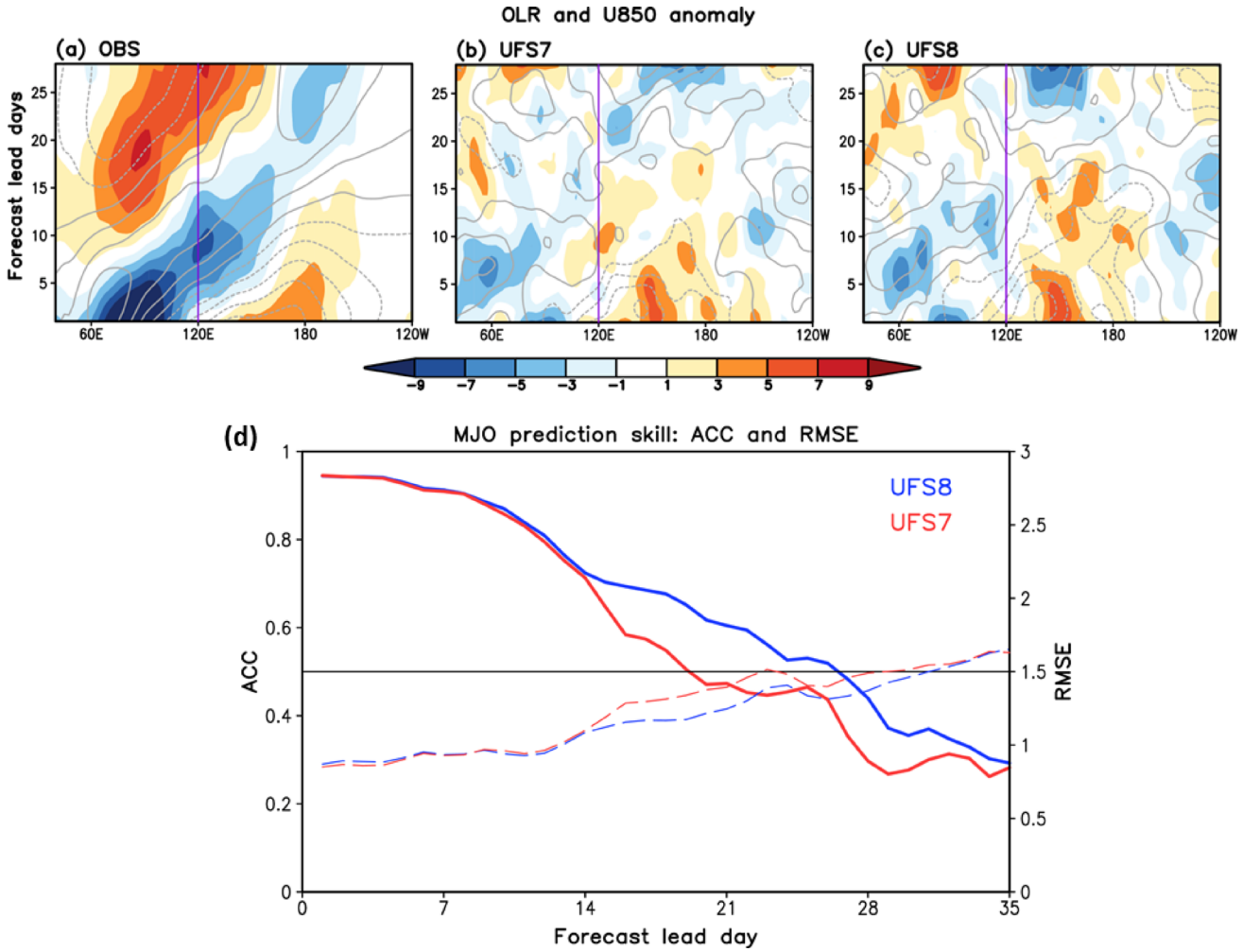

Fig. 1 **a**–**c** Longitude-time composites of OLR (shading; W/m^2) and U850 (contour; interval 0.3 m/s) anomalies averaged over 15° S–15° N for active MJO events in reanalysis, UFS7, and UFS8, respectively. The results are for events initialized during MJO phases 2 and 3. The vertical lines indicate 120°E (approximately the center of the Maritime Continent). A 5-day moving average is applied. **d** MJO prediction skill for UFS7 and UFS8 reforecasts initialized with active MJO events. The prediction skill is evaluated based on ACC (solid lines) and RMSE (dashed lines) between the model and reanalysis RMM indices. The gray solid horizontal line indicates an ACC of 0.5

MJO phases, i.e., initializations during MJO phases 6 and 7, show similar features (not shown).

Figure 1d shows the ACC and RMSE for UFS7 and UFS8 for active (RMM amplitude > 1) MJO events at initialization. As expected, the MJO prediction skill (ACC) decreases and the RMSE becomes larger as the lead time increases. The ACC at 0.5 is used as the threshold to quantify the MJO prediction skill which follows the common standard in previous studies (e.g., Rashid et al. 2011; Vitart 2014; Kim et al. 2018). The lead time when the ACC drops below 0.5 is approximately 19 days for UFS7, while it extends to 27 days for UFS8. This represents a roughly one-week improvement in the latest UFS prototype. UFS8 performs similarly to most high-skill S2S and SubX models as shown in Kim et al. (2018, 2019). This large improvement in MJO prediction skill may benefit the model's ability to predict MJO teleconnections (as will be discussed later). When all MJO events (RMM amplitude > 0) are computed, the lead time when ACC reaches 0.5 reduces slightly to 18 days for UFS7 and 23 days for UFS8.

We next examine the overall MJO teleconnection prediction skill using the STRIPES index for Z500 during Week 2–3. Figure 2a shows the STRIPES index for reanalysis. The STRIPES index for Z500 exhibits a strong maximum over the North Pacific, indicating a significant MJO influence in this region. This is where the PNA pattern is dominant, consistent with previous studies that have shown the PNA to be significantly modulated by the MJO (e.g., Mori and Watanabe 2008; Seo and Lee 2017; Tseng et al. 2019). The biases in UFS7 and UFS8 in capturing the observed patterns in Fig. 2a are given in Fig. 2b and c, respectively. In UFS7, the magnitude of teleconnection response in the

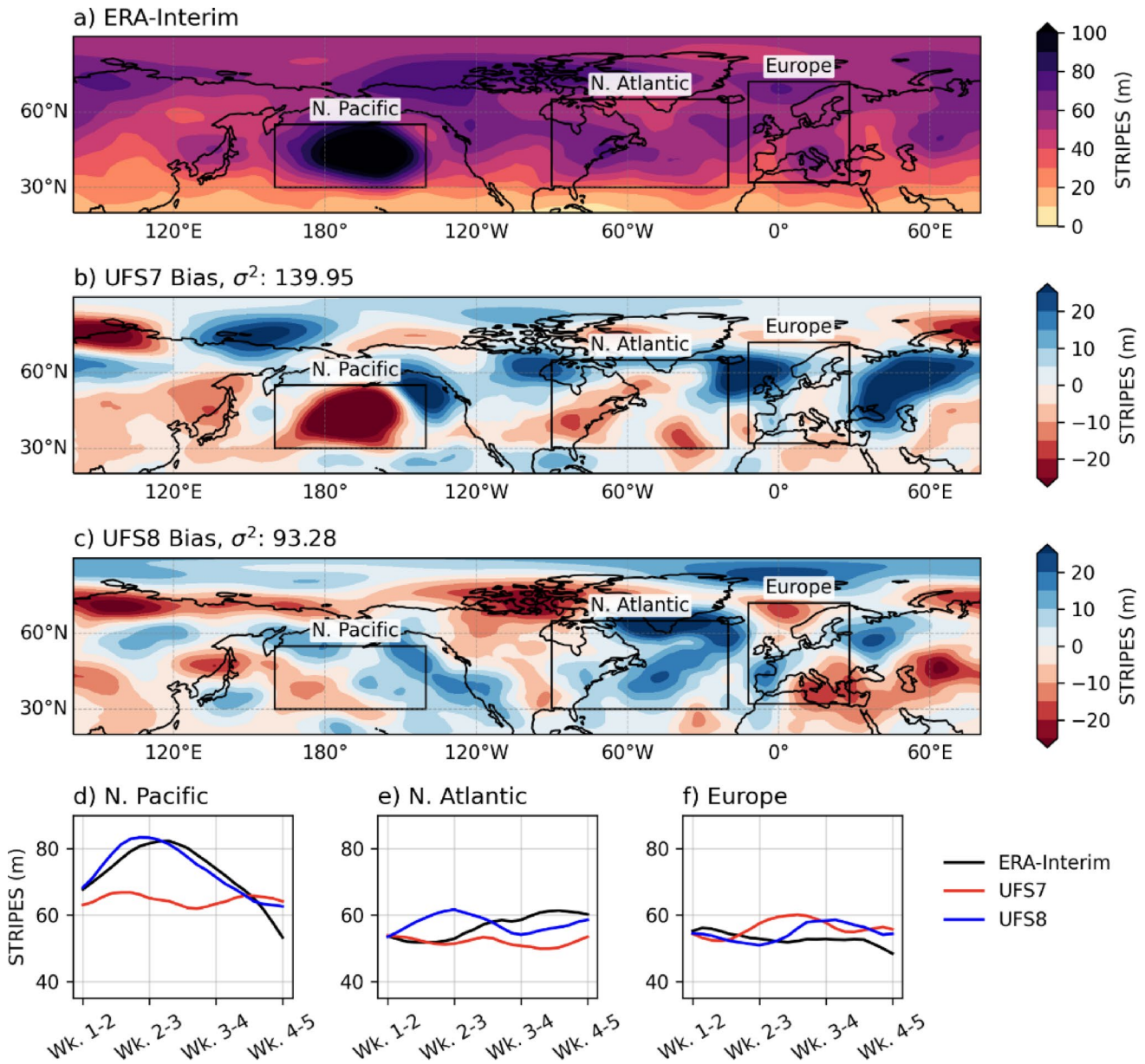

Fig. 2 Week 2–3 STRIPES indices for Z500 for **a** reanalysis, **b** bias in UFS7 (UFS7 minus reanalysis), and **c** bias in UFS8 (UFS8 minus reanalysis). The numbers indicate the spatial variance of bias over 20–90°N. **d**–**f** Evolution of STRIPES index averaged over the North Pacific (160–220°E, 30–55°N), North Atlantic (270–340°E, 30–65°N), and Europe (348–28°E, 32–72°N), respectively, as a function of lead time in reanalysis, UFS7, and UFS8. These three regions are indicated as boxes in (**a**–**c**)

North Pacific is underestimated by more than 20 m. This bias is largely reduced in UFS8 (Fig. 2c) with only slight positive biases to the northeast and negative biases to the southwest. Biases are not reduced everywhere, however. For example, over the North Atlantic region indicated in the figure, a slight underestimation of the STRIPES magnitude in UFS7 is replaced by an overestimation in UFS8. Over the Northern Hemisphere as a whole, the bias of the STRIPES index for Z500 is reduced by ~ 30% if comparing the spatial variance of the UFS7 bias (139.95) to the spatial variance of the UFS8 bias (93.28).

The temporal evolution of 2-week overlapping STRIPES indices over three regions of interest (the North Pacific, North Atlantic, and Europe) as a function of lead time is shown in Fig. 2d–f. Again, the evolution of the STRIPES index for the North Pacific in UFS8 closely follows the reanalysis which shows a peak at Week 2–3, consistent with the time it takes for the Rossby waves to propagate from the tropics to this region (e.g., Hoskins and Karoly 1981), and a weakening of the response afterwards. This is a considerable improvement compared to the systematically weaker response in UFS7. However, differences in bias over the

North Atlantic and Europe are small between the two prototypes. Although there is some improvement from UFS7 to UFS8, it is not consistent throughout the forecast period. For example, over Europe, the bias is reduced in UFS8 during week 1–3 while week 3–5 show a similar overestimation for both prototypes.

Next, we show the lagged composites (week 1–4) of 500 hPa geopotential height anomalies (Z500a) after MJO phases 2 and 3 in reanalysis, UFS7 and UFS8 over the PNA region (Fig. 3). The response after an active MJO (RMM amplitude > 1) located over the Indian Ocean resembles the negative phase of the PNA with a ridge over the North Pacific and Northeastern North America and a trough near Alaska. The evolution from Week 1 to Week 4 is mainly characterized by changes in the magnitude of the response, such as a weakening in the North Pacific ridge. Although the pattern in reanalysis shown in Fig. 3a–d only includes a limited sample size to match the UFS, it is still consistent with previous studies (e.g., Wang et al. 2020). Both UFS7 and UFS8 reproduce the PNA pattern in Week 1 with a pattern correlation greater than 0.95 (Fig. 3a, e, l). However, starting in Week 2, a large difference emerges between the two prototypes. While UFS 7 has large biases such as a notably weaker response over North America, UFS8 is able to capture the evolution of the PNA such as a southeastward extension of the trough towards the North American West Coast. The pattern correlation in Week 2 for UFS8 is greater than 0.8 while it drops below 0.5 for UFS7. However, although UFS8 is able to capture the overall pattern in Week 2, it shows a large overestimation. In Week 3, the pattern correlation for both prototypes decreases, although it remains higher (0.57) for UFS8 than for UFS7 (0.42). The drop in skill mainly comes from the trough which shifts too far eastward in UFS8, while UFS7 begins to develop the wrong-signed response over western North America. The Week 4 PNA pattern is difficult to predict for both prototypes, although it is slightly better predicted in UFS8. In summary, UFS8 generally performs better than UFS7 in predicting the PNA-like pattern that develops following MJO phases 2 and 3, especially in Week 2. This systematic improvement in UFS8 is not found in weak MJO cases (RMM amplitude < 0.5, figure not shown), indicating the better performance is more likely due to improved MJO predictions and subsequent excitation of poleward propagating

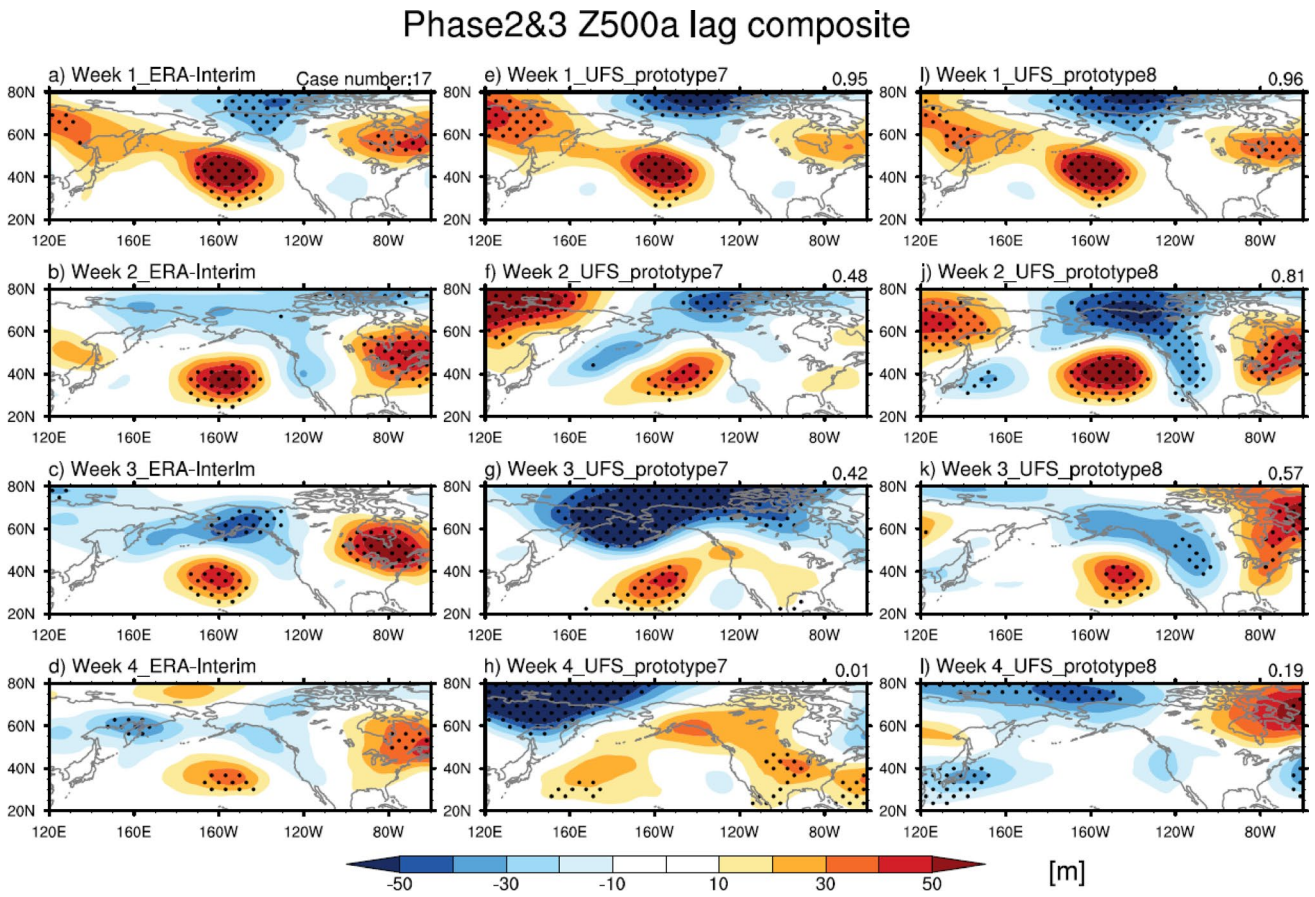

Fig. 3 Weekly averaged phase composites of 500 hPa geopotential height anomalies (Z500a) after MJO phases 2 and 3 for the lead times from week 1 to week 4 in **a**–**d** reanalysis, **e**–**h** UFS7, and **i**–**l** UFS8. The dotted areas indicate the significance at the 0.05 level. The numbers in the upper right corner of each plot indicate the spatial correlation between the model and reanalysis over the region

Rossby waves rather than reduced growth of biases from the atmospheric initial conditions.

Figure 3 shows the Z500a composites, which may be dominated by a few events with very large amplitudes. To further examine the prediction skill of Z500a in UFS prototypes, the pattern correlation over the PNA region is calculated for individual samples and then an average is taken across those samples. The results for MJO phases 2–3 and 6–7 are shown in Fig. 4, along with the skill in predicting the teleconnection amplitude (measured by the relative amplitude between reforecast and reanalysis). The prediction skill in the teleconnection pattern decreases with increasing lead time as expected. Similar to the S2S models (Stan et al. 2022), the pattern correlation drops below 0.5 during Week 2. The skill in both prototypes falls within the range of the S2S models, which is between 8 and 13 days. No systematic improvement across lead times is found from UFS7 to UFS8. Both prototypes show large biases in the teleconnection amplitude. UFS7 overestimates the magnitude of response across most lead times. In UFS8, the teleconnection response is generally weaker than reanalysis for the first two weeks (relative amplitude < 1), but becomes stronger than reanalysis in Week 3–4 (relative amplitude > 1). A difference is found between Fig. 4 and Fig. 2, in that the analysis based on the STRIPES index (Fig. 2) indicates a systematic underestimation of the amplitude in UFS7, whereas the relative amplitude metric (Fig. 4) suggests an overestimation of the amplitude. The possible causes of this discrepancy are investigated by changing the calculation domain and method. The results (not shown) suggest that the inconsistency is most likely due to differences in the calculation method, e.g., the STRIPES index is calculated based on composites while the pattern CC and relative amplitude metrics apply to individual events.

Figure 5 shows the relationship between the MJO and the extratropical storm tracks measured by EKE850 in Week 3–4. During MJO phases 8–1, a strong enhancement in storm track activity is seen near northeastern North America in reanalysis. Both prototypes have difficulty capturing these features and overall, the response in storm activity at Week 3–4 lead times after these two MJO phases is poorly predicted (pattern correlation is 0.14 in UFS7 and 0.06 in UFS8). The bias largely comes from the overestimation of changes over the North Pacific and an eastward shift of the

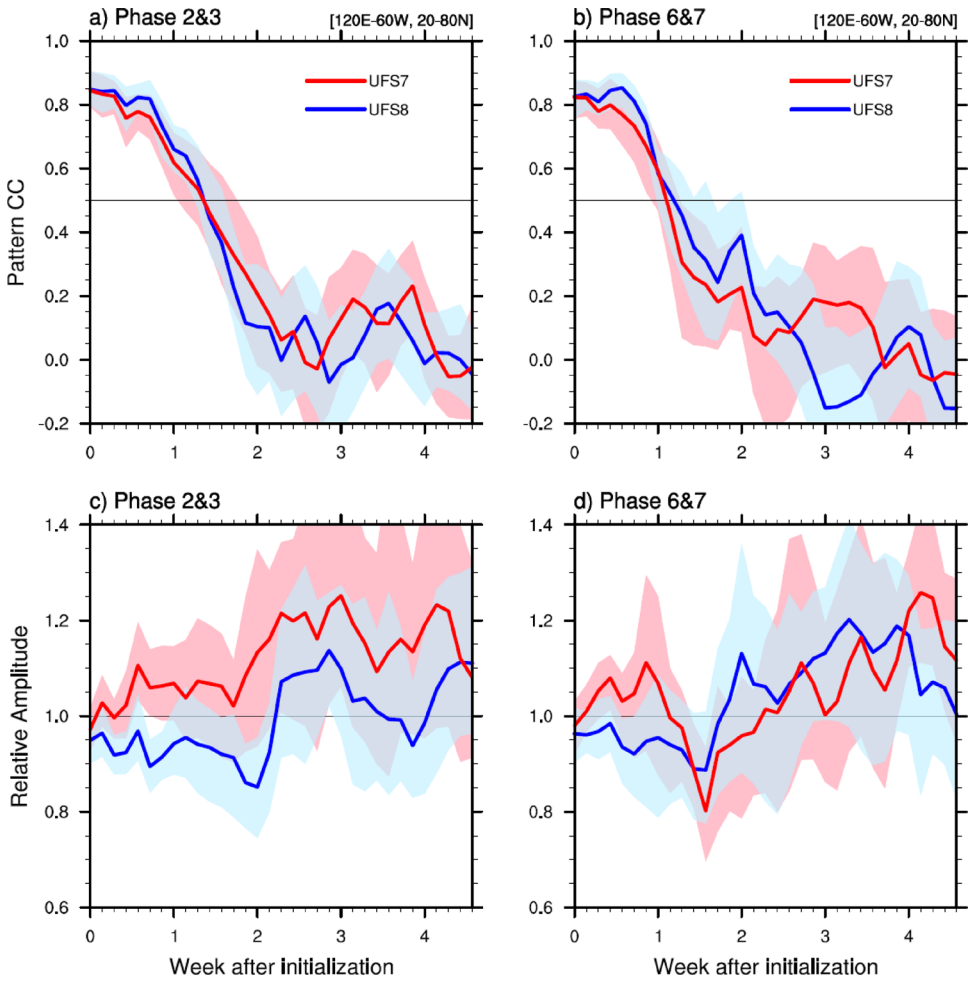

Fig. 4 Pattern correlation coefficient (UFS vs reanalysis) and relative amplitude (UFS/reanalysis) of Z500a over the PNA region (120° E–60° W, 20–80° N) vs forecast lead days for the MJO phases **a**, **c** 2–3 and **b**, **d** 6–7. Horizontal solid lines in (**a**) and (**b**) represent the reference line of pattern correlation at 0.5. Horizontal solid lines in (**c**) and (**d**) represent the reference line above (below) which Z500a is overestimated (underestimated) in the UFS models. The red and blue shadings indicate the significance at the 0.05 level determined by the bootstrapping test. The lower boundary represents the minimum 2.5th percentile of the bootstrapping distribution between the models, and the upper boundary represents the maximum 97.5th percentile distribution between the models

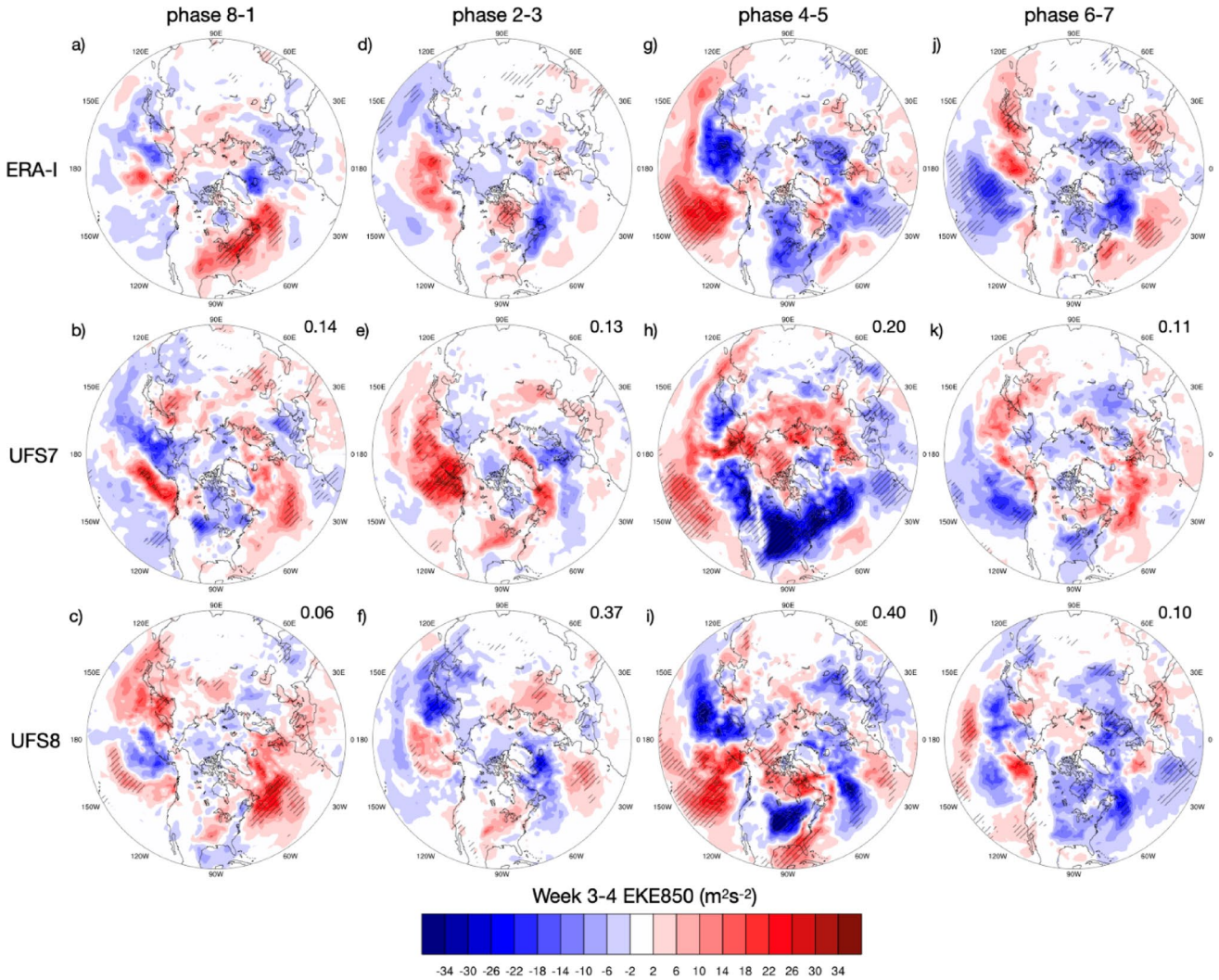

Fig. 5 Week 3–4 composites for anomalous extratropical cyclone activity (EKE850) after different MJO phase combinations in top) reanalysis, middle) UFS7, and bottom) UFS8. The significance at the 0.05 level is indicated by the hashing. The pattern correlation between reanalysis and the model results over the extratropics (20–80° N) is shown in the upper right corner of each plot

increased storm activity over the North Atlantic. When the MJO is located over the Indian Ocean in phases 2–3, the storm activity response in reanalysis is relatively weak, with a slight increase over most of the North Pacific and a slight decrease over the North Atlantic, both of which are generally not significant. UFS8 performs better than UFS7 in capturing the MJO-storm track relationships, especially for the decrease in storm activity over the North Atlantic. The east–west dipole response in storm tracks over the North Pacific is also better captured in UFS8 albeit with large biases in the amplitude. MJO phases 4–5 tend to have the strongest impact on storm tracks which show an east–west dipole response over the North Pacific featuring increased activity in the eastern part of the storm track and decreased activity in the western part, suggesting an eastward shift of the storm track. This response is accompanied by a widespread weakening of storm activity over most of North America and the North Atlantic. Again, UFS8 appears to better capture the pattern shown by reanalysis, and the improvement in prediction skill largely stems from better predictions over the North Pacific. The dipole response is still observed after MJO phases 6–7 in both the North Pacific and North Atlantic, in which case a weakening of storm tracks is expected over most of the North Pacific and a southward shift is seen over the North Atlantic. Both prototypes have difficulty reproducing the relationships between the MJO and storm tracks for MJO phases 6–7, especially for the North Atlantic. Generally, for MJO phases 2–5, UFS8 outperforms UFS7 in capturing the dipole response over the North Pacific. This greater improvement in prediction skill over the North Pacific in comparison to the other regions is consistent with the results from the STRIPES index (Fig. 2).

The sensitivity of Northern Hemisphere precipitation to the MJO forcing is examined with the STRIPES index

applied to precipitation and shown in Fig. 6. The strongest MJO teleconnection signal for extratropical precipitation falls within the North Pacific and North Atlantic storm track regions. Both prototypes underestimate the precipitation response over the ocean and slightly overestimate the response over much of North America. The evolution of the STRIPES index from Week 1–2 to Week 4–5 clearly shows underestimation in the North Pacific, North Atlantic, and Europe regions across lead times. The underestimation, however, is slightly improved in UFS8 compared to UFS7, especially over the North Pacific during the first half of the forecast period.

Figure 7 shows the composites of Week 3 and Week 4 surface temperature anomalies after MJO phases 3 and 7. When the MJO is located over the Indian Ocean (phase 3), significant cold temperature anomalies are observed in the Arctic region, and warm temperature anomalies are present over Europe. In North America, cold anomalies are present over the United States (US) West Coast and warm anomalies to the north. This temperature response is generally persistent out to Week 4, with changes mainly in the amplitude which tends to be stronger and a zonal expansion of the warm anomalies. A comparison between UFS7 and UFS8 indicates that UFS7 better captures the sign, amplitude, and approximate locations of Week 3 temperature

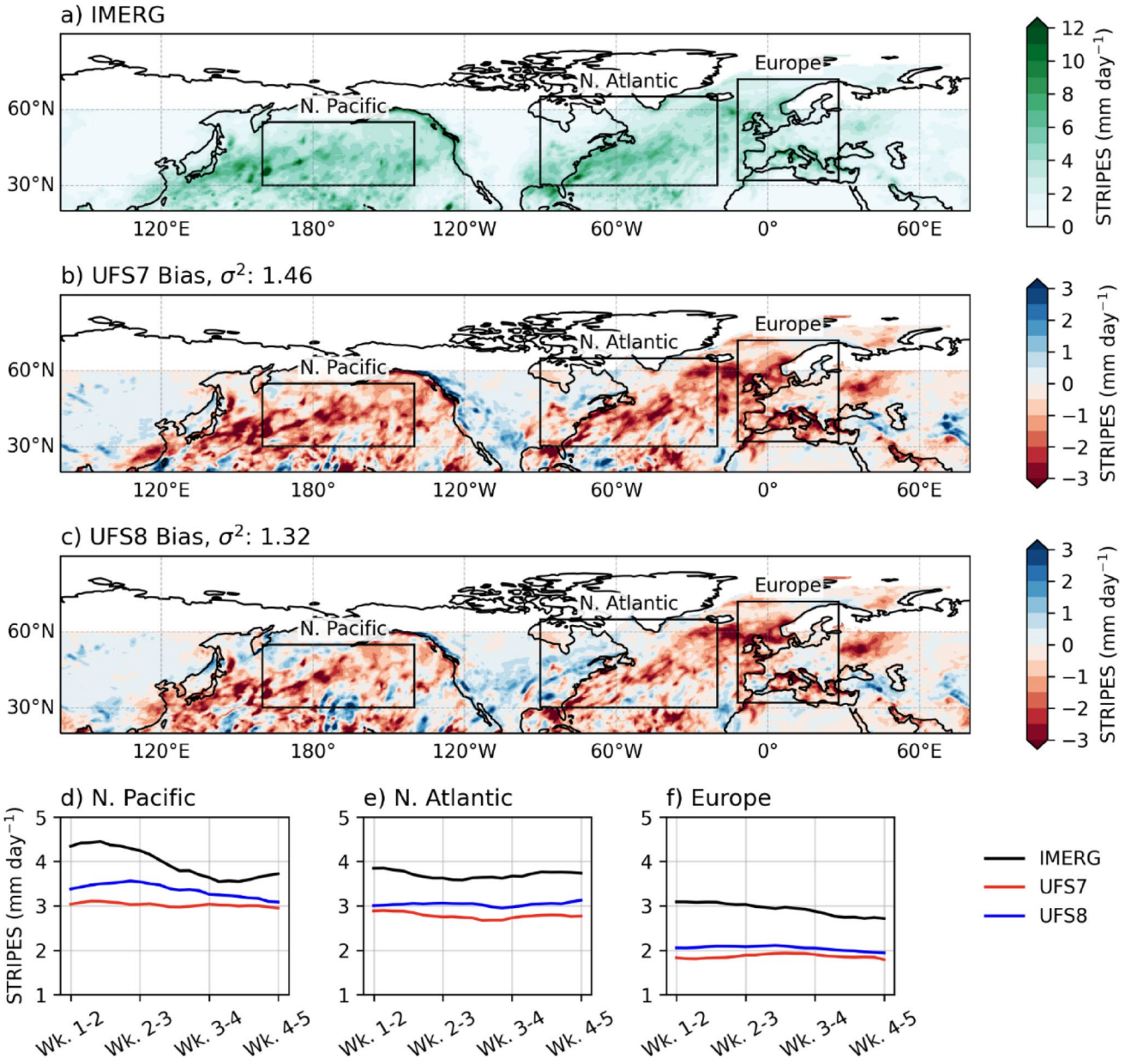

Fig. 6 As in Fig. 2, but for precipitation

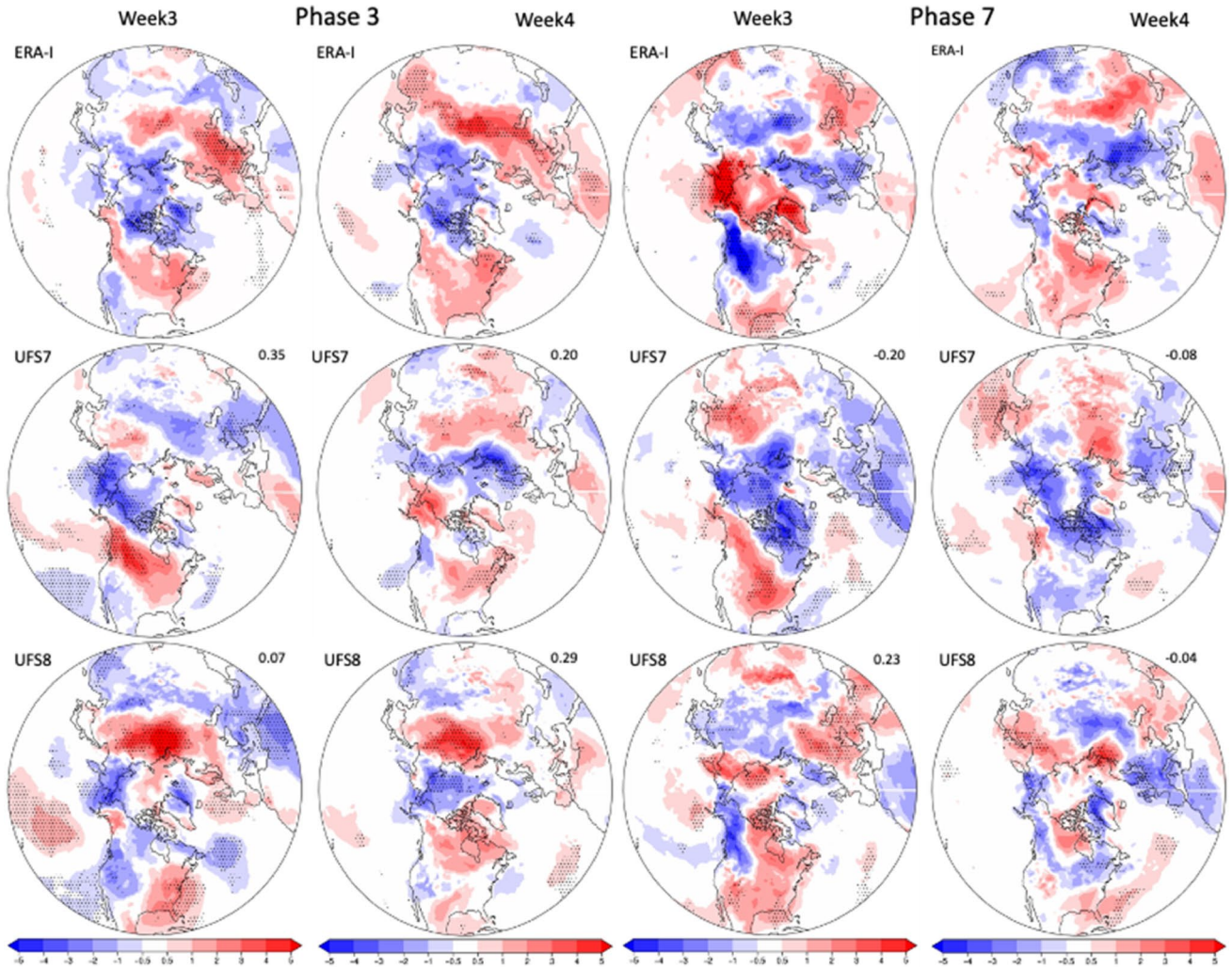

Fig. 7 Composites of 2 m temperature anomalies in Week 3 and Week 4 after MJO (left two columns) phase 3 and (right two columns) phase 7 in reanalysis (top), UFS7 (middle), and UFS8 (bottom). Dotted areas denote anomalies that are statistically significant at the 0.05 level based on a bootstrap resampling calculation. Numbers in the upper right corners show the pattern correlation between reforecasts and reanalysis

anomalies over North America and the Arctic as seen in reanalysis. UFS8, on the other hand, forecasts erroneous too-strong cold anomalies for most of the western US. The degradation of MJO teleconnections seen in UFS8 may be related to the faster propagation of the MJO over the Indian Ocean compared to UFS7 (Fig. 1), which will result in a weak positive NAO phase. Both prototypes fail to capture the positive anomalies over Europe (an issue relevant to the teleconnection stratospheric pathway that will be discussed in the next section) but reproduce the general persistence of temperature anomalies from Week 3 to Week 4.

In reanalysis, the temperature response after MJO phase 7 shows a sign reversal over North America from Week 3 to 4 (from cold anomalies to warm anomalies). Both prototypes struggle to forecast this reversal with the correct sign as they both show an opposite response than reanalysis in Week 3. In Week 4, UFS7 forecasts erroneous negative temperature anomalies over North America, while the negative anomalies shrink to the north in UFS8, and southern North America is replaced by weaker-than-reanalysis positive anomalies.

In summary, with a better MJO prediction in UFS8 (Fig. 1), a noticeable improvement in capturing the geopotential height response with realistic amplitude over the North Pacific (Fig. 2), the PNA-like pattern and its evolution (Fig. 3), a dipole response in storm tracks over the North Pacific after MJO phases 2–5 (Fig. 5) are found in UFS8 than in UFS7. This improvement, however, could be dominated by a few strong events as the pattern correlation of geopotential height response over the PNA region for individual events does not show significant differences between the two prototypes (Fig. 4). UFS8 also has slightly better skill in predicting the precipitation response at all lead

times and the temperature response in Week 4 than UFS7 (Figs. 6, 7).

3.2 Stratospheric pathway

In this section, we discuss how the UFS prototypes capture the MJO stratospheric pathway including upward propagating wave activity from the troposphere into the stratosphere, changes in the polar vortex, downward wave propagation, and the subsequent changes in climate modes.

As a proxy for upward wave propagation from the troposphere to the stratosphere, meridional eddy heat flux ($v\prime T\prime$) from 40° to 80°N is calculated, where v is the meridional wind and T is temperature, and the prime denotes the deviations from the zonal mean. Hence, the eddy heat flux corresponds to the heat transported northward by eddies. Figure 8 shows the wavenumber-1 and wavenumber-2 meridional heat flux anomaly at 500 hPa and 100 hPa during Week 1 to Week 5 after each active MJO phase. In reanalysis, positive anomalies of wavenumber-1 + 2 heat fluxes are present during Week 1–3 at 500 hPa and Week 2–4 at 100 hPa after MJO phase 5, which correspond to increased planetary wave flux entering the stratosphere. Corresponding to the increased heat fluxes in the stratosphere, the stratospheric polar vortex is weakening 4–5 weeks after MJO phase 5. Both prototypes simulate the increased heat fluxes in the troposphere following MJO phases 4–5. UFS8 better captures the increase in heat fluxes in the lower stratosphere after MJO phase 5 than UFS7, whose maximum positive heat flux anomalies are present in Week 3 after MJO phase 4 and are less pronounced than in UFS8 and reanalysis. Comparing the first two rows in Fig. 8, the strong heat flux at 500 hPa is mainly dominated by its wave-1 component in both prototypes and reanalysis. Given the weaker heat fluxes after MJO phases 4–5 in UFS7, the weakening of the polar vortex in UFS7 is less pronounced than in UFS8 and reanalysis (Fig. 8j, k, l), and its negative anomaly is biased to peak in Week 3–4 following MJO phase 4 (Fig. 8k). Besides, both prototypes

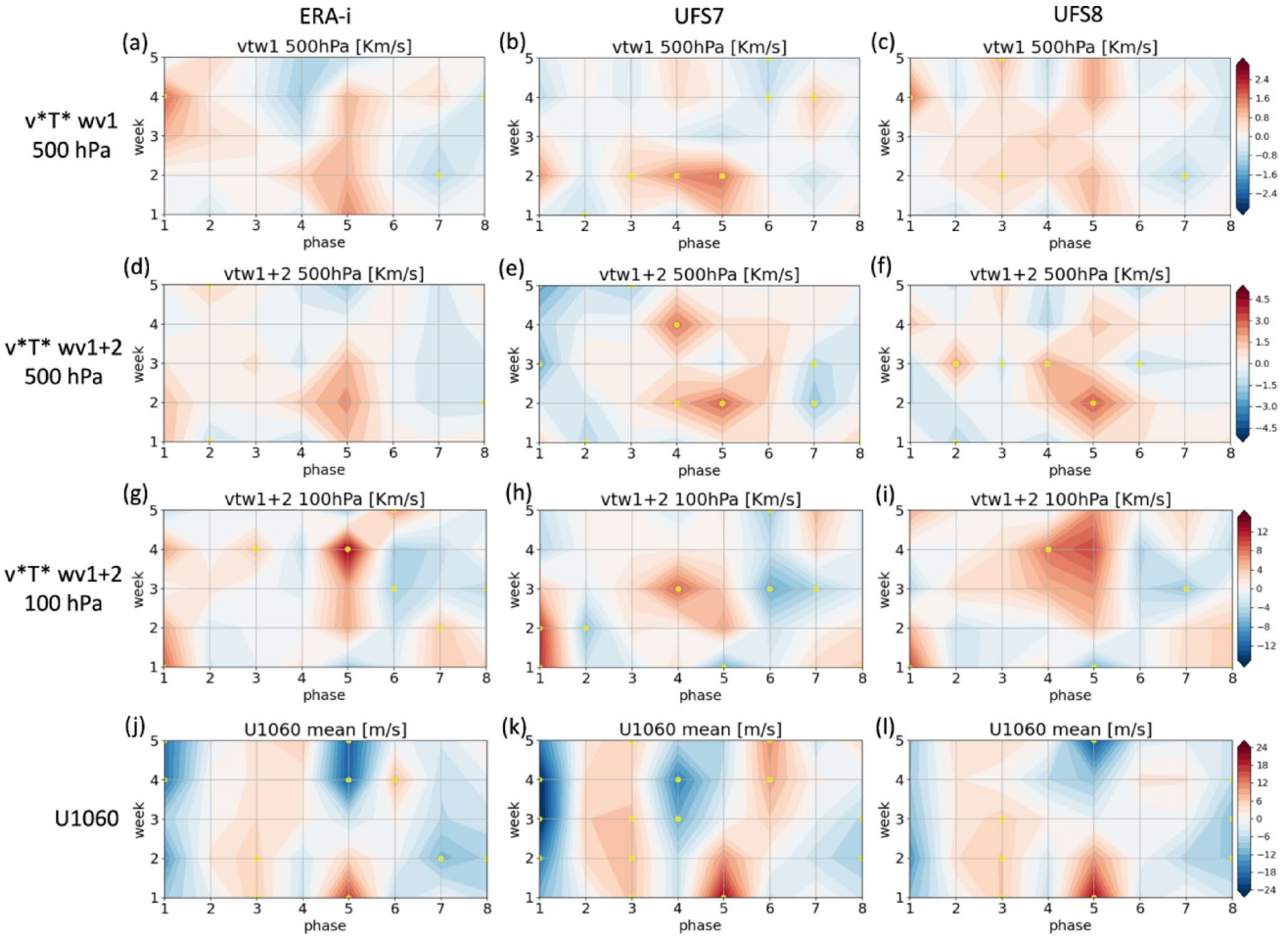

Fig. 8 November–March wavenumber-1 component of meridional heat flux ($v\prime T\prime$) anomalies averaged over 40–80° N at 500 hPa (first row), wavenumber-1 and wavenumber-2 meridional components of meridional heat flux anomalies at 500 hPa (second row) and 100 hPa (third row), and zonal-mean zonal wind anomalies at 10 hPa over 60° N (U1060; fourth row) following MJO phases 1–8 for reanalysis (left), UFS7 (middle), and UFS8 (right). Yellow dots indicate anomalies statistically significant at the 0.1 level based on the bootstrap test

reproduce the positive anomaly of the polar vortex strength in Week 1–5 after MJO phases 2–3. In general, UFS8 better predicts the upward wave propagation and the weakening of the polar vortex, with comparable strength to reanalysis.

To further look into the different responses of the stratospheric polar vortex after MJO phases, Fig. 9 shows the distribution of zonal-mean zonal wind at 10 hPa and 60°N (U1060) during Week 1–2 and Week 3–5 after MJO phases 1–2 (blue) and MJO phases 5–6 (yellow), respectively, as in Stan et al. (2022). The distribution of U1060 following MJO phases 5–6 (yellow) shifts to weaker vortex strength during Week 1–2 and Week 3–5 compared to that following MJO phases 1–2 (blue) in reanalysis, as indicated by the differences in the mean, 5th, and 95th percentile of U1060, consistent with other subseasonal prediction models (Domeisen et al. 2020). This shift indicates that a weakening of the stratospheric polar vortex occurs more frequently after MJO phases 5–6 than after MJO phases 1–2. The distributions of the polar vortex strength in UFS7 and UFS8 are comparable to those of the reanalysis during the first two weeks, as evident from the means and 5th and 95th percentiles of U1060. However, differences in the distributions of U1060 between the two prototypes and reanalysis become larger at longer lead times during Week 3–5. The polar vortex strength simulated by UFS7 is weaker than the reanalysis and UFS8 after MJO phases 1–2, as shown in the mean values of U1060. On the other hand, the polar vortex strength simulated by UFS8 after MJO phases 5–6 in Week 3–5 has smaller variations and overall weaker magnitude (as indicated by the mean values and 95th percentile of U1060) than UFS7 and reanalysis. Together, the differences between the polar vortex strength after MJO phases 5–6 and 1–2 are smaller in UFS7 and larger in UFS8 in Week 3–5 compared to reanalysis. In general, compared to S2S models in Stan et al. (2022), UFS8 tends to better simulate the difference in the mean of polar vortex strength distributions between MJO phases 5–6 and 1–2.

The anomalous conditions of the winter stratospheric polar vortex can have a downward impact on surface weather, particularly over the North Atlantic and European regions. For instance, the weakening of the polar vortex is typically followed by a negative phase of the NAO (e.g., Charlton-Perez et al. 2018; Domeisen 2019). The downward coupling with the tropospheric circulation in response to the MJO is diagnosed with the polar cap and the Euro-Atlantic sector geopotential height responses in Fig. 10. Positive anomalies indicate a negative phase of the Northern Annular Mode (NAM) and NAO, which is often accompanied by anomalously weak polar vortex events such as sudden stratospheric warmings (Baldwin et al. 2021). An increase in polar cap height is observed 3–5 weeks after MJO phase 5 at 10 hPa (Fig. 10a) and in Week 1–5 after MJO phases

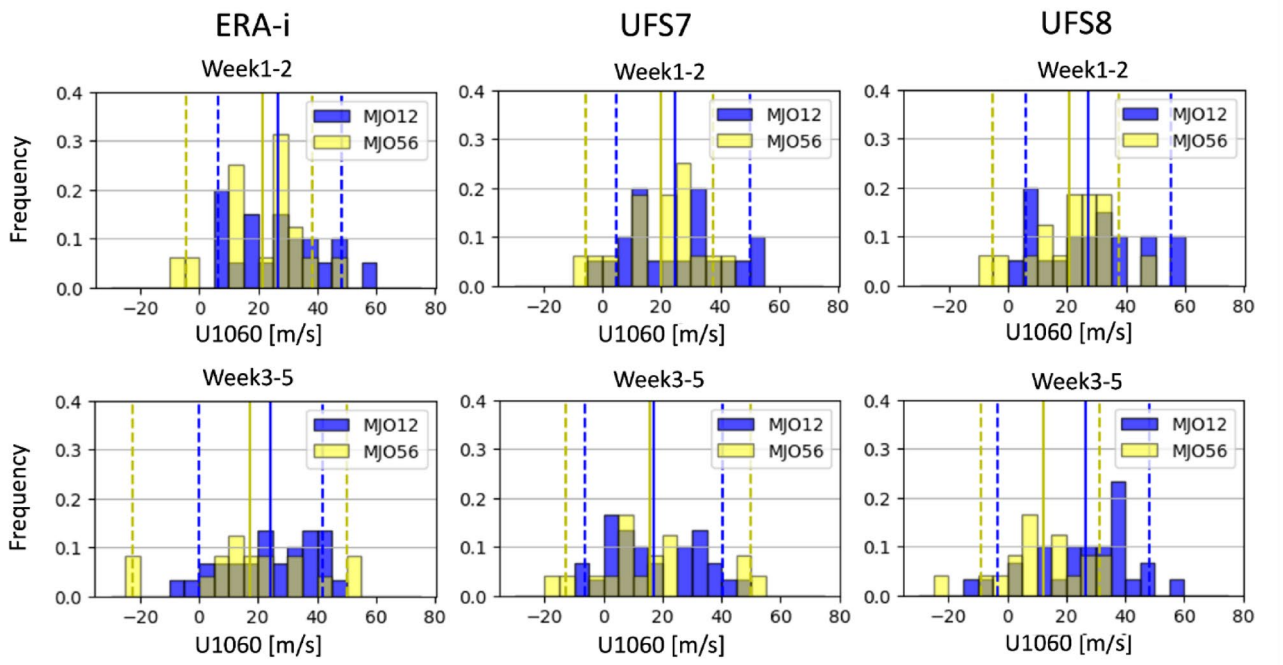

Fig. 9 Histograms of zonal-mean zonal wind at 10 hPa over 60° N (U1060) for forecasts during week 1–2 (top) and week 3–5 (bottom) following MJO phases 1–2 (blue) and phases 5–6 (yellow). The solid blue and yellow lines indicate the mean values of U1060 during phases 1–2 and 5–6, respectively. The dashed blue and yellow lines indicate the 5th and 95th percentile of U1060 during MJO phases 1–2 and 5–6, respectively

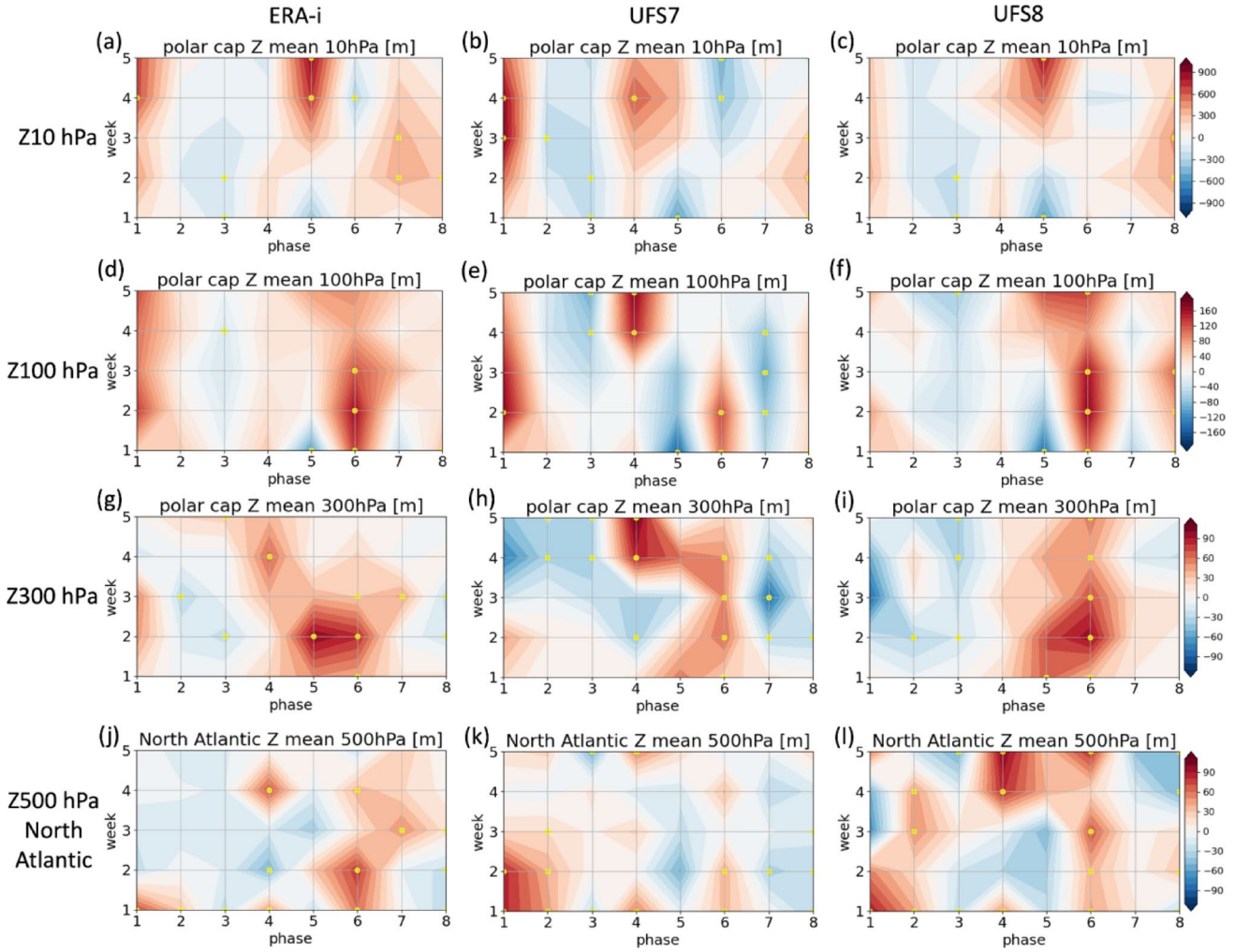

Fig. 10 November-March geopotential height anomalies averaged over the polar cap at 10 hPa (first row), 100 hPa (second row), 300 hPa (third row), and averaged over the subpolar North Atlantic at 500 hPa (fourth row), following MJO phases 1–8 for reanalysis (left), UFS7 (middle), and UFS8 (right). Yellow dots indicate anomalies statistically significant at the 0.1 level based on the bootstrap test

5–6 at 100 and 300 hPa (Fig. 10d and g), corresponding to the weakening of the polar vortex. UFS8 simulates a more realistic polar-cap averaged geopotential height response after MJO phases 5–6 than UFS7 in both the stratosphere and troposphere (Fig. 10c, f, and i), consistent with the downward impacts of the weakening of the polar vortex. While UFS7 simulates the maximum positive polar-cap height anomalies in Week 4–5 in both the stratosphere and troposphere (Fig. 10b, e, and h), which is different from UFS8 and reanalysis, it does capture the negative phase of NAM following MJO phase 6 (Fig. 10b). UFS8 captures the observed pattern well but overestimates the magnitude of the response in Week 1–5 after MJO phase 6 at both 100 and 300 hPa (Fig. 10f and i). In terms of the changes over the Euro-Atlantic sector, which represent changes in the NAO, a negative phase of the NAO is observed 1–4 weeks after MJO phase 6 and 4–5 weeks after MJO phase 5 (Fig. 10j). Both prototypes reproduce the positive response over the North Atlantic sector in the troposphere in Week 1–5 following MJO phase 6 and in Week 5 following MJO phase 5 with biases in the magnitude. UFS8 produces a stronger response in Week 5 following MJO phase 5 and a weaker response in Week 1–3 following MJO phase 6 than in reanalysis (Fig. 10l), while the response is systematically weaker in UFS7 (Fig. 10k). In general, UFS8 better captures the downward propagation of the MJO stratospheric pathway while both prototypes underestimate the MJO impacts on the NAO in Week 1–3. Compared to other prototypes (Garfinkel et al. 2025), the UFS8 performs better in predicting the upward wave propagation, changes in the polar vortex, and downward coupling, and shows comparable skill in predicting the MJO impacts on the NAO.

As indicated by Figs. 8, 9 and 10, both the upward and downward couplings are simulated well by UFS8, although the downward coupling near the surface has large biases in its magnitude. This bias is also reflected in the near-surface

temperature response in Fig. 7: a hallmark of a negative NAM (as is observed in Week 3–4 after MJO phase 7) is a cold temperature anomaly over Northern Europe extending into Siberia, and warm temperatures over Greenland and the Middle East (Butler et al. 2017). Both prototypes miss this response and predict a response opposite to reanalysis for Week 3–4 after MJO phase 3.

Next, we assess whether the bias in the amplitude of the response seen in Fig. 10 is specific to the MJO, or a more general problem that S2S models have in capturing downward coupling from the stratosphere. Figure 11 shows the persistence of the lower stratospheric temperature anomalies by associating polar cap temperature at 100 hPa with temperature at a 2-week lag. Persistent lower stratospheric anomalies are essential for the continuous forcing of the troposphere. UFS7 and UFS8 are found to underestimate the maintenance of the polar cap temperature signals. In addition, the general impacts of the polar vortex on the tropospheric circulation in the Atlantic are examined by comparing polar cap height at 100 hPa with height anomalies

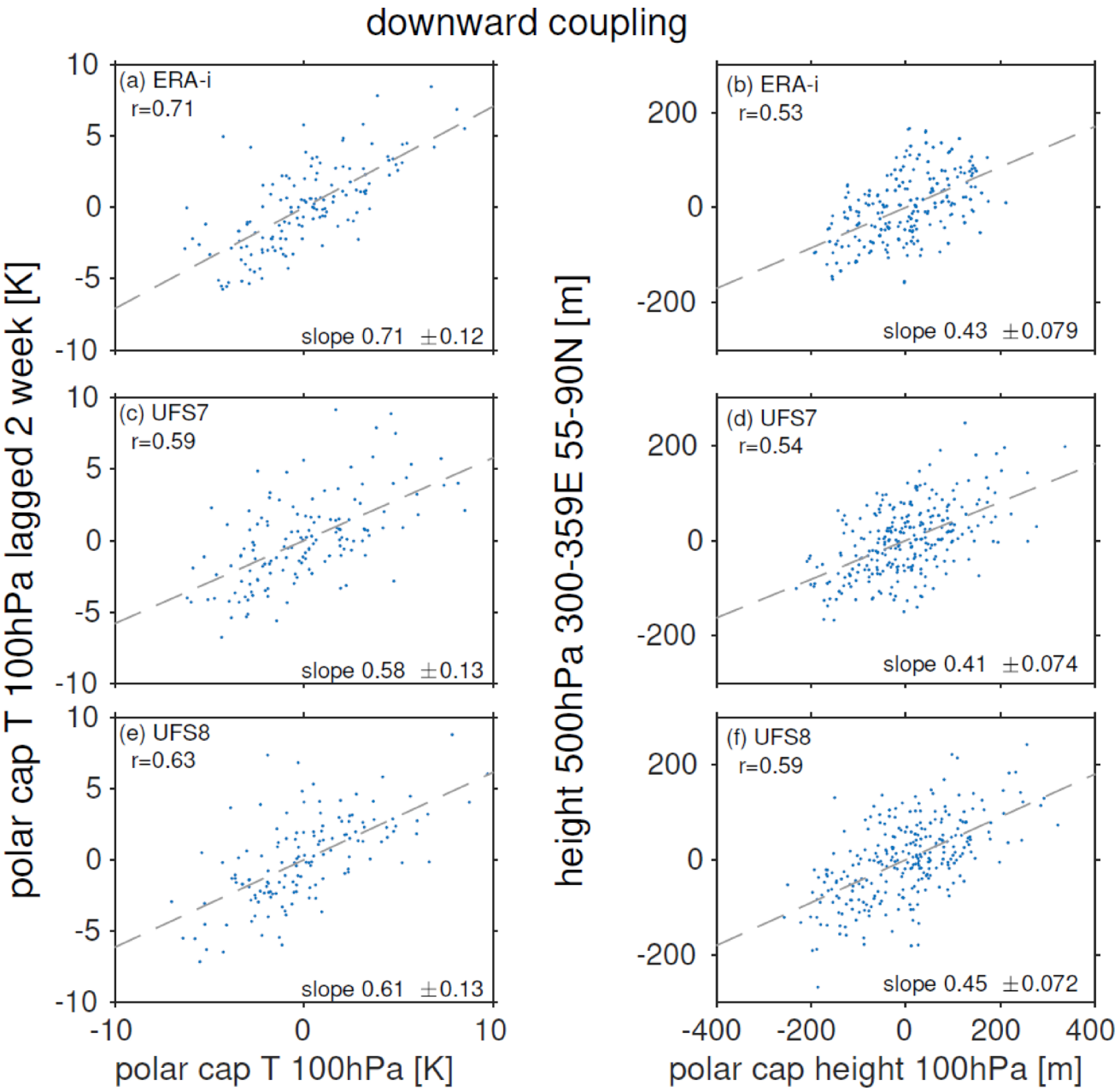

Fig. 11 (left) Association between polar cap temperature anomalies from Week 2 and 3 to the corresponding anomalies 2 weeks after, which indicates the persistence of the lower stratospheric anomalies. (right) Comparison of polar cap height anomalies at 100 hPa with height anomalies at 500 hPa over the North Atlantic during weeks 2–5

Table 1 Summary of model performance in predicting the MJO and MJO teleconnections

Metrics			Comparison
MJO	RMM index		Skillful prediction in ACC increases from 19 days in UFS7 to 27 days in UFS8
	Propagation and amplitude		Faster propagation speed in UFS8 than in UFS7; Weaker MJO in both prototypes
MJO teleconnections-tropospheric pathway	Z500	STRIPES index	The weaker response over the North Pacific in UFS7 is largely improved in UFS8 with realistic amplitude; No systematic improvement over the North Atlantic and Europe
		PNA evolution	Better performance in UFS8 than UFS7
		Pattern CC and relative amplitude	No systematic improvement is found across lead times from UFS7 to UFS8; Both prototypes show large biases in the amplitude
	Extratropical cyclone activity		UFS8 outperforms UFS7 in terms of capturing the dipole response over the North Pacific after phases 2–5
	Precipitation		The underestimation of precipitation response is slightly improved in UFS8 than in UFS7
	T2m		Slightly better performance in Week 4 forecasts in UFS8 than in UFS7, but degradation in week 3 forecasts in UFS8 after phase 3; Both prototypes fail to capture the positive anomalies over Europe
MJO teleconnections-stratospheric pathway	Upward wave propagation		UFS8 better predicts the upward wave propagation and the weakening of the polar vortex than UFS7, with comparable strength to reanalysis
	Downward coupling		UFS8 better captures the downward propagation of the MJO stratospheric pathway while both prototypes underestimate the MJO impacts on the NAO in weeks 1–3

at 500 hPa over the Atlantic. Both models are capable of capturing the general downward coupling. Thus the relatively poor simulation of the NAO response to the MJO in these prototypes is not related to a struggle in predicting the downward response over the North Atlantic region but rather related to the poor simulation of the vortex persistence and/or internal variability.

4 Summary and discussion

The MJO is recognized as an important source of subseasonal predictability due to its dominant variations on intraseasonal timescales, significant impacts on global weather and climate phenomena, and potentially high predictability out to 6 weeks. This study is focused on examining the degree to which an improved MJO prediction may lead to changes in the prediction skill of MJO teleconnections in the recently developed high-resolution fully coupled model (UFS8), which includes atmosphere, land, ocean, sea-ice, wave, and aerosol components, in comparison to the previous model generation (UFS7). This comparison enables a better understanding of how much a more comprehensive coupling process between different model components along with the updated physics packages may influence the model capability of predicting the MJO and its teleconnections, and how the better predicted MJO influences the model skill in capturing MJO teleconnections. This study can also be used as a reference for 1) applying various MJO teleconnection metrics to the operational systems to quantify the prediction skill and 2) a comprehensive update of the recent advances in MJO teleconnection prediction by examining both the tropospheric and stratospheric pathways.

Table 1 provides a summary of the model performance in capturing the MJO and MJO teleconnections in comparison between UFS7 and UFS8. The overall MJO prediction skill measured by the correlation of predicted and observed RMM indices improves from 19 to 27 days from UFS7 to UFS8, which is comparable with the high-skill models in the S2S and SubX datasets. Forecasts of MJO teleconnections are then investigated to examine how this improved MJO prediction skill impacts the skills of MJO teleconnections. UFS8 realistically predicts the evolution of the PNA after an active MJO event, while large biases exist in UFS7. The above improvement shown in the composites is not found when measuring pattern correlation separately using individual MJO events, suggesting that one or two strong events may be contributing to the improvement seen in UFS8. For the stratospheric pathways, UFS8 better captures the upward and downward wave propagation but underestimates the MJO impacts on NAO, which may be due to the fact that it struggles to maintain the polar cap temperature signals and has a large spread in the predicted values (the same in UFS7). Alternatively, even if downward coupling is handled well by a model in general, that does not imply it is capable of distinguishing whether a particular event or set of events will have a tropospheric impact (Nebel et al. 2024). In summary, the results show that an increase in MJO prediction skill tends to translate to a similar increase in prediction skill of the geopotential height and storm track response over the North Pacific, while predictions in the downstream regions (North Atlantic and Europe) show no significant improvement. Predictions of precipitation and temperature also do not discernibly benefit from an improved MJO. In terms of stratospheric processes, a skill increase in MJO-associated upward wave propagation and the following weakening in the polar vortex is detected with an improved MJO prediction. However, the MJO-NAO relationships are

not captured very well by the model, which may be related to the too-fast MJO propagation in the model as implied by the previous studies.

By comparing the results of UFS8 with Zheng et al. (2024) and Garfinkel et al. (2025) who examined the UFS5 and UFS6, further improvement in MJO teleconnections include: (1) Better predicted MJO with higher prediction skill if measured by the RMM indices (about 4 days improvement); (2) More realistic evolution of PNA response and more realistic magnitude of the teleconnection response over the North Pacific; 3) Weakening in polar vortex with strength comparable to reanalysis and correct prediction of the lead time that shows the increased positive heat fluxes in the troposphere. Here, we mainly compare the results across different UFS generations to allow a consistent and fair comparison rather than comparing with S2S or SubX datasets which have different forecast periods and initialization dates.

As part of the upgrades from UFS7 to UFS8, the microphysics packages are updated from GFSv16 Physics with GFDL Cloud Microphysics Scheme in UFS7 to GFSv17 Physics with Thompson Microphysics Scheme in UFS8. The microphysics could be particularly relevant to the MJO and its teleconnections; therefore, how this difference in microphysics may potentially influence MJO teleconnections is tested. We performed single-column model experiments using the two physics packages. Figure S1 illustrates the specific humidity tendency, considering contributions from microphysics alone and from all physics processes (moist convection, microphysics, and others). It is noted that although two different microphysics schemes exhibit some differences, their combined contribution to the total tendencies has no notable impact. This suggests that changes due to different microphysics packages may not have a significant impact on MJO teleconnections. This result is not entirely surprising because parameterizations are tuned to emulate observed bulk statistics.

In addition to the MJO-related stratospheric processes examined in this study, we also investigated troposphere-stratosphere coupling more broadly in the two prototypes. Here, we discuss the upward coupling, its modulation on the polar vortex, and the downward coupling within the stratosphere. Figure S2 compares the wavenumber-1 and wavenumber-2 heat flux at 500 hPa with heat flux 3 days after at 100 hPa. A linear relationship around 0.4 is observed in reanalysis, which is realistically captured in UFS8 but largely underestimated in UFS7. This suggests that the coupling between the stratospheric waves and their precursors is well captured by UFS8. UFS8 also has the best performance in predicting upward coupling when compared with UFS5 and UFS6 (Garfinkel et al. 2025). The influence of the meridional heat flux on the polar vortex is indicated in Fig. S3 by associating the wavenumber 1+2 heat flux at 100 hPa with the tendency in polar cap height. Similar to previous UFS prototypes, UFS8 captures the linear relationship well but overestimates the impact on the vortex (higher value of the slope compared to reanalysis). The downward coupling within the stratosphere is shown in Fig. S3. UFS8 reproduces the weakening of the vortex at 100 hPa following three days after a weakening of the vortex at 10 hPa. In general, UFS8 realistically captures the upward coupling from the troposphere to the stratosphere and downward propagation in the stratosphere but overpredicts the wave impacts on the polar vortex.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-025-07783-9.

Acknowledgements This study was supported by the NOAA/OAR Weather Program Office through grant NA22OAR4590216. H. Kim was supported by the National Research Foundation of Korea grant funded by the Korean government (NRF-RS-2023-00278113) and by the Korea Meteorological Administration Research and Development Program (RS-2025-02313090).

Funding This study was supported by the NOAA/OAR Weather Program Office through grant NA22OAR4590216. H. Kim was also supported by the National Research Foundation of Korea grant funded by the Korean government (NRF-RS-2023-00278113) and by the Korea Meteorological Administration Research and Development Program (RS-2025-02313090).

Data availability ERA-Interim reanalysis is available at https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-interim. UFS reforecast data is available on AWS https://registry.opendata.aws/noaa-ufs-s2s/. Process-based metrics used in this study are integrated into an open-source Python package (MJO-Teleconnections diagnostics package; Stan et al. 2025a, b) which could be applied to both operational and research studies.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Adcroft A et al (2019) The GFDL global ocean and sea ice model OM4. 0: model description and simulation features. J Adv Model Earth Syst 11:3167–3211

Ahn M-S, Kim D, Ham Y-G, Park S (2020) Role of Maritime Continent land convection on the mean state and MJO propagation. J Clim 33:1659–1675

Baldwin MP et al (2021) Sudden stratospheric warmings. Rev Geophys 59:e2020RG000708

Bengtsson L, Dias J, Tulich S, Gehne M, Bao J-W (2021) A stochastic parameterization of organized tropical convection using cellular automata for global forecasts in NOAA's unified forecast system. J Adv Model Earth Syst 13:e2020MS002260. https://doi.org/10.1029/2020MS002260

Bengtsson L, Gerard L, Han J, Gehne M, Li W, Dias J (2022) A prognostic-stochastic and scale-adaptive cumulus convection closure for improved tropical variability and convective gray-zone representation in NOAA's unified forecast system (UFS). Mon Weather Rev 150:3211–3227

Brassington G et al (2015) Progress and challenges in short-to medium-range coupled prediction. J Oper Oceanogr 8:s239–s258

Butler AH, Sjoberg JP, Seidel DJ, Rosenlof KH (2017) A sudden stratospheric warming compendium. Earth Syst Sci Data 9:63–76

Charlton-Perez AJ, Ferranti L, Lee RW (2018) The influence of the stratospheric state on North Atlantic weather regimes. Q J R Meteorol Soc 144:1140–1151

Chen J-H, Lin S-J (2013) Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. J Clim 26:380–398

Chin M. et al. (2003) A global aerosol model forecast for the ACE-Asia field experiment. J Geophys Res Atmos. vol 108

Dee DP et al (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quart J Roy Meteor Soc 137:553–597. https://doi.org/10.1002/qj.828

Domeisen D (2019) Estimating the frequency of sudden stratospheric warming events from surface observations of the north Atlantic oscillation. J Geophys Res Atmos 124:3180–3194. https://doi.org/10.1029/2018JD030077

Domeisen D et al (2020) The Role of the stratosphere in subseasonal to seasonal prediction: 2. predictability arising from stratosphere-troposphere coupling. J Geophys Res Atmos. https://doi.org/10.1029/2019JD030923

Garfinkel CI, Feldstein SB, Waugh DW, Yoo C, Lee S (2012) Observed connection between stratospheric sudden warmings and the Madden-Julian Oscillation. Geophys Res Lett vol 39

Garfinkel CI, Chen W, Li Y, Schwartz C, Yadav P, Domeisen D (2022) The winter North Pacific teleconnection in response to ENSO and the MJO in operational subseasonal forecasting models is too weak. J Clim 35:8013–8030

Garfinkel CI et al (2025) The impact of model levels on the prediction of MJO teleconnections. Part 2: The stratospheric pathways in the UFS global coupled model. Clim Dyn 63:1–17

Gentile ES, Gray SL, Barlow JF, Lewis HW, Edwards JM (2021) The impact of atmosphere–ocean–wave coupling on the near-surface wind speed in forecasts of extratropical cyclones. Bound-Layer Meteorol 180:105–129

Gottschalck J et al (2010) A framework for assessing operational Madden–Julian oscillation forecasts: a CLIVAR MJO working group project. Bull Amer Meteor Soc 91:1247–1258

Green MR, Furtado JC (2019) Evaluating the joint influence of the Madden-Julian oscillation and the stratospheric polar vortex on weather patterns in the Northern hemisphere. J Geophys Res Atmos 124:11693–11709

Han J, Bretherton CS (2019) TKE-based moist eddy-diffusivity mass-flux (EDMF) parameterization for vertical turbulent mixing. Weather Forecast 34:869–886

Han J, Li W, Yang F, Strobach E, Zheng W, Sun R (2021) Updates in the NCEP GFS cumulus convection, vertical turbulent mixing, and surface layer physics.

Harris LM, Lin S-J (2013) A two-way nested global-regional dynamical core on the cubed-sphere grid. Mon Weather Rev 141:283–306

Henderson SA, Maloney ED, Barnes EA (2016) The influence of the Madden–Julian oscillation on Northern Hemisphere winter blocking. J Clim 29:4597–4616. https://doi.org/10.1175/JCLI-D-15-0502.1

Hersbach H et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049

Hoskins BJ, Karoly DJ (1981) The steady linear response of a spherical atmosphere to thermal and orographic forcing. J Atmos Sci 38:1179–1196. https://doi.org/10.1175/1520-0469(1981)038%3c1179:TSLROA%3e2.0.CO;2

Huffman GJ, Bolvin DT, Braithwaite D, Hsu K, Joyce R, Xie P, Yoo SH (2015) NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theor Basis Doc. 4(26):30

Jenney A, Randall D, Barnes E (2019) Quantifying regional sensitivities to periodic events: application to the MJO. J Geophys Res Atmos 124:3671–3683

Jiang Z, Feldstein SB, Lee S (2017) The relationship between the Madden–Julian oscillation and the North Atlantic oscillation. Quart J Roy Meteor Soc 143:240–250

Kim D, Kim H, Lee MI (2017) Why does the MJO detour the Maritime Continent during austral summer? Geophys Res Lett 44:2579–2587

Kim H, Vitart F, Waliser DE (2018) Prediction of the Madden–Julian oscillation: a review. J Clim 31:9425–9443

Kim H, Janiga MA, Pegion K (2019) MJO propagation processes and mean biases in the SubX and S2S reforecasts. J Geophys Res Atmos 124:9314–9331

Kim H, Son S-W, Kim H, Seo K-H, Kang M-J (2023) MJO influence on subseasonal-to-seasonal prediction in the Northern Hemisphere extratropics. J Clim 36:7943–7956

Kwon YC, Hong S-Y (2017) A mass-flux cumulus parameterization scheme across gray-zone resolutions. Mon Weather Rev 145:583–598

Liebmann B, Smith CA (1996) Description of a complete (interpolated) outgoing longwave radiation dataset. Bull Amer Meteor Soc 77:1275–1277. https://doi.org/10.1175/1520-0477-77.6.1274

Lin H, Brunet G, Derome J (2009) An observed connection between the North Atlantic Oscillation and the Madden–Julian oscillation. J Clim 22:364–380. https://doi.org/10.1175/2008JCLI2515.1

Lin H, Brunet G, Fontecilla JS (2010) Impact of the Madden-Julian oscillation on the intraseasonal forecast skill of the North Atlantic Oscillation. Geophys Res Lett 37:L19803. https://doi.org/10.1029/2010GL044315

Lin Y-L, Farley RD, Orville HD (1983) Bulk parameterization of the snow field in a cloud model. J Clim Appl Meteorol 22:1065–1092

Madden RA, Julian PR (1971) Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. J Atmos Sci 28:702–708. https://doi.org/10.1175/1520-0469(1971)%3c0702:DOADOI%3e2.0.CO;2

Madden RA, Julian PR (1972) Description of global-scale circulation cells in the tropics with a 40–50 day period. J Atmos Sci 29:1109–1123. https://doi.org/10.1175/1520-0469(1972)029%3c1109:DOGSCC%3e2.0.CO;2

Mori M, Watanabe M (2008) The growth and triggering mechanisms of the PNA: a MJO-PNA coherence. J Meteor Soc Japan 86:213–236. https://doi.org/10.2151/jmsj.86.213

Mundhenk BD, Barnes EA, Maloney ED, Baggett CF (2018) Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden–Julian oscillation and quasi-biennial oscillation. NPJ Clim Atmos 1:20117. https://doi.org/10.1038/s41612-017-0008-2

Nebel DM, Garfinkel CI, Cohen J, Domeisen DI, Rao J, Schwartz C (2024) The predictability of the downward vs. non-downward propagation of sudden stratospheric warmings in S2S hindcasts. Geophys Res Lett. https://doi.org/10.1029/2024GL110529

Niu GY.et al. (2011) The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. J Geophys Res Atmos. vol 116

Pegion K et al (2019) The subseasonal experiment (SubX): a multimodel subseasonal prediction experiment. Bull Amer Meteor Soc 100:2043–2060

Putman WM, Lin S-J (2007) Finite-volume transport on various cubed-sphere grids. J Comput Phys 227:55–78

Rashid HA, Hendon HH, Wheeler MC, Alves O (2011) Prediction of the Madden–Julian oscillation with the POAMA dynamical prediction system. Clim Dyn 36:649–661

Randles C et al (2017) The MERRA-2 aerosol reanalysis, 1980 onward. Part I: system description and data assimilation evaluation. J Clim 30:6823–6850

Riddle EE, Stoner MB, Johnson NC, L'Heureux ML, Collins DC, Feldstein SB (2013) The impact of the MJO on clusters of wintertime circulation anomalies over the North American region. Clim Dyn 40:1749–1766

Robertson AW, Vitart F, Camargo SJ (2020) Subseasonal to seasonal prediction of weather to climate with application to tropical cyclones. J Geophys Res Atmos 125:e2018JD029375

Sardeshmukh PD, Hoskins BJ (1988) The generation of global rotational flow by steady idealized tropical divergence. J Atmos Sci 45:1228–1251

Schwartz C, Garfinkel CI (2017) Relative roles of the MJO and stratospheric variability in North Atlantic and European winter climate. J Geophys Res Atmos 122:4184–4201

Schwartz C, Garfinkel CI (2020) Troposphere-stratosphere coupling in subseasonal-to-seasonal models and its importance for a realistic extratropical response to the Madden-Julian oscillation. J Geophys Res Atmos 125:e2019JD032043

Seo K-H, Lee H-J (2017) Mechanisms for a PNA-like teleconnection pattern in response to the MJO. J Atmos Sci 74:1767–1781

Seo K-H, Son S-W (2012) The global atmospheric circulation response to tropical diabatic heating associated with the Madden–Julian oscillation during northern winter. J Atmos Sci 69:79–96

Smith GC et al (2018) Impact of coupling with an ice–ocean model on global medium-range NWP forecast skill. Mon Weather Rev 146:1157–1180

Stan C et al (2022) Advances in the prediction of MJO teleconnections in the S2S forecast systems. Bull Am Meteor Soc 103:E1426–E1447

Stan C et al (2025a) A Python diagnostics package for evaluation of MJO-teleconnections in S2S forecast systems. Geosci Model Dev. https://doi.org/10.5194/egusphere-2025-1142

Stan C et al. (2025b) pyMTDG. https://doi.org/10.5281/zenodo.15002614

Stefanova L et al. (2022) Description and results from UFS coupled prototypes for future global, ensemble and seasonal forecasts at NCEP

Straus DM, Domeisen DIV, Lock S-J, Molteni F, Yadav P (2023) Intrinsic predictability limits arising from Indian Ocean Madden–Julian oscillation (MJO) heating: effects on tropical and extratropical teleconnections. Weather Clim Dyn 4:1001–1018

Straub KH (2013) MJO initiation in the real-time multivariate MJO index. J Clim 26:1130–1151

Thompson G, Eidhammer T (2014) A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. J Atmos Sci 71:3636–3658

Tseng K-C, Maloney E, Barnes E (2019) The consistency of MJO teleconnection patterns: an explanation using linear Rossby wave theory. J Clim 32:531–548

Vitart F, Molteni F (2010) Simulation of the MJO and its teleconnections in the ECMWF forecast system. Quart J Roy Meteor Soc 136:842–855. https://doi.org/10.1002/qj.623

Vitart F (2014) Evolution of ECMWF sub-seasonal forecast skill scores. Quart J Roy Meteor Soc 140:1889–1899

Vitart F, Balsamo G, Buizza R, Ferranti L, Keeley S, Magnusson L, Molteni F, Weisheimer A (2014) Sub-seasonal predictions. ECMWF Tech Memo 738:45

Vitart F et al (2017) The subseasonal to seasonal (S2S) prediction project database. Bull Amer Meteor Soc 98:163–173

Vitart F, Balmaseda MA (2024) Sources of MJO teleconnection errors in the ECMWF extended-range forecasts. Quart J Roy Meteor Soc 150:2028–2044

Xie Y-B, Chen S-J, Zhang I-L, Hung Y-L (1963) A preliminarily statistic and synoptic study about the basic currents over southeastern Asia and the initiation of typhoon (in Chinese). Acta Meteor Sin 33:206–217

Wang J, Kim H, Kim D, Henderson SA, Stan C, Maloney ED (2020) MJO teleconnections over the PNA region in climate models. Part I: Performance-and process-based skill metrics. J Clim 33:1051–1067

Wang J, DeFlorio MJ, Guan B, Castellano CM (2023) Seasonality of MJO impacts on precipitation extremes over the western United States. J Hydrometeorol 24:151–166

Wheeler MC, Hendon HH (2004) An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. Mon Wea Rev 132:1917–1932. https://doi.org/10.1175/1520-0493(2004)132%3c1917:AARMMI%3e2.0.CO;2

Xiang B et al (2022) S2S prediction in GFDL SPEAR: MJO diversity and teleconnections. Bull Amer Meteor Soc 103:E463–E484

Yadav P, Straus DM (2017) Circulation response to fast and slow MJO episodes. Mon Wea Rev 145:1577–1596

Yadav P, Straus DM, Swenson ET (2019) The euro-Atlantic circulation response to the Madden-Julian oscillation cycle of tropical heating: coupled GCM intervention experiments. Atmos Ocean 57:161–181

Yadav P, Garfinkel CI, Domeisen DI (2024) The role of the stratosphere in teleconnections arising from fast and slow MJO episodes. Geophys Res Lett 51:e2023GL104826

Yau AMW, Chang EKM (2020) Finding storm track activity metrics that are highly correlated with weather impacts. Part I: frameworks for evaluation and accumulated track activity. J Clim 33:10169–10186. https://doi.org/10.1175/JCLI-D-20-0393.1

Zheng C, Kar-Man Chang E, Kim H-M, Zhang M, Wang W (2018) Impacts of the Madden–Julian oscillation on storm-track activity, surface air temperature, and precipitation over North America. J Clim 31:6113–6134

Zheng C, Chang EKM, Kim H, Zhang M, Wang W (2019) Subseasonal to seasonal prediction of wintertime Northern Hemisphere extratropical cyclone activity by S2S and NMME models. J Geophys Res Atmos 124:12057–12077

Zheng C et al (2024) The impact of vertical model levels on the prediction of MJO teleconnections. Part I: The tropospheric pathways in the UFS global coupled model. Clim Dyn 62:9031–9056

Zhou S, L'Heureux M, Weaver S, Kumar A (2012) A composite study of the MJO influence on the surface air temperature and precipitation over the continental United States. Clim Dyn 38:1459–1471