ADVANCING EARTH AND SPACE SCIENCES

JGR Atmospheres

RESEARCH ARTICLE

10.1029/2026JD046695

Subseasonal Prediction of the Wintertime North Pacific Blocking in AI-Based Weather Models

Jaeyoung Hwang[1,2], Yi Deng[1], Simchan Yook[3], and Zhenyu You[4]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA, [2]Division of Climate Change, Hankuk University of Foreign Studies, Yongin-si, Republic of Korea, [3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [4]Department of Earth and Environmental Sciences, Schiller Institute for Integrated Science and Society, Boston College, Chestnut Hill, MA, USA

Key Points:

- Spatiotemporal evolution and moist dynamics of North Pacific (NP) blocking are well captured by ECMWF IFS, NeuralGCM, and Pangu-Weather
- The subseasonal prediction of NP blocking is skillful for up to 2 weeks, but it varies depending on the occurrence of blocking events
- The subseasonal prediction skill of NP blocking can be enhanced by leveraging MJO activity and ENSO-like SST patterns in AI-based models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Hwang and Y. Deng,
jayhwang@hufs.ac.kr;
yi.deng@eas.gatech.edu

Citation:

Hwang, J., Deng, Y., Yook, S., & You, Z. (2026). Subseasonal prediction of the wintertime North Pacific blocking in AI-based weather models. *Journal of Geophysical Research: Atmospheres*, *131*, e2026JD046695. https://doi.org/10.1029/2026JD046695

Received 15 FEB 2026
Accepted 16 JUL 2026

Author Contributions:

Conceptualization: Jaeyoung Hwang, Yi Deng
Data curation: Jaeyoung Hwang
Formal analysis: Jaeyoung Hwang
Investigation: Jaeyoung Hwang
Methodology: Jaeyoung Hwang, Yi Deng
Supervision: Yi Deng
Validation: Jaeyoung Hwang
Visualization: Jaeyoung Hwang
Writing – original draft: Jaeyoung Hwang, Yi Deng, Simchan Yook, Zhenyu You
Writing – review & editing: Jaeyoung Hwang, Yi Deng, Simchan Yook, Zhenyu You

Abstract

Recent studies have demonstrated that AI-based weather models exhibit reasonably encoded atmospheric dynamics and skill of short-term forecasting comparable to physics-based models. However, it remains unclear whether AI-based weather models can simulate system-oriented dynamics and provide skillful predictions across subseasonal timescales. Here, we assess the simulation of the North Pacific blocking and its subseasonal prediction skill in three weather models with a hierarchical implementation of AI: ECMWF IFS (full-physics), NeuralGCM (hybrid), and Pangu-Weather (data-driven). While all models reproduce blocking frequency well at 1–2 week lead times, biases diverge when the lead time increases to 3–4 weeks. ECMWF IFS overpredicts blocking frequency, related to overestimations of both short-lived blockings (5–6 days) and long-lived blockings (>9 days). Pangu-Weather tends to underpredict blocking frequency, largely due to an underestimation of blockings with 7–8 day durations. NeuralGCM has a dipolar structure of blocking frequency bias, exhibiting mixed characteristics between the ECMWF IFS and Pangu-Weather models. Despite these biases, all three models successfully reproduced blocking evolutions through 3–4 weeks of simulation. Remarkably, they all reasonably captured the moist processes of blocking dynamics despite different model hierarchies. The subseasonal prediction remains skillful for up to 2 weeks across all models. In ECMWF IFS, this skill can be enhanced by leveraging interannual and subseasonal variability originating from the tropical Pacific. This enhancement of skill is also qualitatively observed in NeuralGCM and Pangu-Weather, suggesting that AI-based weather models have the potential to deliver physically reliable predictions of weather extremes across subseasonal timescales.

Plain Language Summary

Recent studies have shown that AI-based weather models reasonably well reproduce atmospheric dynamics and achieve short-term forecasting skill comparable to traditional models. However, it remains unclear whether AI-based models can physically reproduce blocking—a key driver of midlatitude weather and climate extremes and provide skillful predictions of blocking at subseasonal lead times. Here, we assess the simulation of the wintertime North Pacific blocking and its subseasonal prediction skill in three weather models: ECMWF IFS (full-physics), NeuralGCM (hybrid), and Pangu-Weather (data-driven). While all models reproduce blocking frequency well at 1–2 week lead times, biases diverge when the lead time increases to 3–4 weeks. Despite these biases, they still capture important features of how blocking develops and evolves through 3–4 weeks of simulation. Our results also show that North Pacific blocking can be predicted with useful skill for up to about 2 weeks. The prediction skills improve when the initial conditions include tropical Pacific signals, such as El Niño-like conditions or the Madden-Julian Oscillation. These results suggest that AI-based weather models have the potential to deliver physically reliable predictions of weather extremes across subseasonal timescales.

1. Introduction

Atmospheric blocking, often characterized by a quasi-stationary high-pressure system, disrupts or even splits the westerly jet, thereby enhancing the meridional flows in the midlatitudes (Rex, 1950). This persistent circulation anomaly, lasting from several days to weeks, shapes surface weather across the mid- and high latitudes (Coumou & Rahmstorf, 2012; Kautz et al., 2021; Lupo, 2021; Woollings et al., 2018). More specifically, blocking can induce cold spells in its downstream region through an equatorward cold advection (e.g., Hwang et al., 2022; Park et al., 2014; Xie et al., 2017; Sillmann et al., 2011; Tuel & Martius, 2024; Xie et al., 2019). Precipitating systems (e.g., mid-latitude cyclones and atmospheric rivers) along a storm track are anchored or deflected by the blocking circulation, leading to a localized heavy precipitation (e.g., Benedict et al., 2019; Hong et al., 2011; Hu et al., 2023; Lenggenhager & Martius, 2019). Droughts or heatwaves (i.e., warm anomalies) can be intensified around the center of blocking due to clear skies, enhanced subsiding motion, and atmosphere-land interaction (Barriopedro et al., 2023; Domeisen et al., 2023; Rousi et al., 2023; Schaller et al., 2018; White et al., 2023).

Given its importance, reliable prediction of blocking is crucial in operational models (Wang et al., 2026). Earlier studies reported that, while operational models show reasonable model prediction skill when initialized from blocked states, they struggle with predicting the blocking onset, which in turn decreases the overall prediction skill (Kimoto et al., 1992; Tibaldi et al., 1994; Tibaldi & Molteni, 1990; Tracton et al., 1989). Nutter et al. (1998) confirmed that the models tend not to capture the blocking onset as lead time increases. The authors also suggested that preconditioning of planetary-scale waves might be important for long lead-time blocking prediction compared with synoptic-scale activity in the initial state. Similarly, Pelly and Hoskins (2003) reported that since the onset of blocking events is closely related to synoptic-scale activity, the intrinsically limited prediction skill of synoptic-scale variability could lead to a deterministic predictability limit for blocking. This suggests that probabilistic ensemble prediction could improve the blocking prediction (e.g., Watson & Colucci, 2002). Matsueda (2009) reported that skillful prediction of atmospheric blocking extends overall up to ∼9 days of lead time across operational models. Nevertheless, its predictability exhibits regional variations (e.g., Matsueda, 2009; Xavier et al., 2026).

More recently, limited blocking prediction has been reframed in the context of longer timescale prediction, with a particular focus on the winter season. For subseasonal-to-seasonal timescales, where flow regimes and large-scale patterns are emphasized as key forecast targets (Vitart et al., 2017; White et al., 2017), Quinting and Vitart (2019) demonstrated that successful representation of synoptic-scale Rossby wave packets can lead to enhanced blocking prediction skill in operational models. Wandel et al. (2024) showed that misrepresentation of warm conveyor belts (WCBs) is linked to limited prediction skill for blocking, whereas well-predicted blocking cases with up to 2–3 weeks of lead time are accompanied by accurate WCB predictions. Huang et al. (2024) further found that blocking cases driven by strong latent heating from WCBs are overall less predictable than blocking cases accompanied by weak latent heating. At seasonal timescales, Davini et al. (2021) showed that the predicted climatological blocking statistics, especially over the North Pacific and Greenland, improve across successive generations of operational seasonal forecast systems. However, lower predictive skill and signal-to-noise ratio for blocking are still exhibited in the most recent forecast generation. Athanasiadis et al. (2014) and Park et al. (2024) reported that blocking representation and its seasonal predictability can be improved by leveraging subseasonal and/or interannual variability.

The aforementioned studies suggest that the limited prediction skill of blocking may, in part, reflect the intrinsic prediction limit of the atmosphere due to its chaotic nature (Lorenz, 1969), as well as the complex dynamics of blocking (e.g., Charney & DeVore, 1979; Hauser et al., 2023; Nakamura et al., 1997). In addition, the systematic biases in the representation of blocking contribute to the limited blocking prediction, possibly due to coarse model resolution (e.g., Berckmans et al., 2013; Davini et al., 2017), imperfect physical parameterization (e.g., Steinfeld et al., 2020), mean-state bias (e.g., Kleiner et al., 2021), and misrepresentation of other atmospheric processes affecting blocking occurrence (e.g., Dolores-Tesillos et al., 2025; Jung et al., 2012). This highlights the need for a comprehensive assessment of both the physical representation of atmospheric blocking and its prediction skill in weather models.

With advances in artificial intelligence (AI) techniques and computing resources, AI-based weather prediction models have recently emerged (e.g., Bi et al., 2023; Lam et al., 2023; Pathak et al., 2022). In the context of short-to medium-range forecasting (up to ∼2 weeks), these models have reached performance levels comparable to traditional physics-based models (Ben Bouallègue et al., 2024; Bi et al., 2023; Brenowitz et al., 2025; K. Chen et al., 2023; L. Chen et al., 2023; Kochkov et al., 2024; Lam et al., 2023; Lang et al., 2024; Pathak et al., 2022; Price et al., 2025; Rasp et al., 2024). These models have also shown potential for extended-range prediction at subseasonal time scales (e.g., Chen et al., 2024). In addition, AI-based models are able to encode the fundamental dynamics of the atmosphere, suggesting that they can serve as a simple testbed for weather experiments (Hakim & Masanam, 2024). Nevertheless, AI-based models still have room for improvement (e.g., Bonavita, 2024). For instance, the models qualitatively simulate the large-scale structure and evolution of cyclones, but they underestimate the peak intensity, surface impacts, and smaller-scale processes of the systems (Charlton-Perez et al., 2024; Dacre et al., 2025; DeMaria et al., 2025; Liu et al., 2024; Xu et al., 2024). The models also tend to

10.1029/2026JD046695

produce increasingly smooth (or blurred) forecasts at longer lead times (Keisler, 2022; Lam et al., 2023; Rasp et al., 2024).

Building on these studies, the present study investigates the predictability of atmospheric blocking in AI-based models, with a particular focus on the North Pacific region during winter. Beyond the short-to medium-range forecasts where AI-based models have generally shown good performance, we aimed to evaluate whether these models can be utilized in a physically consistent manner for subseasonal blocking prediction. By analyzing two AI-based models with different architectures, namely Pangu-Weather (fully data-driven) and NeuralGCM (hybrid), and comparing them with the Integrated Forecasting System (IFS) at the European Center for Medium-Range Weather Forecasts (ECMWF), we also assess the representation of moist processes in blocking onset across a hierarchy of model architectures. While blocking can occur through dry dynamical processes, moist processes associated with upstream latent heat release often modulate blocking amplification and maintenance, particularly over the North Pacific (Liu & Wang, 2025; Pfahl et al., 2015; Steinfeld & Pfahl, 2019; Yamamoto et al., 2021). This potentially provides an architecture-sensitive diagnostic for evaluating how full-physics, hybrid, and fully data-driven models represent blocking-relevant diabatic processes. Thus, our main questions are as follows:

1. How well do AI-based models reproduce North Pacific blocking statistics and moist dynamics?
2. Can AI-based models provide physically consistent subseasonal predictions of North Pacific blocking?

This study is organized as follows. In Section 2, the data description and methods, including the blocking tracking, diabatic heating metrics, and prediction skill, are detailed. North Pacific blocking statistics and blocking dynamics are assessed in Section 3.1. In Section 3.2, subseasonal prediction skills of North Pacific blocking in three operational models are assessed. Finally, a summary and discussion are provided in Section 4.

2. Data and Methods

2.1. Data

To evaluate the subseasonal predictability of North Pacific blocking, this study employs three different operational models. The hindcasts of ECMWF IFS Cycle 49r1, a physics model coupled with ocean and sea-ice, are integrated with different resolutions of O1280 (∼9 km) for the medium forecast range and O320 (∼36 km) for the subseasonal forecast range. For subseasonal prediction, this model configuration provides 46-day forecasts initialized every 2 days starting from November 2004. Instantaneous atmospheric variables at 00 UTC are available at 24-hr intervals with a 1.5° by 1.5° horizontal resolution at 13 vertical levels. The hindcast data sets were analyzed for 20 winter seasons (DJF) from December 2004 to February 2024, totaling 920 initializations.

Pangu-Weather is a fully data-driven model with a vision transformer architecture developed by Huawei Cloud. The model is trained with 5th generation of ECMWF reanalysis (ERA5; Hersbach et al., 2020) for the period of 1979–2017 (Bi et al., 2023). It provides the forecast output using only atmospheric variables such as horizontal winds, geopotential height, specific humidity, temperature, and surface variables, with 0.25° by 0.25° horizontal resolution on 13 vertical levels. The forecast system employs a hierarchical temporal aggregation strategy using 4 versions of the model trained for 1-hr, 3-hr, 6-hr, and 24-hr forecasts. Here, the 24-hr model configuration is only employed to compare it with the ECMWF IFS.

NeuralGCM is a hybrid model with a differentiable dynamical core and physics parameterization in the form of a learned neural network. Similar to Pangu-Weather, we employed the pre-trained model based on ERA5 data for the period of 1979–2017. NeuralGCM produces forecasts with 37 vertical levels at different horizontal resolutions (i.e., 0.7°, 1.4°, and 2.8°), by prescribing persistent raw sea surface temperature (SST) and sea-ice from ERA5 at the initialization date as lower-boundary conditions. In this study, 1.4° by 1.4° resolution is employed for stable model integration at subseasonal forecasts.

As in ECMWF IFS, the forecasts from Pangu-weather and NeuralGCM models are also produced every 2 days from December 2004 to February 2024, totaling 920 initializations. The models are integrated for up to 46 days, producing 24-hourly instantaneous outputs. By adopting a single-member ensemble across the models, the deterministic prediction skill of North Pacific blocking is examined. For reference, we also utilized ERA5 to construct the same data set as in the model outputs. All atmospheric variables were then interpolated to a 2.5° by

2.5° horizontal resolution prior to analysis. To understand the potential predictor sources for North Pacific blocking, SST (OISSTv2; Huang et al., 2021) and outgoing longwave radiation (OLR) are analyzed.

It should be noted that the models differ not only in their architecture but also in their boundary conditions and horizontal resolution, which may affect their representation of blocking. Thus, although we examine NP blocking predictability in the models by considering a hierarchy of model architecture, the differences in the results shown later cannot be attributed solely to model architecture. Additionally, the analysis should be ideally conducted for the test period (2018–2024) for a fair comparison. However, due to the limited sample size of North Pacific blocking during this period, the analysis is extended to 20 years of hindcasts including the training period, to secure a large sample size of blocking cases. To ensure the validity of this approach, we also performed a sensitivity analysis using the test period. As will be discussed later, the overall results are qualitatively similar when the test period is considered (see Figures S1–S6 in Supporting Information S1).

2.2. Blocking Index

To identify blocking events, this study adopts the 2D-hybrid blocking index (e.g., Barriopedro et al., 2010; Dunn-Sigouin et al., 2013), which integrates the two distinct approaches: the anomaly-based (Dole & Gordon, 1983) and absolute-field-reversal (Tibaldi & Molteni, 1990) methods. A large-scale ($\geq 2.0 \times 10^6$ km^2) positive 500-hPa geopotential height (Z500) anomaly is considered as a blocking candidate if it exceeds the monthly-based amplitude threshold. For each calendar month, the threshold is defined as the 1.3 standard deviations of the Z500 anomaly distribution, calculated using all grid points over mid- and high latitudes (30°N–90°N) centered on a 3-month window. The identified Z500 anomaly, satisfying quasi-stationary (overlap ratio ≥ 0.5 within consecutive days) and persistence (≥ 5 days duration) criteria along with meridional Z500 gradient reversal, is then defined as a blocking event (see details in Hwang et al., 2022). Specifically, the meridional Z500 gradient reversal is diagnosed when Z500 locally increases poleward relative to its value 15° to the south within grid points surrounding the maximum Z500 anomaly. The anomaly is defined as a departure from lead-time-dependent model climatology. The model climatology is obtained by averaging the 20 years of hindcasts from the same calendar-day initializations, minimizing the model drift. A blocking event is classified as a NP blocking when its center (i.e., maximum Z500 anomaly) is located within the North Pacific sector (30°N–70°N, 120°E–250°E) on the onset day. A total of 56 events of NP blocking were identified during the winters of 2004–2024 (16 events for the test period).

2.3. Moist Process in Blocking Onset

To examine the role of the moist process in blocking formation across the models, latent heating was calculated as follows (Emanuel et al., 1987):

$$\dot{\theta}_{\mathrm{LH}} = \omega\left(\frac{\partial\theta}{\partial p} - \frac{\gamma_m}{\gamma_d}\frac{\theta}{\theta_e}\frac{\partial\theta_e}{\partial p}\right) \tag{1}$$

where ω is the vertical velocity (Pa s^{-1}); θ_e is the equivalent potential temperature (K), and γ_d and γ_m denote dry and moist adiabatic lapse rates (K m^{-1}), respectively. Since the vertical velocity is not available in NeuralGCM and Pangu-Weather, the vertical velocity of all data sets, including ERA5, is systematically diagnosed by vertically integrating horizontal divergence (i.e., mass continuity equation). Here, the lower boundary condition is set to zero. Although the diagnosed vertical velocity is unreliable over high-topography regions, this diagnostic is still qualitatively useful over the oceans, where North Pacific blocking frequently occurs (Bonavita, 2024). Then, the contribution of latent heating is translated into geopotential height fields through the quasi-geostrophic geopotential tendency equation (Hwang & Son, 2025; Text S1 in Supporting Information S1).

2.4. Anomaly Correlation

To evaluate the subseasonal prediction skill of North Pacific blocking, two complementary anomaly correlation coefficient (ACC) measures are utilized based on weekly-mean blocking frequency (e.g., DeFlorio et al., 2018; Zheng et al., 2019):

$$\mathrm{ACC}_t(\lambda,\tau) = \frac{\sum_{i=1}^{N_i}(m_i(\lambda,\tau)-c(\lambda,\tau))(o_i(\lambda,\tau)-c(\lambda,\tau))}{\sqrt{\sum_{i=1}^{N_i}(m_i(\lambda,\tau)-c(\lambda,\tau))^2}\sqrt{\sum_{i=1}^{N_i}(o_i(\lambda,\tau)-c(\lambda,\tau))^2}} \tag{2}$$

$$\mathrm{ACC}_s(i,\tau) = \frac{\sum_{\lambda=1}^{N_R}(m_\lambda(i,\tau)-c_\lambda(i,\tau))(o_\lambda(i,\tau)-c_\lambda(i,\tau))}{\sqrt{\sum_{\lambda=1}^{N_R}(m_\lambda(i,\tau)-c_\lambda(i,\tau))^2}\sqrt{\sum_{\lambda=1}^{N_R}(o_\lambda(i,\tau)-c_\lambda(i,\tau))^2}} \tag{3}$$

where, m and o denote the weekly-mean blocking frequency at a given grid point (λ) in the models and ERA5, respectively. i represents initialization index; τ denotes the weekly lead time; N_R indicates the total number of grid points for the target region. c represents the observed climatology of weekly-mean blocking frequency for 2004–2024 at a given grid point.

The ACC_t (i.e., temporal ACC) quantifies the mean prediction skill of weekly-mean blocking frequency at a given lead time and grid point (Equation 2). Thus, it highlights the geographical distribution of the blocking prediction skill across the initializations. The latter shows the spatial ACC (ACC_s) of weekly-mean blocking frequency for a given initialization and lead time (Equation 3). This metric evaluates the spatial pattern of weekly-mean blocking frequency between the forecasts and observations, characterizing the initialization-dependent variability of regional pattern skill.

3. Results

3.1. Statistics and Dynamics of the Predicted North Pacific Blocking

Figure 1 shows the representation of the North Pacific (NP) blocking frequency in three operational models at lead times of 1–2 and 3–4 weeks during boreal winter. Note that blocking events are counted using all initializations. As a result, blocking events in the observations (i.e., ERA5) are counted multiple times (~6 times; see Methods), but this approach still allows at least a qualitative evaluation of blocking statistics in the models. The climatological NP blocking frequency exhibits a northwest–southeast-stretched spatial distribution (contours). This suggests a westward propagating nature of NP blocking (Branstator, 1987; Kim & Kim, 2025; Kushnir, 1987). All models reproduce this pattern well at lead times of 1–2 weeks, with minor biases in blocking frequency (Figures 1a, 1c, and 1e), implying reliable predictions of NP blocking up to a lead time of 2 weeks. However, the biases start to diverge as the lead time increases to 3–4 weeks (Figures 1b, 1d, and 1f). In the ECMWF IFS, the positive blocking-frequency bias over the eastern North Pacific in 1–2 weeks becomes more pronounced in 3–4 weeks, suggesting an overall overestimation of the NP blocking frequency (Figures 1a and 1b). In NeuralGCM, the underestimation of blocking frequency over the western North Pacific in 1–2 weeks strengthens in 3–4 weeks, while a notable positive frequency bias also appears over the eastern North Pacific (Figures 1c and 1d), showing a dipole-like bias pattern. In Pangu-Weather, the negative frequency bias over the western North Pacific shown in 1–2 weeks becomes more pronounced in 3–4 weeks, leading to an overall underestimation of NP blocking in the model. It is also accompanied by a minor positive bias of the blocking frequency over the eastern North Pacific (Figures 1e and 1f).

Figure 2 illustrates the number of NP blocking events occurring at lead times of 1–2 and 3–4 weeks in the three models with respect to blocking duration. At lead times of 1–2 weeks (Figures 2a–2c), the number of NP blocking events, counted across all initializations, decreases exponentially with increasing duration in all models, showing a largely similar blocking distribution. However, all models overestimate the frequency of blockings with durations of 5–6 and 9–10 days compared to ERA5, whereas they underestimate the blocking with durations of 7–8 days. These biases largely offset each other, resulting in weak blocking-frequency biases at 1–2 weeks. At lead times of 3–4 weeks, ECMWF IFS continues to overestimate the blocking events with the duration of 5–6 days and 9–10 days with an underestimation of 7–8-day blocking events, as in 1–2 weeks. However, it additionally overestimates long-lived blocking events (≥11 days of duration), which explains the overall overestimation of blocking frequency (Figure 2d). NeuralGCM exhibits a duration distribution similar to that of ECMWF IFS, explaining the overestimation of NP blocking over the eastern North Pacific (Figure 2e). However, it also simulates slightly fewer blocking events (≥7 days of duration) compared to ECMWF IFS, which might be associated with the negative model bias over the western North Pacific. In Pangu-Weather, although the

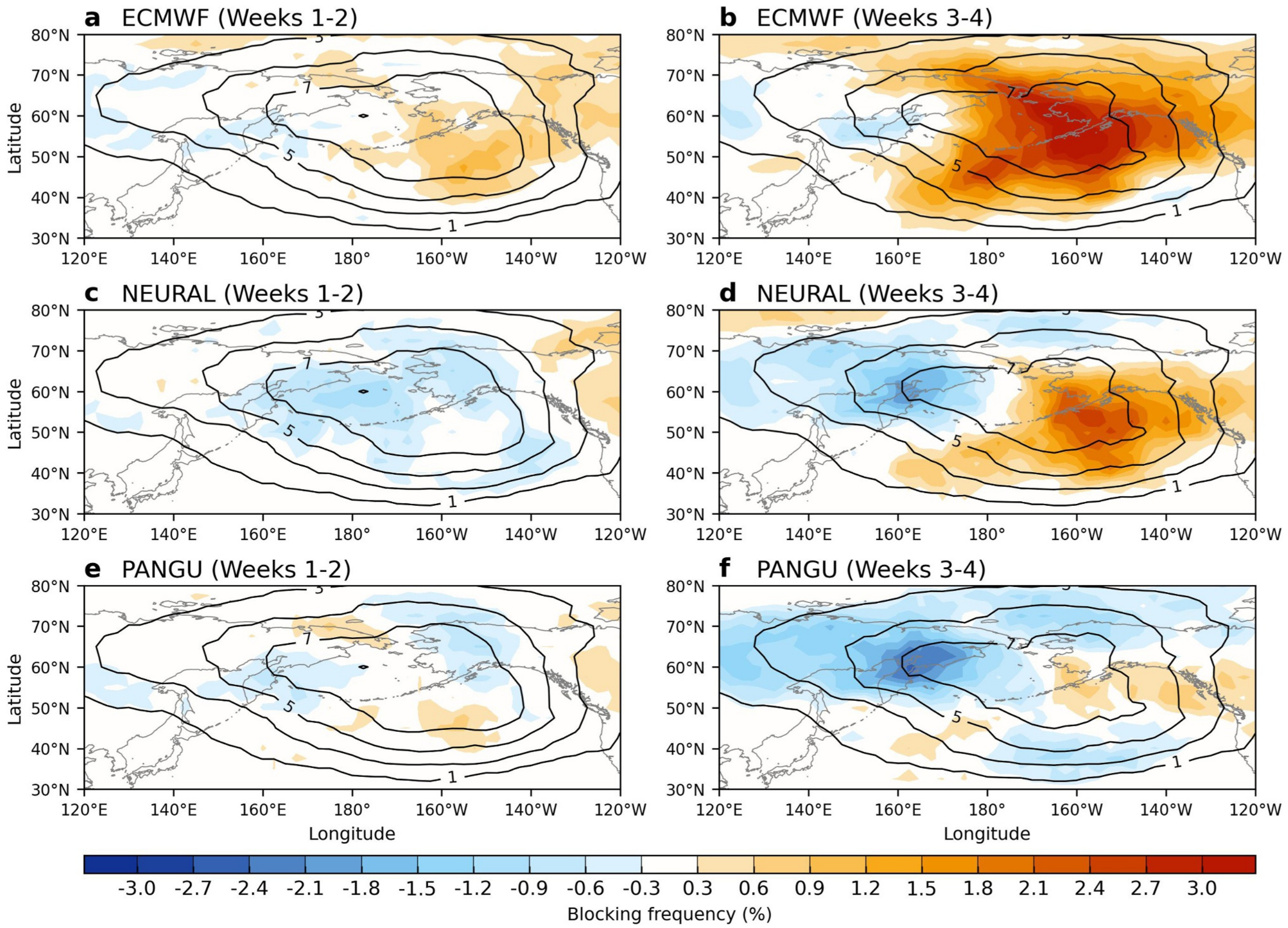

Figure 1. Biases of North Pacific blocking frequency (shadings; %) in (a, b) ECMWF IFS, (c, d) NeuralGCM, and (e, f) Pangu-Weather during (left) 1–2 weeks and (right) 3–4 weeks of simulation. As a reference, climatology of North Pacific blocking frequency in ERA5 is superimposed (contour; 2% interval starting from 1%).

9–10-day blocking events are overestimated, the reduced number of the 7–8-day blocking events largely explains the overall underestimation of blocking frequency (Figure 2f).

The blocking statistics were also assessed for the test period (2018–2024). All models capture the spatial distribution of the NP blocking frequency at lead times of 1–2 weeks (Figures S1a, S1c, S1e in Supporting Information S1), although they exhibit a reduced blocking frequency. At lead times of 3–4 weeks, ECMWF IFS tends to overestimate the NP blocking frequency, consistent with the results for the full period (2004–2024; Figure S1b in Supporting Information S1). Similarly, Pangu-Weather shows a systematic underestimation of NP blocking frequency at subseasonal lead times (Figure S1f in Supporting Information S1). In contrast, NeuralGCM exhibits a slightly different blocking bias (Figure S1d in Supporting Information S1). Unlike the dipole pattern found over the full period, it shows a positive bias in NP blocking frequency during the test period (Figure S1d in Supporting Information S1).

Figure 3 shows the spatiotemporal evolution of NP blocking in ERA5 and the three operational models at lead times of 1–2 weeks. All variables are composited relative to the blocking center on the onset day. In ERA5, NP blocking frequently occurs in the climatological jet exit regions (purple contour; Figure 3a) (e.g., Jiang et al., 2013). The blocking-related Z500 anomaly exhibits a westward-tilting structure, suggesting cyclonic wave breaking (e.g., Masato et al., 2013). During the mature phase (on a day of maximum blocking anomaly), the blocking anomaly overall amplifies with a westward expansion (Figure 3b). The anomaly substantially weakens during the decay phase and propagates westward (Figure 3c). All models qualitatively reproduce the overall life cycle of NP blocking, which occurs at lead times of 1–2 weeks. In the models, NP blocking occurs downstream of the jet and exhibits a westward-tilting structure, confirming a close relationship between cyclonic wave breaking

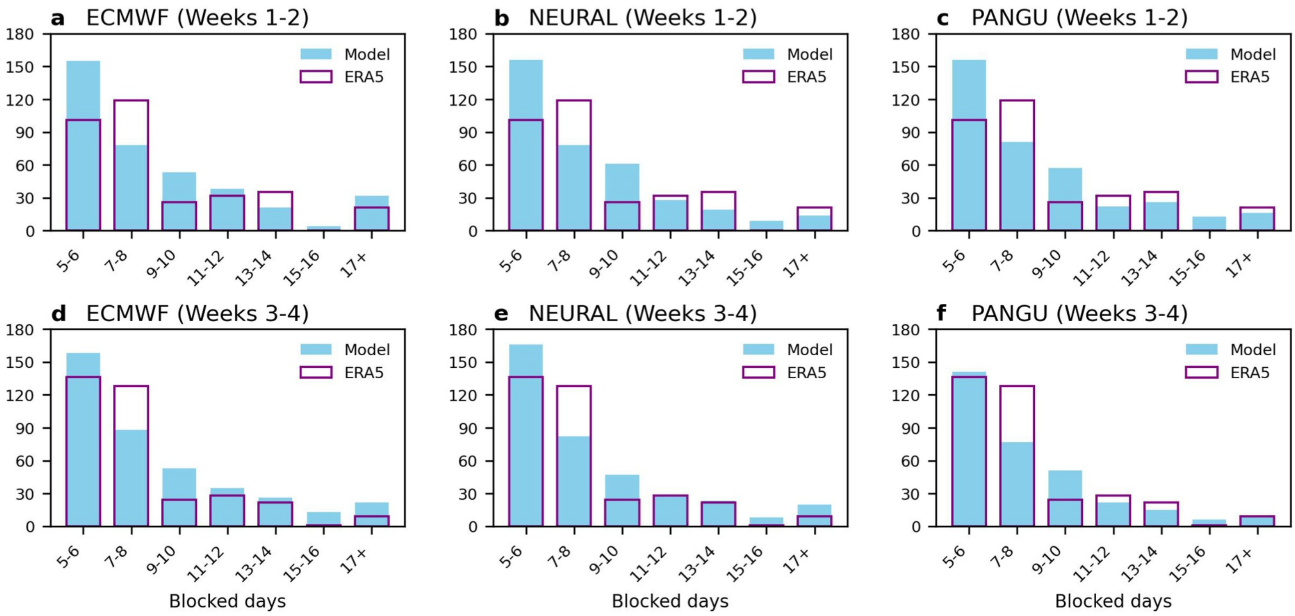

Figure 2. The number of North Pacific blocking events (filled bar) as a function of duration in (a, d) ECMWF IFS, (b, e) NeuralGCM, and (c, f) Pangu-Weather. Here, blocking events are counted when they occur during (top) 1–2 weeks and (bottom) 3–4 weeks of simulation considering all initialization for the period of 2004–2024. As a reference, the blocking events in ERA5 (empty bar) are also counted based on the same data set as in the models.

and NP blocking onset (Figures 3d, 3g, and 3j). Although weak, the blocking signal becomes more pronounced, and the upstream trough deepens during the mature phase, which is consistent with ERA5 (Figures 3e, 3h, and 3k). The blocking anomaly weakens and exhibits westward propagation during the decay phase (Figures 3f, 3i, and 3l).

At lead times of 3–4 weeks, a qualitatively similar NP blocking evolution is found in all models (Figure 4). This suggests that AI-based models can produce physically consistent NP blocking in subseasonal forecast ranges despite the biases of NP blocking in the models. A systematically weaker blocking anomaly is also found across the models during the mature (Figures 4e, 4h, and 4k) and decay phases (Figures 4f, 4i, and 4l). Notably, Pangu-Weather shows a weakened climatological jet at 3–4 weeks, which may reflect the blurred model outputs at longer lead times in a fully data-driven architecture (Figures 4j–4l). A sensitivity analysis further shows that all models reproduce the characteristics of the NP blocking evolution during the test period, which is consistent with the full period (Figures S2 and S3 in Supporting Information S1).

The center-composited blocking evolution suggests that NP blocking formation is closely related to cyclonic wave breaking, implying an important role for moist processes in NP blocking formation (e.g., Orlanski, 2003; Riviere & Orlanski, 2007; Steinfeld & Pfahl, 2019). In this regard, the direct and indirect effects of latent heating on NP blocking formation are further investigated in Figure 5. The direct effect of latent heating, related to the low potential vorticity (PV) transport from the lower to the upper troposphere in the upstream warm conveyor belts (WCBs; Steinfeld & Pfahl, 2019), is translated into geopotential height tendency using the quasi-geostrophic geopotential tendency equation (Equation; see also Hwang & Son, 2025). To focus on the upper-level processes, the attributed height tendency is vertically averaged from 500 to 300 hPa. The indirect effect of latent heating is estimated from horizontal PV advection by divergent flow in the upper troposphere (300 hPa). Here, the divergent flow in the blocking upstream region, driven by ascent associated with latent heating, is obtained via Helmholtz decomposition of the horizontal winds.

In ERA5, a pronounced positive geopotential height tendency is observed upstream of the blocking region (purple contour; Figure 5a). The tendency is elongated northeastward along with the blocking circulation, suggesting a strong upper-level negative PV tendency in the upstream region. A weaker negative geopotential height tendency is also evident over the east of the blocking center, likely related to less condensation associated with anomalous subsidence motion (e.g., Nabizadeh et al., 2021). The divergent flow (vectors) appears centered on the maximum of the positive height tendency. The resulting negative PV advection by divergent flow (shading) is observed in

10.1029/2026JD046695

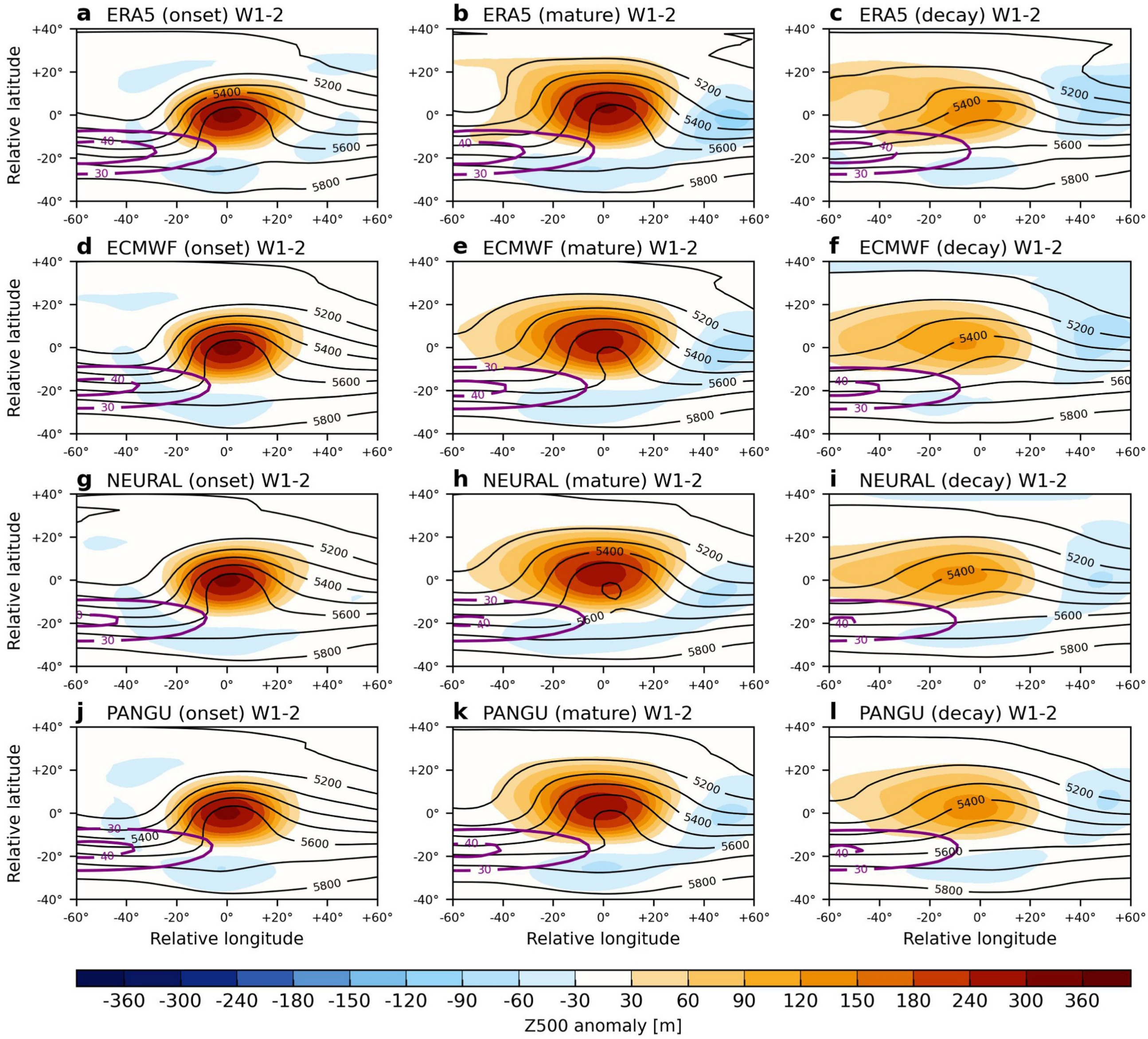

Figure 3. Spatiotemporal evolution of the 500 hPa-geopotential height (black contour; 100 m interval) and its anomaly (shading) associated with the North Pacific blocking during (left) onset, (middle) mature, and (right) decay in (a–c) ERA5, (d–f) ECMWF IFS, (g–i) NeuralGCM, and (j–l) Pangu-Weather. Here, the blocking is considered when it occurs during the periods corresponding to 1–2 weeks of simulation. The composite was conducted based on the blocking center on the onset day. Purple contours indicate the 300 hPa zonal wind and its climatology above 30 m s^{-1} (10 m s^{-1} interval), respectively.

the upstream of blocking, consistent with a westward expansion of the system. All three models qualitatively capture both the direct and indirect latent-heating processes during lead times of 1–2 weeks (Figures 5c, 5e, and 5g). The positive height tendency upstream of the blocking and its peak magnitude are well reproduced. The indirect process, that is, negative PV advection by divergent flow, is also simulated in a qualitatively similar manner. Notably, smoother PV composites are observed across the models compared to ERA5. This is partly because the model composites, averaging over a larger amount of blocking samples across all initializations, reduce small-scale noise. At lead times of 3–4 weeks, similar moist processes were still evident in all models (Figures 5d, 5f, and 5h). Although the peak positive tendency is weaker in ECMWF IFS and Pangu-Weather than in ERA5,the overall results indicate that key moist processes associated with NP blocking are represented in subseasonal forecasts regardless of model architecture.

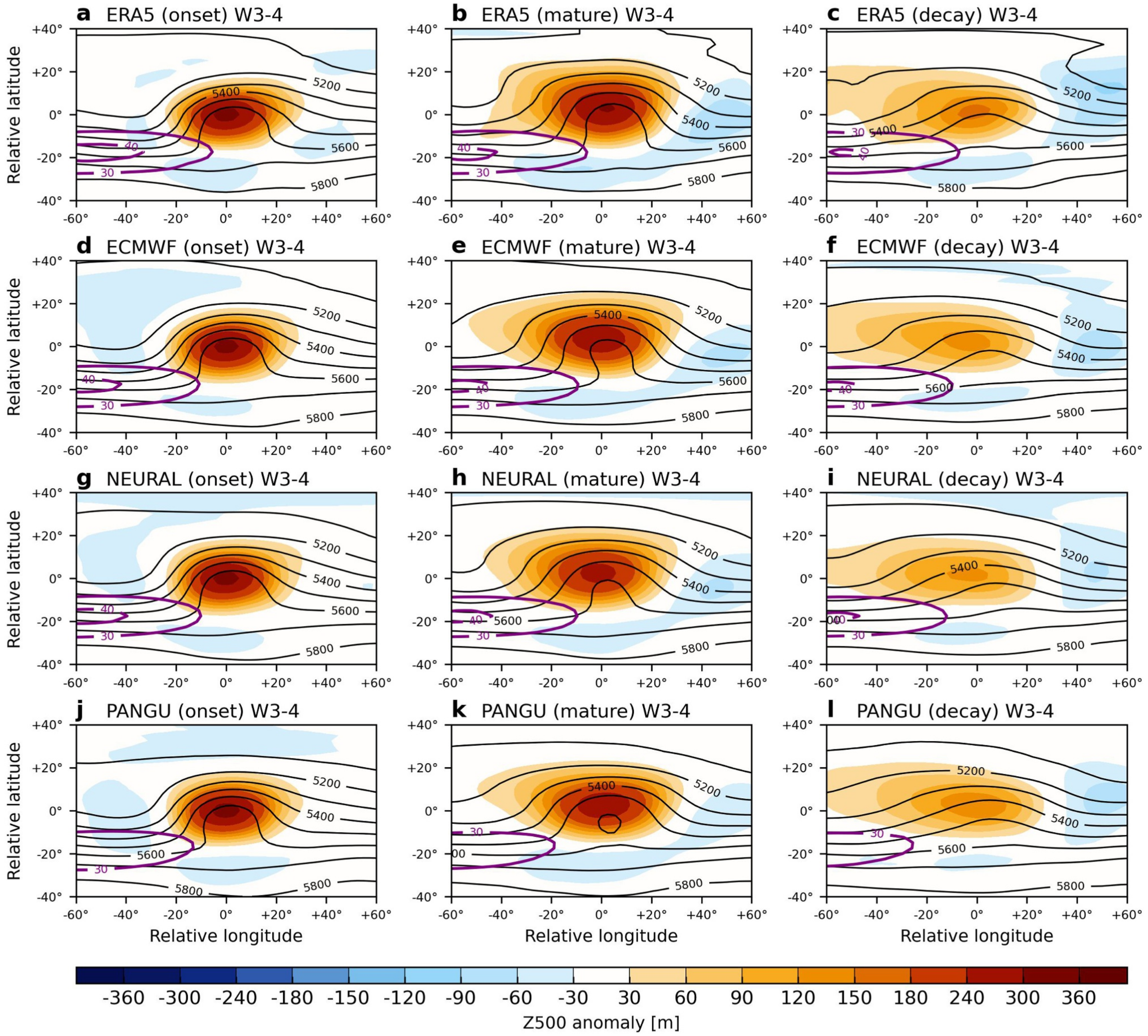

Figure 4. Same as Figure 3 but for 3–4 weeks.

To further quantify the model representation of the direct effect of latent heating, Figure 6 shows box plots of the attributed height tendency (boxed region in Figure 5) for lead times of 1–2 and 3–4 weeks. At 1–2 weeks, the distributions of the height tendency in the three models exhibited mean (dashed) and median (solid) values similar to those in ERA5, showing a pronounced positive skewness (Figure 6a). The positive tails of the distributions (upper whiskers) are also broadly comparable to ERA5 across the models. At 3–4 weeks, qualitatively similar distributions persist (Figure 6b), but the overall spread becomes narrower (as indicated by smaller boxes) in all models. In Pangu-Weather, not only a narrower distribution but also a weakened contribution from moist processes is observed at 3–4 weeks (Figure S4 in Supporting Information S1), partly reflecting the smoother forecasts of data-driven models over time. The models also qualitatively simulate moist processes during the test period, showing consistency with the full-period analysis (Figure S5 and S6 in Supporting Information S1)

3.2. Subseasonal Prediction Skill of the North Pacific Blocking Frequency

The above results show that, although operational models exhibit biases in NP blocking frequency, all models, including the AI-based models, simulate physically consistent dynamics and evolution of NP blocking at

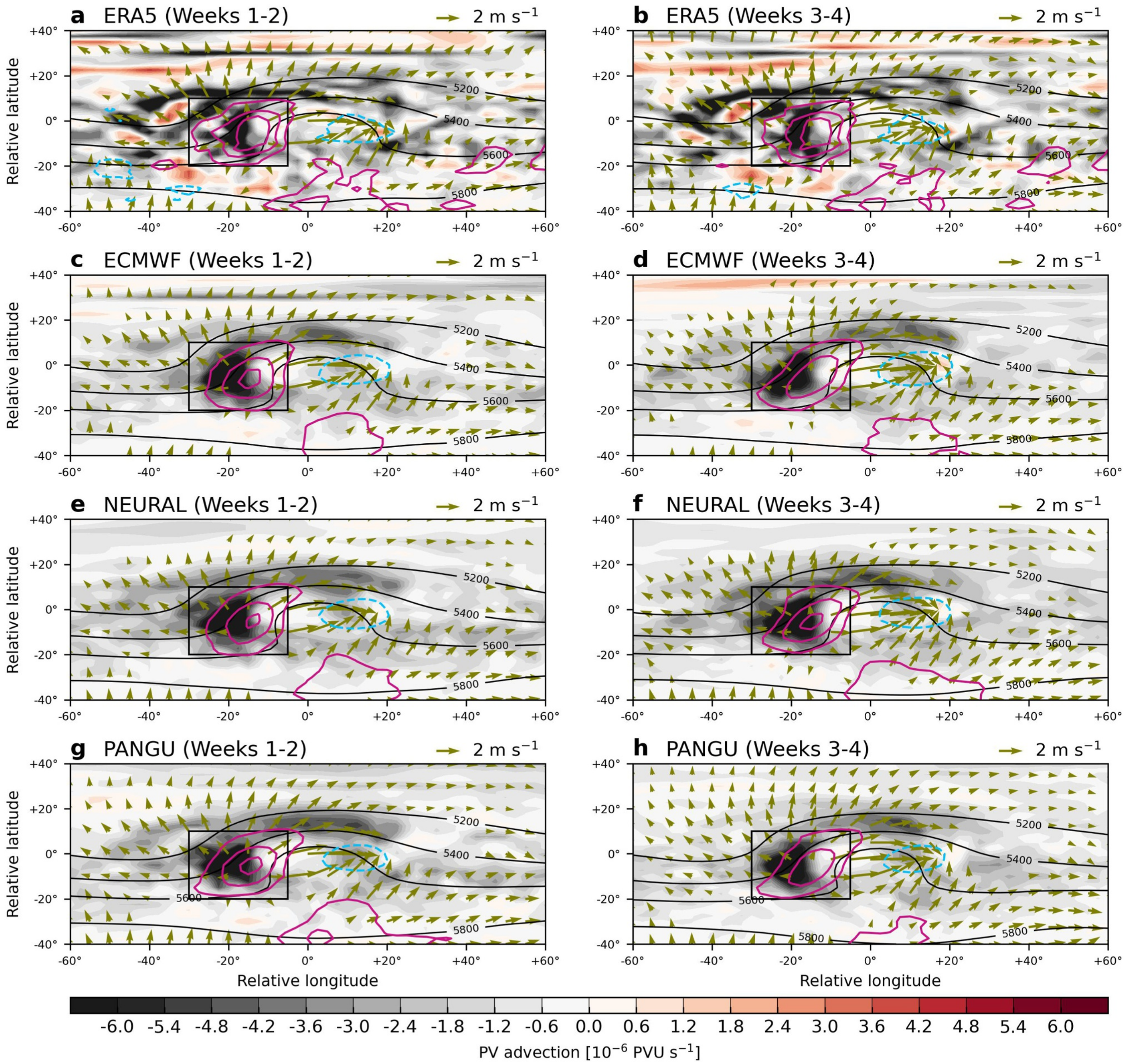

Figure 5. Composites of North Pacific blocking centered on the onset day in (a, b) ERA5, (c, d) ECMWF IFS, (e, f) NeuralGCM, and (g, h) Pangu-Weather. Here, the composite is conducted using blockings when they occur during the periods corresponding to (left) 1–2 weeks and (right) 3–4 weeks of simulation. Black contours indicate geopotential height at 500 hPa (200 m interval). Purple and blue contours show a vertically averaged geopotential height tendency attributed to the moist process (m 6 hr^{-1}) from 500 to 300 hPa. Shading indicates the horizontal advection of potential vorticity (pvu s^{-1}) by divergent winds at 300 hPa. Green vectors denote 300-hPa divergent wind vectors (their magnitude less than 0.7 m s^{-1} is omitted).

subseasonal lead times. In this section, we assess the predictability of NP blocking in the three operational models. For the subseasonal prediction, we evaluate the prediction skill for weekly-mean NP blocking frequency (i.e., the number of blocked days per week).

Figure 7 shows the prediction skill of weekly-mean NP blocking frequency, evaluated using the temporal anomaly correlation coefficient (ACC_t). At a lead time of 1 week, all models show relatively high skill over the North Pacific, reaching ~0.8 in regions where NP blocking occurs frequently (Figures 7a–7c). ACC decreases across the North Pacific at lead times of 2 weeks with a prediction skill up to ~0.5, which suggests a prediction limit of the NP blocking prediction at ~2 weeks (Figures 7d–7f). A simple sensitivity analysis using the mean squared skill score also confirms the appreciable prediction skill of NP blocking up to approximately 2 weeks

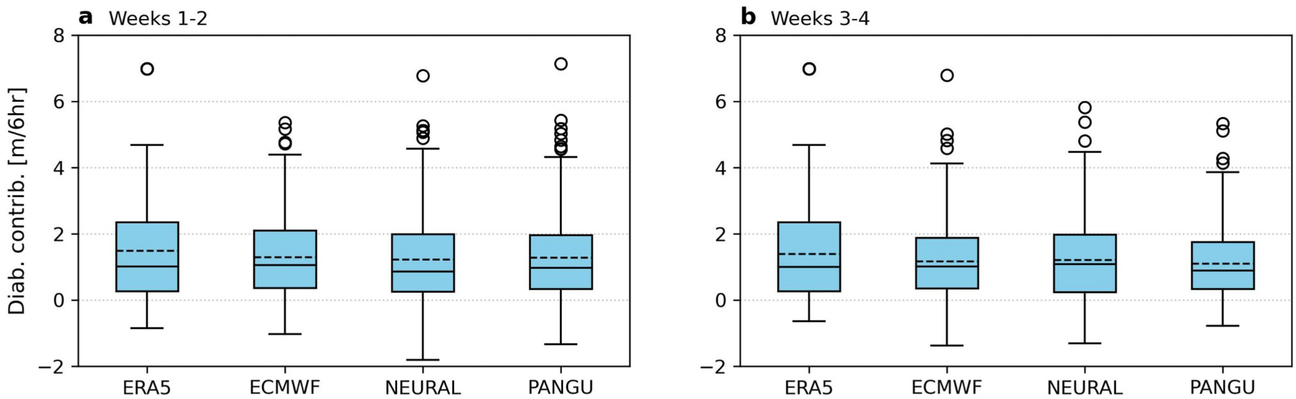

Figure 6. Boxed plot of a vertically averaged geopotential height tendency attributed to a moist process (m 6 hr^{-1}) from 500 to 300 hPa (boxed region in Figure 5) in (a) weeks 1–2 and (b) weeks 3–4 of simulation.

(Figure S8 in Supporting Information S1). At week 3, prediction skill keeps decreasing across the North Pacific to ~0.3 (Figures 7g–7i). Finally, ACC values approach zero over most regions at lead times of week 4 in all models (Figures 7j–7l). These results indicate that the overall NP blocking predictability is qualitatively similar across the model hierarchy. Similar results are also observed using the test period (Figure S6 in Supporting Information S1).

What drives the predictability of NP blocking at subseasonal timescales? To address this issue, we further evaluated NP blocking prediction skill in ECMWF IFS, by distinguishing between cases when NP blocking is present at lead times of 3–4 weeks (blocked cases) and those when it is absent (non-blocked cases) based on ERA5. Figure 8a shows spatial ACC (ACC_s) of weekly-mean NP blocking frequency in ECMWF IFS. Here, prediction skill is quantified by averaging individual initialization's ACC for weekly-mean NP blocking frequency across all grid points in the blocking preferred region (boxed region in Figure 7a). As indicated in Figure 7, prediction skill decreases with an increasing lead time (green line). For blocked cases (sky blue), while the prediction skill is comparable to all cases during the lead time of 1 week, it decreases sharply with increasing lead time. In contrast, non-blocked cases (pink) show higher prediction skill compared to blocked cases across lead times, showing ~0.6 at lead times of 4 weeks. Given that blocking is a rare event, the higher skill observed in non-blocked cases is not surprising, as the model captures well the normal states (i.e., blocking absent states). Figure 8b further shows that the prediction skill of the top 30% blocked cases (navy) remains appreciable (~0.6) even at lead times of 3–4 weeks (Figure 8b). For the non-blocked cases, the skill of the well-predicted (top 30%) subset shows an ACC of 1, indicating that the model successfully captures the normal states at lead times of 3–4 weeks (Figure 8c).

To identify potential sources of enhanced predictability for blocked cases (i.e., cases in which blocking occurs at lead times of 3–4 weeks), the initial conditions are compared between all blocked cases and well-predicted blocked cases in Figure 9. Figure 9a shows the composite of 300-hPa zonal wind at initialization in all blocked cases. A poleward-shifted jet is observed in the North Pacific region in comparison to the climatological jet (contours). This suggests the presence of reduced meridional wind shear, which is a preferred condition for NP blocking occurrence (e.g., Deng & Jiang, 2011; Nakamura & Huang, 2018). To understand the jet change, equatorial Pacific SST anomalies and 10–100-day bandpass-filtered OLR anomalies were further investigated (Figure 9b). The result shows that cold SST anomalies are pronounced over the equatorial Pacific (−0.27 K in the Niño 3.4 region). Positive OLR anomalies appear over the cold SSTs with the negative OLR anomalies over the western equatorial Pacific, suggesting a typical La Niña-related OLR response. These initialized states with enhanced NP blocking activity at lead times of 3–4 weeks suggest that blocked cases are largely related to La Niña-like SST conditions (Renwick & Wallace, 1996). Figure 9c shows the composite of the mid-latitude jet for the well-predicted blocked cases. The poleward-shifted NP jet is more evident than in all blocked cases (compare Figures 9a and 9c). Interestingly, this jet change is not accompanied by the same equatorial SST signal (Figure 9d). The cold equatorial SST anomaly is not apparent in well-predicted cases (−0.07 K in the Niño 3.4 region). Instead, a positive OLR anomaly is pronounced over the east of the Maritime Continent. This suggests that modes of subseasonal variability in the tropics, such as the Madden–Julian Oscillation (MJO), may serve as a potential source of predictability for subseasonal prediction of NP blocking.

10.1029/2026JD046695

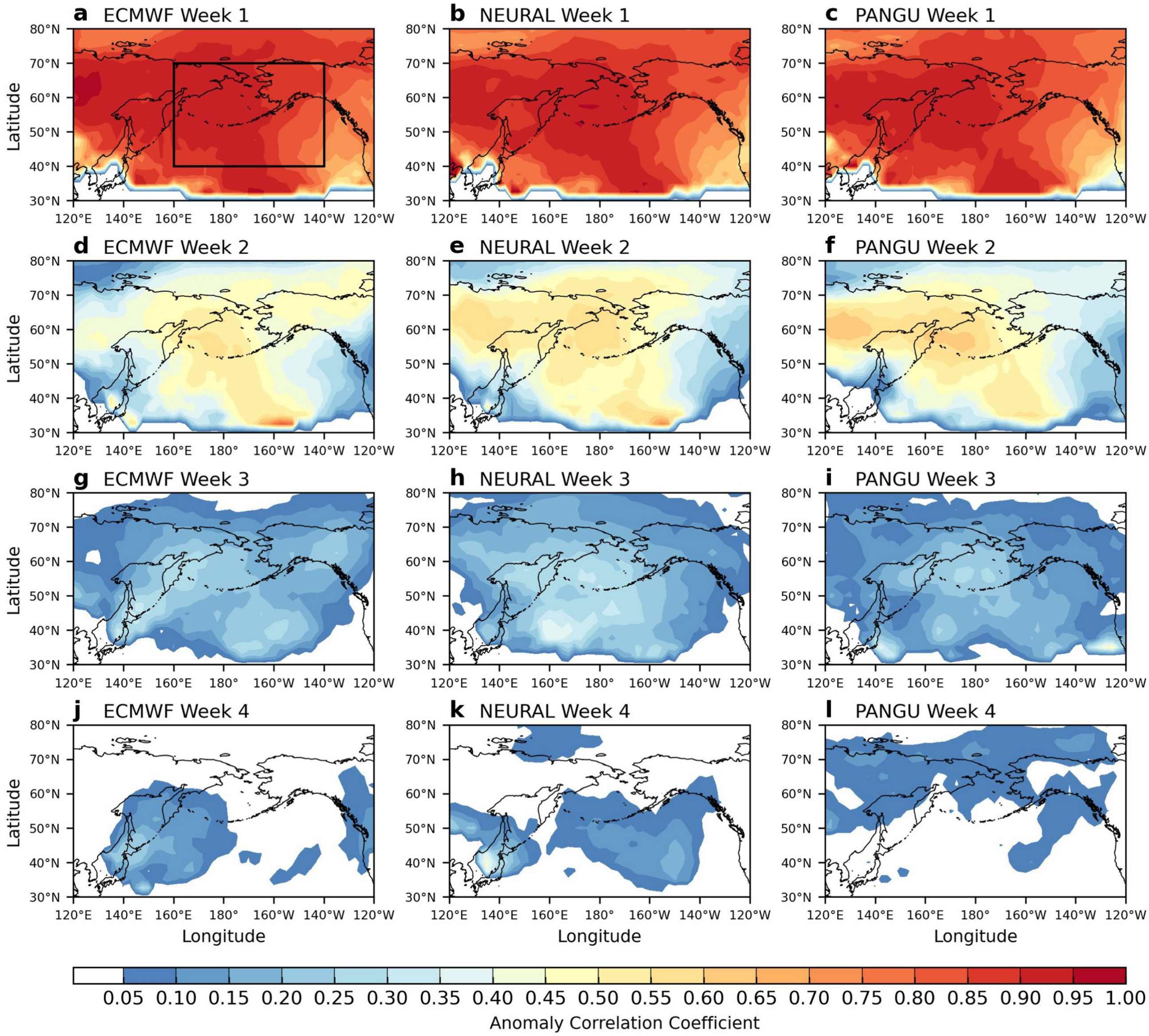

Figure 7. Prediction skill, estimated by temporal anomaly correlation coefficient (ACC_t), of weekly frequency of North Pacific blocking in (left) ECMWF IFS, (middle) NeuralGCM, and (right) Pangu-Weather.

The same analysis was repeated for non-blocked cases (Figure 10). Figure 10a shows the 300-hPa zonal wind composite at initialization for non-blocked cases. In contrast to blocked cases, the NP jet is more zonally elongated relative to the climatology, leading to enhanced meridional wind shear. The warm equatorial SST anomalies (+0.29 K in the Niño 3.4 region) and the corresponding OLR dipole response over the equatorial Pacific indicate that initializations are related to El Niño-like conditions, explaining the elongated NP jet and reduced blocking activity (Figure 10b; Renwick & Wallace, 1996). The well-predicted cases show a more zonally elongated jet (Figure 10c), with stronger warm SST anomalies (+0.49 K in the Niño 3.4 region) and the corresponding OLR response (Figure 10d). This suggests that, when initial conditions are favorable for suppressed blocking activity (e.g., warm equatorial SST anomalies), the model successfully produces fewer blocking events, thereby leading to higher NP blocking predictability.

To further assess the influence of tropical subseasonal variability on NP blocking prediction, Figure 11 examines the relative frequency of MJO phases at initialization for blocked and non-blocked cases. Here, relative frequency is defined as the ratio of the MJO phase frequency in each subset to its wintertime climatological frequency for the

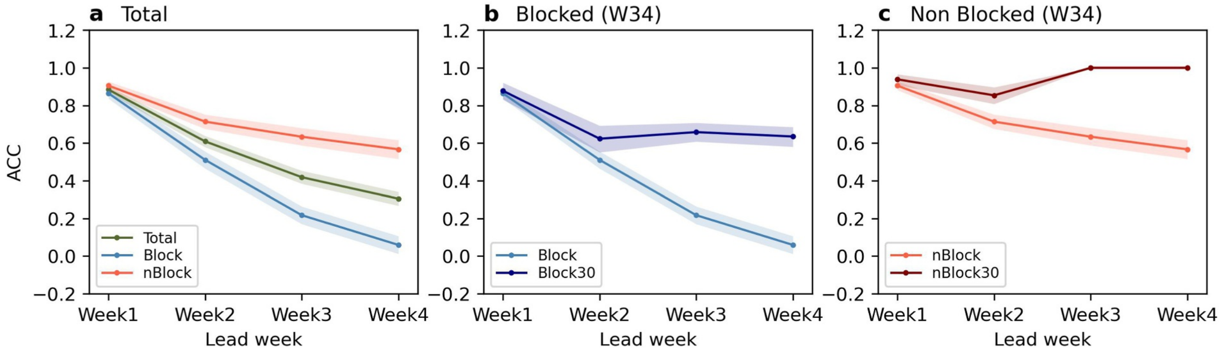

Figure 8. (a) Spatial anomaly correlation coefficient (ACC_s) of weekly North Pacific blocking frequency in ECMWF IFS (boxed region in Figure 7a) for all initializations (green), and for initializations classified as observed blocked (sky blue) or non-blocked (pink) during 3–4 weeks. Shading indicates the 90% confidence intervals estimated from 2,000 bootstrap resamples. Panels (b, c) highlight the top 30% of ACC for observed blocked (navy) and non-blocked (red) cases during 3–4 weeks, respectively. For reference, the ACCs for all observed blocked and non-blocked cases shown in panel (a) are reproduced in panels (b, c), respectively, to facilitate comparison with the corresponding top 30% cases.

period of 2004–2024. MJO phases are considered only when the amplitude of the real-time Multivariate MJO (RMM) index exceeds 1 (Wheeler & Hendon, 2004). The result shows that MJO phases 1–3 occur more frequently in blocked cases, whereas MJO Phase 4 is less frequent (Figure 11a). A different phase distribution is observed between all cases and well-predicted cases. Well-predicted blocked cases are associated with increased occurrences of MJO phases 8 and 2 compared to all blocked cases, explaining the poleward-shifted jet in the initial states through MJO teleconnection (Figure 9c). The appreciable prediction of MJO and the resulting teleconnection can influence the North Pacific extratropical circulations for 3–4 weeks (e.g., Kim et al., 2023),

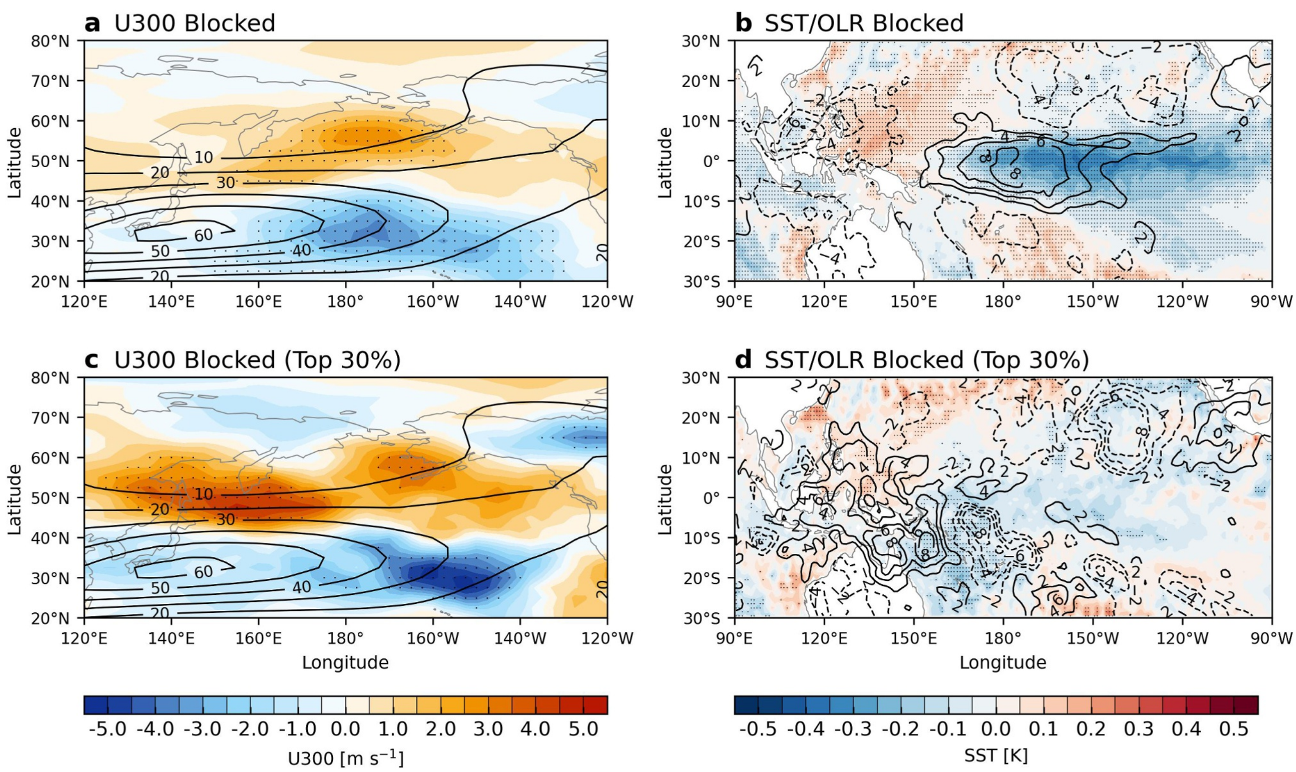

Figure 9. Composites of (a, c) 300-hPa zonal wind anomalies (U300; shading; m s^{-1}) and (b, d) sea surface temperature anomalies (shading; K) and 10–100 day bandpass filtered OLR (contours; 2 W m^{-2} intervals) for the initialization dates of (a, b) all blocked cases and (c, d) the top 30% ACC blocked cases during 3–4 weeks. In panels (a, c), the climatology of U300 is contoured (10 m s^{-1} interval) as a reference. Zero lines are omitted. Statistical significance was evaluated using Monte Carlo resamples (2,000 iterations). Regions that are statistically significant at the 90% confidence level are dotted.

10.1029/2026JD046695

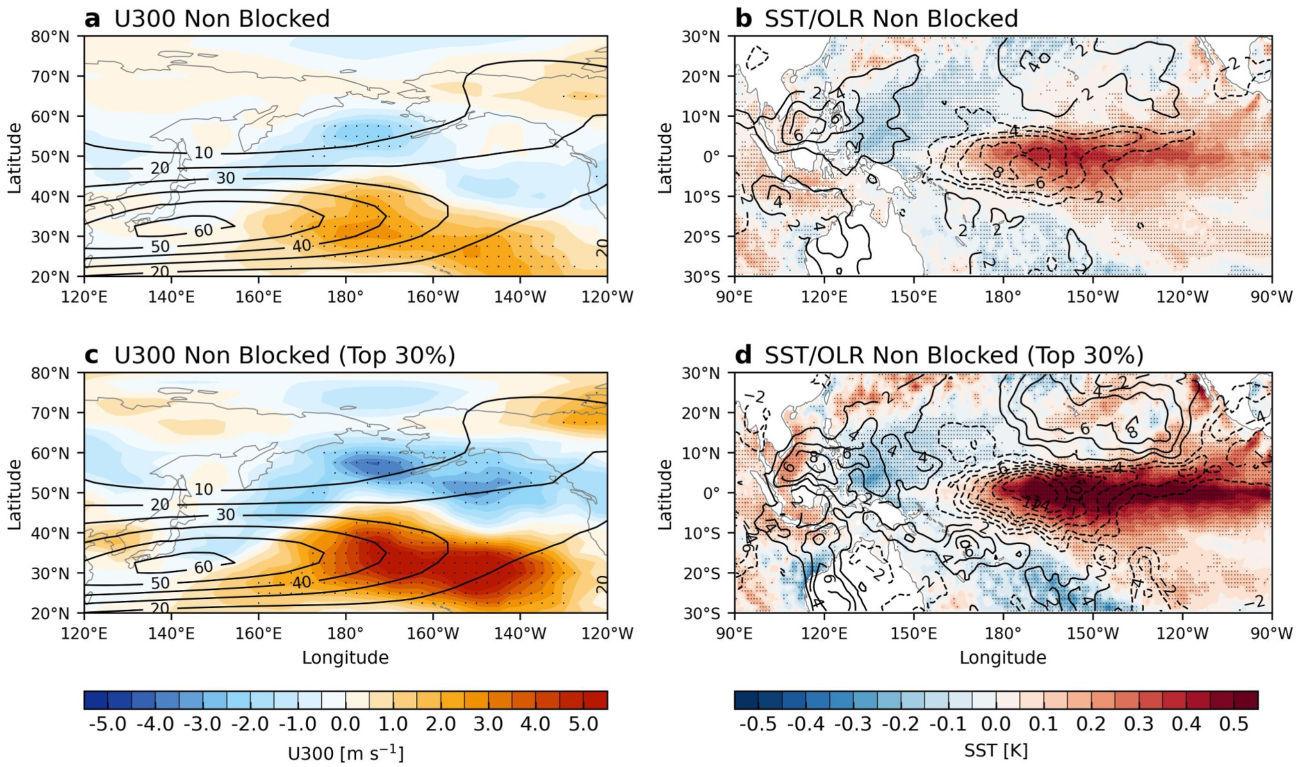

Figure 10. Same as Figure 9 but for non-blocked cases.

thereby modulating NP blocking predictability. Conversely, a reduced occurrence of MJO phases 1–3 with increased MJO Phase 4 is evident in non-blocked cases. For non-blocked cases (Figure 11b), the MJO phase distributions are overall similar between all cases and well-predicted cases, suggesting a secondary role of MJO in these cases.

Overall, these results suggest that interannual and subseasonal variability from the tropical Pacific may serve as potential sources of predictability for subseasonal NP blocking prediction in ECMWF IFS. We next tested whether leveraging this tropical variability can also enhance subseasonal NP blocking prediction in AI-based models. Figure 12 evaluates NP blocking prediction in NeuralGCM and Pangu-Weather using the well-

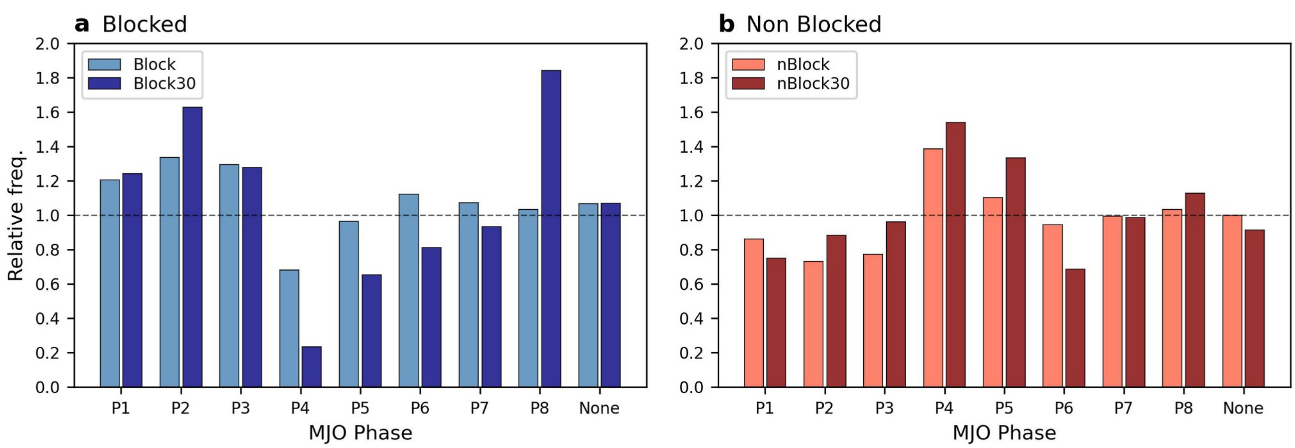

Figure 11. Relative frequency of MJO phases on the initialization dates for (a) blocked and (b) non-blocked cases. In each panel, all initializations (light color) are compared with the top 30% ACC cases (dark color). Relative frequency is defined as the ratio of the MJO phase frequency in each subset to its wintertime climatological frequency for the period of 2004–2024. MJO phases are considered only when the amplitude of the RMM index exceeds 1.

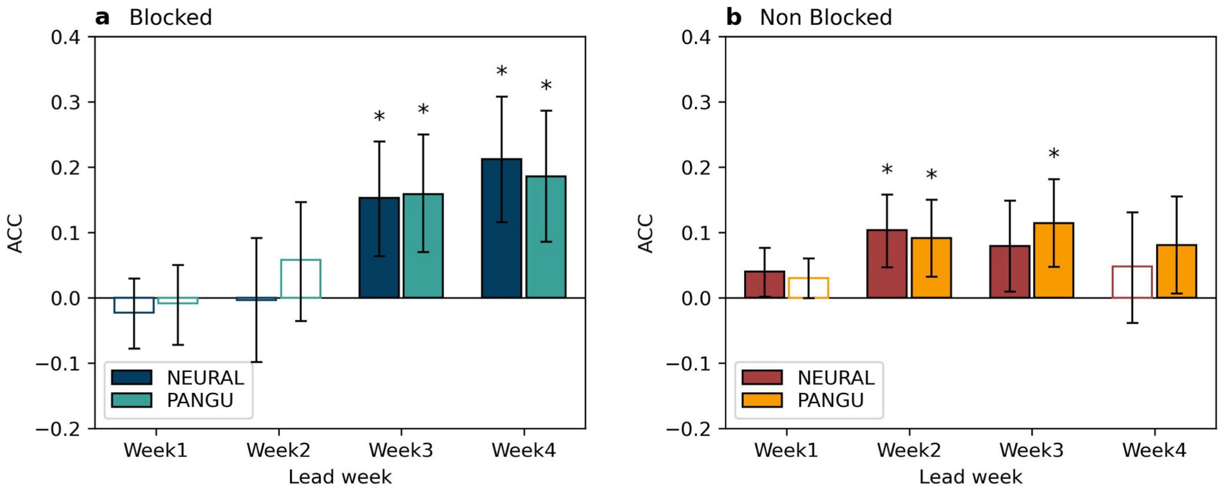

Figure 12. Differences in spatial ACC (ACC_s) between well-predicted and all initializations for (a) blocked and (b) non-blocked cases in NeuralGCM and Pangu-Weather. Error bars indicate 90% confidence intervals estimated from 2,000 bootstrap resamples. Filled bars indicate significant differences at 90% confidence level, and asterisks denote significance at the 95% level. Here, the top 30% ACC cases are selected based on ECMWF IFS.

predicted initializations identified in the ECMWF IFS. For blocked cases, both NeuralGCM and Pangu-Weather showed enhanced NP blocking prediction skill at lead times of weeks 3–4 (Figure 12a). Similarly, prediction skill increases for non-blocked cases in both models (Figure 12b). This result indicates that AI-based models may utilize tropical sources of predictability to enhance blocking prediction (e.g., Park et al., 2024) even when the information of the predictability sources is implicitly embedded only in the atmospheric initial states, as in the Pangu-Weather.

4. Summary and Discussion

This study evaluates the representation and subseasonal predictability of wintertime NP blocking across a hierarchy of operational prediction systems, namely ECMWF IFS (physics-based), NeuralGCM (hybrid), and Pangu-Weather (data-driven). All three models successfully reproduced the spatial distribution of the NP blocking frequency at forecast lead times of 1–2 weeks, exhibiting relatively weak biases. However, the forecasted blocking biases start to diverge substantially at lead times of 3–4 weeks. ECMWF IFS demonstrates an overestimation of NP blocking frequency over the eastern North Pacific, partly due to the overestimation of long-lived events. NeuralGCM exhibits a negative-positive dipolar bias pattern across the North Pacific. Pangu-Weather overall underestimates the NP blocking frequency due to the fewer events with 7–8-day durations being predicted.

Our results indicate that the three models predict distinct blocking statistics at long lead time. However, it is hard to conclude that these differences stem purely from variations in model architecture (i.e., from physics-based to data-driven) since the three models also differ in other aspects. For instance, the ECMWF IFS is fully coupled to the ocean and sea ice, whereas NeuralGCM utilizes fixed sea ice and SST as boundary conditions throughout the model integration. In contrast, Pangu-Weather is an atmosphere-only model without explicit knowledge of the sea ice and SST states. Furthermore, while ECMWF IFS and Pangu-Weather are both run at a horizontal resolution of 0.25° by 0.25° in our prediction experiments, NeuralGCM is simulated at a coarser resolution (1.4° by 1.4°). These differences in boundary conditions and model resolution can affect model representation of blocking. Additionally, the number of NP blocking events in the observation is limited even considering the period used for the training of AI-based models. Therefore, the exact sources of the blocking frequency biases found in the predictions across the three operational models should be interpreted with caution.

Despite the prediction biases identified, the blocking-center composites indicate that the overall NP blocking life cycle is qualitatively captured across the model hierarchy at both lead times of 1–2 and 3–4 weeks. NP blocking occurs in the climatological jet exit region and exhibits a westward-tilting structure consistent with cyclonic wave breaking. During the mature phase, the NP blocking anomaly is further intensified with a westward expansion. The blocking anomaly then weakens and futher propagates westward during the decay phase. All models

10.1029/2026JD046695

reproduce these dynamic features, although the blocking anomaly is systematically weaker than in ERA5, particularly at lead times of 3–4 weeks. Pangu-Weather additionally shows a weakened climatological jet at 3–4 weeks, partly due to blurred model forecasts at longer lead times in a data-driven model architecture.

The moist processes of NP blocking onset are qualitatively represented in all models despite variations in the model hierarchy. The direct effect of latent heating, estimated through upper-tropospheric positive geopotential (i.e., negative PV) tendency attributed to ascending motion within WCBs, is observed in the upstream of NP blocking. The indirect effect, horizontal PV advection by upper-tropospheric divergent flow, is also qualitatively captured in all three models. The distributions of the height tendency attributed to latent heating are overall comparable to ERA5 at lead times of 1–2 weeks. They remain qualitatively similar but narrower at lead times of 3–4 weeks, especially for Pangu-Weather.

Deterministic prediction skill for weekly-mean NP blocking frequency, measured by temporal and spatial anomaly correlation coefficients, declines rapidly with lead time in all models, showing a prediction limit of NP blocking near 2 weeks. The ERA5-based separation between blocked and non-blocked cases at 3–4 weeks further shows different prediction skill of NP blocking in ECMWF IFS. Non-blocked cases exhibit an overall higher skill score at 3–4 weeks compared to blocked cases. This is partly due to the rare occurrence of the blocking events, making it easier for models to simulate non-blocking states (i.e., normal states).

Composite analyses further show that blocked cases are preferentially initialized under La Niña-like background states consistent with a poleward-shifted NP jet. However, a robust subseasonal tropical signal associated with MJO phases 8 and 2 with a weak dependence on the equatorial SST anomalies improves the blocking prediction for blocked cases at lead times of 3–4 weeks. Conversely, non-blocked cases are closely related to the initializations with El Niño-like background states (e.g., a zonally elongated NP jet). The model simulations initialized with stronger El Niño-like background states exhibit a higher skill score across non-blocked cases. Using the well-predicted initializations identified in the ECMWF IFS, both NeuralGCM and Pangu-Weather demonstrated enhanced prediction skill for blocked and non-blocked cases at lead times of 3–4 weeks. This indicates that the AI-based models can be utilized for physically consistent NP blocking prediction at subseasonal forecast ranges by leveraging the sources of predictability associated with modes of subseasonal and interannual variability in the tropical Pacific. Nevertheless, it remains unclear whether the enhanced prediction skill arises primarily from the intrinsic predictability of the tropical forcing or from the models' ability to exploit the forcing from tropics. Further analysis is therefore needed to distinguish these factors more clearly and to better understand the source of enhanced predictability.

This study utilized initializations from the ECMWF IFS hindcast period from 2004 to 2024, which partly overlaps with the training period of the AI-based models, to ensure a sufficiently large sample of blocking events. To address this limitation, we additionally conducted a sensitivity test utilizing only the period outside the training data (i.e., test period). Unlike ECMWF IFS, which shows biases of comparable magnitude between the full period and the test period, the blocking frequency biases in the AI-based models are less consistent or become more pronounced during the test period. More specifically, NeuralGCM exhibits an overall enhanced blocking frequency at weeks 3–4, whereas the underestimation becomes more pronounced in Pangu-Weather. This difference may partly result from the limited sample size of observed NP blocking events (16 events identified) during the test period. Another possible factor is that the AI-based models may have greater difficulty reproducing the NP blocking statistics under the out-of-sample conditions. In particular, during the training period, especially for Pangu-Weather, the influence of long-memory systems (e.g., the MJO and SST) on the atmosphere may be overfitted, which could lead to a better reproduction of blocking statistics compared to the test period. Despite the larger or less consistent biases in the blocking statistics, the AI-based models qualitatively reproduce the evolution of blocking and the moist processes associated with blocking formation. The AI-based models exhibit appreciable prediction skill for NP blocking up to ∼2 weeks during the test period, which is consistent with the full period. In addition, the relationship between NP blocking and other internal variability was tested during the test period, particularly focusing on Pangu-Weather. The changes in ACC for weekly mean NP blocking frequency with respect to MJO and the differences in NP blocking frequency related to tropical SST variability are qualitatively reproduced as in ECMWF IFS (Figures S9 and S10 in Supporting Information S1). These results suggest the potential use of AI-based models for physically reliable prediction of weather extremes on subseasonal timescales despite the larger bias in blocking statistics.

It is also noted that the present study primarily evaluates subseasonal blocking prediction skill using deterministic forecasts. Given the diversity of blocking and its complex mechanisms (e.g., Woollings et al., 2018), the uncertainty arising from the internal variability in the model can play a crucial role in evaluating blocking prediction skill. In this regard, ensemble-based diagnostics can provide additional information. For instance, assessing the potential predictability of blocking in the models would help estimate the gap between realized prediction skill and potential predictability, thereby indicating the possible room for further improvement in blocking prediction (e.g., Admasu & White, 2025; Becker et al., 2014). Given the available data set, lagged initializations may provide an alternative way to generate ensembles. However, this approach can be practically limited in subseasonal forecasting (e.g., Toth & Kalnay, 1993). In addition, different ensemble-generation methods may lead to different estimates of blocking prediction skill (e.g., Matsueda et al., 2011), indicating the importance of applying a consistent ensemble-generation framework across models. Therefore, it remains an important next step to develop ensemble techniques that are tailored to subseasonal forecasting and can also be implemented consistently across a model hierarchy, from full-physics systems to purely data-driven (AI-based) models.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

ERA5 data are available from the Copernicus Climate Data Store (Copernicus Climate Change Service, 2023). The OLR data set is available from NOAA NCEI (Lee, 2025). The OISSTv2 data set was obtained from NOAA PSL (2025). The RMM index for MJO phases is available from the Bureau of Meteorology (BOM, 2026). ECMWF IFS hindcast data are freely available from the ECMWF S2S database (ECMWF, 2024). The pre-trained versions of Pangu-Weather and NeuralGCM were obtained from Huawei (2024) and Google DeepMind (2025), respectively.

Acknowledgments
J. H. was supported by Hankuk University of Foreign Studies Research Fund. Y. D. was supported by the U.S. National Science Foundation (NSF) through Grant AGS-2032532 and by the U.S. National Oceanic and Atmospheric Administration (NOAA) through Grant NA22OAR4310606.

References

Admasu, L., & White, R. H. (2025). Subseasonal prediction skill of winter quasi-stationary waves in the Northern Hemisphere. *Journal of Geophysical Research: Atmospheres*, *130*(19), e2025JD043337. https://doi.org/10.1029/2025jd043337

Athanasiadis, P. J., Bellucci, A., Hermanson, L., Scaife, A. A., MacLachlan, C., Arribas, A., et al. (2014). The representation of atmospheric blocking and the associated low-frequency variability in two seasonal prediction systems. *Journal of Climate*, *27*(24), 9082–9100. https://doi.org/10.1175/jcli-d-14-00291.1

Barriopedro, D., García-Herrera, R., Ordóñez, C., Miralles, D. G., & Salcedo-Sanz, S. (2023). Heat waves: Physical understanding and scientific challenges. *Reviews of Geophysics*, *61*(2), e2022RG000780. https://doi.org/10.1029/2022rg000780

Barriopedro, D., García-Herrera, R., & Trigo, R. M. (2010). Application of blocking diagnosis methods to general circulation models. Part I: A novel detection scheme. *Climate Dynamics*, *35*(7), 1373–1391. https://doi.org/10.1007/s00382-010-0767-5

Becker, E., den Dool, H. V., & Zhang, Q. (2014). Predictability and forecast skill in NMME. *Journal of Climate*, *27*(15), 5891–5906. https://doi.org/10.1175/jcli-d-13-00597.1

Ben Bouallegue, Z., Clare, M. C., Magnusson, L., Gascon, E., Maier-Gerber, M., Janoušek, M., et al. (2024). The rise of data-driven weather forecasting: A first statistical assessment of machine learning–based weather forecasts in an operational-like context. *Bulletin of the American Meteorological Society*, *105*(6), E864–E883. https://doi.org/10.1175/bams-d-23-0162.1

Benedict, J. J., Clement, A. C., & Medeiros, B. (2019). Atmospheric blocking and other large-scale precursor patterns of landfalling atmospheric rivers in the North Pacific: A CESM2 study. *Journal of Geophysical Research: Atmospheres*, *124*(21), 11330–11353. https://doi.org/10.1029/2019jd030790

Berckmans, J., Woollings, T., Demory, M. E., Vidale, P. L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: Influences of mean state.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, *619*(7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3

BOM. (2026). Real-time multivariate MJO index [Dataset]. Retrieved from http://www.bom.gov.au/climate/mjo/graphics/rmm.74toRealtime.txt

Bonavita, M. (2024). On some limitations of current machine learning weather prediction models. *Geophysical Research Letters*, *51*(12), e2023GL107377. https://doi.org/10.1029/2023gl107377

Branstator, G. (1987). A striking example of the atmosphere's leading traveling pattern. *Journal of the Atmospheric Sciences*, *44*(16), 2310–2323. https://doi.org/10.1175/1520-0469(1987)044<2310:aseota>2.0.co;2

Brenowitz, N. D., Cohen, Y., Pathak, J., Mahesh, A., Bonev, B., Kurth, T., et al. (2025). A practical probabilistic benchmark for ai weather models. *Geophysical Research Letters*, *52*(7), e2024GL113656. https://doi.org/10.1029/2024gl113656

Charlton-Perez, A. J., Dacre, H. F., Driscoll, S., Gray, S. L., Harvey, B., Harvey, N. J., et al. (2024). Do AI models produce better weather forecasts than physics-based models? A quantitative evaluation case study of storm Ciarán. *npj Climate and Atmospheric Science*, *7*(1), 93. https://doi.org/10.1038/s41612-024-00638-w

Charney, J. G., & DeVore, J. G. (1979). Multiple flow equilibria in the atmosphere and blocking. *Journal of the Atmospheric Sciences*, *36*(7), 1205–1216. https://doi.org/10.1175/1520-0469(1979)036<1205:mfeita>2.0.co;2

Chen, K., Han, T., Gong, J., Bai, L., Ling, F., Luo, J. J., & Ouyang, W. (2023). Fengwu: Pushing the skillful global medium-range weather forecast beyond 10 days lead. *arXiv preprint arXiv:2304.02948*.

Chen, L., Zhong, X., Li, H., Wu, J., Lu, B., Chen, D., et al. (2024). A machine learning model that outperforms conventional global subseasonal forecast models. *Nature Communications*, *15*(1), 6425. https://doi.org/10.1038/s41467-024-50714-1

Chen, L., Zhong, X., Zhang, F., Cheng, Y., Xu, Y., Qi, Y., & Li, H. (2023). FuXi: A cascade machine learning forecasting system for 15-day global weather forecast. *npj Climate and Atmospheric Science*, *6*(1), 190. https://doi.org/10.1038/s41612-023-00512-1

Copernicus Climate Change Service, C. D. S. (2023). ERA5 hourly data on pressure levels from 1940 to present [Dataset]. Retrieved from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-pressure-levels?tab=overview

Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. *Nature Climate Change*, *2*(7), 491–496. https://doi.org/10.1038/nclimate1452

Dacre, H. F., Charlton-Perez, A. J., Driscoll, S., Gray, S. L., Harvey, B., Harvey, N. J., & Volonté, A. (2025). Northern hemisphere midlatitude cyclone intensity biases in machine learning weather prediction models. *Bulletin of the American Meteorological Society*, *BAMS-D*.

Davini, P., D'Andrea, F., Park, S. B., & Gentine, P. (2017). Coherent structures in large-eddy simulations of a nonprecipitating stratocumulus-topped boundary layer. *Journal of the Atmospheric Sciences*, *74*(12), 4117–4137. https://doi.org/10.1175/jas-d-17-0050.1

Davini, P., Weisheimer, A., Balmaseda, M., Johnson, S. J., Molteni, F., Roberts, C. D., et al. (2021). The representation of winter Northern Hemisphere atmospheric blocking in ECMWF seasonal prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *147*(735), 1344–1363. https://doi.org/10.1002/qj.3974

DeFlorio, M. J., Waliser, D. E., Guan, B., Lavers, D. A., Ralph, F. M., & Vitart, F. (2018). Global assessment of atmospheric river prediction skill. *Journal of Hydrometeorology*, *19*(2), 409–426. https://doi.org/10.1175/jhm-d-17-0135.1

DeMaria, M., Franklin, J. L., Chirokova, G., Radford, J., DeMaria, R., Musgrave, K. D., & Ebert-Uphoff, I. (2025). An operations-based evaluation of tropical cyclone track and intensity forecasts from artificial intelligence weather prediction models. *Artificial Intelligence for the Earth Systems*, *4*(4), 240085. https://doi.org/10.1175/aies-d-24-0085.1

Deng, Y., & Jiang, T. (2011). Intraseasonal modulation of the North Pacific storm track by tropical convection in boreal winter. *Journal of Climate*, *24*(4), 1122–1137. https://doi.org/10.1175/2010jcli3676.1

Dole, R. M., & Gordon, N. D. (1983). Persistent anomalies of the extratropical Northern Hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. *Monthly Weather Review*, *111*(8), 1567–1586. https://doi.org/10.1175/1520-0493(1983)111<1567:paoten>2.0.co;2

Dolores-Tesillos, E., Martius, O., & Quinting, J. (2025). On the role of moist and dry processes in atmospheric blocking biases in the Euro-Atlantic region in CMIP6. *Weather and Climate Dynamics*, *6*(2), 471–487. https://doi.org/10.5194/wcd-6-471-2025

Domeisen, D. I., Eltahir, E. A., Fischer, E. M., Knutti, R., Perkins-Kirkpatrick, S. E., Schär, C., et al. (2023). Prediction and projection of heatwaves. *Nature Reviews Earth & Environment*, *4*(1), 36–50. https://doi.org/10.1038/s43017-022-00371-z

Dunn-Sigouin, E., Son, S. W., & Lin, H. (2013). Evaluation of Northern Hemisphere blocking climatology in the global environment multiscale model. *Monthly Weather Review*, *141*(2), 707–727. https://doi.org/10.1175/mwr-d-12-00134.1

ECMWF. (2024). S2S, ECMWF, CMA, CNRM, HMCR reforecasts datasets [Dataset]. Retrieved from https://apps.ecmwf.int/datasets/data/s2s/levtype=sfc/type=cf/

Emanuel, K. A., Fantini, M., & Thorpe, A. J. (1987). Baroclinic instability in an environment of small stability to slantwise moist convection. *Journal of the Atmospheric Sciences*, *44*(12), 1559–1573. https://doi.org/10.1175/1520-0469(1987)044<1559:biiaeo>2.0.co;2

Google DeepMind. (2025). NeuralGCM documentation [Software]. Retrieved from https://neuralgcm.readthedocs.io/en/latest/

Hakim, G. J., & Masanam, S. (2024). Dynamical tests of a deep learning weather prediction model. *Artificial Intelligence for the Earth Systems*, *3*(3). https://doi.org/10.1175/aies-d-23-0090.1

Hauser, S., Teubler, F., Riemer, M., Knippertz, P., & Grams, C. M. (2023). Towards a holistic understanding of blocked regime dynamics through a combination of complementary diagnostic perspectives. *Weather and Climate Dynamics*, *4*(2), 399–425. https://doi.org/10.5194/wcd-4-399-2023

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly journal of the royal meteorological society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hong, C. C., Hsu, H. H., Lin, N. H., & Chiu, H. (2011). Roles of European blocking and tropical-extratropical interaction in the 2010 Pakistan flooding. *Geophysical Research Letters*, *38*(13). https://doi.org/10.1029/2011gl047583

Hu, Y., Lin, Y., Deng, Y., & Bao, J. (2023). Summer extreme rainfall over the middle and lower reaches of Yangtze River: Role of synoptic patterns in historical changes and future projections. *Journal of Geophysical Research: Atmospheres*, *128*(24), e2023JD039608. https://doi.org/10.1029/2023jd039608

Huang, B., He, Y., Li, D., Zhang, B., & Jia, J. (2024). The influence of latent heating on the sub-seasonal prediction of winter blocking over Northern Hemisphere. *Climate Dynamics*, *62*(11), 10167–10184. https://doi.org/10.1007/s00382-024-07442-5

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., et al. (2021). Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1. *Journal of Climate*, *34*(8), 2923–2939. https://doi.org/10.1175/jcli-d-20-0166.1

Huawei. (2024). Pangu-weather documentation [Software]. Retrieved from https://github.com/ecmwf-lab/ai-models-panguweather

Hwang, J., & Son, S. W. (2025). How well does the climate model simulate the North Pacific blocking dynamics? *Journal of Geophysical Research: Atmospheres*, *130*(15), e2025JD044322. https://doi.org/10.1029/2025jd044322

Hwang, J., Son, S. W., Martineau, P., & Barriopedro, D. (2022). Impact of winter blocking on surface air temperature in East Asia: Ural versus Okhotsk blocking. *Climate Dynamics*, *59*(7), 2197–2212. https://doi.org/10.1007/s00382-022-06204-5

Jiang, T., Deng, Y., & Li, W. (2013). Local kinetic energy budget of high-frequency and intermediate-frequency eddies: Winter climatology and interannual variability. *Climate Dynamics*, *41*(3), 961–976. https://doi.org/10.1007/s00382-013-1684-1

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N., Achuthavarier, D., et al. (2012). High-resolution global climate simulations with the ECMWF model in Project Athena: Experimental design, model climate, and seasonal forecast skill. *Journal of Climate*, *25*(9), 3155–3172. https://doi.org/10.1175/jcli-d-11-00265.1

Kautz, L. A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., & Woollings, T. (2021). Atmospheric blocking and weather extremes over the Euro-Atlantic sector–a review. *Weather and Climate Dynamics Discussions*, *2021*, 1–43.

Keisler, R. (2022). Forecasting global weather with graph neural networks. *arXiv preprint arXiv:2202.07575*.

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. *Journal of Climate*, *36*(22), 7943–7956. https://doi.org/10.1175/jcli-d-23-0139.1

Kim, S. H., & Kim, B. M. (2025). Characterizing blocking mobility and its role in Northern Hemisphere cold extremes. *Environmental Research Letters*, *20*(10), 104078. https://doi.org/10.1088/1748-9326/ae0957

Kimoto, M., Mukougawa, H., & Yoden, S. (1992). Medium-range forecast skill variation and blocking transition. A case study. *Monthly Weather Review*, *120*(8), 1616–1627. https://doi.org/10.1175/1520-0493(1992)120<1616:mrfsva>2.0.co;2

Kleiner, N., Chan, P. W., Wang, L., Ma, D., & Kuang, Z. (2021). Effects of climate model mean-state bias on blocking underestimation. *Geophysical Research Letters*, *48*(13), e2021GL094129. https://doi.org/10.1029/2021gl094129

Kochkov, D., Yuval, J., Langmore, I., Norgaard, P., Smith, J., Mooers, G., et al. (2024). Neural general circulation models for weather and climate. *Nature*, *632*(8027), 1060–1066. https://doi.org/10.1038/s41586-024-07744-y

Kushnir, Y. (1987). Retrograding wintertime low-frequency disturbances over the North Pacific Ocean. *Journal of the Atmospheric Sciences*, *44*(19), 2727–2742. https://doi.org/10.1175/1520-0469(1987)044<2727:rwlfdo>2.0.co;2

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416–1421. https://doi.org/10.1126/science.adi2336

Lang, S., Alexe, M., Chantry, M., Dramsch, J., Pinault, F., Raoult, B., & Rabier, F. (2024). AIFS--ECMWF's data-driven forecasting system. *arXiv preprint arXiv:2406.01465*.

Lee, H.-T. (2025). NOAA Climate Data Record (CDR) of Daily Outgoing Longwave Radiation (OLR), version 2.0 [Dataset]. *NOAA National Centers for Environmental Information*. https://doi.org/10.25921/9pt5-by77

Lenggenhager, S., & Martius, O. (2019). Atmospheric blocks modulate the odds of heavy precipitation events in Europe. *Climate Dynamics*, *53*(7), 4155–4171. https://doi.org/10.1007/s00382-019-04779-0

Liu, C. C., Hsu, K., Peng, M. S., Chen, D. S., Chang, P. L., Hsiao, L. F., et al. (2024). Evaluation of five global AI models for predicting weather in Eastern Asia and Western Pacific. *npj Climate and Atmospheric Science*, *7*(1), 221. https://doi.org/10.1038/s41612-024-00769-0

Liu, Z., & Wang, L. (2025). Blocking diversity causes distinct roles of diabatic heating in the Northern Hemisphere. *Nature Communications*, *16*(1), 5613. https://doi.org/10.1038/s41467-025-60811-4

Lorenz, E. N. (1969). The predictability of a flow which possesses many scales of motion. *Tellus*, *21*(3), 289–307. https://doi.org/10.3402/tellusa.v21i3.10086

Lupo, A. R. (2021). Atmospheric blocking events: A review. *Annals of the New York Academy of Sciences*, *1504*(1), 5–24. https://doi.org/10.1111/nyas.14557

Masato, G., Hoskins, B. J., & Woollings, T. (2013). Wave-breaking characteristics of Northern Hemisphere winter blocking: A two-dimensional approach. *Journal of Climate*, *26*(13), 4535–4549. https://doi.org/10.1175/jcli-d-12-00240.1

Matsueda, M. (2009). Blocking predictability in operational medium-range ensemble forecasts. *Inside Solaris*, *5*, 113–116. https://doi.org/10.2151/sola.2009-029

Matsueda, M., Kyouda, M., Toth, Z., Tanaka, H. L., & Tsuyuki, T. (2011). Predictability of an atmospheric blocking event that occurred on 15 December 2005. *Monthly Weather Review*, *139*(8), 2455–2470. https://doi.org/10.1175/2010mwr3551.1

Nabizadeh, E., Lubis, S. W., & Hassanzadeh, P. (2021). The 3D structure of Northern Hemisphere blocking events: Climatology, role of moisture, and response to climate change. *Journal of Climate*, *34*(24), 9837–9860.

Nakamura, H., Nakamura, M., & Anderson, J. L. (1997). The role of high-and low-frequency dynamics in blocking formation. *Monthly Weather Review*, *125*(9), 2074–2093. https://doi.org/10.1175/1520-0493(1997)125<2074:trohal>2.0.co;2

Nakamura, N., & Huang, C. S. (2018). Atmospheric blocking as a traffic jam in the jet stream. *Science*, *361*(6397), 42–47. https://doi.org/10.1126/science.aat0721

NOAA Physical Sciences Laboratory. (2025). NOAA Optimal Interpolation (OI) SST version 2 [Dataset]. Retrieved from https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html

Nutter, P. A., Mullen, S. L., & Baumhefner, D. P. (1998). The impact of initial condition uncertainty on numerical simulations of blocking. *Monthly Weather Review*, *126*(9), 2482–2502. https://doi.org/10.1175/1520-0493(1998)126<2482:tioicu>2.0.co;2

Orlanski, I. (2003). Bifurcation in eddy life cycles: Implications for storm track variability. *Journal of the Atmospheric Sciences*, *60*(8), 993–1023. https://doi.org/10.1175/1520-0469(2003)60<993:bielci>2.0.co;2

Park, M., Johnson, N. C., Hwang, J., & Jia, L. (2024). A hybrid approach for skillful multiseasonal prediction of winter North Pacific blocking. *npj Climate and Atmospheric Science*, *7*(1), 227. https://doi.org/10.1038/s41612-024-00767-2

Park, T. W., Ho, C. H., & Deng, Y. (2014). A synoptic and dynamical characterization of wave-train and blocking cold surge over East Asia. *Climate Dynamics*, *43*(3), 753–770. https://doi.org/10.1007/s00382-013-1817-6

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., & Anandkumar, A. (2022). Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214*.

Pelly, J. L., & Hoskins, B. J. (2003). How well does the ECMWF ensemble prediction system predict blocking? *Quarterly Journal of the Royal Meteorological Society*, *129*(590), 1683–1702. https://doi.org/10.1256/qj.01.173

Pfahl, S., Schwierz, C., Croci-Maspoli, M., Grams, C. M., & Wernli, H. (2015). Importance of latent heat release in ascending air streams for atmospheric blocking. *Nature Geoscience*, *8*(8), 610–614. https://doi.org/10.1038/ngeo2487

Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T. R., El-Kadi, A., Masters, D., et al. (2025). Probabilistic weather forecasting with machine learning. *Nature*, *637*(8044), 84–90. https://doi.org/10.1038/s41586-024-08252-9

Quinting, J. F., & Vitart, F. (2019). Representation of synoptic-scale Rossby wave packets and blocking in the S2S prediction project database. *Geophysical Research Letters*, *46*(2), 1070–1078. https://doi.org/10.1029/2018gl081381

Rasp, S., Hoyer, S., Merose, A., Langmore, I., Battaglia, P., Russell, T., et al. (2024). WeatherBench 2: A benchmark for the next generation of data-driven global weather models. *Journal of Advances in Modeling Earth Systems*, *16*(6), e2023MS004019. https://doi.org/10.1029/2023ms004019

Renwick, J. A., & Wallace, J. M. (1996). Relationships between North Pacific wintertime blocking, El Niño, and the PNA pattern. *Monthly Weather Review*, *124*(9), 2071–2076. https://doi.org/10.1175/1520-0493(1996)124<2071:rbnpwb>2.0.co;2

Rex, D. F. (1950). Blocking action in the middle troposphere and its effect upon regional climate. *Tellus*, *2*(4), 275–301. https://doi.org/10.1111/j.2153-3490.1950.tb00339.x

Riviere, G., & Orlanski, I. (2007). Characteristics of the Atlantic storm-track eddy activity and its relation with the North Atlantic Oscillation. *Journal of the Atmospheric Sciences*, *64*(2), 241–266. https://doi.org/10.1175/jas3850.1

Rousi, E., Fink, A. H., Andersen, L. S., Becker, F. N., Beobide-Arsuaga, G., Breil, M., et al. (2023). The extremely hot and dry 2018 summer in central and northern Europe from a multi-faceted weather and climate perspective. *Natural Hazards and Earth System Sciences*, *23*(5), 1699–1718. https://doi.org/10.5194/nhess-23-1699-2023

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., & Russo, S. (2018). Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles. *Environmental Research Letters*, *13*(5), 054015. https://doi.org/10.1088/1748-9326/aaba55

Sillmann, J., Croci-Maspoli, M., Kallache, M., & Katz, R. W. (2011). Extreme cold winter temperatures in Europe under the influence of North Atlantic atmospheric blocking. *Journal of Climate*, *24*(22), 5899–5913. https://doi.org/10.1175/2011jcli4075.1

Steinfeld, D., Boettcher, M., Forbes, R., & Pfahl, S. (2020). The sensitivity of atmospheric blocking to changes in upstream latent heating–numerical experiments. *Weather and Climate Dynamics Discussions*, *2020*, 1–32.

Steinfeld, D., & Pfahl, S. (2019). The role of latent heating in atmospheric blocking dynamics: A global climatology. *Climate Dynamics*, *53*(9), 6159–6180. https://doi.org/10.1007/s00382-019-04919-6

Tibaldi, S., & Molteni, F. (1990). On the operational predictability of blocking. *Tellus*, *42*(3), 343–365. https://doi.org/10.3402/tellusa.v42i3.11882

Tibaldi, S., Tosi, E., Navarra, A., & Pedulli, L. (1994). Northern and Southern Hemisphere seasonal variability of blocking frequency and predictability. *Monthly Weather Review*, *122*(9), 1971–2003. https://doi.org/10.1175/1520-0493(1994)122<1971:nashsv>2.0.co;2

Toth, Z., & Kalnay, E. (1993). Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, *74*(12), 2317–2330. https://doi.org/10.1175/1520-0477(1993)074<2317:efantg>2.0.co;2

Tracton, M. S., Mo, K., Chen, W., Kalnay, E., Kistler, R., & White, G. (1989). Dynamical extended range forecasting (DERF) at the National Meteorological Center. *Monthly Weather Review*, *117*(7), 1604–1635. https://doi.org/10.1175/1520-0493(1989)117<1604:derfat>2.0.co;2

Tuel, A., & Martius, O. (2024). Persistent warm and cold spells in the Northern Hemisphere extratropics: Regionalisation, synoptic-scale dynamics and temperature budget. *Weather and Climate Dynamics*, *5*(1), 263–292. https://doi.org/10.5194/wcd-5-263-2024

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., et al. (2017). The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, *98*(1), 163–173. https://doi.org/10.1175/bams-d-16-0017.1

Wandel, J., Büeler, D., Knippertz, P., Quinting, J. F., & Grams, C. M. (2024). Why moist dynamic processes matter for the sub-seasonal prediction of atmospheric blocking over Europe. *Journal of Geophysical Research: Atmospheres*, *129*(8), e2023JD039791. https://doi.org/10.1029/2023jd039791

Wang, L., Lu, J., Breeden, M. L., Chen, G., Henderson, S. A., Narinesingh, V., & Lubis, S. W. (2026). Gaps and ways forward in atmospheric blocking and extreme weather research. *Nature Communications*.

Watson, J. S., & Colucci, S. J. (2002). Evaluation of ensemble predictions of blocking in the NCEP global spectral model. *Monthly Weather Review*, *130*(12), 3008–3021. https://doi.org/10.1175/1520-0493(2002)130<3008:eoepob>2.0.co;2

Wheeler, M. C., & Hendon, H. H. (2004). An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, *132*(8), 1917–1932. https://doi.org/10.1175/1520-0493(2004)132<1917:aarmmi>2.0.co;2

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J., Lazo, J. K., Kumar, A., et al. (2017). Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, *24*(3), 315–325. https://doi.org/10.1002/met.1654

White, R. H., Anderson, S., Booth, J. F., Braich, G., Draeger, C., Fei, C., et al. (2023). The unprecedented Pacific Northwest heatwave of June 2021. *Nature Communications*, *14*(1), 727. https://doi.org/10.1038/s41467-023-36289-3

Woollings, T., Barriopedro, D., Methven, J., Son, S. W., Martius, O., Harvey, B., et al. (2018). Blocking and its response to climate change. *Current Climate Change Reports*, *4*(3), 287–300. https://doi.org/10.1007/s40641-018-0108-z

Xavier, A. K., Hamilton, O., Faranda, D., & Vannitsem, S. (2026). Predictability of North Pacific blocking events: Analogue-based analysis of historical MIROC6 simulations. *Quarterly Journal of the Royal Meteorological Society*, *152*(776), e70104. https://doi.org/10.1002/qj.70104

Xie, Z., Black, R. X., & Deng, Y. (2019). Planetary and synoptic-scale dynamic control of extreme cold wave patterns over the United States. *Climate Dynamics*, *53*(3), 1477–1495. https://doi.org/10.1007/s00382-019-04683-7

Xu, H., Duan, Y., & Xu, X. (2024). Evaluating AI's capability to reflect physical mechanisms: A case study of tropical cyclone impacts on extreme rainfall. *Environmental Research Letters*, *19*(10), 104006. https://doi.org/10.1088/1748-9326/ad6fbb

Yamamoto, A., Nonaka, M., Martineau, P., Yamazaki, A., Kwon, Y. O., Nakamura, H., & Taguchi, B. (2021). Oceanic moisture sources contributing to wintertime Euro-Atlantic blocking. *Weather and Climate Dynamics*, *2*(3), 819–840. https://doi.org/10.5194/wcd-2-819-2021

Zheng, C., Chang, E. K. M., Kim, H., Zhang, M., & Wang, W. (2019). Subseasonal to seasonal prediction of wintertime Northern Hemisphere extratropical cyclone activity by S2S and NMME models. *Journal of Geophysical Research: Atmospheres*, *124*(22), 12057–12077. https://doi.org/10.1029/2019jd031252