Received: 5 May 2025 | Revised: 10 November 2025 | Accepted: 19 November 2025

DOI: 10.1002/qj.70090

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Windows of opportunity in subseasonal weather regime forecasting: A statistical–dynamical approach

Fabian Mockert[1,2] | **Christian M. Grams**[1] | **Sebastian Lerch**[2,3] | **Julian Quinting**[1]

[1]Institute of Meteorology and Climate Research Troposphere Research (IMKTRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

[2]Institute of Statistics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

[3]Heidelberg Institute for Theoretical Studies, Heidelberg, Germany

Correspondence
Fabian Mockert, Institute of Statistics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.
Email: fabian.mockert@kit.edu

Present address
Christian M. Grams, Federal Office of Meteorology and Climatology, MeteoSwiss, Zurich-Airport, Switzerland; Sebastian Lerch, Department of Mathematics and Computer Science, University of Marburg, Marburg, Germany; and Julian Quinting, Institute of Geophysics and Meteorology, University of Cologne, Cologne, Germany

Funding information
Vector Stiftung: Young investigator Group "Artifical Intelligence for Probabilistic Weather Forecasting"; Helmholtz Association, Grant/Award Number: VH-NG-1243; KIT Centre MathSEE; European Union: ERC, ASPIRE, Grant/Award Number: 101077260

Abstract

The Madden–Julian Oscillation (MJO) and stratospheric polar vortex (SPV) are prominent sources of subseasonal predictability in the extratropics. It has been shown that the joint interaction of the MJO and the SPV can modulate the preferred phase of the North Atlantic Oscillation (NAO) and the occurrence of weather regimes. However, improving numerical weather prediction (NWP) at 3-week lead times remain underexplored. This study investigates how MJO and SPV phases affect Greenland Blocking (GL) activity and integrates atmospheric state information into a neural network to enhance week 3 weather regime activity forecasts. We define a weather regime activity metric using European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis and reforecasts. In reanalyses we find increased GL activity following MJO phases 7, 8, and 1, as well as weak SPV phases, indicating climatological windows of opportunity in line with previous studies. However, ECMWF forecast skill improves only in MJO phases 8 and 1 and weak SPV phases, identifying somewhat different model windows of opportunity. Next, we explore using these findings in postprocessing tools. Climatological forecasts based on MJO/SPV–NAO relationships provide a purely statistical approach to subseasonal GL activity forecasting, independent of NWP models. Notably, MJO-conditioned climatological forecasts show clear signals when evaluated against observed GL activity. Statistical–dynamical models, using neural networks that combine historical atmospheric state data with NWP-derived weather regime metrics, improve weather regime activity forecasts across all regimes considered, achieving an absolute accuracy increase of 5.8 percentage points in forecasting the dominant weather regime compared with ECMWF. This is particularly beneficial to blocking in the European domain, where NWP models often underperform. Atmospheric conditioned and neural network forecasts serve as valuable decision-support tools alongside NWP models, enhancing the reliability of subseasonal predictions.

KEYWORDS

Madden–Julian Oscillation, neural networks, stratospheric polar vortex, weather regimes, windows of opportunity

1 | INTRODUCTION

The demand for accurate subseasonal forecasting—extending beyond 10 days and up to 2 months—has grown over the past decade, driven by socio-economic needs. Decision-makers in sectors such as public health, agriculture, and energy rely on accurate forecasts several weeks ahead. For example, in the United Kingdom, winter cold spells can increase hospital admissions due to respiratory diseases (Charlton-Perez *et al.*, 2019; Elliot *et al.*, 2008). In agriculture, improved drought forecasting can save millions of dollars by mitigating crop shortages (Salient & Atlas, 2023). Additionally, large-scale weather patterns can serve as indicators of peak winter energy demand, which can place immense stress on power systems (Bloomfield *et al.*, 2020; Millin *et al.*, 2024). Moreover, in the context of renewable energy forecasting, Bloomfield *et al.* (2021) show that pattern-based methods yield higher forecast skill than grid-point-based methods for lead times beyond 12 days.

Large-scale quasi-stationary, recurrent, and persistent weather patterns are commonly referred to as weather regimes (Hannachi *et al.*, 2017; Michelangeli *et al.*, 1995; Vautard, 1990). They represent the complex large-scale circulations through a finite set of states, facilitating the interpretation and categorisation of the prevailing or forecasted atmospheric flow. For Europe, four seasonal weather regimes are common; namely, the two phases of the North Atlantic Oscillation (NAO+/NAO−), Atlantic Ridge (AR), and Blocking (e.g., Cassou, 2008; Ferranti *et al.*, 2015; Michelangeli *et al.*, 1995). These regimes are associated with distinct impacts on surface variables, including 2-m temperature, 10-m wind, and radiation (Grams *et al.*, 2017; Yiou & Nogaj, 2004). Often, the weather regimes are identified on the basis of the 500 hPa geopotential height field (Z500). The seasonal regimes have been refined recently, for the European region, by a year-round definition of seven weather regimes (Grams *et al.*, 2017) that can be formalised with a seven-dimensional normalised weather regime index (IWR). Two of these seven weather regimes, Zonal Regime (ZO) and Greenland Blocking (GL), closely resemble the positive and negative phase of the NAO, respectively; see Beerli and Grams (2019) for more details. We focus on the seven weather regimes, as they provide a more nuanced view of surface weather impacts, particularly during winter (Beerli & Grams, 2019; Mockert *et al.*, 2023). The forecast skill of the GL and ZO is generally the best among the seven weather regimes, whereas the European and Scandinavian Blocking regimes are the worst predicted. This has been demonstrated using reforecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) model alone (Büeler *et al.*, 2021) and in combination with the National Centers for Environmental Prediction and the UK Met Office models (Osman *et al.*, 2023), with the forecast skill evaluated via the Brier skill score applied to the categorical IWR. These results are further supported when assessing a four-regime definition: the anomaly correlation between NAO+/NAO− to blocking/anti-blocking across six models indicates similar relative skill differences (Ferranti *et al.*, 2018).

Blocking regimes are especially important for the prediction of renewable power generation in winter, as periods of low wind and solar power output occur mainly within these regimes (Drücke *et al.*, 2021; Otero *et al.*, 2022; Wiel *et al.*, 2019). GL is further connected with colder-than-average winter temperature anomalies in central and northern Europe, leading to increased heating demand (Mockert *et al.*, 2023). Improvements in the predictions for these three regimes, along with the other four regimes, in winter would greatly benefit the energy sector, particularly energy trading companies and transmission grid operators.

Previous studies based on reanalysis data suggest links between the atmospheric state at initialisation time and the occurrence of weather regimes weeks later (e.g., Cassou, 2008). Though much weaker, these connections are reflected in the skill of numerical weather prediction (NWP) models in forecasting these regimes (e.g., Vitart, 2017). Two major atmospheric modes of variability influencing GL (or NAO−) are the Madden–Julian Oscillation (MJO; Madden & Julian, 1971), and the stratospheric polar vortex (SPV). Cassou (2008) and Lin *et al.* (2009) showed a lagged connection between MJO phases and the preferred flow pattern in the North Atlantic–European region in reanalysis data. The occurrence of NAO− is anomalously high 2 weeks after MJO phases 6–8, whereas NAO+ is anomalously more frequent following MJO phases 3 and 4. Similarly, although considerably weaker than in reanalysis, reforecasts by the ECMWF show positive NAO forecasts for lead times of 11–15 days following phase 3 and negative NAO forecasts following phase 7 (Vitart, 2017). Ferranti *et al.* (2018) describe an asymmetric connection of the MJO to forecasts of the NAO phases. Forecasts of NAO− tend to have more (less) skill when the MJO is active (inactive) at initialisation time, whereas little sensitivity of the forecast skill to the MJO activity at initial time is found for the NAO+ phase. The underlying mechanisms have been explored with idealised model experiments, showing that tropical MJO forcing excites Rossby waves that modulate the Pacific jet and synoptic eddy activity, which in turn favours the development of NAO+ after phase 3 and NAO− after phase 6 (Fromang & Rivière, 2020). Despite the MJO teleconnection to Europe shown, NWP still struggles exploiting this potential source of subseasonal predictability. A potential

reason is the multiscale interaction with other modes of subseasonal variability and the chaotic growth of error (Roberts *et al.*, 2023).

Most importantly, the relationship between the MJO and European weather regimes is further modulated by the quasi-biennial oscillation (QBO), El Niño Southern Oscillation (ENSO), and the SPV. In this context, Feng and Lin (2019) investigate the role of the QBO, Lee *et al.* (2019) and Roberts *et al.* (2023) focus on ENSO-related modulation and its representation in forecasts, and Kent *et al.* (2023) explore where skill improvements for the MJO–NAO link might be achieved.

The MJO influences the NAO through both a tropospheric pathway and a stratospheric pathway involving the polar vortex. Lee *et al.* (2019) show that the Rossby-wave-driven tropospheric pathway linking the MJO to NAO+ is strongly enhanced during El Niño, whereas the stratospheric pathway connecting the MJO to NAO− is strengthened during La Niña. When the SPV is in a state that supports the expected NAO response, the likelihood of the NAO being in the corresponding state following active MJO periods increases by up to 30% (Barnes *et al.*, 2019). Reanalysis data show that GL (NAO−) occurs with an enhanced probability of up to 2 months following weak SPV conditions (e.g., Beerli & Grams, 2019) due to persistent anomalies in the stratosphere (White *et al.*, 2020). Further, Domeisen *et al.* (2020c) state, based on reanalysis data, that sudden stratospheric warmings (SSWs) are also frequently followed by an Atlantic Trough (AT) in the weeks after. GL is most likely if a blocking situation over western Europe and the North Sea prevailed, whereas the AT is most likely when GL was present around the SSW onset.

The connection between the atmospheric state at initialisation time and the frequency of occurrence of weather regimes, as well as the enhanced forecast skill after forecast initialisation in certain atmospheric states, indicates that the phases of the MJO and SPV could potentially be considered as windows of opportunity for subseasonal predictions. Windows of opportunity can generally be defined as specific atmospheric states in which either the frequency of occurrence of an event is increased, referred to as a climatological window of opportunity, or the forecast of an event is improved, referred to as a model window of opportunity—following the definition of Specq and Batté (2022) applying windows of opportunity on the MJO and heavy tropical precipitation events.

This study explores windows of opportunity linking the MJO and SPV to weather regimes during the extended winter period (November to March). A key objective is to determine whether the established links between these atmospheric modes of variability and weather regimes, typically observed up to 2 weeks, extend into forecast week 3.

To investigate this, we first focus on GL as an illustrative example. GL is chosen because it plays a critical role in European weather impacts in wintertime; for instance, it is associated with cold Dunkelflauten (Mockert *et al.*, 2023), causing strong stress scenarios for the power system in wintertime. Using GL as an example, we investigate whether a climatological forecast conditioned on the atmospheric state at initialisation can provide meaningful indications of GL occurrence beyond week 2. Building on this, we then broaden the analysis using statistical–dynamical models, particularly neural networks (NNs) that combine prior atmospheric knowledge with NWP outputs. In this context, we generalise beyond GL to predict the full set of weather regimes, since understanding the joint activity across all regimes is essential for anticipating societal impacts.

Before answering these questions, data and methods are introduced in Section 2. This study provides a detailed analysis of windows of opportunity conditioned on the MJO and SPV for predicting the activity of GL on a lead time of 3 weeks (Sections 3.1 and 3.2). That information is then further used to provide an atmospheric-based climatology conditioned on the state of the MJO or SPV as a tool alongside an NWP model (Section 3.3). In a final step, all information from the NWP and the atmospheric state at initialisation time are joined by an NN to improve the forecasts of the activity of weather regimes (Section 3.4). Section 4 summarises and discusses the main findings.

2 | DATA AND METHODS

2.1 | ECMWF reforecast and reanalysis

This study utilises subseasonal to seasonal reforecast data of the ensemble prediction system of the ECMWF, provided through the Subseasonal-to-Seasonal Prediction Project Database (Vitart *et al.*, 2017). To increase the number of forecast initial dates, we include forecasts from two consecutive model cycles, Cy46R1 and Cy47R1 (ECMWF, 2018). The forecasts from these two model cycles are treated together and weighted equally; differences between the cycles are minor compared with larger changes in other model cycles and are not expected to substantially affect the results. These reforecasts are initialised twice a week (Mondays and Thursdays) from ECMWF Reanalysis v5 (ERA5) reanalysis data (Hersbach *et al.*, 2020) for the past 20 years and consist of 11 ensemble members. The forecasts cover a forecast lead time of 0–46 days at a native horizontal grid spacing of 16 km up to day 15 and 32 km from day 15 onwards. Forecast data were remapped from their native resolution to a regular latitude–longitude grid of $1° \times 1°$ grid spacing. The two

model cycles were operational from June 11, 2019, to May 11, 2021, with a cycle change from Cy46R1 to Cy47R1 on June 30, 2020. As a result, for the period between May 11 and June 30, reforecasts are only available from Cy46R1. Our reforecast dataset comprises a total of 4,000 initial dates, each with 11 ensemble members. For forecast evaluation, ERA5 is used as the verification dataset by aligning its data with each initialisation date and lead time (cf. Wandel *et al.*, 2024). Additionally, ERA5 data are remapped from their native grid to match the reforecast grid spacing.

2.2 | Weather regimes

In this study, we use the seven year-round North Atlantic–European weather regimes introduced by Grams *et al.* (2017) adapted to ERA5 reanalysis as described in Hauser *et al.* (2023a, 2023b) and applied to reforecasts following the approach of Büeler *et al.* (2021) and Osman *et al.* (2023). These weather regimes represent the most common large-scale circulation patterns in the North Atlantic–European region (30–90° N, 80° W–40° E).

For completeness, the key steps of the regime computation are repeated here. First, we compute 500 hPa geopotential height anomalies (Z500 anomalies) relative to a 91-day running climatology, defined as the multi-year mean for each calendar day smoothed with a 91-day window, within the domain of weather regimes. As in previous studies (e.g., Michel & Rivière, 2011) we suppress variability on time-scales shorter than 10 days, by applying a Lanczos low-pass filter with a 10-day cut-off period. These filtered Z500 anomalies are then seasonally normalised[1] to ensure that the anomalies from different seasons are comparable in magnitude, allowing for a consistent year-round classification of weather regimes. The resulting six-hourly Z500 anomaly fields (1979–2019), after filtering and normalisation, are used as input for an empirical orthogonal function analysis. We then apply a k-means clustering algorithm on the first seven EOFs and set $k = 7$. These seven clusters represent the seven weather regimes—for a visualisation of the weather patterns, see Mockert *et al.* (2024, fig. S1)—with three cyclonic regime types (AT, ZO, and Scandinavian Trough [ScTr]) and four anticyclonic regime types (AR, European Blocking [EuBL], Scandinavian Blocking [ScBL], and GL).

The projection of instantaneous anomalies onto the mean regime patterns, whether in reanalysis or reforecast, is determined by a seven-dimensional IWR following the approach of Michel and Rivière (2011). As this study relies on the dataset computed in Mockert *et al.* (2024), we refer the interested reader to their publication for a detailed explanation and only briefly outline the computation of the IWR. Here, we specifically use the Z500 bias-corrected IWR reforecasts and reanalysis from Mockert *et al.* (2024):

1. Compute Z500 anomalies for reforecasts/ERA5 relative to a 91-day running mean model/ERA5 climatology (1999–2015 and 1979–2019, respectively).
2. Apply low-pass filtering and normalisation to obtain filtered and standardised Z500 anomalies.
3. Project filtered and standardised Z500 anomalies onto seven cluster mean Z500 anomalies of distinct weather regimes.
4. Compute IWR by normalising the projection against its climatological mean and standard deviation.

2.3 | Weather regime activity

This study focuses on the predictability of weather regimes at 3 weeks forecast lead time. Owing to the reduced skill of daily weather regime forecasts at this lead time (Büeler *et al.*, 2021) and the practical needs of decision-makers, we consider weekly aggregated forecast information based on daily data. To achieve this, we introduce the concept of weather regime activity, which quantifies the presence of a specific weather regime over a given period. The weather regime activity is defined in two complementary ways.

2.3.1 | Weekly mean weather regime activity

The weekly mean weather regime activity ($\mathrm{WRact}_{\mathrm{mean}}$), suited for deterministic decision-making, considers a weather regime active in ERA5 or an ensemble member ($\mathrm{WRact}_{\mathrm{mean}} = 1$) in the respective week, if the weekly mean of daily IWR[2] exceeds a predefined activity threshold of $\mathrm{IWR}_{\mathrm{min}} = 1$, otherwise $\mathrm{WRact}_{\mathrm{mean}} = 0$. Applied to the forecasts across 11 ensemble members, $\mathrm{WRact}_{\mathrm{mean}}$ takes values between 0 and 1 in increments of 1/11, representing a probability in the ensemble. Hence, for the $\mathrm{WRact}_{\mathrm{mean}}$, ERA5 and the individual ensemble members provide a deterministic measure of weather regime activity, whereas the ensemble mean (across 11 members) offers a probabilistic forecast of this deterministic outcome.

2.3.2 | Aggregated daily weather regime activity

In contrast, the aggregated daily weather regime activity ($\mathrm{WRact}_{\mathrm{agg}}$) provides a more detailed view by measuring the fraction of days within a given week where the daily IWR exceeds $\mathrm{IWR}_{\mathrm{min}}$. Applied to ERA5 or an

ensemble member, this results in values between 0 and 1, with discrete steps of 1/7, representing the fraction of days per week with active regime conditions. When applied to ensemble forecasts across 11 ensemble members, $WRact_{agg}$ yields values between 0 and 1 in steps of 1/77, allowing for a probabilistic perspective, representing the fraction of days per week with active regime conditions in the ensemble.

2.3.3 | Computation process

Figure 1 illustrates the computation process. Seven consecutive daily IWR values from ERA5 reanalysis, shown as turquoise circles in Figure 1a, are averaged to compute the weekly mean IWR (blue lines). This value (here, 2.1 in week 3) is then compared with the activity threshold IWR_{min} (black dashed line; here, 1.0), determining whether the regime is considered active (1.0) or inactive (0.0). The corresponding $WRact_{mean}$ is represented as blue bar in Figure 1b. The $WRact_{agg}$ follows a similar approach, but instead assesses each daily IWR value individually against IWR_{min}. The fraction of days exceeding IWR_{min} defines the $WRact_{agg}$, visualised by the turquoise bar in Figure 1b.

The same procedure applies to ensemble forecasts, where the IWR forecast for each ensemble member, initialised at the beginning of the period shown, is processed in the same manner (green lines in Figure 1a). The resulting $WRact_{mean}$ for the ensemble is displayed as purple bars in Figure 1b, whereas the $WRact_{agg}$ is shown in green. This methodology enables a comprehensive assessment of weather regime activity across both reanalysis and forecasts. Since this process is applied individually to each weather regime, multiple regimes can be considered active simultaneously, reflecting the inherent variability in large-scale atmospheric patterns.

2.3.4 | Maximum weather regime activity

To further simplify the interpretation of a weekly weather regime activity, we introduce the maximum weather regime activity ($WRact_{max}$). This metric assigns a single dominant regime to each day, independent of any IWR threshold, by selecting the regime with the highest

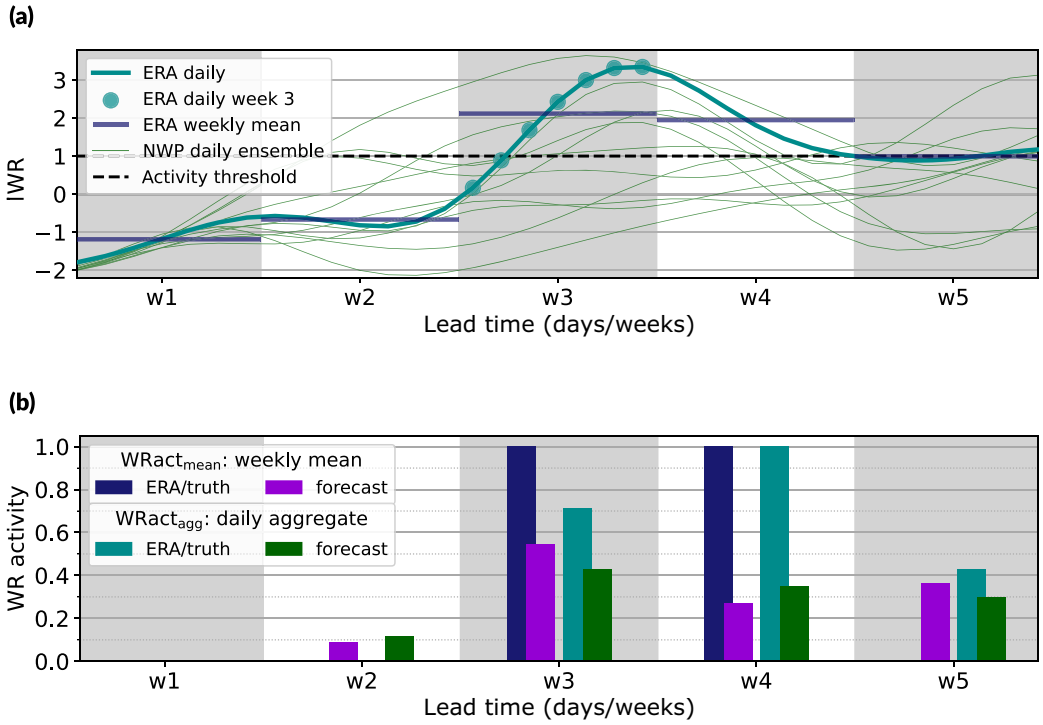

FIGURE 1 Visualisation of $WRact_{mean}$ and $WRact_{agg}$ by using the weather regime (WR) index (IWR). (a) The ensemble forecast (green lines) and the verification data (European Centre for Medium-Range Weather Forecasts Reanalysis v5 [ERA5], turquoise line) of the daily IWR. Further, the dark blue horizontal lines indicate the mean weekly IWR in ERA5 and the turquoise dots the daily IWR in ERA5 for forecast week 3. The latter is used for the computation of the $WRact_{agg}$ by comparing them against the activity threshold (black dashed line). (b) The $WRact_{mean}$ (dark blue bars) and $WRact_{agg}$ (turquoise bars) in ERA5 and in forecasts (purple and green bars, respectively) for each forecast week. By definition, $WRact_{mean}$ (dark blue bars) for ERA5 takes values of either 0.0 or 1.0. [Colour figure can be viewed at wileyonlinelibrary.com]

aggregated activity ($WRact_{agg}$, computed with the IWR of the surrounding 7-day window, ±3 days from the day of interest). In case multiple weather regimes share the highest $WRact_{agg}$ at a particular time, the regime persisting the longest is chosen. By designating the most active regime, $WRact_{max}$ provides a practical and intuitive summary, ensuring that each atmospheric scenario is linked to the prevailing large-scale circulation pattern. With this definition, we aim to offer a simplified yet informative representation of regime dominance, making it accessible to users beyond the meteorological sector, such as decision-makers who rely on large-scale weather variability information.

2.4 | Contingency table and metrics

We define windows of opportunity and validate the $WRact_{mean}$ on the basis of contingency tables. We utilise both traditional and probabilistic contingency tables—following the methodology in Gold *et al.* (2020)—as well as verification metrics that are derived from these.

A traditional contingency table (Table 1, text) represents the relationship between forecasted and observed categorical events (e.g., whether a weather regime is active or not). The contingency table classifies outcomes into four categories: hit (or true positive, TP), where weather regime activity is forecasted and the weather regime activity occurs; miss (or false negative, FN), where no weather regime activity is forecasted but a weather regime activity occurs; false alarm (or false positive, FP), where weather regime activity is forecasted but no weather regime activity occurs; and correct rejection (or true negative, TN), where no weather regime activity is forecasted and no weather regime activity occurs. These categories form the basis for several verification metrics, which are detailed in Table 2. The key metrics include base rate (BR), representing the overall occurrence of a weather regime activity; forecast rate (FR), which is the proportion of forecasts predicting a weather regime activity; hit rate (HR, also known as recall or accuracy), which is the proportion of actual weather regime activities correctly predicted; false alarm rate (FAR), the proportion of incorrect predictions of weather regime activities; Peirce skill score (PSS), which is the difference between the HR and the FAR; precision, representing the proportion of true-positive forecasts out of all positive predictions; and F_1 score, the harmonic mean of precision and HR, balancing both metrics.

TABLE 1 General classification terms of the traditional 2 × 2 contingency table (text in cells) and probabilistic equations for the occurrence of weather regime activities (equations in cells).

	Weather regime active/inactive	
Forecast	**Yes**	**No**
Yes	Hit/True positive (TP) $E_{11} = \mathbf{p} \cdot \mathbf{k}$	False alarm/False positive (FP) $E_{12} = \mathbf{p} \cdot (\mathbf{1} - \mathbf{k})$
No	Miss/False negative (FN) $E_{21} = (\mathbf{1} - \mathbf{p}) \cdot \mathbf{k}$	Correct rejections/True negative (TN) $E_{22} = (\mathbf{1} - \mathbf{p}) \cdot (\mathbf{1} - \mathbf{k})$

Note: **p** represents a vector of probabilities for a weather regime activity to occur and **k** is a vector of binary indicators whether the weather regime activities have occurred at the given dates (see Section 2.4). The operator "·" denotes the inner product of two vectors, and "**1**" is the unit vector. The table is adapted from Gold *et al.* (2020) and modified.

TABLE 2 Collection of verification metrics originating from the contingency table in Table 1 used to identify windows of opportunity and to validate the skill of a categorical $WRact_{max}$ forecast.

Metric	Formula
Base rate	$BR = \frac{TP+FN}{TP+FN+FP+TN}$
Forecast rate	$FR = \frac{TP+FP}{TP+FN+FP+TN}$
Hit rate (recall, accuracy)	$HR = \frac{TP}{TP+FN}$
False alarm rate	$FAR = \frac{FP}{FP+TN}$
Peirce skill score	$PSS = HR - FAR$
Precision	$Prec = \frac{TP}{TP+FP}$
F_1 score	$F_1 = 2 \times \frac{Prec \times HR}{Prec+HR}$

Abbreviation: FN: false negative; FP: false positive; TN: true negative; TP: true positive.

An alternative to the traditional contingency table is the probabilistic contingency table (Table 1, equations), which incorporates probabilistic forecasts (Gold *et al.*, 2020). In this framework, the vector $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ represents the forecast probabilities for an event (weather regime active) occurring, where n is the total number of dates in the sample. Each p_i is the mean of the binary predictions from all ensemble members for date i, with M being the number of ensemble members; in this study, $M = 11$. Specifically, each ensemble member m provides a binary prediction $p_{i,m}$ (which is either 0 or 1), and the overall probability for date i is calculated as

$$p_i = \frac{1}{M} \sum_{m=1}^{M} p_{i,m}. \quad (1)$$

Similarly, the vector $\mathbf{k} = (k_1, k_2, \ldots, k_n)$ represents the actual binary outcomes, where $k_i = 1$ if the weather regime activity occurs and $k_i = 0$ if it does not occur. The scalar product of the forecast probabilities p_i and the binary outcomes k_i generates the probabilistic outcomes, which are analogous to the traditional contingency table.

In addition to the verification metrics derived from the contingency tables, we use the mean-squared error

(MSE) to assess the quality of the forecasts (verifying $\mathrm{WRact}_{\mathrm{agg}}$). The MSE quantifies the average squared differences between the predicted values and the observed values across all data points:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{2}$$

where y_i is the forecasted value and $\hat{y}_i$ is the observed value for the ith data point. To compare the MSE of a forecast with that of a reference forecast, we compute the MSE score (MSESS):

$$\mathrm{MSESS} = 1 - \frac{\mathrm{MSE}_{\mathrm{forecast}}}{\mathrm{MSE}_{\mathrm{reference}}}. \tag{3}$$

Here, the MSESS represents the normalised performance of the forecast relative to the reference, which in this case is the NWP forecast. A higher MSESS indicates better forecast skill compared with the reference forecast.

2.5 | Windows of opportunity

Windows of opportunity integrate knowledge of the atmospheric state at the time of initialisation with the likelihood of a particular event (specifically, weather regime activity) occurring after a set time period. The atmospheric state at initialisation time can be characterised by atmospheric indicators, in this work, the MJO and the SPV. Following Specq and Batté (2022), we categorise each 3-week forecast by the prevalent state of the atmospheric indicators at initialisation:

MJO: Unlike the commonly used real-time multivariate MJO index (RMM) by Wheeler and Hendon (2004), we adopt the outgoing long-wave radiation (OLR) MJO index (OMI) (NOAA, 2025). The OMI is derived solely from OLR anomalies, without incorporating zonal winds at 850 hPa and 200 hPa. Similar to the RMM index, it is based on a pair of empirical orthogonal functions computed from near-equatorial (15° N–15° S) satellite-derived OLR data. The temporal evolution of the MJO signal by the OMI is smoother (measured by the variability of the Euclidean distance between two consecutive dates) compared with the RMM (not shown). The MJO's position and strength can be identified in a two-dimensional phase space and split into nine categories: eight active phases plus an inactive phase. For the MJO, the inactive phase serves as the reference state.

SPV: The strength of the SPV is quantified using the zonal-mean zonal wind at 10 hPa and 60° N, referred to as the SPV index. Daily anomalies of this index are calculated relative to a 91-day running-mean climatology over the period 1979–2019. Based on these anomalies during the extended winter season, the SPV index is categorised into weak, neutral, and strong states. The categories are defined using terciles of the anomaly distribution, with the lower tercile corresponding to weak, the upper tercile to strong, and the middle tercile to neutral SPV states. The neutral phase serves as the reference state. This anomaly-based classification ensures an approximately equal distribution among the three categories while accounting for seasonality (see also Supporting Information Figure S1). To highlight more extreme states of the SPV, we additionally analyse the strongest 10% and weakest 10% SPV states.

For each atmospheric state, we then calculate verification metrics (BR, HR, PSS) for $\mathrm{WRact}_{\mathrm{mean}}$, as defined in Tables 1 and 2. Once the verification metrics are calculated, we categorise windows of opportunity into two main types: climatological and model windows of opportunity, with the model windows of opportunity further divided into three subcategories.

2.5.1 | Climatological windows of opportunity

A climatological window of opportunity is based on the historical relationship between the atmospheric state at initialisation time and the likelihood of a weather regime to occur 3 weeks later as derived from ERA5 reanalysis. A certain atmospheric state is considered to represent a climatological window of opportunity if the frequency of weather regime occurrence (referred to as the BR) is higher after a specific state than after the reference state (top row in Table 3). By identifying these climatological windows of opportunity, a forecaster can estimate the likelihood of a weather regime activity occurring simply by relying on past observational data corresponding to a given atmospheric condition at initialisation.

2.5.2 | Model windows of opportunity

In contrast, a model window of opportunity takes into account both the BR and the ability of the forecasting system (here the ECMWF reforecasts) to detect and benefit from the atmospheric signal. For a model window of opportunity to be useful, it might not be enough for the

system to simply identify a higher BR and HR. The system should also demonstrate a reliable forecast quality, which is where the PSS comes into play. The PSS assesses the difference between the HR and FAR, allowing us to distinguish between forecasts that are genuinely improved and those that may have increased HRs at the cost of more false alarms. Within the model window of opportunity category, there are three subtypes (see also Table 3):

Type 1: This is characterised by an increase in both the BR and the HR, but without an improvement in the PSS due to an excessive FAR. In this case, though the forecast may indicate a higher likelihood of the event, it comes with a higher number of false alarms, reducing overall forecast reliability.

Type 2: This shows improvements in the BR, HR, and the PSS. This subtype indicates a higher quality forecast where the HR improves while minimising false alarms, resulting in a more accurate and reliable forecast.

Type 3: This exhibits a positive anomaly in both the HR and the PSS compared with the neutral phase, even if this improvement cannot be explained by a climatological signal, due to no increase in the BR. These phases are still considered valuable because they reflect an improvement in forecast performance over what would be expected without the active phase.

2.5.3 | Testing for confidence and significance

To assess the confidence of the weather regime activity BRs and FRs for each atmospheric state category, we apply a bootstrap approach. Specifically, the dates associated with each category are resampled 1000 times with replacement. Alongside the non-resampled rates, we report the 5th and 95th percentiles as confidence bounds.

We also evaluate whether differences in BR, FR, and PSS between individual states and the reference state, required for the computation of model windows of opportunity, are statistically significant. For each verification metric and state category, we derive the 90% confidence intervals (5th–95th percentiles) from 1,000 bootstrap resamples with replacement. The difference to the reference state is considered statistically significant if the corresponding metric for the reference state lies outside this interval.

2.6 | Atmospheric-conditioned climatology

The existence of windows of opportunity under certain atmospheric conditions naturally motivates the computation of an ERA5 climatology conditioned on the atmospheric state at initialisation time, which we refer to as an atmospheric-conditioned climatology. Instead of averaging over all historical data (an "unconditioned climatology"), we calculate averages separately for different atmospheric situations (an "atmospheric-conditioned climatology"). Since the relationship between the atmospheric conditioning variables and weather regimes is sensitive to seasonality, all climatologies are computed using a 91-day sliding window. For each atmospheric state, we select all historical dates within the sliding window that meet the state condition at initialisation (see Supporting Information Figure S2, left column).

This approach gives us the climatological average of the variable $\mathrm{WRact}_{\mathrm{agg}}$ at a given time lag, but specifically under the chosen atmospheric conditions (e.g., a certain MJO phase or SPV state), rather than over all conditions mixed together. These climatologies offer insight into how frequently weather regime activity occurs under different atmospheric conditions (see Section 3.3).

2.7 | Atmospheric-conditioned climatology forecast

Building on the atmospheric-conditioned climatology, we construct an atmospheric-conditioned

TABLE 3 Decision logic for climatological and model windows of opportunity based on anomalies (denoted with a prime) to the reference state of the base rate (BR′), hit rate (HR′), and Peirce skill score (PSS′).

Window type	BR′ > 0	HR′ > 0	PSS′ > 0	Practical meaning
Climatological	Yes	N/A	N/A	Higher climatological likelihood
Model Type 1	Yes	Yes	No	Model predicts more events correct but with many false alarms
Model Type 2	Yes	Yes	Yes	Best: increased likelihood and model skill
Model Type 3	No	Yes	Yes	Model gains skill predicting events even though climatological likelihood not higher

climatological forecast by selecting the corresponding atmospheric-conditioned climatology for each initialisation date based on the prevailing atmospheric condition. Although this simple, condition-based selection is straightforward and effective, it can lead to abrupt changes in forecasted $WRact_{agg}$ when transitioning between atmospheric states. To address this issue, we introduce a sliding forecast, which smooths these transitions. This approach differs depending on whether the conditioning variable is two-dimensional (e.g., the MJO) or one-dimensional (e.g., in our case the SPV index). In the following, we detail the method for the MJO and highlight modifications required for one-dimensional variables (see Supporting Information Figure S2, right column for a schematic overview).

The sliding atmospheric-conditioned climatological forecast for the MJO is generated as follows (illustrated in Figure 2). First, all historical OMI1 and OMI2 pairs (the first two principal components of the OLR anomaly empirical orthogonal function) within a 91-day sliding climatological window (grey dots) centred on the day of year of the current initialisation date (red diamond) are collected. Rather than selecting only the single MJO phase for the current date, we include all historical dates whose OMI1–OMI2 pairs fall within a 45° arc (blue dots) around the current OMI pair in the phase space. Dates with an OMI amplitude less than 1 (inner circle), representing the MJO inactive phase, are excluded. Additionally, to ensure independence from recent events all events in a windows of ±45 days (orange dots) are removed from consideration. Finally, the mean $WRact_{agg}$ is computed from the remaining historical dates and used as the MJO-conditioned climatological forecast for the current initialisation.

For one-dimensional atmospheric variables, such as the SPV index, the sliding forecast approach is adjusted accordingly. Instead of using a 45° arc in a two-dimensional phase space, we select the 5% of historical values closest to the current SPV value from within the 91-day sliding climatological window, excluding dates from the current seasonal cycle. The mean $WRact_{agg}$ from these selected historical dates is then used to produce the SPV-conditioned climatological forecast.

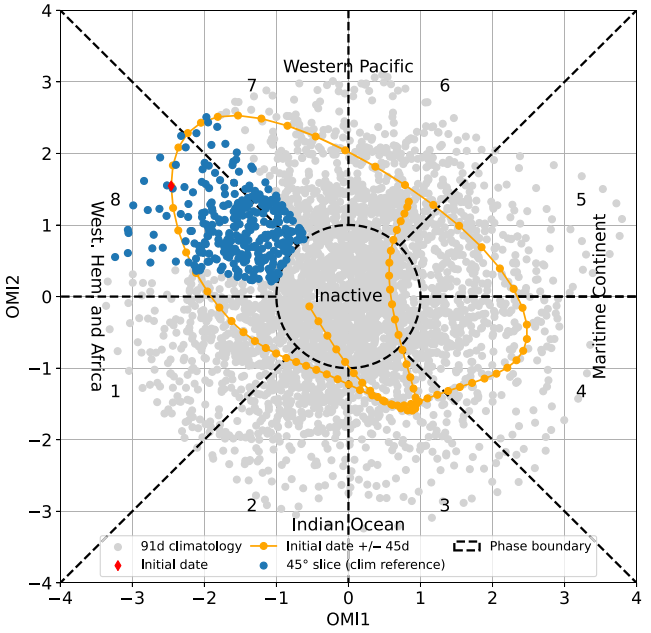

FIGURE 2 Phase diagram of the outgoing long-wave radiation Madden–Julian Oscillation (MJO) index. The first and second principal components derived from empirical orthogonal function analysis of filtered outgoing long-wave radiation data are given on the *x*- and *y*-axes. Further, the phases of the MJO are indicated with dashed lines and labelling. For visualisation of how the atmospheric-conditioned climatological forecast works, one initial date, February 10, 2018, is marked with a red diamond and all values in the 45° arc are used for the MJO-conditioned climatological forecast of that initialisation date (blue dots). Values in the ±45-day window around the initial date (orange dots) are excluded from the data used for the computation. With grey dots, all other data points of the 91-day running climatology are indicated. For a visualisation of the OLR MJO index (OMI) phase space similar to the real-time multivariate MJO index (RMM1/2), OMI1 is multiplied by −1 and OMI1 and OMI2 are exchanged with each other. [Colour figure can be viewed at wileyonlinelibrary.com]

2.8 | Neural networks

To combine information from (atmospheric-conditioned) climatological forecasts, NWP forecasts, and recent $WRact_{agg}$, we employ a statistical–dynamical approach represented with fully connected dense NNs. With these networks, we aim to identify complex relationships between predictors and enhance the predictability of $WRact_{agg}$, particularly for strong weather regime events in forecast week 3. For readers unfamiliar with machine-learning techniques and terminology, we recommend the machine-learning tutorials by Chase *et al.* (2022, 2023). A fully connected dense NN set-up is selected with a relative simple architecture of two hidden layers with 64 and 16 nodes, each using a rectified linear unit activation function, and a dropout layer with a rate of 0.2 is inserted between the hidden layers to reduce overfitting. The output layer has a single node with a sigmoid activation function, producing a probability-like output representing the $WRact_{agg}$ for a specific weather regime. The models are trained separately for each weather regime and lead time. The final architecture and hyperparameters were selected after testing different network depths, loss functions, and output configurations (not shown). Overall, model performance was influenced more strongly by the choice of predictors than by the hyperparameters.

Owing to computational expenses and data availability, the hyperparameter tuning was not fully exploited. Each predictor time series is individually normalised using min–max scaling based on the training dataset, resulting in values ranging from 0 to 1 within the training set, whereas values in the test set may exceed this range.

The network is trained for up to 50 epochs, but early stopping is applied to prevent overfitting. Early stopping monitors the validation loss with a patience of 10 epochs, halting training if the validation loss does not improve. Additionally, a learning-rate scheduler reduces the learning rate, which starts at 0.001, by a factor of 0.5 after five epochs without improvement, with a minimum learning rate of 10^{-6}. During training, the model uses a batch size of 32 and a validation split of 20%, without shuffling the data to preserve temporal dependencies. To robustly assess model performance, we apply fourfold cross-validation. The 1,720 data points from the extended winter period (November–March, NDJFM) are split into four consecutive subsets of 430 points each. Each fold trains on 1,290 data points (data includes the validation split of 20%) and tests on the remaining 430. To ensure results are not dependent on a single random initialisation parameter setting, 10 fully independent models with identical configurations though different random initialisation parameters are trained simultaneously, and their ensemble mean is used as the final prediction. The model predicts the $WRact_{agg}$ for a single weather regime and one specific lead time, such as GL activity for forecast week 3.

To improve model performance while keeping model complexity low, we apply a stepwise feature selection procedure, restricting ourselves to the most informative predictors and keeping the signal-to-noise ratio low. The stepwise selection begins by training the model using each predictor individually and recording the MSE from cross-validation. In subsequent steps, the best-performing predictor from the previous step is retained and additional predictors are tested one by one. The set of predictors with the lowest cross-validated MSE is selected. This procedure is performed separately for each lead time and weather regime, tailoring the predictor set to the specific forecast horizon and regime dynamics. The final models are trained from scratch, though using these preselected predictors which could be considered as a minor information leakage. However, owing to the data availability, selecting the features on a fully separate set of training data is not feasible and we consider the leakage through the predictor selection as small (without providing a quantitative measure for our statement).

To ease interpretability, we group all available predictors into four distinct subcategories. Climatology predictors include a 91-day mean $WRact_{agg}$ climatology, day-of-year, and atmospheric-conditioned climatological forecasts based on large-scale drivers such as the MJO and SPV. Atmospheric state predictors describe the current state of the atmosphere, including indices representing large-scale circulation patterns such as the MJO and SPV, as well as sea-surface temperature anomalies in the North Atlantic. NWP predictors capture forecast-based $WRact_{agg}$ indicators, including trends and variability across different forecast weeks. Recent weather regime predictors represent recent observed $WRact_{agg}$ and IWR from reanalysis data. A comprehensive description of these predictors is provided in Supporting Information Tables S1 and S2.

Using these predictor subcategories, we train three separate NNs with distinct predictor pools to explore the contributions of different predictor types (Table 4). The climatological NN, NN_{noNWP}, uses no data of the NWP model. As it does not rely on forecast data, this model is trained on a larger dataset, consisting of daily data from 1979 to 2020, though it is evaluated on the same test set as the other models for fair comparison. The NWP and weather regime NN, NN_{NWP+WR}, uses only NWP and weather-regime-based predictors, such as NWP weather regime forecast evolutions and recent past $WRact_{agg}$ and IWR from ERA. This model is trained exclusively on the reforecast period. Finally, the all-predictors NN, NN_{all}, combines all available predictors, integrating (atmospheric-conditioned) climatology, NWP, and recent past weather regime information. This model provides insight into the added value of integrating all sources of predictability.

3 | RESULTS

We approach our investigation of windows of opportunity in subseasonal weather regime forecasting with an analysis of the modulation of occurrence and forecast quality for GL by different states of the MJO and SPV (Sections 3.1 and 3.2), chosen because GL exhibits strong links to both the MJO and SPV in forecasting weeks 1–2

TABLE 4 Overview of predictor pools used as input predictors for the three neural network (NN) configurations: NN_{all} (all predictors), NN_{NWP+WR} (numerical weather prediction [NWP] and weather regime [WR] predictors), and NN_{noNWP} (excluding NWP predictors). An "X" indicates inclusion of the respective predictor pool.

Predictor pools	NN_{all}	NN_{NWP+WR}	NN_{noNWP}
Climatology	X		X
Atmospheric state	X		X
NWP	X	X	
Recent WR	X	X	X

and it is highly relevant for European winter impacts such as cold Dunkelflauten. Next, we utilise climatological knowledge of the atmospheric state to produce an atmospheric-conditioned climatological forecast for the occurrence of GL (Section 3.3). Finally, we integrate information from NWP regarding $\mathrm{WRact}_{\mathrm{agg}}$ with knowledge about the atmospheric state at initialisation, employing a fully connected dense NN to enhance activity forecasts and expand our analysis on other weather regimes (Section 3.4). The research presented here focuses exclusively on the extended winter period (NDJFM) and a forecast lead time of 3 weeks.

3.1 | Climatological windows of opportunity

In our setting, a window of opportunity indicates whether $\mathrm{WRact}_{\mathrm{mean}}$ is more likely to occur or to be forecasted 3 weeks after the initialisation date, conditioned on the state of the MJO phase or SPV state at forecast initialisation time. To identify these opportunities, we analyse the BR of $\mathrm{WRact}_{\mathrm{mean}}$ following a given phase, along with the HR and PSS (see Section 2.5).

As a preliminary step, we compare the BR and FR of weekly GL activity 3 weeks after specific states of the MJO and SPV (Figure 3). For an analysis across all forecasting weeks, see Supporting Information Figure S3. On average, throughout the analysis period, the BR and the FR of $\mathrm{WRact}_{\mathrm{mean,GL}}$ are approximately 15%, with the FR slightly lower than the BR, indicating a slight negative bias in $\mathrm{WRact}_{\mathrm{mean,GL}}$ and, thus, in GL frequency, as found earlier (e.g., Osman *et al.*, 2023). An analysis conditioned on the MJO phases reveals that, during the inactive MJO phase, the BR and FR align closely, both at approximately 14%. The BR is generally higher for MJO phases 5–8 and 1, whereas it is lower for phases 2–4. Therefore, phases 5–8 and 1 can be considered as climatological windows of opportunity (phases 7, 8 and 1 are identified as significant for GL, according to the bootstrap test; see Section 2.5 and the comparison of the confidence intervals to the BR of the inactive MJO phase in Figure 3). Notably, FRs are consistently closer to the climatological BR and the inactive phase than the respective BRs. This discrepancy may be attributed to the 3-week forecast lead time, as subseasonal forecasts tend to regress toward climatological values. Interestingly, the FRs following strong and weak SPV states (particularly those following within the strongest and weakest 10%) deviate more strongly from climatology than the corresponding BRs (Figure 3, right side). A more detailed discussion and possible explanation of this behaviour are provided in Section 3.2. The BR following a weak (weakest 10%) SPV state is 5.5 (13.0) percentage points higher, whereas the BR following a strong (strongest 10%) SPV state is 3.7 (6.6) percentage points lower than for the neutral SPV state. These results align with previous research conducted on the influence of the SPV on forecast skill for the European region (e.g., Tripathi *et al.*, 2015a, 2015b; Scaife *et al.*, 2016; King *et al.*, 2019; Domeisen *et al.*, 2020b). The positive BR anomaly of $\mathrm{WRact}_{\mathrm{mean,GL}}$ following weak (and weakest 10%) SPV states indicates that these periods represent a climatological window of opportunity. Although we do not consider these events further, one could also consider the strong (and strongest 10%) SPV state, as well as the MJO phases 2–4, as a climatological window of opportunity but for a reduced occurrence of GL, which is also captured by the FR.

Overall, Figure 3 suggests that, though the forecasted frequency of weather regime activity aligns with observed occurrences, discrepancies remain in capturing the observed $\mathrm{WRact}_{\mathrm{mean,GL}}$ 3 weeks after a specific atmospheric state. Large differences exist particularly when separating by MJO phases at initialisation, with the mean frequency of $\mathrm{WRact}_{\mathrm{mean,GL}}$ ranging from 6% following MJO phase 3 to 28% following MJO phase 8. The absolute values of the BRs and FRs should be treated with care owing to the data size, hence the bootstrapping approach. The sample size for individual MJO phases, where an NWP reforecast is initialised, is small (1720 reforecasts considered across 21 winter seasons; sample size per phase is indicated at the bottom of Figure 3).

MJO phases 7, 8, and 1 as well as the weak (and weakest 10%) SPV can be considered as climatological windows of opportunity, with the MJO phases having significantly different BRs compared with their neutral phases, marking them as promising phases for further analysis to identify whether these climatological signals also translate into model windows of opportunity.

3.2 | Model windows of opportunity

Following the approach described in Section 2.5, we now analyse the model windows of opportunity by computing the BR, HR, and PSS for each state of the MJO and SPV. We then compute rate and skill score anomalies (denoted with a prime) relative to the inactive MJO and neutral SPV, respectively (Figure 4): $R'_{\mathrm{state}} = R_{\mathrm{state}} - R_{\mathrm{ref}}$, where $R \in$ BR, HR, PSS and "ref" indicates either the inactive or neutral phase.

The HR′ and PSS′ values closely align for almost all MJO phases, suggesting that the FARs in these sub-selections are similar to those of the neutral phases (not shown). In contrast, the weak (and weakest 10%) SPV state has considerably smaller PSS′ than HR′, indicating that the FAR is greater than for the neutral state. As shown

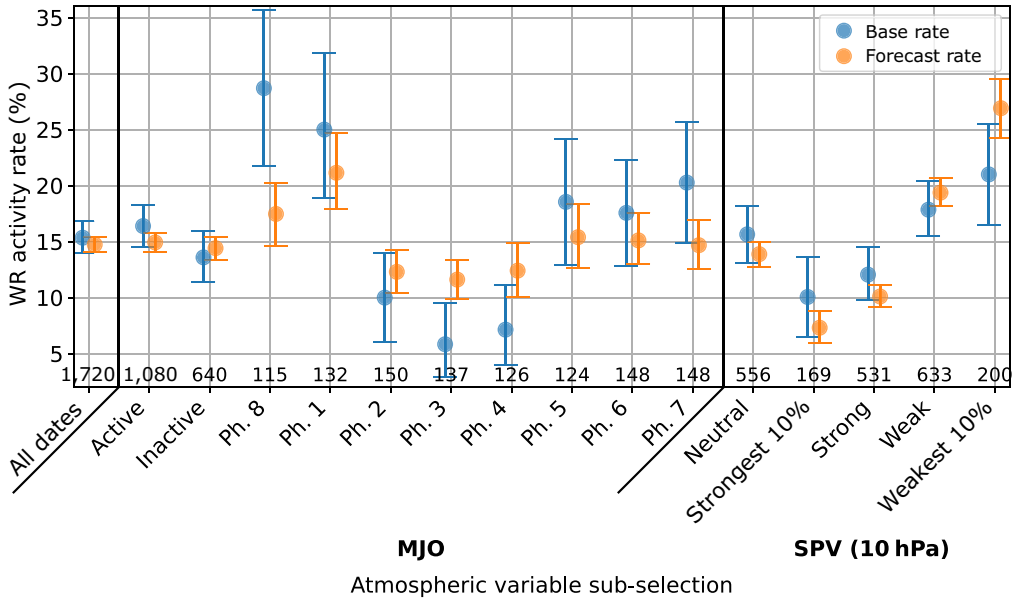

FIGURE 3 Base rate (blue dots) and forecast rate (orange dots) for weekly mean Greenland Blocking (GL) activity ($WRact_{mean,GL}$) in the extended winter period (November–March, NDJFM). These rates are split into subsets based on phases of the Madden–Julian Oscillation (MJO) or stratospheric polar vortex (SPV) with a lead time of 3 weeks. The size of the subset is indicated at the bottom of the respective phase column. Slight deviations from an equal distribution among strong, neutral, and weak SPV states arise because the thresholds are defined based on the climatological period (NDJFM 1979–2019, all European Centre for Medium-Range Weather Forecasts Reanalysis v5 dates), whereas the base and forecast rates are computed only for dates with available reforecasts (NDJFM 1999–2020). For each subset, the error bars indicate the 5th to 95th percentiles, computed by resampling each subset 1000 times with replacement. [Colour figure can be viewed at wileyonlinelibrary.com]

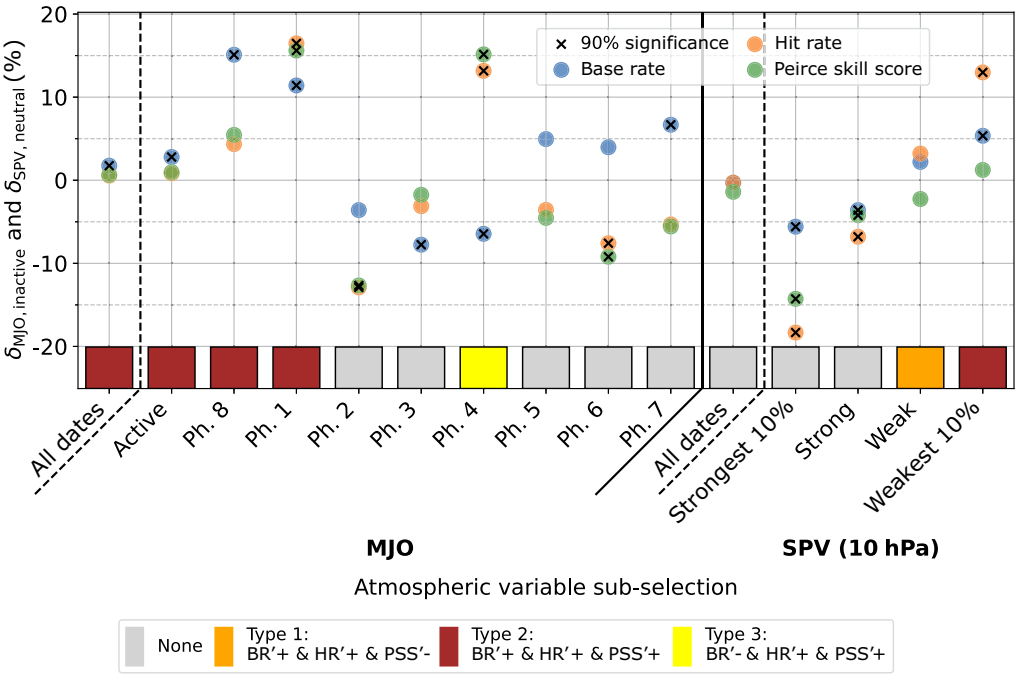

FIGURE 4 Indication of windows of opportunity for Greenland Blocking activity in week 3 by the base rate, hit rate, and Peirce skill score anomalies and colour-coded indication. Similar to Figure 3, the metrics are separated by the state of the Madden–Julian Oscillation (MJO) and the stratospheric polar vortex (SPV). Here, the base rate, hit rate, and Peirce skill score (blue, orange, and green dots, respectively) are anomalies (δ in per cent) with respect to the neutral phases of the respective atmospheric variables. The combination of positive and negative anomalies (denoted with a prime) for the base rate (BR′), hit rate (HR′) and Peirce skill score (PSS′) are responsible for the model window of opportunity type indicated by the coloured boxes at the bottom with its legend below. Further crosses in the dots for anomalies indicate whether the rates in the specific categories are significantly different than their neutral phases with a 90% confidence interval. [Colour figure can be viewed at wileyonlinelibrary.com]

in Figure 3, the BR′ values are positive for MJO phases 5–8 and 1. Furthermore, for phases 8 and 1, both HR′ and PSS′ are positive, indicating that these phases act as a type 2 model window of opportunity, where BR′, HR′, and PSS′ are all positive. Note that only for MJO phase 1 are all three anomalies significantly different to the neutral phase at a 90% confidence interval. MJO phase 4 is classified as a type 3 model window of opportunity for $WRact_{mean,GL}$ with a 3-week lead time. In this case, HR′ and PSS′ are distinctly positive (significant difference to neutral phase), despite the BR′ being negative. This suggests that, although $WRact_{mean,GL}$ is significantly less frequent following phase 4, the forecast accurately predicts these fewer occurrences, with an even lower FAR than in the neutral phase. The signals associated with strong and weak SPV states are less distinct than for MJO phases 8 and 1, likely due to the broad classification into only three categories (strong, neutral, and weak SPV) that encompasses a wide range of individual cases. To address this, we additionally consider the weakest 10% and strongest 10% SPV states. Further, our choice of a scalar definition for the SPV index excludes information about its future temporal evolution. At 3 weeks after a weak SPV, $WRact_{mean,GL}$ is increased (positive BR′), and the forecast also predicts more frequent activity, leading to a higher HR. However, the FAR also increases, which reduces PSS′. In fact, PSS′ is negative, indicating that the HR/FAR ratio is worse than for the neutral SPV state, thereby classifying this as a type 1 model window of opportunity. The BR′, HR′, and PSS′ values for the weakest 10% SPV state are all positive and show stronger signals than for the weak SPV state, marking it as a type 2 model window of opportunity. In contrast, strong SPV states are followed by less frequent $WRact_{mean,GL}$, along with a decrease in HR, PSS, and FAR, ultimately resulting in no identifiable window of opportunity type. The positive and negative BR′ for the weak and strong SPV states, respectively, indicate that there is a link between the SPV and the occurrence of a weather regime activity. The negative PSS′ (and therefore the higher FAR) following the weak SPV state indicates that the model has trouble predicting GL activities correctly following an anomalous state of the SPV. The model shows a GL activity response too often following a weak SPV state, even when it does not realise. A plausible explanation for the excessive GL activity following weak SPV states is an overly strong downward lower stratosphere–troposphere coupling in the model, as noted by Kolstad *et al.* (2020) and Garfinkel *et al.* (2025), which can lead models to overamplify surface responses to weak vortex conditions.

In summary, the frequency of $WRact_{mean,GL}$ (BR) varies across different MJO phases and SPV states. It is particularly high (significantly different to the neutral phase) for MJO phases 8 and 1. The magnitude of the FR is generally too low for MJO phases. MJO phases 8 and 1 emerge as type 2 model windows of opportunity, meaning that not only is the BR increased for $WRact_{mean,GL}$ 3 weeks after these phases, but both the HR and PSS are also improved compared with the neutral phase. All three values are significantly different to the neutral phase for MJO phase 1. Forecasts initialised during a weak SPV state show an improved HR, but the FAR also increases compared with the neutral phase, classifying this as a type 1 model window of opportunity, whereas for the weakest 10% SPV states, the FAR′ is positive, indicating a type 2 model window of opportunity. Finally, MJO phase 4 is a type 3 model window of opportunity owing to better predicting the less frequent GL activity.

3.3 | Atmospheric-conditioned climatological weather regime activity forecasts

The previous section and existing literature (e.g., Beerli & Grams, 2019) suggest that the $WRact_{mean,GL}$ 3 weeks after forecast initialisation is influenced by the prevailing states of the MJO and SPV. To better understand this relationship, an atmospheric-conditioned climatology is introduced with a 91-day sliding window over extended winter. This climatology expresses the likelihood of $WRact_{agg}$ 3 weeks after a given MJO phase or SPV state, demonstrating that phase-specific climatological frequencies of GL occurrence differ significantly from the overall climatology and exhibit substantial intraseasonal variability (Figure 5).

During the extended winter period, the 91-day running mean of the $WRact_{agg,GL}$ fluctuates between 15% and 18% (black lines). When categorised by MJO phases (Figure 5a), notable deviations from this baseline emerge. The most prominent signal appears for MJO phase 8, where $WRact_{agg,GL}$ increases from 12% in December to 30% in March (brown line). Similarly, MJO phase 1 (red line) exhibits an increase from 14% in mid December to 23% in January, maintaining an elevated activity level of approximately 20% until March. For SPV states (Figure 5b, based on the 10 hPa zonal-mean zonal wind at 60° N), the separation is even clearer. The neutral state (grey line) largely follows the overall climatological cycle, whereas conditions following a weak SPV (blue line) show an overall increase in $WRact_{agg,GL}$. This increase becomes even more pronounced after the 10% weakest SPV states (cyan line). Conversely, when a strong SPV (red line) is present, $WRact_{agg,GL}$ decreases relative to the climatological cycle, with an even stronger reduction following the strongest 10% SPV states. The latter is particularly evident in early winter, when the strongest 10% SPV states are rare; and,

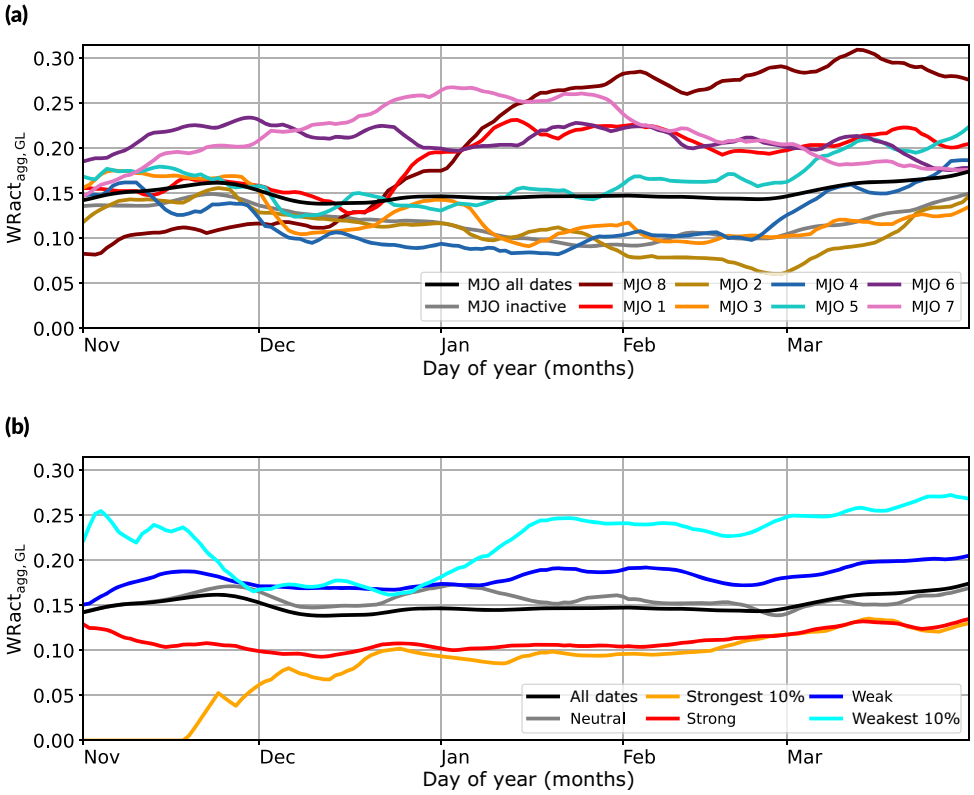

FIGURE 5
Atmospheric-conditioned climatology for WRact$_{agg,GL}$ during the extended winter period (November–March), conditioned on the state of (a) the Madden–Julian Oscillation (MJO) and (b) the stratospheric polar vortex (SPV; 10 hPa, 60° N zonal-mean wind speeds) 3 weeks before the regime occurrence. The *x*-axis indicates the timing of regime activity. The black lines in (a) and (b) indicate the 91-day running climatology without any atmospheric-conditioned sub-selection. [Colour figure can be viewed at wileyonlinelibrary.com]

when they do occur, they are seldom followed by GL activity.

To use the atmospheric-conditioned climatology as a forecast tool, a slight modification is introduced to the computation method. Instead of strictly categorising the atmospheric-conditioned climatology by discrete phases, a sliding window approach is applied. This adjustment ensures a smoother transition between categorical phases, preventing abrupt shifts in activity forecast (see Section 2.7 for further details).

As a proof of concept, we select periods of active GL and analyse atmospheric-conditioned climatological and NWP forecasts leading up to these GL activities. Though this approach may seem selective, as it focuses only on active GL periods, it enables a targeted examination of two key forecasting questions:

1. Do single forecast runs of the NWP model and atmospheric-conditioned climatological forecasts indicate the onset of WRact$_{agg,GL}$ in week 3 and the subsequent evolution of WRact$_{agg,GL}$?
2. Do consecutively initialised week 3 forecasts show an increasing signal leading up to the WRact$_{agg,GL}$?

To address these questions, two mean composite plots are computed (Figure 6).

For both analyses, the actual GL activity onset is defined as the first time when the ERA5 WRact$_{mean,GL}$ reaches 1. For the first question, individual forecasts initialised 3 weeks prior to GL activity onset (or as close as possible, given the availability of NWP reforecasts) are analysed with lead times ranging from 0 to 36 days (Figure 6a). For the second question, only those forecasts at a fixed lead time of 3 weeks (represented by 18 days, the midpoint of week 3) are considered. Thus, instead of a single forecast per event, consecutive forecasts leading up to the GL activity are analysed to track the signal of GL activity over time (Figure 6b). The mean composite of forecasts in Figure 6 is computed centring all events around the GL onset. For the individual events, see Supporting Information Figures S4 and S5; and for a case study example of the prolonged GL activity in the winter of 2009–2010 see Supporting Information Section S1.

The true (ERA5) WRact$_{agg,GL}$ (black line, right *y*-axis) is identical in Figure 6a,b. As expected, there is a clear increase in WRact$_{agg,GL}$ around the onset (vertical black dashed line), peaking 3 days after the onset. Additionally, a secondary local maximum appears around 12 days prior to the GL onset, indicating that, in some of the scenarios, two GL events occur in close succession (e.g., GL event indices 129, 137, 181 in Supporting Information Figure S4). This secondary maximum may also indicate that the state of weather regime activity prior to, at, and after initialisation carries valuable predictive information. Owing to the centring of the data around the activity onset, the mean composite shows actual WRact$_{agg,GL}$ reaching up to 85%,

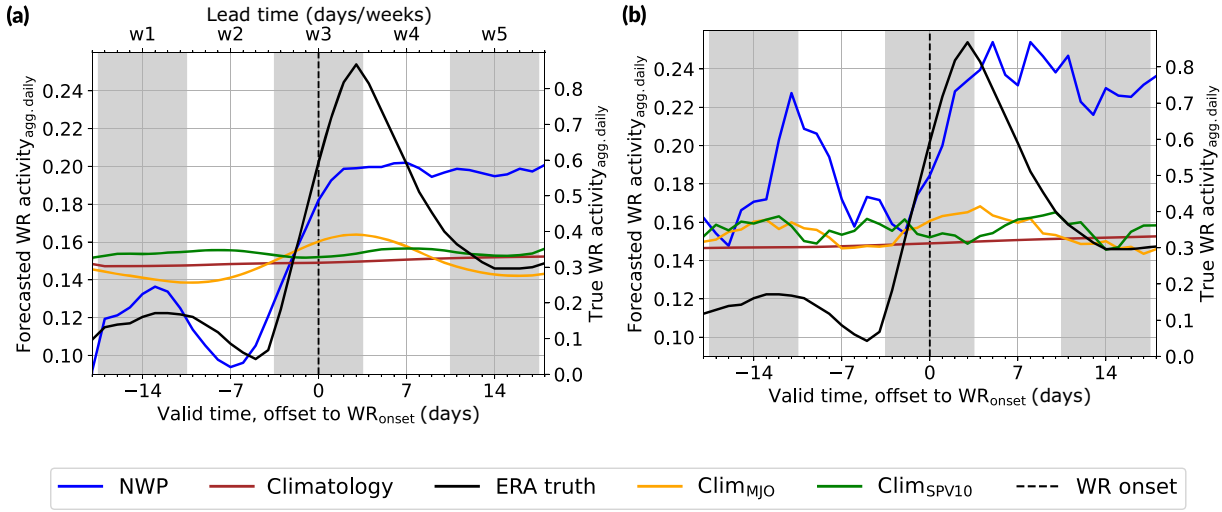

FIGURE 6 Mean composites of forecasted and observed Greenland Blocking (GL) activity centred on observed GL onsets (black dashed line). Forecasted $WRact_{agg}$ (left axis) is shown together with the corresponding European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) values (right axis). (a) Forecasts initialised 3 weeks (18 days) before GL onsets, illustrating the evolution of forecast lead times up to and beyond the onset. (b) Forecasts with a fixed 3-week (18-day) lead time, evaluated at valid times within ±18 days of the onset, showing the evolution of consecutive forecasts around the onset. Numerical weather prediction (NWP) forecasts are compared with climatological and Madden–Julian Oscillation (MJO)/stratospheric polar vortex (SPV)-conditioned baselines. All composites are restricted to the extended-winter period 1999–2020, when NWP forecasts are available. A GL onset is defined as the time when $WRact_{mean,GL}$ transitions from 0 to 1. WR: weather regime. [Colour figure can be viewed at wileyonlinelibrary.com]

whereas forecasts only reach up to 25% of $WRact_{agg,GL}$ in the mean composite. Therefore, the $WRact_{agg,GL}$ forecast values are displayed on a separate *y*-axis for better visualisation (left *y*-axis).

The unconditional climatology (brown line) serves as a reference, showing no distinct signal apart from the seasonal BR of approximately 15% $WRact_{agg,GL}$. Comparing the MJO- and SPV-conditioned climatological forecasts (orange and green lines, respectively) to the NWP model (blue line), the NWP forecast exhibits the strongest $WRact_{agg,GL}$ signal. This result is not surprising, as the NWP model incorporates full dynamical atmospheric information, whereas the atmospheric-conditioned climatological forecasts rely solely on the MJO phase or SPV state at initialisation. Thus, rather than comparing absolute values, the temporal changes of the forecast signals relative to the unconditional climatology are the primary focus.

For forecasts initialised 3 weeks prior to GL onsets (Figure 6a), the NWP forecast shows a strong increase in $WRact_{agg,GL}$ beginning approximately 7 days before onset, mirroring the observed activity trend. An additional increase appears at earlier lead times, specifically 11–17 days before onset, aligning with the secondary maximum observed in the ERA5 $WRact_{agg,GL}$. At longer lead times (7–18 days after onset), the NWP forecast suggests a prolonged high $WRact_{agg,GL}$, even when the observed activity declines. This behaviour is likely an artefact caused by forecast timing errors, where some forecasts predict $WRact_{agg,GL}$ too late (e.g., GL indices 142, 166, 202 in Supporting Information Figure S4). As a result, in the mean composite, $WRact_{agg,GL}$ appears misleadingly prolonged in the forecast.

The MJO-conditioned climatological forecast impressively captures the key characteristics of $WRact_{agg,GL}$ evolution. At earlier lead times (up to 5 days before onset), $WRact_{agg,GL}$ is lower than the unconditional climatology. However, from 4 days before to 10 days after onset, forecasted $WRact_{agg,GL}$ exceeds the unconditioned climatology, indicating an increased likelihood of $WRact_{agg,GL}$. This result reinforces previous findings that MJO phases modulate $WRact_{agg,GL}$ 2–4 weeks later, validating the teleconnection mechanism.

In contrast, the SPV-conditioned climatological forecast produces a less distinct signal. Forecasted $WRact_{agg,GL}$ remains higher than the unconditional climatology across all lead times, yet the highest values appear at early lead times and 5–10 days following the onset, with minimal values around the actual weather regime onset. One possible explanation is that the one-dimensional SPV index lacks the complexity of the two-dimensional MJO index (amplitude and phase angle) in our analysis. Previous studies (e.g., Domeisen *et al.*, 2020c) suggest that sudden stratospheric warming events (indicated by low westerly wind speeds at 60° N and 10 hPa or even a reversal of the wind direction) can trigger concurrent GL activity, but not all GL activities are linked to sudden stratospheric warming events. Since the SPV-conditioned climatological forecast

does not account for changes in the index over time, it merges scenarios where the atmosphere is both approaching and departing from a sudden stratospheric warming event, leading to a diffuse climatological signal. Incorporating temporal changes in the SPV index could improve the clarity of the forecast signal and better represent the dependence between the SPV index at initialisation and the $WRact_{agg,GL}$ with a lag in time. To account for this, we introduced a two-dimensional index that incorporates the SPV change over the past 7 days. However, integrating this into the atmospheric-conditioned climatological forecast did not yield noticeable improvements (and therefore is not shown).

When analysing consecutive forecasts with a fixed 3-week lead time (Figure 6b), similar signals emerge, though with distinct differences. The NWP forecast still predicts GL activity onsets, indicated with an activity increase starting for valid times 2 days prior to the onset. In contrast to the single forecast runs, the $WRact_{agg,GL}$ forecasts prior to the weather regime onset are already higher than climatology. The delayed signal increase compared with ERA5 (2 days rather than 5 days prior to the onset) and the later maximum of the $WRact_{agg,GL}$ (5 days rather than 3 days after the onset) suggest a systematic delay in predicted onset timing. Additionally, the previously active GL activity is still visible in the forecasts by the NWP (7–14 days prior to the GL onset). This could also indicate that the NWP model favours the persistence of a GL activity rather than the transition into another weather regime. The atmospheric-conditioned climatological forecasts continue to provide meaningful signals. The MJO-conditioned climatological forecast exhibits a clear increase relative to the unconditional climatology, beginning 3 days before the onset and persisting for up to 10 days after the onset. The SPV-conditioned climatological forecast, though more diffuse, shows an increase multiple days prior to the weather regime onset and from 3 to 10 days after the onset. The first days of the GL onset are not well captured by the SPV-conditioned climatological forecasts. For both set-ups, the single forecast runs and the consecutive forecasts around the GL onset, the atmospheric-conditioned climatological forecasts are above the unconditioned climatological values. However, the extent of the differences is clearer for the MJO-conditioned climatological forecast than for the SPV-conditioned climatological forecast. The less distinct signal for SPV-conditioned climatological forecasts also holds when analysing the same SPV index, but computed on the 100 hPa rather than 10 hPa level (not shown). However, these results could be due to the choice of the computation of the forecasts, and a more complex method considering, for example, time evolutions to indicate sudden stratospheric warming events could potentially better indicate GL activity.

These results demonstrate that atmospheric-conditioned climatologies offer a valuable framework for forecasting $WRact_{agg,GL}$. The MJO-conditioned climatological forecast exhibits a robust increase in $WRact_{agg,GL}$, aligning well with observed $WRact_{agg,GL}$ and confirming the role of the MJO in modulating $WRact_{agg,GL}$ 2–4 weeks later. The SPV-conditioned climatological forecast also indicates a weak modulation of GL activity, though it is more diffuse and less clear to interpret. Though the NWP model predicts $WRact_{agg,GL}$ well, it exhibits a systematic delay of predicting the onset. Overall, these findings support the use of atmospheric-conditioned climatological forecasts as a complementary tool to dynamical NWP models. However, further refinements, particularly incorporating temporal changes in the SPV index, could enhance predictive accuracy. It is important to emphasise that these findings provide only a partial interpretation of the forecast signals, as the analysis is limited to mean composites centred around actual ERA5 $WRact_{agg,GL}$ onsets. The forecasts within this subset exhibit inherent variability, and the composites do not account for scenarios where, for example, the NWP model predicts $WRact_{agg,GL}$ that does not verify.

3.4 | Statistical–dynamical approach

Given the promising results observed in $WRact_{agg,GL}$ forecasts conditioned on the MJO and SPV, as well as the performance of the NWP model, we explore whether these forecasts (along with additional atmospheric variables) can be combined to enhance the NWP model's skill in predicting $WRact_{agg}$ 3 weeks in advance. To achieve this, we employ statistical–dynamical models using fully connected NNs (as described in Section 2.8) to predict $WRact_{agg}$ at week 3.

To systematically assess the impact of different information sources, we construct three NNs based on distinct predictor pools (see Supporting Information Tables S1 and S2 for a complete list of predictors). The "NWP and weather regime" NN (NN_{NWP+WR}) combines information from the NWP model with recent $WRact_{agg}$. The "no NWP" NN (NN_{noNWP}) incorporates all available predictors except those derived from the NWP model. The all-predictors NN (NN_{all}) utilises the full set of available predictors, including climatological forecasts, atmospheric indicators at initialisation, NWP-based weather regime forecasts and temporal changes, and $WRact_{agg}$ history.

Each NN undergoes a stepwise feature selection process (see Section 2.8) to identify the most relevant predictors. Once the predictors yielding the highest skill are determined, models are retrained using *k*-fold cross-validation, ensuring a robust evaluation across multiple forecast instances.

3.4.1 | Stepwise predictor selection

Our analysis begins with the predictor selection for the NN_{all} in predicting $WRact_{agg,GL}$ for GL, followed by a comparison of the predictor order across all weather regimes. The most influential predictor for $WRact_{agg,GL}$ in NN_{all} is the NWP model's week 3 mean IWR forecast, followed closely but not chosen by its $WRact_{agg}$ forecast of the NWP model for the same week, which is inherently derived from the IWR forecast (Supporting Information Figure S6). The second most important predictor is, somewhat unexpectedly, the 10 hPa SPV-conditioned climatological forecast. Though this may appear surprising given the weaker SPV signal compared with the MJO discussed in the previous section (Figure 6), it is consistent with prior findings that the SPV can serve as an indicator of GL activity. This suggests that, although the SPV-conditioned climatological forecast provides a less distinct signal than the MJO, its combination with the NWP forecast adds more predictive value than combining the NWP forecast with MJO-related predictors. The third strongest predictor is the NAO index at initialisation. Given the strong correlation between GL and NAO−, this likely reflects the persistence of GL activity. Additional key predictors include the QBO-, MJO-, and oceanic Niño index-conditioned climatological forecasts, the Pacific–North American index, as well as the trend of the IWR forecast for GL in forecasting weeks 1 and 2.

Extending the predictor selection analysis to all seven weather regimes (Figure 7) reveals that, except for GL and AR, the NWP forecast of the $WRact_{agg}$ (target variable) remains the dominant predictor. NWP-derived predictors designed to mimic a human forecaster's analysis (such as the temporal change of the IWR across consecutive forecasts or its evolution in week 2) play a minor role in the NNs. Similar to GL, the ZO network (representing the NAO+ phase, the counterpart to GL/NAO−) includes information from the MJO, QBO, and SPV, though differences in the actual predictors show. The ZO network learns from the phase of the MJO and the 100 hPa SPV index rather than the MJO and 10 hPa SPV-conditioned climatological forecast.

An unexpected but recurrent predictor is the Antarctic Oscillation (AAO). Its index ranks second for EuBL and ScBL, fourth for AT, highlighting its potential link to $WRact_{agg}$. The AAO has been associated with its

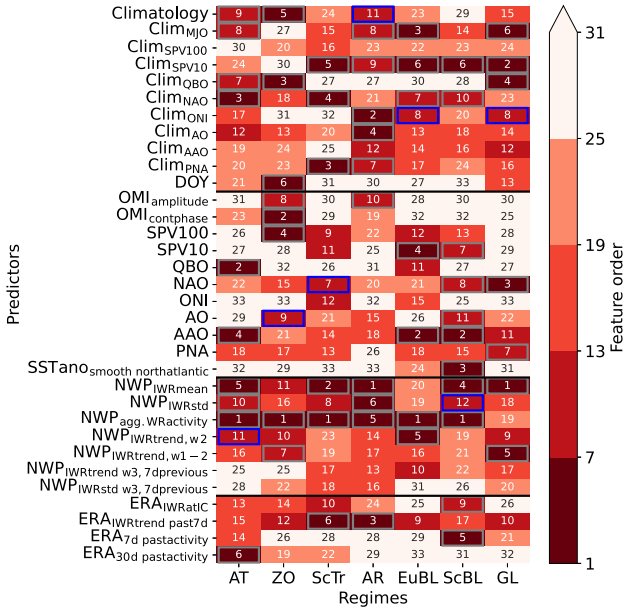

FIGURE 7 Summary of the predictor order for all-predictor neural networks (NN_{all}). The predictor order is shown for all seven weather regimes. Blue boxes indicate the last predictor which is considered due to the best mean-squared error for the predictor combination of all predictors with a lower predictor order number (indicated with grey boxes). A detailed explanation of the predictors can be found in Supporting Information Tables S1 and S2. AAO: Antarctic Oscillation; AO: Arctic Oscillation; AR: Atlantic Ridge; AT: Atlantic Trough; DOY: day of year: ERA: European Centre for Medium-Range Weather Forecasts Reanalysis; EuBL: European Blocking; GL: Greenland Blocking; IWR: weather regime index; MJO: Madden–Julian Oscillation; NAO: North Atlantic Oscillation; NWP: numerical weather prediction; OMI: outgoing long-wave radiation MJO index; ONI: oceanic Niño index; PNA: Pacific–North American; QBO: quasi-biennial oscillation; ScBL: Scandinavian Blocking; ScTr: Scandinavian Trough; SPV: stratospheric polar vortex; SST: sea-surface temperature; WR: weather regime; ZO: Zonal Regime. [Colour figure can be viewed at wileyonlinelibrary.com]

Northern Hemisphere counterpart, the Arctic Oscillation (AO), which is closely linked to the NAO. Tachibana *et al.* (2018) demonstrate that the AAO and AO exhibit correlations, particularly in October and February, whereas Song *et al.* (2009) find that negative AAO phases correspond to anomalously high 300 hPa geopotential heights over the North Atlantic–European region, with a lag of 25–40 days. This is consistent with the AAO being a good predictor for EuBL and ScBL in our NNs. Despite the potential connections between the AAO and EuBL/ScBL, our results only establish the existence of a relationship, without clarifying whether it reflects correlation or causality. Rather than reflecting a direct dynamical connection, the relationship between the AAO and EuBL/ScBL may instead arise from a mutual dependence on a common atmospheric mode, such as the QBO, ENSO, or MJO.

The highly variable predictor selection across NNs for different weather regimes highlights the diverse sources of predictability. Key predictors stem from the Arctic (AO-related predictors for AR, ZO, and ScBL), the stratosphere (SPV-related predictors for ZO, ScTr, AR, EuBL, ScBL, and GL; QBO-related predictors for AT, ZO, and GL), and even the Southern Hemisphere (AAO-related predictors for AT, EuBL, and ScBL; El Niño-related predictors for AR, EuBL, and GL). This diversity underscores the complexity of weather regime forecasting in the North Atlantic–European region, where no single factor (aside from NWP predictors) dominates $WRact_{agg}$ across all regimes.

3.4.2 | Performance evaluation

Following the predictor selection, we evaluate the NNs' performance in forecasting the $WRact_{agg}$ (Figure 8). Comparing the MSE skill score) across all forecasts in the extended winter season provides key insights into the statistical–dynamical model performance relative to both the NWP model and climatology. The climatological forecast, based on a standard 91-day running mean of $WRact_{agg}$, serves as a baseline for assessing NWP skill. Consistent with previous research (e.g., Büeler *et al.*, 2021; Osman *et al.*, 2023), the NWP model struggles to outperform climatology for ScBL and, particularly, EuBL at week 3 but performing well for ZO and GL. All three NN set-ups generally outperform the NWP model for EuBL and ScBL, with NN_{all} showing the most pronounced improvements of 11.0%. Notably, NN_{noNWP} surpasses the NWP model for EuBL and ScBL, demonstrating that, even without NWP-derived inputs, it achieves lower MSEs in $WRact_{agg}$ forecasts. Meanwhile, NN_{NWP+WR} provides only minor improvements compared with the NWP model, suggesting that it mainly corrects biases rather than introducing new sources of skill. The most striking result comes from NN_{all}, which consistently outperforms the NWP model across all weather regimes, achieving MSE reductions of 3.1–11.0%, with the largest gains observed for ScBL activity. Furthermore, NN_{all} is the only neural network that enhances the already high forecast skill of $WRact_{agg,GL}$. These improvements go beyond simple error reduction; as illustrated in the $WRact_{agg}$ forecasts (Supporting Information Figure S7), some NNs better capture key $WRact_{agg}$ patterns than the NWP model alone. Notably, NN_{noNWP} successfully captures the prolonged GL activity during the winters of 2009–2010 and 2010–2011. Additionally, for forecasts with the largest x% differences ($x \in (0, 100]$) in predicted activity between NN_{all} and NWP, NN_{all} consistently achieves lower MSE, regardless of the magnitude of the difference (not shown here).

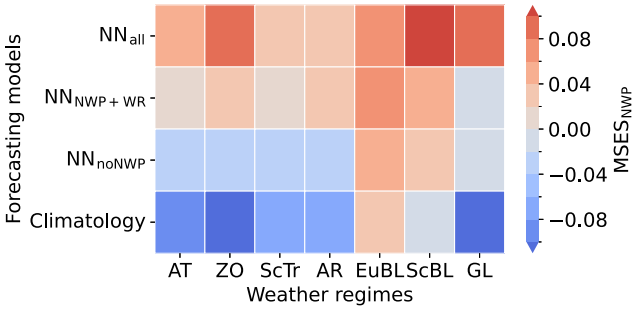

FIGURE 8 Performance of the different neural network and climatological forecasts in comparison with the numerical weather prediction (NWP) forecast for $WRact_{agg}$. As performance measure, the mean-squared error (MSE) is calculated for each model and weather regime (WR) and put into perspective against the MSE of the NWP model (MSE skill score, MSESS). AT: Atlantic Trough; ZO: Zonal Regime; ScTr: Scandinavian Trough; AR: Atlantic Ridge; EuBL: European Blocking; ScBL: Scandinavian Blocking; GL: Greenland Blocking. [Colour figure can be viewed at wileyonlinelibrary.com]

In the final stage of our analysis, we evaluate whether the improvements observed for individual $WRact_{agg}$ forecasts translate into better predictions of the $WRact_{max}$ 3 weeks after forecast initialisation. We determine the weather regime with the maximum predicted activity for both the NN_{all} and the NWP model and then compare their performance using verification metrics derived from the contingency table (Tables 1 and 2), including accuracy (HR), FAR, precision, and F_1 score (Figure 9). Across all metrics, a rather uniform image emerges: the NN_{all} performs equal or better for all weather regimes except the ScTr. The HR improves for all weather regimes except ScTr, with the overall accuracy increasing from 28.7% for the NWP model to 34.5% for the NN_{all}, representing a relative improvement of 20%. FARs decrease for ZO, ScTr, EuBL, ScBL, and GL, and they increase for AT and AR. The precision and F_1 score improve for all weather regimes except the ScTr, which is due to the decrease in the HR. The stepwise predictor selection for ScTr already indicates its limited predictive skill, as the two strongest predictors, $NWP_{agg.,WRactivity}$ and $NWP_{IWRmean}$, are closely related. Moreover, adding further predictors leads to only marginal reductions in MSE compared with the predictor selections for the other weather regimes (not shown).

In conclusion, NNs can extract additional predictive information beyond what is provided by the NWP model alone. Though the NWP-based $WRact_{agg}$ forecast (or IWR forecast for AR and GL) remains the most influential predictor, added value comes from atmospheric indicators such as the SPV, QBO, MJO, and AAO indices. The NN_{all} not only outperforms the NWP model for individual weather regime forecasts (based on the MSE) but

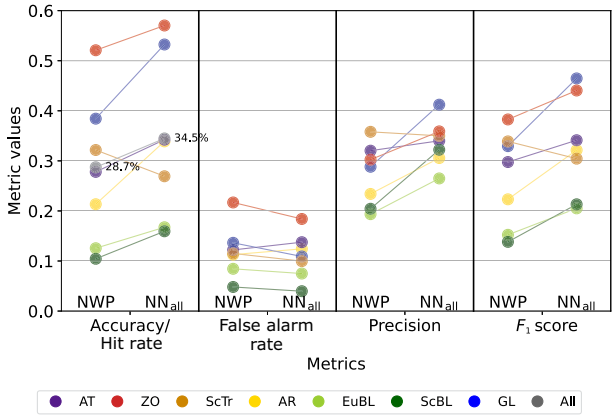

FIGURE 9 Verification metrics comparing the skill of the numerical weather prediction (NWP) model and the all-predictors neural network in predicting the $WRact_{max}$ at a lead time of 3 weeks. Four metrics, the hit rate, the false alarm rate, the precision, and the F_1 score, are visualised (x-axis). Further, together with the hit rate, the accuracy for the full time series is given and indicated by percentages in text. For each metric, on the left the NWP model and on the right the all-predictors neural network (NN_{all}) is represented, joined by a thin line that serves as a guiding line for the reader to directly see whether the respective metric increases or decreases for the neural network in comparison with the NWP model. AT: Atlantic Trough; ZO: Zonal Regime; ScTr: Scandinavian Trough; AR: Atlantic Ridge; EuBL: European Blocking; ScBL: Scandinavian Blocking; GL: Greenland Blocking. [Colour figure can be viewed at wileyonlinelibrary.com]

also achieves a 20% relative improvement in predicting the dominant weather regime at week 3. Furthermore, the NN_{noNWP}, which excludes NWP-derived inputs, performs equally well or better for three of the four blocked regimes (EuBL, ScBL, GL), highlighting the importance of non-NWP sources of predictability. These findings underscore the potential of hybrid forecasting approaches that integrate dynamical and statistical models to enhance subseasonal $WRact_{agg}$ predictions.

4 | CONCLUSIONS AND DISCUSSIONS

Various studies documented the connection between subseasonal sources of predictability (e.g., MJO or SPV) and the large-scale circulation in the North Atlantic–European region (e.g., Cassou, 2008; Lee *et al.*, 2019, 2020; Domeisen *et al.*, 2020c; Roberts *et al.*, 2023). However, only a few studies so far could demonstrate its practical applicability to improve subseasonal NWP forecasts (e.g., Scaife *et al.*, 2022).

In this study, we investigate the influence of the state of the MJO and SPV at initialisation on the occurrence and forecast of seven year-round North Atlantic European weather regimes, with a focus on the activity of GL, 3 weeks later. Furthermore, we leverage information describing the atmospheric state prior to and at initialisation to enhance NWP week 3 weather regime activity forecasts using NNs.

We compute a weather regime activity metric ($WRact_{mean}$ and $WRact_{agg}$), representing the fraction of daily IWR values in a given week exceeding a given threshold, based on reanalysis and reforecasts. Our findings show that GL $WRact_{mean,GL}$ is significantly enhanced following MJO phases 7, 8 and 1, as well as weak (and weakest 10%) SPV state (though not significant), suggesting a climatological window of opportunity (BR anomaly to neutral state is positive). However, ECMWF reforecast HRs show improvement over the respective neutral phase only in MJO phases 1 (significant) and 8 and in the weak SPV state (not significant). A positive PSS in MJO phases 8 and 1, as well as in the weakest 10% SPV state, classifies them as a model window of opportunity type 2—according to Specq and Batté (2022), with a positive anomaly in BR, HR, and PSS with respect to neutral state—whereas the weak SPV state corresponds to a model window of opportunity type 1 (BR and HR anomalies to neutral state are positive and PSS anomaly is negative). GL is rare following the MJO phase 4, though the NWP model is performing well in correctly predicting these rare GLs, classifying MJO phase 4 as a window of opportunity type 3 (HR and PSS anomalies are positive and the BR anomaly is negative). We demonstrate the potential skill for predicting GL activity in week 3 purely derived from reanalysis data by constructing an atmospheric-conditioned climatological forecast that utilises historical information of the MJO state (amplitude and phase) or SPV state to estimate the mean historical $WRact_{agg,GL}$ for the prevailing situation. Incorporating this information (alongside MJO, SPV, and other atmospheric variability indices) with NWP-derived $WRact_{agg}$ metrics in a statistical–dynamical model using NNs improves $WRact_{agg}$ forecasts across all seven weather regimes. When aggregating individual $WRact_{agg}$ forecasts to determine the dominant regime ($WRact_{max}$), the best-performing NN achieves accuracy increases for all regimes individually, except the ScTr. The overall accuracy increase of the neural network is 5.8 percentage points, representing a relative improvement of 20% compared with the NWP model.

Our findings on windows of opportunity align with previous studies on the MJO–NAO (Cassou, 2008; Ferranti *et al.*, 2018; Vitart, 2017) and SPV–NAO (Beerli & Grams, 2019; Büeler *et al.*, 2020) connection. In particular, we find an increased climatological occurrence of GL following weak SPV states, a relationship that is well captured by the ECMWF forecast model. This supports the results of Spaeth *et al.* (2024), who show that ensemble forecasts exhibit greater confidence in predicting the

dominant regime (typically GL) after a weak SPV state compared with neutral or strong states. However, our results also reveal a downside: the NWP model tends to overpredict GL in these situations, leading to (as we can show) frequent false alarms. In contrast, the NWP model shows poor performance in predicting the relatively rare GL occurrence during strong SPV states. These limitations are likely related to the findings of Domeisen *et al.* (2020c), who show that the tropospheric response to a sudden stratospheric warming (i.e., a weak SPV state) varies depending on the tropospheric conditions at the time of the warming. It is plausible that the model lacks the ability to distinguish between these different pathways.

A further limitation of our analysis lies in the choice of the SPV index itself. We rely on the 10 hPa zonal-mean zonal wind at 60° N as a one-dimensional diagnostic, but this definition may not fully capture the complexity of stratosphere–troposphere coupling. More sophisticated predictors could provide a better representation of the SPV state. For example, a combination of indices on different vertical levels (10 and 100 hPa), a combination of temperature and potential vorticity on a specific level (Baldwin *et al.*, 2024), or predictors that reflect the time evolution of the SPV, either in the form of extended one-dimensional indices or time series inputs, may be more suitable. For the NNs, we provide the SPV index on the 10 and 100 hPa level, though the 10 hPa level dominates.

The SPV-conditioned climatological forecast shows only a weak signal for GL onsets, especially when compared with the signal by the MJO. However, in the NN framework, the SPV emerges as a more important predictor alongside the NWP forecasts. This suggests that the NN is able to extract complex, nonlinear dependencies between the SPV, other predictors, and NWP information in ways that are not easily accessible to purely statistical analyses.

To our knowledge, this study is the first to demonstrate how the MJO/SPV–NAO relationship can be harnessed as a decision-support tool for subseasonal GL forecasts, based on purely statistical information and independent of NWP models. Additionally, our NN approach shows particular value for predicting blocking over Europe (EuBL and ScBL). Consistent with previous findings that NWP models struggle with these regimes—often performing worse than simple climatological forecasts (e.g., Büeler *et al.*, 2021; Osman *et al.*, 2023; Wandel *et al.*, 2024)—we find that two of our NNs, one excluding and one including NWP information, both outperform the raw NWP forecasts for these blocked regimes. These results highlight that predictability sources should not be approached with a "one-size-fits-all" strategy but rather tailored to specific weather regimes.

This perspective may have uncovered previously unknown sources of skill, such as the AAO index's potential relevance for EuBL and ScBL. To our knowledge, this connection has not been reported in the literature, and it raises important questions about the underlying mechanisms. Though previous studies suggest some links between the AAO and Northern Hemisphere circulation (e.g., Song *et al.*, 2009; Tachibana *et al.*, 2018), the pathways through which the AAO could influence EuBL/ScBL remain unclear. One possibility is that the AAO exerts a delayed influence on North Atlantic circulation by modulating planetary wave propagation into the stratosphere, thereby affecting the strength and variability of the polar vortex. Through this pathway, changes in the Southern Hemisphere circulation could indirectly alter EuBL/ScBL probabilities. Alternatively, the statistical importance of the AAO may reflect an indirect relationship: the AAO and EuBL/ScBL could both respond to another large-scale mode of variability, such as the QBO, ENSO, or MJO, rather than being causally linked. Our results should be interpreted as revealing a statistical connection rather than demonstrating a direct dynamical mechanism. Future work will be needed to disentangle the influence of the AAO on European weather regimes. Either way, the apparent relevance of the AAO for blocking prediction highlights the potential of data-driven approaches to uncover predictors that are not readily anticipated from a purely dynamical perspective.

Though our analysis highlights statistical links between the MJO, SPV, AAO, and European weather regimes, we do not examine the underlying dynamical pathways, as this lies beyond the scope of the article. For the MJO and SPV, we instead refer the reader to studies that provide detailed discussions of the relevant mechanisms (e.g., Cassou, 2008; Stan *et al.*, 2017, 2022; Domeisen *et al.*, 2020a, 2020b, 2020c; Fromang & Rivière, 2020).

Further research should focus on refining predictor selection for individual weather regimes to establish robust connections. One approach could be incorporating time-series analysis of indices in more complex machine-learning models, such as transformers, to capture temporal dependencies. Additionally, conducting case studies on specific forecasts and their associated atmospheric states may provide deeper insights into the mechanisms driving predictability. Another important step is extending the analysis to longer lead times within the extended range to evaluate how different predictors contribute to forecast skill over increasing time horizons. By enhancing our understanding of regime-specific sources of predictability, we can further improve subseasonal forecasting and develop more reliable decision-support tools for subseasonal prediction.

These tools have the potential to benefit a wide range of weather-sensitive sectors. For example, anticipating prolonged cold spells is crucial for managing energy demand, and forecasts of temperature and precipitation extremes can support agricultural planning during sensitive stages such as harvest. Similarly, reliable information on the likelihood of floods, cold air outbreaks, or heatwaves can enhance risk management and preparedness. The interpretability of regime forecasts makes them particularly valuable for stakeholders who require actionable guidance rather than raw model output.

To conclude, our study demonstrates the potential of statistical–dynamical forecasting with NNs for the subseasonal prediction of North Atlantic–European weather regimes, providing valuable improvements to NWP performance, particularly in situations where traditional NWP models struggle.

ACKNOWLEDGEMENTS

FM has received funding from the KIT Centre for Mathematics in Sciences, Engineering and Economics under the seed funding programme. The contribution of JQ was partly funded by the European Union (ERC, ASPIRE, 101077260). The work of JQ and CMG was funded by the Helmholtz Association as part of the Young Investigator Group, "Sub-seasonal Predictability: Understanding the Role of Diabatic Outflow" (SPREADOUT, grant VH-NG-1243). SL gratefully acknowledges support by the Vector Stiftung through the Young Investigator Group "Artificial Intelligence for Probabilistic Weather Forecasting". Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Code is available on the GitHub repository https://github.com/fmockert/windowsofopportunity_WR. ERA5 data can be obtained from the ECMWF data catalogue at: https://doi.org/10.24381/cds.f17050d7 (Hersbach *et al.*, 2023). Weather regime data are available from zenodo at: https://zenodo.org/records/17080146 (Grams *et al.*, 2025). The original S2S database is hosted at ECMWF as an extension of the TIGGE database.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ENDNOTES

[1]The seasonal normalisation accounts for the inherently stronger variability in winter compared with summer. This is achieved by dividing the filtered Z500 anomaly fields at each grid point by a calendar-day-dependent normalisation factor. The factor is computed as the spatial average (across the weather regime domain) of the 31-day running standard deviation of the Z500 anomalies for the corresponding calendar day.

[2]Note that the weekly mean of daily IWR equals the IWR of the weekly mean geopotential height field (after filtering and normalisation).

ORCID

Fabian Mockert https://orcid.org/0000-0002-3222-6667
Christian M. Grams https://orcid.org/0000-0003-3466-9389
Julian Quinting https://orcid.org/0000-0002-8409-2541

REFERENCES

Baldwin, M.P., Birner, T. & Ayarzagüena, B. (2024) Tropospheric amplification of stratosphere-troposphere coupling. *Quarterly Journal of the Royal Meteorological Society*, 150, 5188–5205. Available from: https://doi.org/10.1002/qj.4864

Barnes, E.A., Samarasinghe, S.M., Ebert-Uphoff, I. & Furtado, J.C. (2019) Tropospheric and stratospheric causal pathways between the MJO and NAO. *Journal of Geophysical Research: Atmospheres*, 124(16), 9356–9371. Available from: https://doi.org/10.1029/2019JD031024

Beerli, R. & Grams, C.M. (2019) Stratospheric modulation of the large-scale circulation in the Atlantic-European region and its implications for surface weather events. *Quarterly Journal of the Royal Meteorological Society*, 145, 3732–3750. Available from: https://doi.org/10.1002/qj.3653

Bloomfield, H.C., Suitters, C.C. & Drew, D.R. (2020) Meteorological drivers of European power system stress. *Journal of Renewable Energy*, 2020, 1–12. Available from: https://doi.org/10.1155/2020/5481010

Bloomfield, H.C., Brayshaw, D.J., Gonzalez, P.L. & Charlton-Perez, A. (2021) Pattern-based conditioning enhances sub-seasonal prediction skill of European national energy variables. *Meteorological Applications*, 28, 7. Available from: https://doi.org/10.1002/met.2018

Büeler, D., Beerli, R., Wernli, H. & Grams, C.M. (2020) Stratospheric influence on ECMWF sub-seasonal forecast skill for energy-industry-relevant surface weather in European countries. *Quarterly Journal of the Royal Meteorological Society*, 146, 3675–3694. Available from: https://doi.org/10.1002/qj.3866

Büeler, D., Ferranti, L., Magnusson, L., Quinting, J.F. & Grams, C.M. (2021) Year-round sub-seasonal forecast skill for Atlantic-European weather regimes. *Quarterly Journal of the Royal Meteorological Society*, 147, 4283–4309. Available from: https://doi.org/10.1002/qj.4178

Cassou, C. (2008) Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature*, 455, 523–527. Available from: https://doi.org/10.1038/nature07286

Charlton-Perez, A.J., Aldridge, R.W., Grams, C.M. & Lee, R. (2019) Winter pressures on the UK health system dominated by the Greenland Blocking weather regime. *Weather and Climate Extremes*, 25, 100218. Available from: https://doi.org/10.1016/j.wace.2019.100218

Chase, R.J., Harrison, D.R., Burke, A., Lackmann, G.M. & McGovern, A. (2022) A machine learning tutorial for operational meteorology. Part I: traditional machine learning. *Weather and Forecasting*, 37(8), 1509–1529. Available from: https://doi.org/10.1175/WAF-D-22-0070.1

Chase, R.J., Harrison, D.R., Lackmann, G.M. & McGovern, A. (2023) A machine learning tutorial for operational meteorology. Part II: neural networks and deep learning. *Weather and Forecasting*, 38(8), 1271–1293. Available from: https://doi.org/10.1175/WAF-D-22-0187.1

Domeisen, D.I., Butler, A.H., Charlton-Perez, A.J., Ayarzagüena, B., Baldwin, M.P., Dunn-Sigouin, E. et al. (2020a) The role of the stratosphere in subseasonal to seasonal prediction: 1. predictability of the stratosphere. *Journal of Geophysical Research: Atmospheres*, 125, 1. Available from: https://doi.org/10.1029/2019JD030920

Domeisen, D.I., Butler, A.H., Charlton-Perez, A.J., Ayarzagüena, B., Baldwin, M.P., Dunn-Sigouin, E. et al. (2020b) The role of the stratosphere in subseasonal to seasonal prediction: 2. predictability arising from stratosphere-troposphere coupling. *Journal of Geophysical Research: Atmospheres*, 125(1). Available from: https://doi.org/10.1029/2019JD030923

Domeisen, D.I.V., Grams, C.M. & Papritz, L. (2020c) The role of North Atlantic-European weather regimes in the surface impact of sudden stratospheric warming events. *Weather and Climate Dynamics*, 1, 373–388. Available from: https://doi.org/10.5194/wcd-1-373-2020

Drücke, J., Borsche, M., James, P., Kaspar, F., Pfeifroth, U., Ahrens, B. et al. (2021) Climatological analysis of solar and wind energy in Germany using the Grosswetterlagen classification. *Renewable Energy*, 164, 1254–1266. Available from: https://doi.org/10.1016/j.renene.2020.10.102

ECMWF. (2018) *Forecast Ensemble (ENS) - Rationale and construction*. https://confluence.ecmwf.int/display/FUG/Section+5+Forecast+Ensemble

Elliot, A.J., Cross, K.W. & Fleming, D.M. (2008) Acute respiratory infections and winter pressures on hospital admissions in England and Wales 1990–2005. *Journal of Public Health*, 30(1), 91–98. Available from: https://doi.org/10.1093/pubmed/fdn003

Feng, P.N. & Lin, H. (2019) Modulation of the MJO-Related Teleconnections by the QBO. *Journal of Geophysical Research: Atmospheres*, 124, 12022–12033. Available from: https://doi.org/10.1029/2019JD030878

Ferranti, L., Corti, S. & Janousek, M. (2015) Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, 141, 916–924. Available from: https://doi.org/10.1002/qj.2411

Ferranti, L., Magnusson, L., Vitart, F. & Richardson, D.S. (2018) How far in advance can we predict changes in large-scale flow leading to severe cold conditions over Europe? *Quarterly Journal of the Royal Meteorological Society*, 144, 1788–1802. Available from: https://doi.org/10.1002/qj.3341

Fromang, S. & Rivière, G. (2020) The effect of the Madden-Julian Oscillation on the North Atlantic Oscillation using idealized numerical experiments. *Journal of the Atmospheric Sciences*, 77, 1613–1635. Available from: https://doi.org/10.1175/JAS-D-19-0178.1

Garfinkel, C.I., Lawrence, Z.D., Butler, A.H., Dunn-Sigouin, E., Erner, I., Karpechko, A.Y. et al. (2025) A process-based evaluation of biases in extratropical stratosphere-troposphere coupling in subseasonal forecast systems. *Weather and Climate Dynamics*, 6, 171–195. Available from: https://doi.org/10.5194/wcd-6-171-2025

Gold, S., White, E., Roeder, W., McAleenan, M., Kabban, C.S. & Ahner, D. (2020) Probabilistic contingency tables: an improvement to verify probability forecasts. *Weather and Forecasting*, 35(2), 609–621. Available from: https://doi.org/10.1175/WAF-D-19-0116.1

Grams, C.M., Beerli, R., Pfenninger, S., Staffell, I. & Wernli, H. (2017) Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, 7(8), 557–562. Available from: https://doi.org/10.1038/NCLIMATE3338

Grams, C.M., Hauser, S. & Bueeler, D. (2025) *Year-round North Atlantic-European Weather Regimes in ERA5 reanalyses. Zenodo*, 9. https://doi.org/10.5281/zenodo.17080146

Hannachi, A., Straus, D.M., Franzke, C.L., Corti, S. & Woollings, T. (2017) Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere. *Reviews of Geophysics*, 55(3), 199–234. Available from: https://doi.org/10.1002/2015RG000509

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

Hauser, S., Teubler, F., Riemer, M., Knippertz, P. & Grams, C.M. (2023b) *Life cycle dynamics of Greenland blocking from a potential vorticity perspective. EGUsphere Preprint repository*. https://doi.org/10.5194/egusphere-2023-2945

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(7), 1999–2049. Available from: https://doi.org/10.1002/qj.3803

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Sabater, J.M. et al. (2023) *ERA5 hourly data on single levels from 1940 to present*. https://doi.org/10.24381/cds.adbb2d47

Kent, C., Scaife, A.A. & Dunstone, N. (2023) What potential for improving sub-seasonal predictions of the winter NAO? *Atmospheric Science Letters*, 24(4), e1146. Available from: https://doi.org/10.1002/asl.1146

King, A.D., Butler, A.H., Jucker, M., Earl, N.O. & Rudeva, I. (2019) Observed relationships between sudden stratospheric warmings and European climate extremes. *Journal of Geophysical Research: Atmospheres*, 124, 13943–13961. Available from: https://doi.org/10.1029/2019JD030480

Kolstad, E.W., Wulff, C.O., Domeisen, D.I. & Woollings, T. (2020) Tracing north Atlantic oscillation forecast errors to stratospheric origins. *Journal of Climate*, 33, 9145–9157. Available from: https://doi.org/10.1175/JCLI-D-20-0270.1

Lee, R.W., Woolnough, S.J., Charlton-Perez, A.J. & Vitart, F. (2019) ENSO modulation of MJO teleconnections to the North Atlantic and Europe. *Geophysical Research Letters*, 46, 13535–13545. Available from: https://doi.org/10.1029/2019GL084683

Lee, J.C.K., Lee, R.W., Woolnough, S.J. & Boxall, L.J. (2020) The links between the Madden-Julian Oscillation and European weather regimes. *Theoretical and Applied Climatology*, 141, 567–586. Available from: https://doi.org/10.1007/s00704-020-03223-2

Lin, H., Brunet, G. & Derome, J. (2009) An observed connection between the North Atlantic oscillation and the Madden-Julian oscillation. *Journal of Climate*, 22, 364–380. Available from: https://doi.org/10.1175/2008JCLI2515.1

Madden, R.A. & Julian, P.R. (1971) Detection of a 40-50 Day Oscillation in the Zonal Wind in the Tropical Pacific. *Journal of the Atmospheric Sciences*, 28(3), 702–708. Available from: https://doi.org/10.1175/1520-0469(1971)028

Michel, C. & Rivière, G. (2011) The link between Rossby wave breakings and weather regime transitions. *Journal of the Atmospheric Sciences*, 68(8), 1730–1748. Available from: https://doi.org/10.1175/2011JAS3635.1

Michelangeli, P., Vautard, R. & Legras, B. (1995) Weather regimes: recurrence and quasi stationarity. *Journal of the Atmospheric Sciences*, 52, 1237–1256. Available from: https://doi.org/10.1175/1520-0469(1995)052

Millin, O.T., Furtado, J.C. & Malloy, C. (2024) The impact of North American winter weather regimes on electricity load in the central United States. *NPJ Climate and Atmospheric Science*, 7, 12. Available from: https://doi.org/10.1038/s41612-024-00803-1

Mockert, F., Grams, C.M., Brown, T. & Neumann, F. (2023) Meteorological conditions during periods of low wind speed and insolation in Germany: the role of weather regimes. *Meteorological Applications*, 30, 7. Available from: https://doi.org/10.1002/met.2141

Mockert, F., Grams, C.M., Lerch, S., Osman, M. & Quinting, J. (2024) Multivariate post-processing of probabilistic sub-seasonal weather regime forecasts. *Quarterly Journal of the Royal Meteorological Society*, 150, 4771–4787. Available from: https://doi.org/10.1002/qj.4840

NOAA. (2025) *Daily MJO index time series*, 2. https://psl.noaa.gov/mjo/mjoindex/

Osman, M., Beerli, R., Büeler, D. & Grams, C.M. (2023) Multi-model assessment of sub-seasonal predictive skill for year-round Atlantic–European weather regimes. *Quarterly Journal of the Royal Meteorological Society*, 7, 2386–2408. Available from: https://doi.org/10.1002/qj.4512

Otero, N., Martius, O., Allen, S., Bloomfield, H. & Schaefli, B. (2022) Characterizing renewable energy compound events across Europe using a logistic regression-based approach. *Meteorological Applications*, 29, 9. Available from: https://doi.org/10.1002/met.2089

Roberts, C.D., Balmaseda, M.A., Ferranti, L. & Vitart, F. (2023) Euro-Atlantic weather regimes and their modulation by tropospheric and stratospheric teleconnection pathways in ECMWF reforecasts. *Monthly Weather Review*, 151, 2779–2799. Available from: https://doi.org/10.1175/MWR-D

Salient, M.D. & Atlas, A. (2023) *On the Money: How Salient's Novel S2S Forecasts Can Save the Agriculture Industry Millions of Dollars, 5.* https://www.salientpredictions.com/blog/on-the-money-how-salients-novel-s2s-forecasts-can-save-the-agriculture-industry-millions-of-dollars

Scaife, A.A., Karpechko, A.Y., Baldwin, M.P., Brookshaw, A., Butler, A.H., Eade, R. et al. (2016) Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, 17(1), 51–56. Available from: https://doi.org/10.1002/asl.598

Scaife, A.A., Hermanson, L., Niekerk, A.V., Andrews, M., Baldwin, M.P., Belcher, S. et al. (2022) Long-range predictability of extratropical climate and the length of day. *Nature Geoscience*, 15(10), 789–793. Available from: https://doi.org/10.1038/s41561-022-01037-7

Song, J., Zhou, W., Li, C. & Qi, L. (2009) Signature of the Antarctic oscillation in the northern hemisphere. *Meteorology and Atmospheric Physics*, 105, 55–67. Available from: https://doi.org/10.1007/s00703-009-0036-5

Spaeth, J., Rupp, P., Osman, M., Grams, C.M. & Birner, T. (2024) Flow-dependence of ensemble spread of subseasonal forecasts explored via North Atlantic-European weather regimes. *Geophysical Research Letters*, 51(7), e2024GL109733. Available from: https://doi.org/10.1029/2024GL109733

Specq, D. & Batté, L. (2022) Do subseasonal forecasts take advantage of Madden–Julian oscillation windows of opportunity? *Atmospheric Science Letters*, 23(4), e1078. Available from: https://doi.org/10.1002/asl.1078

Stan, C., Straus, D.M., Frederiksen, J.S., Lin, H., Maloney, E.D. & Schumacher, C. (2017) Review of tropical-extratropical teleconnections on intraseasonal time scales. *Reviews of Geophysics*, 55(4), 902–937. Available from: https://doi.org/10.1002/2016RG000538

Stan, C., Zheng, C., Chang, E.K., Domeisen, D.I., Garfinkel, C.I., Jenney, A.M. et al. (2022) Advances in the prediction of MJO-teleconnections in the S2S forecast systems. *Bulletin of the American Meteorological Society*, 103, E1426–E1447. Available from: https://doi.org/10.1175/bams-d-21-0130.1

Tachibana, Y., Inoue, Y., Komatsu, K.K., Nakamura, T., Honda, M., Ogata, K. et al. (2018) Interhemispheric synchronization between the AO and the AAO. *Geophysical Research Letters*, 45, 13,477–13,484. Available from: https://doi.org/10.1029/2018GL081002

Tripathi, O.P., Baldwin, M., Charlton-Perez, A., Charron, M., Eckermann, S.D., Gerber, E. et al. (2015a) The predictability of the extratropical stratosphere on monthly time-scales and its impact on the skill of tropospheric forecasts. *Quarterly Journal of the Royal Meteorological Society*, 141(4), 987–1003. Available from: https://doi.org/10.1002/qj.2432

Tripathi, O.P., Charlton-Perez, A., Sigmond, M. & Vitart, F. (2015b) Enhanced long-range forecast skill in boreal winter following stratospheric strong vortex conditions. *Environmental Research Letters*, 10(10), 104007. Available from: https://doi.org/10.1088/1748-9326/10/10/104007

Vautard, R. (1990) Multiple weather regimes over the North Atlantic: analysis of precursors and successors. *Monthly Weather Review*, 118, 2056–2081. Available from: https://doi.org/10.1175/1520-0493(1990)118

Vitart, F. (2017) Madden—Julian Oscillation prediction and teleconnections in the S2S database. *Quarterly Journal of the Royal Meteorological Society*, 143, 2210–2220. Available from: https://doi.org/10.1002/qj.3079

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. Available from: 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(4), e2023JD039791. Available from: https://doi.org/10.1029/2023JD039791

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, 1917–1932. Available from: https://doi.org/10.1175/1520-0493(2004)132

White, I.P., Garfinkel, C.I., Gerber, E.P., Jucker, M., Hitchcock, P. & Rao, J. (2020) The generic nature of the tropospheric response to sudden stratospheric warmings. *Journal of Climate*, 33, 5589–5610. Available from: https://doi.org/10.1175/JCLI-D-19

Wiel, K.V.D., Stoop, L.P., Zuijlen, B.R.V., Blackport, R., Broek, M.A.V.D. & Selten, F.M. (2019) The influence of weather regimes on European renewable energy production and demand. *Environmental Research Letters*, 14(9). Available from: https://doi.org/10.1088/1748-9326/ab38d3

Yiou, P. & Nogaj, M. (2004) Extreme climatic events and weather regimes over the North Atlantic: when and where? *Geophysical Research Letters*, 31(4), 2003GL019119. Available from: https://doi.org/10.1029/2003GL019119

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mockert, F., Grams, C.M., Lerch, S. & Quinting, J. (2026) Windows of opportunity in subseasonal weather regime forecasting: A statistical–dynamical approach. *Quarterly Journal of the Royal Meteorological Society*, e70090. Available from: https://doi.org/10.1002/qj.70090