DISSERTATION

APPLICATION OF NEURAL NETWORKS TO SUBSEASONAL TO SEASONAL PREDICTABILITY IN PRESENT AND FUTURE CLIMATES

Submitted by

Kirsten J. Mayer

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2022

Doctoral Committee:

Advisor: Elizabeth A. Barnes

James W. Hurrell
Eric D. Maloney
Charles Anderson

ABSTRACT

APPLICATION OF NEURAL NETWORKS TO SUBSEASONAL TO SEASONAL PREDICTABILITY IN PRESENT AND FUTURE CLIMATES

The Earth system is known for its lack of predictability on subseasonal to seasonal timescales (S2S; 2 weeks to a season). Yet accurate predictions on these timescales provide crucial, actionable lead times for agriculture, energy, and water management sectors. Fortunately, specific Earth system states – deemed *forecasts of opportunity* – can be leveraged to improve prediction skill. Our current understanding of these opportunities are rooted in our knowledge of the historical climate. Depending on societal actions, the future climate could vary drastically, and these possible futures could lead to varying changes to S2S predictability. In recent years, neural networks have been successfully applied to weather and climate prediction. With the rapid development of neural network explainability techniques, the application of neural networks now provides an opportunity to further understand our climate system as well. The research presented here demonstrates the utility of explainable neural networks for S2S prediction and predictability changes under future climates.

The first study presents a novel approach for identifying forecasts of opportunity in observations using neural network confidence. It further demonstrates that neural networks can be used to gain physical insight into predictability, through neural network explainability techniques. We then employ this methodology to explore S2S predictability differences in two future scenarios: under anthropogenic climate change and stratospheric aerosol injection (SAI). In particular, we explore subseasonal predictability and forecasts of opportunity changes under anthropogenic warming compared to a historical climate in the CESM2-LE. We then investigate how future seasonal predictability may differ under SAI compared to a future without SAI deployment in the ARISE-SAI simulations. We find differences in predictability between the historical and future climates

and the two future scenarios, respectively, where the largest differences in skill generally occur during forecasts of opportunity. This demonstrates that the forecast of opportunity approach, presented in the first study, is useful for identifying differences in future S2S predictability that may not have been identified if examining predictability across all predictions. Overall, these results demonstrate that neural networks are useful tools for exploring subseasonal to seasonal predictability, its sources, and future changes.

ACKNOWLEDGEMENTS

The National Science Foundation Graduate Research Fellowship (grant 006784) funded Kirsten J. Mayer for the entirety of the research presented in this dissertation. This work was also funded by the National Science Foundation Harnessing the Data Revolution (grant 1934668; Chapter 2), the Regional and Global Model Analysis program area of the Department of Energy's Office of Biological and Environmental Research as part of the Program for Climate Model Diagnosis and Intercomparison Project (Chapter 3), and by the Defense Advanced Research Projects Agency HR00112290071 (Chapter 4).

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction

In the past few decades, there has been a community effort to bridge the gap between weather and climate prediction [1–4], particularly focusing on subseasonal to seasonal timescales (S2S; 2 weeks to a season). These timescales are often referred to as a "predictability desert", as generally neither the atmospheric initial conditions nor the slower varying oceanic states provide ample information for skillful predictions [2,5,6]. However, accurate predictions on these timescales can provide crucial anticipatory information for public and private sectors including the agriculture, energy, and water management sectors [1,4].

One approach to improve S2S prediction skill is to leverage earth system states that are known to provide enhanced predictability, referred to as "forecasts of opportunity" [2]. These opportunities include the Madden-Julian Oscillation (MJO) [7, 8], the El Niño Southern Oscillation (ENSO) [9–12], stratospheric phenomenon (e.g. sudden stratospheric warming events and the Quasi-Biennial Oscillation) [13, 14], the East Asian summer Monsoon [15, 16], soil moisture [17, 18], and others. Two sources of S2S predictability in particular (MJO and ENSO) are located in the tropics and can influence midlatitude variability and predictability on S2S timescales through tropical-extratropical teleconnections [19–25].

The MJO is composed of an east-west oriented dipole of enhanced and suppressed convection that propagates from the Indian Ocean into the central tropical Pacific over about 20-90 days [26–28]. Through convective heating, the MJO excites quasi-stationary Rossby waves that are then steered by the Pacific subtropical jet [29], and these waves can modulate midlatitude circulation, on S2S timescales [19,20,22,30,31]. Certain phases (i.e. location) of the MJO have been shown to provide a more consistent modulation of midlatitude circulation and subsequently, lead to enhanced midlatitude S2S prediction skill [7].

Another source of S2S predictability comes from ENSO [2], an interannual coupled ocean-atmosphere mode (3-8 years) in the tropical Pacific Ocean. It is typically defined by two phases: La Niña and El Niño, characterized by anomalously cold and warm sea surface temperatures in

the eastern tropical Pacific, respectively [32]. ENSO has been shown to alter the MJO and MJO teleconnections through its impact on the tropical basic state and the location and strength of the subtropical jet [23,33–37]. These impacts have been shown to influence atmospheric blocking frequency [23] and teleconnection consistency [35], as well as enhance S2S prediction under certain MJO-ENSO conditions [11]. Additionally, ENSO has its own teleconnections that can influence variability on longer timescales (e.g. Pacific North American pattern) [38–40], ultimately impacting seasonal predictability [9, 10].

Previous work has demonstrated the utility of empirical models for S2S prediction [8, 41–43] and shown that statistical methods, such as linear inverse models, can identify forecasts of opportunity [44,45]. Recently, neural networks have also been successfully applied to weather and climate prediction [46–50]. Neural networks are useful statistical methods for extracting nonlinear relationships [51]; however, previously the decision making process of the network had been relatively enigmatic. With the development of explainability techniques, or explainable artificial intelligence (XAI) [52], we are now able to extract and visualize what the neural network uses to make its predictions. As a result, XAI provides a means to gauge trust in the network's predictions as well as an opportunity for scientists to further improve our understanding of the climate system [53–58].

Given the success of statistical models for S2S prediction and forecast of opportunity identification, the rapidly growing successful applications of neural networks to the atmospheric sciences, and the recent advances in XAI, it raises the question as to whether neural networks could be used to identify physically meaningful S2S forecasts of opportunity. The first research chapter of this dissertation (Chapter 2) aims to address this question by using neural networks to examine a known opportunistic relationship between the MJO and circulation over the North Atlantic. Through this application, we demonstrate that neural networks can identify subseasonal forecasts of opportunity in observations by using the network's confidence in a prediction, and through an explainability technique, confirm the network is identifying physically relevant regions for enhanced prediction in the North Atlantic. These findings demonstrate the utility of neural networks for forecast of op-

portunity identification, and therefore, provide a framework for future applications of explainable neural networks to S2S prediction.

Our knowledge of the utility of phenomena like the MJO and ENSO for S2S prediction is rooted in our current understanding of the climate system. However, without extensive mitigation, the climate is projected to continue warming [59], and this can subsequently impact the MJO [60] and ENSO [61] as well as their teleconnections [62–69]. This suggests that the role of phenomena like the MJO and ENSO in S2S predictability may also change in the future. Chapter 3 of this dissertation explores possible subseasonal predictability changes under climate change using a global climate model large ensemble simulation (CESM2-LE). To do so, we use neural networks and the approach presented in Chapter 2 to quantify prediction skill across all predictions and during forecasts of opportunity. Overall, we demonstrate that neural networks are useful for evaluating predictability changes under future climate scenarios. Furthermore, we find that largest changes in future subseasonal predictability occur during forecasts of opportunity and minimal differences in skill are seen when comparing across all predictions. These results further demonstrate the value of the network-based forecast of opportunity approach for subseasonal predictability analyses.

In recent years, various forms of solar radiation modification have been proposed to reduce the impact of anthropogenic climate change [70]. One of the most well studied ways to reflect sunlight back into space, and thereby cool the planet, is through stratospheric aerosol injection (SAI): the injection of sub-micron sized reflective particles into the stratosphere. This method has shown promise for reaching global mean temperature targets in climate models, but it may also have climate impacts beyond surface temperature, such as on tropical precipitation [71, 72]. Given the global importance of tropical precipitation [29, 30], midlatitude S2S variability and predictability could be impacted. Chapter 4 of this dissertation investigates how future seasonal variability and predictability may differ under SAI implementation compared to a climate scenario without SAI, using the Assessing Responses and Impacts of Solar climate intervention on the Earth system under SAI (ARISE-SAI) simulations [72]. We find higher seasonal variability throughout the Northern Hemisphere and differences in ENSO teleconnections to the northwest coast of North America.

Motivated by ENSO teleconnection differences, we apply the framework presented in Chapter 2 to explore predictability changes across network confidence. We find that seasonal predictability over the northwest coast of North America is higher under SAI, again largest at higher confidence values.

Overall, this dissertation demonstrates the utility of neural networks for S2S prediction through a forecast of opportunity lens. Specifically, we present a neural network-based approach for identifying forecasts of opportunity, and then further demonstrate how this approach can be used to identify S2S predictability changes under future climates.

The following chapters of this dissertation are organized into three research chapters and a conclusion chapter. Chapter 2 and 3 of this work are published in Geophysical Research Letters [73, 74] and therefore, have been included in this dissertation without changes. Chapter 4 is in preparation to be submitted for publication soon after the submission of this dissertation. The last chapter provides a summary of the three research chapters outlined above and future directions for research at the intersection of machine learning and S2S prediction.

Chapter 2: Subseasonal Forecasts of Opportunity Identified by an Explainable Neural Network

2.1 Introduction

Subseasonal timescales (2 weeks - 2 months) are known for their lack of predictability [75], yet reliable and actionable information on these timescales are required for decision making in many sectors such as public health and water management [1, 5]. Over the past decade, there has been a substantial research effort to improve prediction on these timescales [5, 6, 76, 77]. One area of subseasonal prediction research focuses on forecasts of opportunity, the idea that certain earth system conditions provide opportunities for enhanced subseasonal prediction skill [2]. When these opportunities arise, the information provided by the earth system's state can then be leveraged to improve forecast skill. For example, when the Madden-Julian Oscillation (MJO) [26, 27], a propagating tropical convective phenomenon, is active, its convective heating can lead to the excitation of quasi-stationary Rossby waves [29] that subsequently modulate the midlatitude circulation over the first few weeks following MJO activity [19, 20, 22, 30, 31]. When opposing convective anomalies are located over the Indian Ocean and western Pacific (defined as phases 2, 3, 6, and 7), the MJO has been shown to lead to more coherent and consistent modulations of midlatitude weather on subseasonal timescales and consequently, enhanced prediction skill [7]. Using the strength and location of tropical convective activity of the MJO to identify periods of enhanced midlatitude prediction skill is, therefore, an example of forecast of opportunity identification. Mundhenk et al. (2018) also show that an empirical model, which solely uses information about the state of the MJO and the Quasi-Biennial Oscillation, outperforms a state-of-the-art numerical prediction model for prediction of atmospheric river activity on subseasonal timescales. This highlights the importance of statistical models for enhancing subseasonal prediction.

Albers and Newman (2019) demonstrate a technique for forecast of opportunity identification through the utilization of expected skill from a linear inverse model. The study demonstrates

the ability of the linear statistical model to identify forecasts of opportunity, and raises the question of whether other statistical models, such as artificial neural networks (ANNs), can identify forecasts of opportunity for subseasonal prediction. ANNs are very good at nonlinear function estimation [51], and thus, may be able to identify both linear and nonlinear relationships that lend predictability. Recently, ANNs have been successfully applied to seasonal prediction of meteorological variables such as monthly rainfall [78] and surface temperature [79] as well as yearly prediction of the El Niño Southern Oscillation [47], suggesting ANNs may be useful for identifying subseasonal forecasts of opportunity as well.

In this paper, we test whether an ANN can be used for subseasonal forecast of opportunity identification. To do so, we input tropical outgoing longwave radiation (OLR) anomalies into an ANN and task the network to predict the sign of 500 hPa geopotential height (z500) anomalies in the North Atlantic (40°N, 325°E) 22 days later (e.g. Week 4). Tropical OLR is used to explore the ability of an ANN to identify known relationships between the MJO and the North Atlantic via tropical-extratropical teleconnections [19, 42]. We demonstrate that an ANN can identify subseasonal forecasts of opportunity related to tropical OLR, and through an ANN explainability technique, demonstrate that the ANN identifies these known MJO-like OLR patterns. In addition, we find a possible new tropical OLR pattern associated with predictable behavior of the North Atlantic circulation on subseasonal timescales.

2.2 Data and Methods

2.2.1 Data

We use daily mean OLR (1979-2019) from the National Center for Atmospheric Research/National Oceanic and Atmospheric Administration (NCAR/NOAA) [80] and daily mean z500 (1979-2019) from the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim reanalysis (ERA-I) [81]. MJO teleconnections tend to be stronger during boreal winter [82], and therefore, the extended boreal winter months (November-February) are used for the OLR fields. Since we

task the network to predict the sign of the z500 anomaly 22 days following a given OLR field, March is also included in the z500 analysis (see Text A.2 for reasoning behind the choice of lead).

The annual cycle is removed from both the z500 and OLR data. For z500, the annual cycle is removed by subtracting the daily climatology over the record (1979-2019). A Fast Fourier Transform (FFT) high-pass filter is then applied to the z500 anomalies to remove seasonal oscillations (frequencies smaller than $\frac{1}{120days}$) to ensure the network focuses on subseasonal anomalies. The median of the z500 anomalies for the training data (see 2.2.1) is subtracted to obtain an equal number of positive and negative values. These anomalies are then converted into 0s and 1s depending on the sign (negative or positive, respectively). To filter the testing data (see 2.2.1), z500 anomalies from 2017-2019 are appended to the unfiltered z500 anomalies from 1979-2016 and another FFT high pass filter is applied to all years. The now filtered 2017-2019 data are then subset and used as testing data. Initially, the FFT analysis is not applied to the full dataset to ensure the network has no information about the testing data during training. The median of the z500 anomalies for the training data (see 2.2.1) is then subtracted and the anomalies are converted into 0s and 1s. For OLR, the annual cycle is removed by subtracting the first 3 harmonics of the daily climatology from the raw field. The first 3 harmonics are used instead of the daily mean because OLR is a noisier field than z500.

2.2.2 Methods

Artificial Neural Network Architecture

A two layer ANN (Figure 2.1) is tasked to ingest tropical OLR and predict the *sign* of the z500 anomaly over the North Atlantic (40°N, 325°E; red dot in Figure 2.1) 22 days later. The North Atlantic is chosen for this analysis since the MJO is known to force circulation anomalies over this region on subseasonal timescales and thus allows us to explore the utility of an ANN in the context of a well known problem [19, 37, 42]. In addition, we find that this grid point is representative of a larger area within the North Atlantic (see supplemental Figure A.1).

Each input sample to the ANN consists of vectorized daily anomalous OLR from 30°N to 20°S and 45 to 210°E, where the number of input nodes is equal to the number of OLR grid points (N = 1407). The ANN then outputs two values that describe the categorical prediction, positive or negative sign of z500, given the initial OLR input image. The softmax activation function is applied to this final layer and transforms the two output values such that they sum to 1. The output then represents an estimation of the likelihood that an input belongs to a particular category, where the predicted category is defined by a likelihood greater than 0.5. We refer to this estimation of likelihood as “model confidence". A more confident prediction will, therefore, have a predicted category value closer to 1. We define forecasts of opportunities as the top 10% most confident predictions by the network, although we explore alternative percentages as well.

The ANN architecture consists of two hidden layers of 128 and 8 nodes, respectively, and both use the rectified linear activation function. The final layer includes 2 nodes and uses the rectified linear and softmax activation function. Categorical cross entropy is used for the loss function. This architecture is chosen because it was found to consistently lead to reasonably high accuracies across many combinations of training/validation sets, but our ANN approach should be equally applicable to both shallow and deep networks. The batch size is set to 256 samples (i.e. OLR vectorized images) and the ANN is trained for 50 epochs unless the validation loss increases for two epochs in a row. If this occurs, the ANN stops training early and restores the model’s best weights to reduce overfitting. It is found that 50 epochs is sufficient for training as the ANN rarely completes all 50 epochs. A more detailed explanation of ANNs is provided in the supplemental material for reference along with a comparison of this ANN approach to multinomial logistic regression.

The data used to train and test the ANN is composed of three groups: training, validation, and testing. Training and validation data are used during training, where training data is used to update the weights and biases of the ANN and the validation data is used to evaluate the model. The testing data is data that has never been “seen" by the ANN to evaluate the ability of the ANN to generalize to new data. To create the testing data, we assume that the years 2017-2019 have not

yet occurred when training the model. In this way, these years act as true testing data for the ANN. While the specific accuracies likely would change with different testing data, the main point of this paper is to introduce a method to identify forecasts of opportunity and then to further identify the associated relevant regions for the enhanced prediction skill, not to provide the most accurate model for this scenario.

For this analysis, the ANN validation data is from November 2007 through February 2011 (N = 481) and the testing data is from November 2017 through February 2019 (N = 240). The remaining extended boreal winter (NDJF) data are used for training (November 1979 - February 2007 and November 2011 - February 2016; N = 4450; see supplemental Figure A.2). All data is standardized for each grid point by the years used for training and validation. To choose a model for the following analysis, ANN training is repeated for a variety of validation years. Different consecutive four-year chunks are removed from the training data and set aside to use as validation. For each of the nine four-year chunks, the ANN was trained 20 times with random initialized weights. We find that our conclusions are robust to our choice in training period and do not change with variations in random initialization weights. We present one model with reasonably high accuracy here and using the training, validation, and testing groups outlined above.

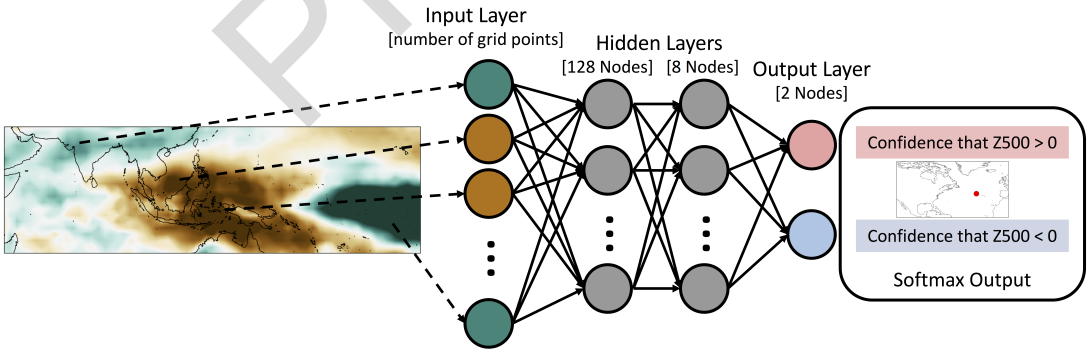

Figure 2.1: Artificial neural network architecture for prediction of the sign of z500 anomalies over the North Atlantic 22 days following tropical OLR anomalies. The neural network consists of two hidden layers of 128 and 8 nodes, respectively, and an output layer of two nodes (one node for each sign). The output layer uses the softmax activation function.

Layer-Wise Relevance Propagation (LRP)

While ANNs are a useful tool for making predictions, in doing so, they are learning *how* to make accurate predictions. Therefore, understanding the inner workings of a trained ANN can provide valuable information for improving prediction skill and understanding, as well as increasing user confidence in the results. Here, we utilize a relatively new neural network explainability technique to the geosciences called layer-wise relevance propagation (LRP) [83,84]) to extract and visualize the features the trained ANN employs to make accurate predictions. While Toms et al. 2020 describes the use of LRP for geoscience applications in detail, we briefly provide a high-level description here (see supplemental material for a more detailed explanation). After network training is completed, a single sample is passed through the network and a prediction is made (in our case, two output values are predicted). Our implementation of LRP then takes the highest of these values (i.e. the winning category) and back-propagates this value through the network via a series of predefined rules, ultimately distributing it across the input nodes (i.e. input gridpoints). What results is a heat map of "relevance" across the input space, where input nodes that are more relevant for the network's specific prediction for that sample are given higher relevance. This process is then repeated for every sample of interest, resulting in a unique relevance heat map for each sample. In our study, since the input layer consists of maps of OLR anomalies, the LRP heat maps are maps of the relevant tropical OLR patterns for each prediction of the circulation in the North Atlantic (40°N, 325°E). These maps are discussed in detail in Section 2.3.2.

2.3 Results

2.3.1 Identifying Forecasts of Opportunity

ANNs with the architecture shown in Figure 2.1 are trained 100 times with random initialized weights to predict the sign of the z500 anomalies 22 days following the tropical OLR anomalies. Figure 2.2a shows the distribution of the testing prediction accuracy for all 100 models, where dark teal represents the distribution of all predictions and light teal represents the distribution of the 10% most confident predictions. The corresponding colored vertical dashed lines indicate

a threshold for what is expected by random chance. To calculate the random chance accuracy threshold, 100,000 randomly generated groups (N=240 for all and N=24 for 10% most confident predictions) of zeros and ones are used to create a distribution of accuracies, and the 90^{th} percentile of this distribution is used as the random chance threshold. In Figure 2.2a, the top 10% most confident prediction accuracies (light teal) are shifted towards higher accuracies compared to the distribution with all predictions (dark teal). This shift in the distributions demonstrates that in general, higher model confidence leads to substantially enhanced prediction accuracy.

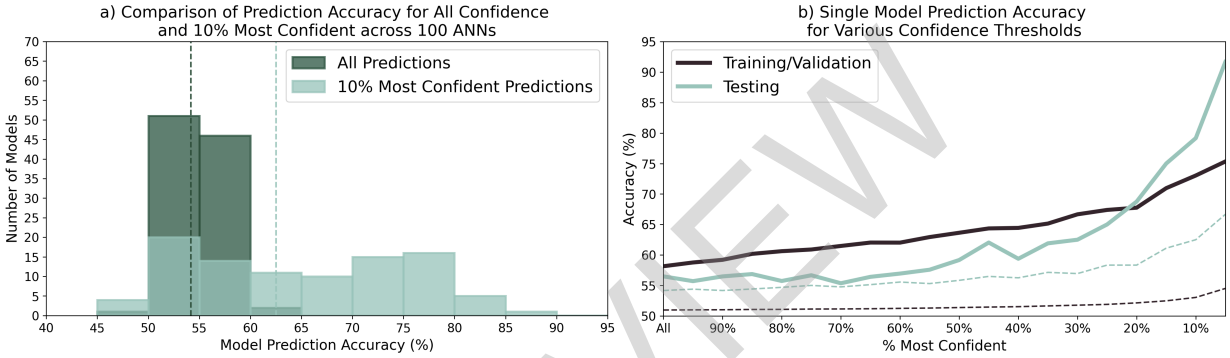

Figure 2.2: (a) Histograms of testing prediction accuracy for 100 trained ANNs. The dark teal represents the histogram of all prediction accuracies and the light teal represents the histogram for the 10% most confident prediction accuracies. The dark teal and light teal dashed lines in (a) are the maximum accuracies expected by random chance at the 90% confidence level for the corresponding colored histogram (see text for details). (b) Accuracy of one particular model as a function of the percent most confident predictions for training and validation (black) and testing (light teal) data. The dashed lines indicate the maximum accuracies expected by random chance at the 90% confidence level for the corresponding colored lines (see text for details).

We chose one model from Figure 2.2a to further understand how accuracy varies when a different percent model confidence is used (Figure 2.2b). The solid lines represent the accuracy across various model confidence values for training and validation (black) and testing (light teal) data sets. Figure 2.2b shows that the testing accuracy (light teal line) barely outperforms the random chance 90% confidence bound (light teal dashed line) for all predictions ("all") while the skill is substantially larger than random chance for the top 10% of predictions. Accuracy increasing with increasing model confidence is also apparent in the training and validation data. Together, Figure 2.2a and b illustrate that model confidence and prediction accuracy generally increase together

and therefore, can be used to identify forecasts of opportunities, or periods of enhanced prediction skill. From this analysis, the 10% most confident predictions are chosen to define forecasts of opportunity since this threshold has one of the largest accuracy differences from random chance while still retaining 10% of the samples.

When evaluating the network with the training and validation data, the prediction accuracy for all predictions is 58% and for the top 10% most confident predictions is 73%. For the testing data, the prediction accuracy for all predictions is 56% and for the top 10% most confident predictions is 79%. The ANN predictions as a function of time are detailed in Figure A.2, and additional skill metrics are provided in Figure A.3 and Table A.4.

2.3.2 Tropical Sources of Predictability

We have shown that ANNs can identify forecasts of opportunity using model confidence; however, understanding where this enhanced skill originates is critical for improving physical understanding as well as gaining trust in the network's predictions. To do so, layer-wise relevance propagation is used to identify the OLR patterns that lead the ANN to make correct predictions (see Section 2.2.2). The correct 10% most confident predictions from the training, validation and testing data sets are combined for this LRP analysis. All three sets of data are used instead of only testing data because all data sets have similar accuracies and LRP values (not shown). Thus, including all the data increases the sample sizes for the analysis. The shading in Figure 2.3c-h shows the regions the network found relevant, on average, to make confident and correct positive (Figure 2.3c,e,g) and negative (Figure 2.3d,f,h) z500 predictions. The contours correspond to the average OLR anomalies for these confident and correct predictions.

The average LRP heat map for the correct forecasts of opportunity of positive sign predictions (Figure 2.3c) indicates four hot spots, one over the southern Indian Ocean into the southern Maritime Continent (20-0°S, 70-130°E), one over the western Pacific (20-0°S, 155°E-170°E), another northwest of Hawaii (25°N, 170°W), and the fourth over Saudi Arabia (30°N, 40-60°E). The average LRP heat map for the correct forecasts of opportunity of negative sign predictions (Figure 2.3d)

indicates four hot spots, one over the Maritime Continent (20-0°S, 110-150°E), one in the western and central Pacific Ocean (20-0°S, 155°E-170°W), another to the west of Hawaii (20°N, 170°W), and the fourth over Saudi Arabia (30°N, 40-60°E).

For both sign predictions, the hot spots over the Maritime Continent and the western Pacific have opposing signed OLR anomalies (contours) that straddle 150°E. These dipoles of convection over the Indian Ocean into the Maritime Continent and over the western Pacific have similar structures to phase 4-5 and phase 1,7-8 of the MJO [85]. This structure of OLR is consistent with previous research of MJO teleconnections over the North Atlantic for average lead times of 10-14 and 15-19 days [7, 19, 23, 42]. In addition, this dipole structure is known to lead to higher pattern consistency of teleconnections in the midlatitudes [86], which has been shown to lead to enhanced prediction skill [7]. Rossby waves initiated by the MJO tend to be quasi-stationary, which suggests that these OLR anomalies may also correspond to 22 day leads as well. This Maritime Continent and western Pacific Ocean dipole highlighted in part by LRP is therefore consistent with previous research and demonstrates that the ANN has learned physically relevant structures.

To test the robustness of these average LRP results for this particular ANN, we calculated the frequency of occurrence of average relevance hotspots greater than 0.5 for models with testing accuracies greater than 70% (Figure 2.3a,b, n = 42 models). We find that all of the hotspots (i.e. the MJO-like structure, the hot spot over Saudi Arabia and the hot spot west of Hawaii) are robust features for enhanced subseasonal prediction throughout these 42 models. In the next section, we hypothesize that the hot spot over Saudi Arabia is associated with the two-way relationship between the North Atlantic Oscillation (NAO) and the MJO [87]. On the other hand, the hot spot west of Hawaii in both sign predictions is discussed as a possible new region relevant for enhanced subseasonal prediction.

K-means Clustering of LRP Maps

To further distinguish the relevant regions for the ANN's predictions, k-means clustering [88] (see supplemental material for more information) is applied to the LRP maps (Figure 2.3e-h).