Statistical predictability of surface wind components

by

Yiwen Mao
B.Sc., University of British Columbia, 2010
M.Sc., University of Alberta, 2013

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the School of Earth and Ocean Sciences

Statistical predictability of surface wind components

by

Yiwen Mao
B.Sc., University of British Columbia, 2010
M.Sc., University of Alberta, 2013

Supervisory Committee

Dr. Adam Monahan, Supervisor
(School of Earth and Ocean Sciences)

Dr. Alex Cannon, Outside Member
(Environment and Climate Change Canada)

Dr. Lucinda Leonard, Departmental Member
(School of Earth and Ocean Sciences)

Dr. Bill Merryfield, Departmental Member
(School of Earth and Ocean Sciences)

ABSTRACT

Predictive anisotropy is a phenomenon referring to unequal predictability of surface wind components in different directions. This study addresses the question of whether predictive anisotropy resulting from statistical prediction is influenced by physical factors or by types of regression methods (linear vs nonlinear) used to construct the statistical prediction. A systematic study of statistical predictability of surface wind components at 2109 land stations across the globe is carried out. The results show that predictive anisotropy is a common characteristic for both linear and nonlinear statistical prediction, which suggests that the type of regression method is not a major influential factor. Both strong predictive anisotropy and poor predictability are more likely to be associated with wind components characterized by relatively weak and non-Gaussian variability and in areas characterized by surface heterogeneity. An idealized mathematical model is developed separating predictive signal and noise between large-scale (predictable) and local (unpredictable) contributions to the variability of surface wind, such that small signal-to-noise ratio (SNR) corresponds to low and anisotropic predictability associated with non-Gaussian local variability. The comparison of observed and simulated statistical predictability by Regional Climate models (RCM) and reanalysis in the Northern Hemisphere indicates that small-scale processes that cannot be captured well by RCMs contribute to poor predictability and strong predictive anisotropy in observations. A second idealized mathematical model shows that spatial variability in specifically the minimum directional predictability, resulting from local processes, is the major contributor to predictive anisotropy.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

This research was supported by the Discovery Grant program of the Natural Sciences and Engineering Research Council of Canada. I gratefully acknowledge guidance and advice from my supervisor, Dr. Adam Monahan, through the course of this work. I also would like to thank the rest of my exam committee for their insightful comments. Finally, I would like to thank my parents, Liangzong Mao and Zhiliang Huang, for their encouragement and support over the years, and my grandmother, Yihua Huang, for her unconditional love.

Chapter 1

Introduction

1.1 Overview of statistical prediction of surface winds

Surface winds are an important climate field. High-resolution, site-specific information is required for many applications involving surface winds, such as wind energy production, climate change prediction, and mitigation of airborne pollutants. Large-scale physically based prognostic models, such as general circulation models (GCMs), may not model surface winds with good skill due to limits in model resolution. The resolution of standard GCMs is typically on the order of 100 km (Church et al., 2013), but surface winds are influenced by physical processes ranging from microscale to planetary-scale, not all of which can be represented in any single physically based prognostic model. An alternative approach to modeling surface winds is through statistical prediction in which a transfer function (TF) is constructed based on empirical relationships between surface winds (the predictands) and climate fields that can influence surface winds (the predictors). Having established a statistical relation, predictions can be made when provided with new values of the predictors. The term prediction here refers to using the state of the atmosphere in one region to estimate the state elsewhere at the same time, rather than the prediction of future states. Statistical prediction has the drawback of not being physically based, but it is computationally less expensive to implement than physics-based models, such as regional climate models (RCMs). There are three essential requirements for accurate statistical prediction:

1. adequate representation of predictor variables,

2. strong relationship between the predictors and predictands,

3. adequate modeling of the statistical relationship by the TF.

One example of statistical prediction is statistical downscaling (SD) which attempts to resolve scale discrepancy between large-scale climate change scenarios and small-scale resolution required for prediction of local fields (e.g., surface winds) based on the assumption that large-scale weather exhibits a strong influence on local-scale weather (Maraun et al., 2010). Specifically, the TF of SD relates the local-scale climate variables to observed, reanalyzed or GCM-derived large-scale climate descriptors by a statistical function(Schoof, 2013).

In this study, large-scale climate fields in the mid-troposphere are used as predictors to predict surface wind components. Requirement (1) is considered satisfied as predictors can be taken from GCMs or reanalysis products which have been shown to have good skill in modeling large-scale climate features. Requirement (2) is equivalent to the assumption that synoptic-scale processes have a strong influence on local-scale weather. Whether requirement (3) can be satisfied or not largely depends on the techniques chosen to build TFs. A wide range of techniques are available, and they can be classified into two basic categories according to whether the methodology of the technique is based on regression or classification (i.e., weather type). In regression based techniques, a regression model is developed to model the statistical relationship between large-scale predictors and small scale predictands. The weather type-based approach requires long historical series in order to search for a pattern matching the GCM simulated pattern corresponding to the small-scale input (Schoof, 2013). This study only considers regression based statistical prediction because of the potential limit of available data length.

Statistical prediction of scalar variables such as temperature and precipitation is well studied. However, not much work has been done on statistical prediction of vector variables (e.g., surface winds). The vector nature introduces the possibility of new phenomena as consequences of statistical prediction, and one such phenomenon is predictive anisotropy of surface wind components, that is variability of predictability with the direction of components.

1.2 Literature review

Surface wind speed is the variable of interest in many SD applications (e.g., wind power generation, assessment of storm related infrastructure; Curry et al. 2012), and various quantities related to surface wind speed have been used as predictands. For example, Sailor et al.

(2000) developed a neural network based SD to predict a 10-year averaged daily surface wind speed at three U.S stations. Davy et al. (2010) predicted a sub-daily index of surface wind variability over a period of three years using a random forest SD model at a station in Australia. Cheng et al. (2012) employed SD based on three different regression methods (i.e., multiple stepwise regression, orthogonal regression, and autocorrelation correction regression) to predict daily wind gusts in Ontario, Canada. Some studies have employed SD to model the parameters of the probability distribution of surface wind speed instead of time series of variables related to surface wind speed. For example, Pryor et al. (2005) modeled annual surface wind speed in Northern Europe assuming that the wind speed distribution is well represented by a two-parameter Weibull distribution. The TF used in their study was built by a multi-linear regression. They argued that this method generates accurate depictions of wind climate in terms of probability distribution of surface wind speed in the region of interest. However, the work of Curry et al. (2012) suggests that the good predictability showed by the results of Pryor et al. (2005) is not robust as the apparent skill comes from the dominant influence of region-wide seasonality. By using interannually varying data of predictors and predictands, aggregated into seasons, Curry et al. (2012) showed that applying a similar probabilistic approach of SD with a TF built by a multi-linear regression to predict surface wind speed can only achieve limited success in British Columbia, Canada.

Recent studies suggest that modeling wind components by statistical prediction can usually achieve better predictability than wind speed. van der Kamp et al. (2012) applied SD to predict monthly mean surface wind observations (both mean wind speed and vector wind projections around the compass) at stations in Western Canada and buoys in the Northeast Pacific Ocean. The TF was also based on a multi-linear regression. The predictors were developed from Empirical Orthogonal Function (EOF) analysis of large-scale variables taken from the NCEP-2 reanalysis. Three significant results were derived from their study:

1. the predictability of wind speed measured by cross-validated squared multiple correlation R^2 is generally low and rarely exceeds 0.5;

2. the predictability of wind components is significantly better than that of wind speed;

3. the predictability of wind components is generally anisotropic (i.e., varying with the direction of projected of surface winds).

Result (3) is also termed predictive anisotropy, which is a key phenomenon explored in this study. The results were also found in other recent studies. Monahan (2012), Culver

and Monahan (2013) and Sun and Monahan (2013) all employed SD with multiple linear regression based TFs to predict both wind speed and wind projections from observation stations respectively in the Subarctic Northeast Pacific, central Canada, and in the global ocean. Results (1) and (2) were shown by all three studies in their region of prediction. Monahan (2012), Culver and Monahan (2013) and Sun and Monahan (2013) also showed that predictability of wind speed is poor because wind speed on a given averaging timescale is highly sensitive to the subaveraging time-scale vector wind variability, which is generally poorly predicted, and the predictability of the best predicted wind component provides an upper bound for the prediction skill of wind speed. Salameh et al. (2009) employed an SD prediction of near-surface wind over complex terrain in the Rhone Valley of the French Alps with a TF built by generalized additive models. They found that the predictability of zonal and meridional winds is not equal, and in general, only the component parallel to the significant topographical feature (e.g., valleys, mountain sidewalls) is predicted with good skill, and the other component that crosses the topographical feature is poorly predicted. This phenomenon can also be observed in western Canada where mountainous landscapes are common; specifically, the best predicted wind component is usually aligned with the significant relief, although there are a few exceptions to this general rule (van der Kamp et al., 2012).

Previous studies applying SD to predict surface wind related quantities suggest that the efficiency of SD potentially depends on a number of factors, such as season, model complexity, averaging time scale, and geographic locations. All of these factors were taken into account when developing models for statistical predictions in this study. As discussed by Curry et al. (2012), data with strong seasonality can inflate predictability and render multi-season SD non-robust. This study avoids this problem by removing seasonal cycles of all data considered.

The influence of model complexity on the efficiency of SD is not straightforward. With increased model complexity (i.e., increased number of statistical model parameters), statistical models require a larger amount of data of higher quality to achieve robustness (Culver and Monahan, 2013). If the data set of predictors and predictands is not large enough, arbitrarily high predictability resulting from model fitting the data noise (i.e., overfitting) can result. A range of regression methods can be used to build TFs in statistical prediction, and a basic classification divides regression-based TFs into linear and nonlinear categories. Both categories have been applied to predict surface winds. There is no consensus on which type is more efficient in predicting surface winds. Linear regression based statistical prediction has the advantage of being computationally inexpensive and statisti-

cally robust; it requires fewer model parameters than more complex nonlinear methods, thereby substantially reducing the chances of overfitting (Hsieh, 2009). However, if the predictor-predictand relationship is intrinsically nonlinear, the resultant predictability by linear regression might be limited and there can be benefits to using nonlinear regression models (taking care to minimize the effects of overfitting).

Previous studies show varying results on the relationship between the time-scale of averaging and the predictability of SD. For example, Salameh et al. (2009) predicted six-hourly, daily and weekly averaged zonal and meridional wind components by SD. They found that although the TF is not sensitive to the averaging timescale, longer averaging windows reduce the scatter of the plot of modeled data vs. observed data. Specifically, weekly-averaged data are best predicted in their study. On the contrary, Monahan (2012) and Sun and Monahan (2013) found that the predictability of daily averaged wind variability for daily averaged predictors is better than that of monthly averaged wind variability from monthly averaged predictors. Yuval and Hsieh (2002) provided one example of how linear predictability of climate data can increase with a longer time scale of averaging. Their study showed that averaging data over a longer period could reduce the nonlinearity of the empirical predictor-predictand relationship shown in a shorter timescale of averaging, thereby increasing the linear correlation between observed and modeled results. Their argument was based on the central-limit theorem which states that the sampling distribution of the means of independent random variables y and x connected by a function $y = f(x)$ with arbitrary distributions approaches the normal distribution as the sample size increases, provided that $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ are independent and identically distributed random vectors with finite variance (Yuval and Hsieh, 2002). However, the results of Monahan (2012) suggest that for the relationship between surface wind components and mid-tropospheric predictors the variability on short time scales may contain valuable predictive information which can be lost when the data are averaged over a longer time scale.

Most SD studies have considered prediction of surface wind speed; relatively fewer studies in SD focus on surface wind as a vector variable. However, since surface wind is a vector variable, it is natural to consider predicting the vector form of surface wind. The study described in this thesis was motivated by the interesting characteristics of predictive anisotropy shown in previous studies (Salameh et al., 2009; van der Kamp et al., 2012; Monahan, 2012; Culver and Monahan, 2013; Sun and Monahan, 2013) on predicting surface wind components by SD. In general, most studies applying statistical prediction (both linear and nonlinear) to predict surface winds are based on limited stations and small geographic regions. No previous research has been done on characterizing the statisti-

cal predictability of surface wind components in a systematic approach in order to clarify how statistical predictability is related to various factors such as linear versus nonlinear regression-based TFs, influences of surface homogeneity in different geographic regions, time scale of averaging and potential others; such an analysis is the focus of this study.

1.3 Research question and objective of the study

In general, no comprehensive explanation of predictive anisotropy exists and only some qualitative descriptions of possible causes have been suggested. For example, Salameh et al. (2009) attributed the difference in predictability of zonal and meridional wind components to the role of topographic steering in generating along-valley winds and cross-valley winds; specifically, along-valley winds are strongly coupled with the synoptic-scale flow. Topographic influence on predictive anisotropy is further complicated by the results from Culver and Monahan (2013) and Sun and Monahan (2013). The prediction regions in these two studies are respectively central Canada and the global ocean, both of which feature relatively homogeneous surface conditions rather than mountainous terrains. However, predictive anisotropy is still observed, which indicates that topographic complexity is not the only control on the limitation of predictability. Besides possible physical controls of predictive anisotropy, such as topographic complexity, characteristics of statistical predictability can potentially be influenced by whether the TF is linear or nonlinear.

The overall question addressed by the thesis is what factors can influence characteristics of statistical predictability, such as anisotropy of predictability of surface wind components projected onto compass directions. In particular, are the factors physical or related to the types of regression methods used as transfer functions in the statistical prediction? If they are physical, what can be said about their origin? In order to find the answer to the research question, this study aims at providing a global characterization of statistical predictability of surface wind components and building a general framework to explain characteristics of statistical predictability with an emphasis on predictive anisotropy.

1.4 Outline of the research

I will approach the objective from two directions. The first direction is to examine whether predictive anisotropy is due to physical factors. In this study, the ratio of minimum to maximum directional predictability is used as a measure of predictive anisotropy

at 2109 land stations across the globe (Fig. 1.1). Most stations are concentrated in mid-latitudes of the Northern Hemisphere due to data availability. In general, it is difficult

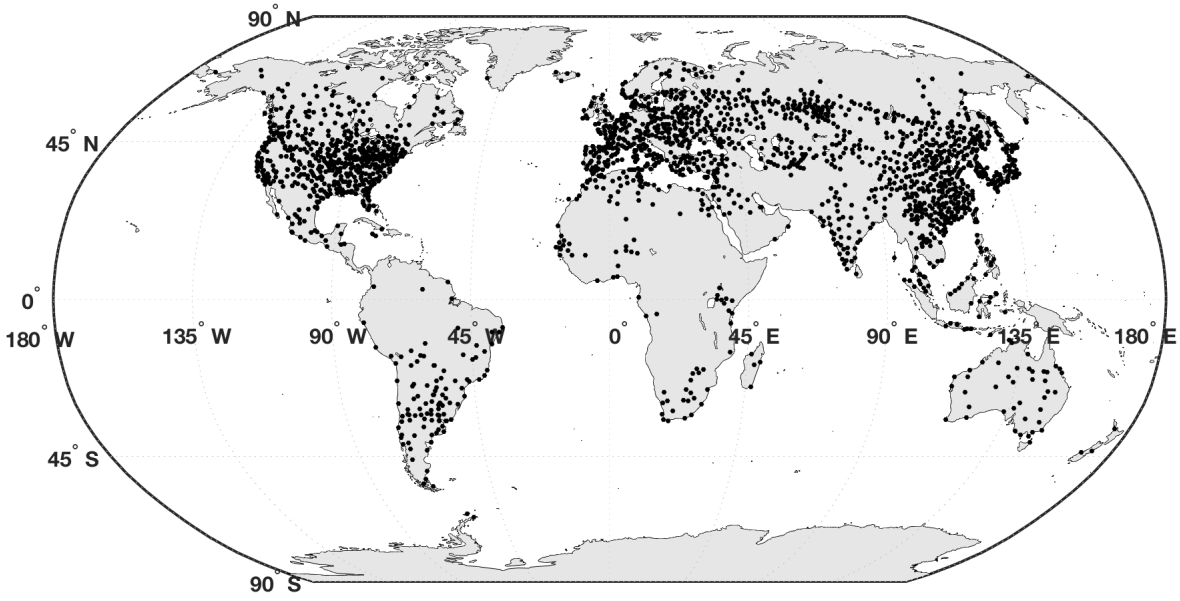

Figure 1.1: The locations of the 2109 land stations used for statistical prediction of surface winds.

to directly identify the physical factors that can contribute to predictive anisotropy, other than topographic complexity which has been discussed in some previous studies. As the TF relates the large-scale free-tropospheric state to small-scale surface winds, I focus on the statistical properties that can characterize some aspects of the empirical relationship between predictors and predictands instead of directly considering the physical mechanisms that can influence predictive anisotropy. First, the standard deviation of surface wind components is considered because it is natural to hypothesize that the overall variability of surface wind components contains both predictable "signal" and unpredictable "noise" associated with the large-scale climate and small-scale local wind systems. Second, the predictor-predictand relationship will be optimally modeled by a linear regression given that both the predictand and predictors are Gaussian, and non-Gaussianity is a potential indication of reduced linear predictability (Yuval and Hsieh, 2002). Since values of kurtosis are related to the Gaussianity of the data, I also consider kurtosis of surface wind components as a factor in influencing characteristics of predictability. Third, since the surface wind components are vectors, I also consider statistical properties related to the direction of predictability. In this regard, the orientations of the mean wind vector and the maximum or minimum directional predictability may be determined by some common underlying

physical mechanisms.

Furthermore, the difference in observed and simulated predictability by RCMs and reanalysis products can indicate how well the physical processes controlling this statistical relationship are simulated. Finally, idealized mathematical models for the statistical predictability of surface wind components are also developed to provide a conceptually organizing perspective on the factors that can influence statistical predictability of surface wind components.

Besides possible physical factors, the predictability of wind components could also be influenced by the regression methods used to develop the TFs. With the exception of Salameh et al. (2009), the studies addressing predictive anisotropy of surface wind components by statistical prediction discussed in 1.2 are all based on multi-linear regression. Therefore, the second direction of my research is to examine whether statistical predictions with nonlinear TFs are more effective in predicting surface wind components in terms of improved predictability and weakening of predictive anisotropy. This question is addressed by systematically comparing linear and nonlinear statistical prediction of surface wind components at the 2109 stations across the globe. If the results show that nonlinear regressions have no distinct advantage over linear regression, this will indicate that limited predictability and/or predictive anisotropy are not related to the use of linear or nonlinear regression methods for developing the TF. On the other hand, substantial differences would indicate that some directions of projections are characterized by stronger nonlinear predictor-predictand relationships than others; in this case, I would conclude that predictive anisotropy is caused by the use of linear regression which limits nonlinear predictability along some directions of projections.

1.5 Structure of the thesis

The thesis is constructed as follows.

- Chapter 2: Summary of the data of predictands and predictors as well as the methods of deriving predictability in this study.
- Chapter 3: The characteristics of the statistical predictability of surface wind components by linear regression at 2109 land stations across the globe are summarized. The relationships between characteristics of statistical predictability and topographic complexity as well as some statistical properties of surface wind components are

investigated. Specifically, the statistical properties of interest are the standard deviation and kurtosis of surface wind components, which measure the variability and Gaussianity of surface wind components respectively. The relationships between predictability and the statistics of distribution of surface wind components are described using an idealized statistical model of large-scale and local influences on surface winds. This study has been published as: Mao and Monahan (2017): "Predictive anisotropy of surface winds by linear statistical prediction." *Journal of Climate* 30.16 (2017): 6183-6201.

- Chapter 4: Statistical predictability by three common nonlinear regression methods (neural networks, support vector machines, and random forests) at the 2109 land stations across the globe is compared with the results of linear regression. This study has been submitted to *Climate Dynamics* with the title "Linear and nonlinear regression prediction of surface wind components".

- Chapter 5: Observed statistical predictability of surface wind components at the 2109 land stations is compared with that simulated by RCMs and a reanalysis product in three domains (North America, Europe-Mediterranean Basin, and East Asia). An idealized mathematical model is used to characterize the predictability of wind components along arbitrary axes in order to conceptually organize the influences of large- and local-scale physical processes on characteristics of statistical predictability. This study has been submitted to *Journal of Applied Meteorology and Climatology* with the title "Comparison of linear predictability of surface wind components from observations with simulations from RCMs and reanalysis".

- Chapter 6: Conclusions regarding the factors that can influence characteristics of statistical predictability and contribute to the predictive anisotropy of surface wind components are given, and an outlook for future research is also provided.

Contributions from Co-author I formulated the analysis, carried out the computations, analyzed results and wrote the papers. Dr. Monahan provided advice for all analysis throughout the study. Moreover, he formulated the first idealized descriptive model described in Section 3.4.2.

Chapter 2

Data and methods of statistical prediction of surface wind components

2.1 Predictand data

I consider statistical predictions of observed surface wind components at a network of 2109 land stations over the 1980-2012 time period (Fig. 1.1). While station data are available from all continents, they are concentrated in the middle latitudes of the Northern Hemisphere. The predictands (the surface wind components) at each station are derived from observational data of hourly wind speed (w) and direction (ϕ) at 10 m above the ground during a 2-minute period ending at the beginning of the hour. Direction is where the wind comes from, measured clockwise from north. The time period of the data is from 1980/01/01 to 2012/12/31. These data were obtained using the WeatherData function of Mathematica 9.0 (Wolfram, 2016), which includes a wide range of data sources. Chief among these sources are the National Weather Service of the National Oceanic and Atmospheric Administration (NOAA), the United States National Climatic Data Center, and the Citizen Weather Observer program. Only stations with fewer than 10% missing data for the period under consideration are considered. Specifically, zonal (u) and meridional winds (v) are derived from the original hourly wind speed and direction data as following:

$$u = -w\sin(\phi), \tag{2.1}$$

$$v = -w\cos(\phi). \tag{2.2}$$

The wind component projected onto direction θ is then calculated as

$$U(\theta) = u\sin(\theta) + v\cos(\theta), \tag{2.3}$$

with θ varying from 0° to 170°. In total, there are 36 surface wind components as predictands at each station. Only 18 of these are distinct, because $U(\theta) = -U(\theta + 180^\circ)$. The mean seasonal cycle of $U(\theta)$ at each station is subtracted before the statistical prediction models are constructed. The seasonal cycles are obtained using the harmonic fit:

$$Y_s(t) = B_0 + B_1\cos(\omega t) + B_2\sin(\omega t) + B_3\cos(2\omega t) + B_4\sin(2\omega t) + B_5\cos(3\omega t) + B_6\sin(3\omega t), \tag{2.4}$$

where $\omega = 2\pi/P$ with $P = 365$ days for daily averaged time series (after removing data of Feb 29th for convenience), and $P = 12$ months for monthly averaged time series. Including a larger number of harmonics in the seasonal cycle has essentially no effect on the resulting regression models (not shown). In order to minimize the influence of any remaining seasonality on the statistical relationship, statistical models are fit separately for the winter and summer seasons. Winter is defined as December, January, and February for Northern Hemisphere stations and June, July, August for Southern Hemisphere stations. The assignment of these months is reversed in summer.

2.2 Predictor data

The four mid-tropospheric meteorological fields of temperature T, geopotential height Z, zonal wind U, and meridional wind V at 500 hPa are chosen as the predictors. They are obtained from NCEP-Reanalysis 2 data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from http://www.esrl.noaa.gov/psd/ (Kanamitsu et al., 2002). Previous studies (Culver and Monahan, 2013; Monahan, 2012) have shown that the correlation structures between surface wind vectors and large-scale free tropospheric climate variables are often spread across a large area surrounding a station, such that the locations of highest predictability aloft are often not directly above the surface station. For computational efficiency, mid-tropospheric predictors at a given station are selected from a domain of fixed size centered at the station. In order to assess the influence of domain size on the statistical predictability of surface wind components, I compare the predictions from a $40^\circ \times 40^\circ$ domain with that from a smaller domain $22.5^\circ \times 22.5^\circ$ for a random sample of 100 stations, using all four regression methods considered (linear regression denoted LR; neural network

denoted NN, support vector machine denoted SVM and random forest denoted RF). For all regression models, the difference between the two domains is negligible for daily averaged prediction but is sometimes substantial for monthly averaged prediction. The difference is reasonable since the spatial scale of correlation between surface fields and atmospheric flow aloft is larger for longer averaging time scales. For daily averages, the correlation structures are on the synoptic scale, while for monthly averages they have the larger scale of atmospheric low-frequency variability. Based on these test results, predictors of statistical predictions are derived from the reanalysis fields within a $22.5^\circ \times 22.5^\circ$ grid box centered at the location of each weather station for daily averaged prediction, and a larger $40^\circ \times 40^\circ$ grid box is used to derive predictors for monthly averaged prediction. Since the resolution at which tropospheric variables are available from the NCEP II reanalysis is $2.5^\circ \times 2.5^\circ$, the prediction domain consists of 81 grid points (i.e. 9 grid points on each side) for daily averaged predictions and 256 grid points (i.e. 16 grid points on each side) for monthly averaged predictions. For each grid point in the domain, time series of $(U_{ij}, V_{ij}, T_{ij}, Z_{ij})$ at 500 hPa are used to predict $U(\theta)$ at the station. The seasonal cycles of the time series of predictors estimated at each grid point from the harmonic function Eq. (2.4) are subtracted before the statistical prediction models are constructed. The deseasonalized time series are then scaled by their individual standard deviations in order to obtain standardized predictors.

The true length scale of the structure of correlation between surface wind components and the large-scale predictors is unknown and may vary with locations, and it is possible that the length scales in the zonal and meridional directions differ from each other. I choose to fix the size of the domain rather than selecting a different domain for each station because it is desirable to have a transparent and systematic way of estimating predictability that can be applied to all stations in order to minimize the chances of overfitting. We also test whether Lasso regression would be more efficient as it can automatically select predictors within a large domain. The number of predictors selected by Lasso is generally large. The selected predictors are likely to be highly correlated with each, which could be a source of inflated predictability. There is no guarantee that the automatically selected predictors are truly relevant factors in explaining variability of surface wind components since Lasso may not distinguish true predictors from irrelevant variables but highly correlated with predictors with any amount of data and any amount of regularization Zhao and Yu (2006). Moreover, fitting a nonlinear model with such a large number of predictors is even more prone to overfitting. Therefore, the method of automatically selecting predictors in a large domain is not considered in this study.

2.3 Measure of predictability

Predictions in this study are obtained using the modeled statistical relationship between deseasonalized predictands and predictors at each grid point using both linear and nonlinear regression methods. Statistical predictability is assessed using leave-one-year-out cross-validation. Specifically, for all four models considered in this study (linear regression, neural network, support vector machine, and random forest), for each year of observations the prediction model is estimated using data from the other 32 years. The process is repeated 33 times to obtain time series of predicted surface wind components. The resulting predictability at each grid point (i,j) is measured by squared correlation ${R^2}_{ij}$,

$$ {R^2}_{ij}(\theta) = |corr(U(\theta), \widehat{U}_{ij}(\theta))|^2 \tag{2.5} $$

where $\widehat{U}_{ij}(\theta)$ is the predicted time series using the four predictors (U,V,T and Z) at the grid point (i,j). A single measure of predictability across the domain is then computed for each method, denoted Π:

$$ \Pi(\theta) = < {R^2}_{ij}(\theta) > . \tag{2.6} $$

The average calculated in Eq. (2.6) is taken over the two (for daily prediction) and four (for monthly prediction) grid points with the top values of ${R^2}_{ij}(\theta)$ within the prediction domain (corresponding to 2% of the grid points in the domain). In computing Eq. (2.6), only grid points for which $corr(U(\theta), \widehat{U}_{ij}(\theta)) > 0$ are considered. Occasionally, the cross-validated correlation between the predicted and observed wind components can become negative. Negative correlations between prediction and predictand may result from large sampling fluctuations and low intrinsic predictability (non cross-validated R^2 values are small in these cases), so I exclude them from the analysis.

The directional predictability values obtained from LR, NN, SVM, and RF are denoted $\Pi_{LR}(\theta)$, $\Pi_{NN}(\theta)$, $\Pi_{SVM}(\theta)$ and $\Pi_{RF}(\theta)$ respectively. The overall predictability at a station is measured by the mean of Π over all directions, denoted $\overline{\Pi(\theta)}$, and predictive anisotropy is measured by

$$ \alpha(\Pi) = \frac{min(\Pi)}{max(\Pi)}, \tag{2.7} $$

where $min(\Pi)$ and $max(\Pi)$ are respectively the minimum and maximum $\Pi(\theta)$ over the 36 values of θ. Values of $\alpha(\Pi)$ range between 0 to 1, such that lower $\alpha(\Pi)$ indicates a stronger degree of anisotropy.

Chapter 3

Predictive anisotropy of surface winds by linear statistical prediction

3.1 Introduction

Near surface winds are important in problems such as air quality, engineering design and renewable energy. Near surface winds are influenced by processes on spatial scales from the microscale to planetary scales, not all of which are resolved (or simulated well) by any given physically-based prognostic model. For example, atmospheric-ocean general circulation models (AOGCMs) are useful in assessing large-scale climate systems in response to changes in natural or anthropogenic forcing. However, the resolution of AOGCMs is rarely finer than $1^\circ \times 1^\circ$; therefore, they are unable to explicitly resolve small-scale processes such as those related to local topography (Schoof, 2013). Therefore, it is useful to explore the statistical predictability of surface winds. The effectiveness of statistical prediction depends on the strength of the statistical relationship between small-scale surface winds and large-scale free-tropospheric climate variables. One example of the use of such models is statistical downscaling (SD) based on the assumption that synoptic-scale weather has a strong influence on local-scale weather (Maraun et al., 2010). For statistical prediction, a transfer function is built to link predictands (e.g. surface winds) to predictors (e.g. free-tropospheric climate variables). A range of statistical and machine learning methods can be used to derive the transfer function such as linear regression, generalized linear and additive models, and various nonlinear regression models. Relatively few studies have focused on using statistical prediction to model physically important vector variables such as surface winds, considering in particular the directional structure of predictability. The

focus of this chapter is to investigate the statistical relationship between components of surface winds (predictands) and free-tropospheric climate variables (predictors) in order to assess the linear statistical predictability of near surface wind vectors. To this end, I use linear regression to derive the transfer function between observed surface wind components and mid-tropospheric climate variables at a large number of observational stations across the world, and the strength of the statistical relationship between predictands and predictors is assessed by the resulting predictability. Predictive anisotropy, which represents the unequal strength of the predictor-predictand relationships for surface wind components projected onto different directions, is of particular interest.

A few previous studies have considered the predictive anisotropy of surface winds. For example, Salameh et al. (2009) applied a generalized additive model as the transfer function to predict surface zonal (u) and meridional wind components (v) from stations located in valleys of the French Alps, and found that in general only one of u and v can be predicted well. Other studies have used linear regression based transfer functions to predict surface winds in Western and Central Canada (van der Kamp et al., 2012; Culver and Monahan, 2013) and at buoys located over the ocean (Monahan, 2012; Sun and Monahan, 2013). These studies found that the predictability of wind components projected onto different compass directions generally exhibits predictive anisotropy. In addition, the best or worst predicted wind component is not always the conventional zonal or meridional component. Knowledge of the predictability of u and v alone is not sufficient in general to assess the predictive anisotropy and the potential utility of statistical prediction at a station; it is necessary to know the predictability of wind components all all directions from 0° to 179° (as the projection along θ is negative that along $180^\circ + \theta$). Note that in this discussion of the predictability of wind components, I am referring to projections of wind vectors onto a coordinate axis rather than wind coming from a specified direction. Although it is possible to estimate the conditional predictability of wind based on the direction of flow (e.g. the predictability of northerly or southerly winds), such an analysis is of limited practical utility as this conditional predictability assumes the direction of wind is known but not the speed. Furthermore, while any vector component can be expressed as a linear combination of the zonal and meridional components, the best-fit statistical predictive model for this combination will not generally be the corresponding linear combination (with the same weights) of the individual regression models of the components.

The present study has two objectives. The first is to characterize the predictor-predictand relationship by applying linear regression based statistical prediction to a large dataset of station-based surface winds. The second is to explore the relationship between statistical

predictability of surface wind components and some potential influential factors. I consider three types of factors in this study: (1) topographic complexity (2) statistical properties of wind component fluctuations and (3) directions of mean wind vectors. The statistical properties here refer to the magnitude and shape of the probability distributions of wind components, respectively measured by the standard deviation and kurtosis. A fundamental question of this study is whether the characteristics of predictability of surface wind components at a station can be associated with these factors.

This study considers empirical relationships between wind component predictability and potential explanatory factors, rather than the physical mechanism responsible for these relationships. As surface heterogeneity is a natural candidate cause of predictive anisotropy, it is natural to consider the influence of topography. While the mean wind vector cannot directly relate to predictability of its components (as the linear regression models are based on fluctuations of anomalies with the means subtracted), common underlying physical mechanisms may determine the orientation of the mean wind and the anisotropy of predictability. The second factor considered in this study is standard deviation of surface wind components based on the hypothesis that the overall variability of surface wind components will contain both predictable "signal" and unpredictable "noise" associated with the relative influences of large-scale and local atmospheric circulations. Kurtosis of surface wind components is considered because linear regression models should be optimal when the predictand and predictors are all Gaussian (Yuval and Hsieh, 2002), so non-Gaussianity might reduce linear predictability.

Previous studies have shown that predictive anisotropy is observed in regions characterized by complex terrain, such as mountainous regions (Salameh et al., 2009; van der Kamp et al., 2012). van der Kamp et al. (2012) argue that there is no straightforward relationship between directions of best predicted wind components and the frequency distribution of wind directions in Western Canada. The results of these studies are derived from a small number of stations located within a limited geographic region. My results will be based on a much larger number of stations over a broader geographic range. However, these operational meteorological stations are not uniformly distributed across the land surface and are most densely concentrated in the Northern Hemisphere extratropics.

This chapter is organized as follows. Section 3.2 presents statistical methods used in analyzing statistical relationships between predictability and potential factors. Section 3.3 explores the characterization of statistical predictability and the aforementioned factors for all stations considered. Section 3.4 presents a discussion, and introduces a simplified statistical model synthesizing the results of this analysis. Conclusions are given in Section

3.5.

3.2 Methods of statistical analysis with potential factors

3.2.1 Measures of topographic complexity

For my purpose, I seek the relationship between characteristics of predictability and variability of terrain. The terrain data in this study are obtained from 1 arc-minute global relief data H from the ETOPO1 Global Relief Model downloaded from `https://www.ngdc.noaa.gov/mgg/global/global.html` (Amante and Eakins, 2009). Different measures of topographic complexity have been proposed, none of which is clearly optimal (Lu, 2008). Therefore, I choose a simplified approach, using statistics of local relief to represent topographic complexity.

Two common measures of topographic variability in a specified domain are the difference between the maximum and minimum elevations $\Delta H = H_{max} - H_{min}$, and standard deviation $\sigma(H)$ over the domain (Lu, 2008). In this study, a circular region with radius of 0.2° arc length centered at each weather station is chosen as the domain used to characterize topographic complexity. The small radius 0.2° arc length is chosen subjectively from empirical tests, which show that the resultant measure of topographic complexity in this study is not particularly sensitive for domain sizes smaller than 0.5°. One example is shown in Fig. 3.1.

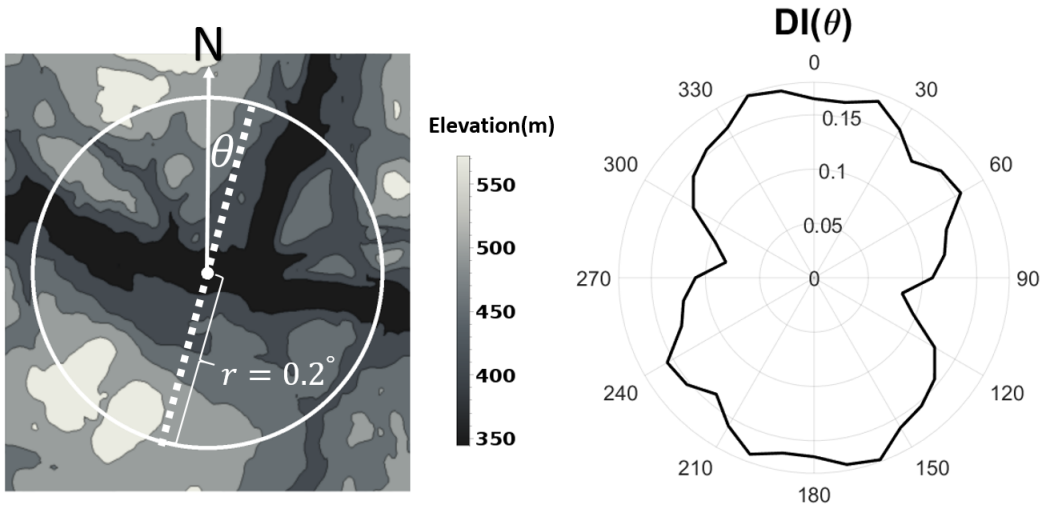

Figure 3.1: Schematic illustrating the quantification of topographic complexity at a representative station indicated by the white dot in the center of the left panel. The station is located at Kamloops Airport in BC, Canada (50.70°N, 120.44°W). The grayscale indicates the topographic relief (in meters). The dotted line represents all elevation points along the diameter for a chosen θ from north within a circle of 0.2° radius, where θ indicates compass directions ranging from 0° to 170°. The right panel shows the corresponding polar plot of directional index of topographic complexity (Eq. (3.3)).

The index for measuring topographic complexity I_{TP} at each station I will use is calculated by combining ΔH and σ_H within the domain. First, I normalize ΔH and σ_H for all 2109 stations to the range of 0 to 1; specifically, ΔH is normalized as follows:

$$N_{\Delta H} = \frac{\Delta H - \Delta H_{min}}{\Delta H_{max} - \Delta H_{min}}, \tag{3.1}$$

where ΔH_{min} and ΔH_{max} refer to the largest and smallest values among all 2109 stations. The quantity N_σ is obtained in the same way. The index I_{TP} at an individual station is then computed by averaging the corresponding $N_{\Delta H}$ and N_σ:

$$I_{TP} = \frac{N_{\Delta H} + N_\sigma}{2}. \tag{3.2}$$

I_{TP} ranges between 0 and 1 with larger I_{TP} corresponding to more complex topography surrounding a station. Although simple, values of I_{TP} can distinguish mountainous regions from relatively flat regions.

Besides the overall topographic complexity at a station represented by I_{TP}, the variabil-

ity of terrain along each compass direction surrounding a station is also considered. The approach to quantifying directional variability of topography is the same as that used to calculate I_{TP}, and I use $DI(\theta)$ to denote topographic complexity along each compass direction θ. For each direction of θ from 0° to 170° at each station, I first obtain elevation data on the diameter oriented along θ within the domain, denoted $[H(\theta,r),H(\theta+\pi,r)]$ where r is the distance from the station to a point along θ within the domain, and r is between 0° and 0.2° (Fig. 3.1). $\Delta H(\theta)$ and $\sigma_H(\theta)$ refer to the difference along the range of r considered between maximum and minimum elevation as well as standard deviation of $[H(\theta,r),H(\theta+\pi,r)]$. The normalized values $N_{\Delta H}(\theta)$ and $N_\sigma(\theta)$ along each direction θ obtained by equations similar to Eq. (3.1) are then used to calculate the directional variability of topographic complexity denoted $DI(\theta)$ by Eq. (3.3),

$$DI(\theta) = \frac{N_{\Delta H}(\theta) + N_\sigma(\theta)}{2}. \tag{3.3}$$

The polar plot of $DI(\theta)$ shown in Fig. 3.1 clearly characterizes the topographic features of this example station. In particular, the orientation of the valley in which the station is situated is the direction of minimum $DI(\theta)$, and larger values of $DI(\theta)$ correspond to across-valley directions along which terrain is more variable.

Besides the index of topographic complexity, I will consider the standard deviation and kurtosis of the wind component fluctuations (respectively $\sigma(\theta)$ and $kurt(\theta)$) and the orientation of the mean wind vector $(\overline{u},\overline{v})$ as potential influential factors. The factors potentially influencing $U(\theta)$ that I consider are not necessarily independent because they may be linked by some common physical origin. Anisotropy in fluctuations of the wind components could conceivably be related to the mean wind; both of these are expected to be influenced by local topography. However, to simplify the objectives in this study, I will not explore the dependency among different factors, but will assess the influence of each individual factor. The influence of interaction among these factors, such as mean wind flow and topography, on predictability of surface wind components is an interesting direction for future study.

I use two approaches to assess the relationship between predictability and potential influential factors in this study. First, relationships between predictability and any given factor are investigated using kernel density estimates of the conditional probability distribution of quantities related to predictability. By definition, the conditional distribution of

the random variable Y given X is

$$P(Y|X) = \frac{P(Y,X)}{P(X)}, \tag{3.4}$$

where $P(Y,X)$ is the joint probability density, and $P(X)$ is the probability density of X. In this study, Y stands for quantities related to predictability and X represents quantities related to chosen factors. Distributions of $P(X)$ are also plotted with $P(Y|X)$ to provide an indication of the number of stations entering the estimate of $P(Y|X)$.

Second, for each individual station, directional relationships between predictability and chosen factors for all 36 wind components are assessed using either the Spearman rank correlation coefficient $\rho(\Pi(\theta),X(\theta))$ (where $X(\theta)$ refers to a directional factor) or the dot product between two unit vectors related to factors and predictability. The Spearman rank correlation is used to assess the directional relationship between predictability and the factors $DI(\theta)$, $\sigma(\theta)$ and $kurt(\theta)$, while the dot product is used to assess the directional relationship with the mean surface wind vectors. The quantity ρ measures the degree of similarity between the orientation of predictability and the chosen factors of 36 wind components at a station. Positive (negative) $\rho(\Pi(\theta),X(\theta))$ means that directions of larger (smaller) predictability correspond to larger values of the factor. A larger magnitude of rank correlation coefficient indicates a stronger directional relationship between predictability and the chosen factor of wind components at a station. $\rho(\Pi(\theta),X(\theta)) = 0$ means that there is no directional relationship between predictability and factors of wind components. Values of the dot product of two unit vectors range between 0 (orthogonal) and 1 (parallel). I do not consider negative values of dot products, as the component along θ is just the negative of the component along $\theta + 180^\circ$.

3.3 Results

This section first displays characteristics of the statistical predictability of surface wind components. An exploratory analysis of relationships between predictability and the three chosen factors is then presented and discussed.

3.3.1 Geographic distribution of predictability

In the following analysis, magnitudes of predictability will be characterized by $min(\Pi)$ and $max(\Pi)$ and the corresponding anisotropy $\alpha(\Pi)$ defined by Eq. (2.7). The correspond-

ing directions of maximum predictability represented by unit vector $\hat{e}_{\Pi_{max}}$ are used to show directional characteristics of predictability.

Magnitude of predictability

Maps of the quantities $min(\Pi)$, $max(\Pi)$, and $\alpha(\Pi)$ for all stations considered are displayed in Fig. 3.2 and Fig. 3.3 (for daily and monthly averaging timescales respectively). Inspection of these figures indicates that the characteristics of predictability considered are spatially organized.

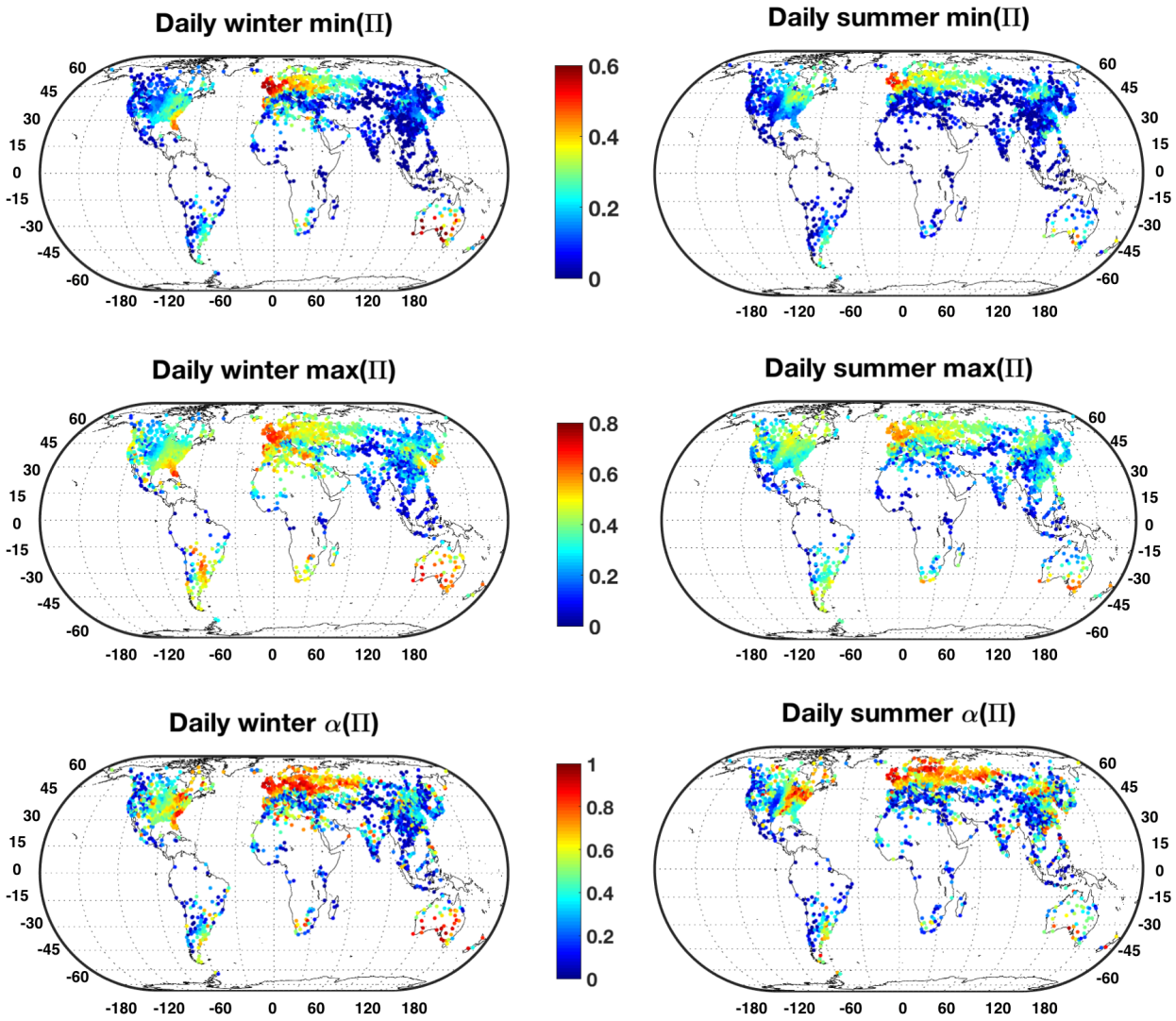

Figure 3.2: Minimum, maximum and anisotropy of predictability resulting from multiple linear regression based statistical prediction for daily averaged winter and summer observations.

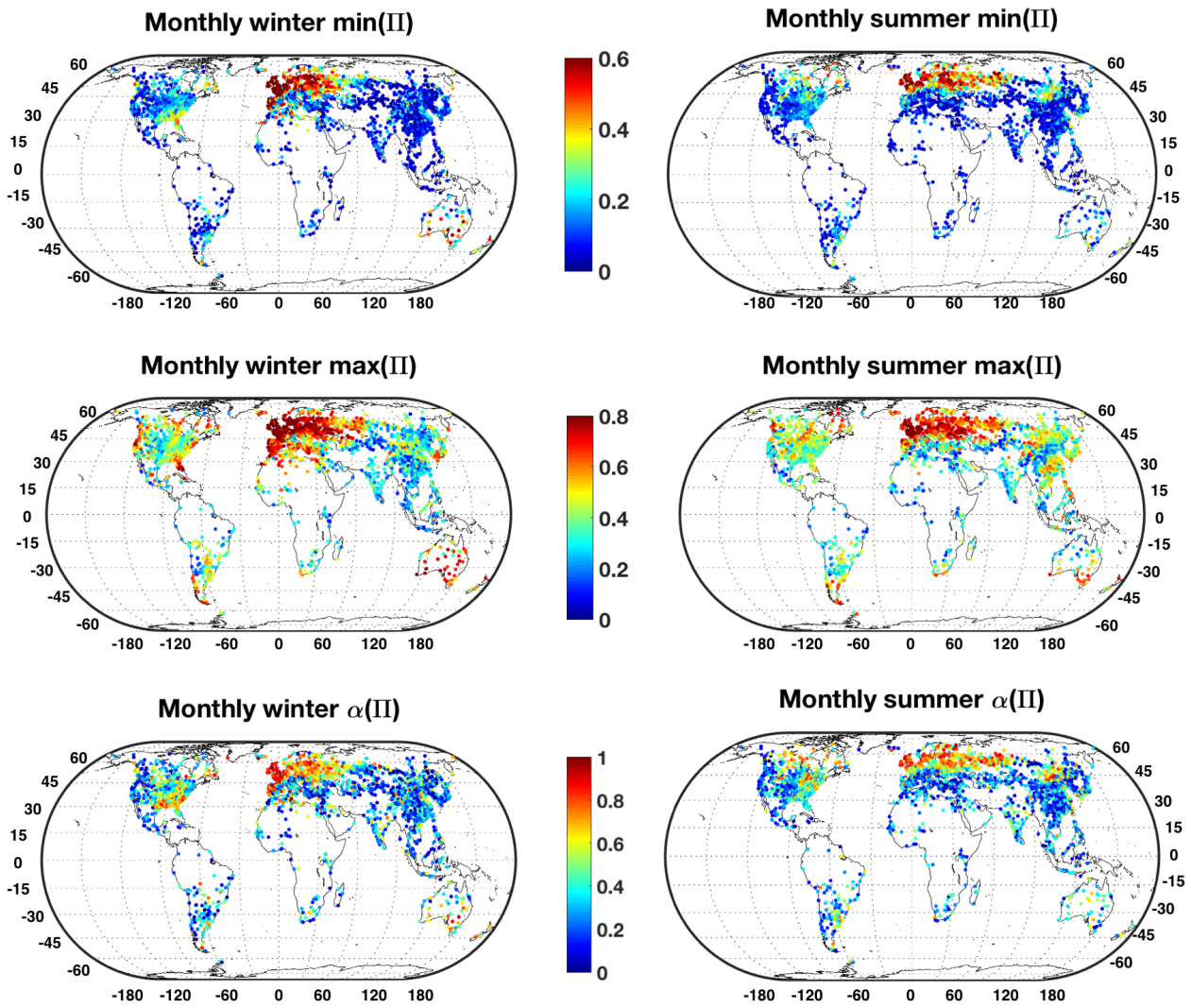

Figure 3.3: As in Fig. 3.2 for monthly averaged winter and summer observations

Table 3.1 summarizes the number of stations with strong predictive anisotropy (i.e. $\alpha < 0.3$) for different cases of prediction in each region. One can see that strong predictive anisotropy is a common feature across the globe with the interesting exception of a band in Eurasia stretching from 10°W to 90°E where only a small fraction of stations show strong predictive anistoropy for all cases of prediction.

Figures. 3.2 and 3.3 show evidence of some relation of these measures to topographic complexity. For instance, the distribution of predictability across the continent of North America demonstrates that lower predictability and stronger predictive anisotropy are more commonly found in the mountainous regions of the West relative to the rest of the continent. Similarly, the predictability along the west coast of South America (dominated by the

Table 3.1: Number of stations with strong predictive anisotropy (i.e. $\alpha < 0.3$) in different regions

regions	total # regional stations	daily summer	daily winter	monthly summer	monthly winter
North America	570	185	116	240	193
South America	105	58	53	51	48
Africa	134	93	67	84	65
Eurasia	390	30	33	51	71
East Asia	518	233	325	294	328
Australia &New Zealand	36	11	3	9	6

Andes) is lower and the anisotropy is stronger than in the rest of South America. However, low or anisotropic predictability also occurs well away from mountainous regions.

The comparison of monthly and daily averaged predictions shows that there are more stations with higher overall monthly predictability than daily predictability. This is observed in both summer and winter results. Specifically, $min(\Pi)$ is larger for monthly averaged data than for for daily averaged data at approximately 60% of the stations. Similarly, about 80% of stations have larger $max(\Pi)$ for monthly averaged data than for daily averaged data. Features related to prediction using winter and summer data are similar.

Direction of predictability

The directional characteristics of predictability also vary by region. In some regions, the orientation of predictability shows no coherent spatial structure while in others it is evidently organized. For instance, the directions of $max(\Pi)$ (denoted $\hat{e}_{\Pi_{max}}$, generally orthogonal to the direction of $min(\Pi)$) at some stations in eastern North America, show evidence of large-scale organization which changes from summer to winter (Fig. 3.4).

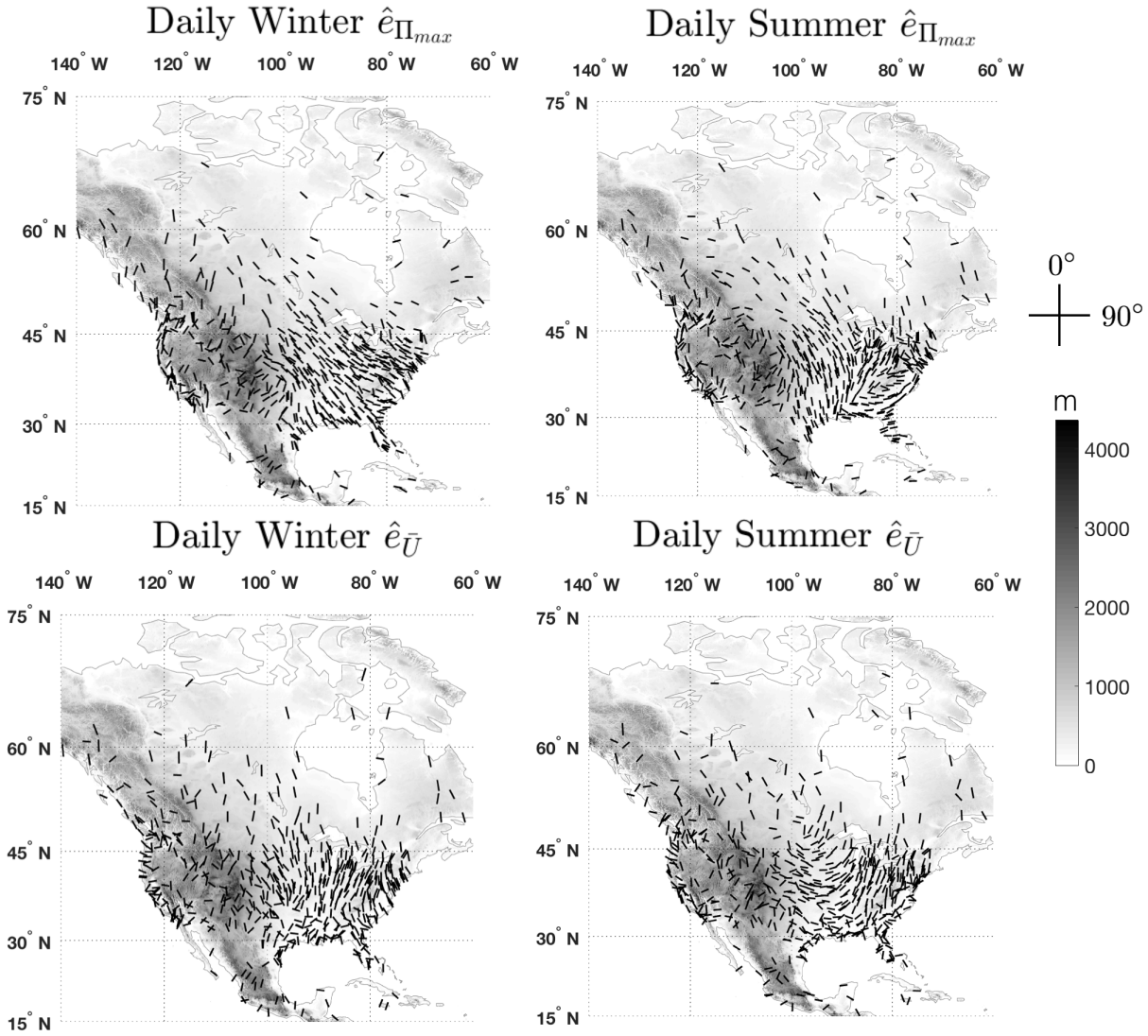

Figure 3.4: Directions of maximum predictability indicated by $\hat{e}_{\Pi_{max}}$ and directions of mean surface wind vectors indicated by $\hat{e}_{\bar{U}}$ for daily winter and summer prediction in North America. The grayscale indicates elevation of topography. The coordinates indicate the compass directions of 0° and 90°.

In contrast, the orientation of wind predictability in the mountainous western North America shows no evidence of large-scale organized spatial structure (away from the western coastlines). The orientations of the time-averaged wind vectors $\hat{e}_{\bar{U}}$ are also shown in Fig. 3.4. While maximum predictability is aligned along the mean wind direction in some circumstances (such as from Central Canada along the Mississippi River to the Gulf of Mexico in summer), in other circumstances these vectors are almost orthogonal (as in the

Great Plains in summer) or the relative orientations show no systematic pattern (as in the Western Cordillera). The global-scale relationship between the orientation of predictability and the mean wind will be discussed in Section 3.3.4.

While large-scale organization of the orientation of predictability is not as evident over other continents as in North America, some features related to orientation of predictability can still be identified. For example, directions of $max(\Pi)$ at stations located in near-coastal areas are often nearly parallel or perpendicular to the coastline, and this feature is more evident for daily averaged summer predictions.

3.3.2 Case studies

To illustrate the range of relationships between predictability of surface wind components and potential explanatory factors in different settings, the predictability of wind components is considered for five stations (Fig. 3.5 and Table 3.2) representative of different characteristics of predictability, together with their surrounding terrain features and the statistical properties of surface wind components, $\sigma(\theta)$ and $kurt(\theta)$.

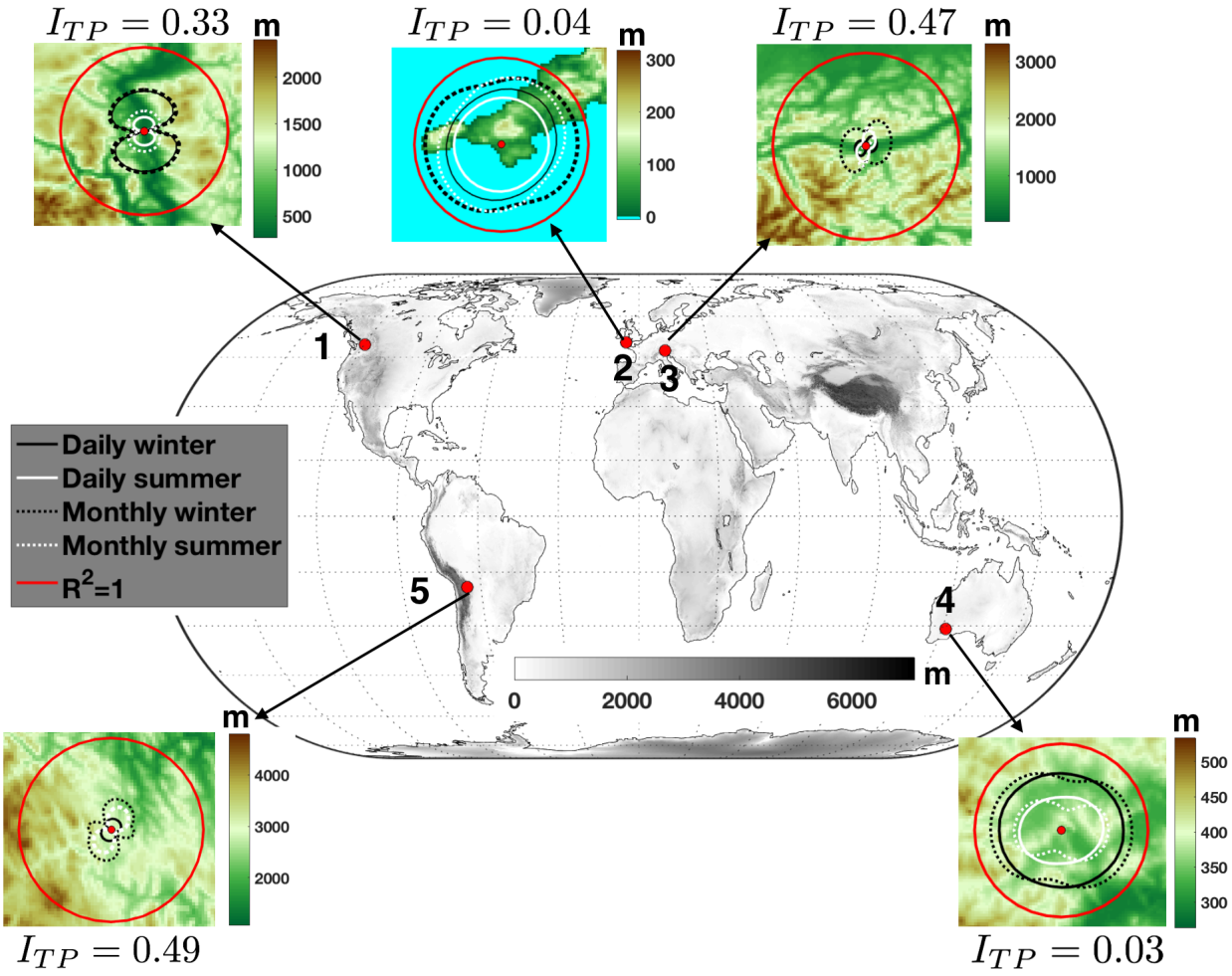

Figure 3.5: Predictability of wind components $\Pi(\theta)$ derived from linear regression based statistical prediction for five representative stations. The polar plots of $\Pi(\theta)$ are superimposed on terrain relief with blue color indicating water bodies. The overall topographic complexity measured by I_{TP} (Eq. (3.2) is also labeled at each station.

Table 3.2: Meteorological stations used for case studies. The name refers to the international ID of the stations

Station	Name	Location	Latitude(°)	Longitude(°)	Elevation(m)	Topographic complexity (I_{TP})
1	CYYF	Penticton Airport,Canada	49.463	-119.602	344	0.33
2	EGDR	Culdrose Airport, UK	50.084	-5.257	82	0.04
3	LOWI	Innsbruck Airport, Austria	47.26	11.344	581	0.47
4	YPKG	Kalgoorlie-Boulder Airport, Australia	-30.785	121.453	365	0.03
5	SLPO	Captain Nicolas Rojas Airport,Bolivia	-19.007	-65.289	3934	0.49

Stations 1, 3 and 5 are located in mountainous terrain. These three stations are characterized by strong predictive anisotropy; in particular, stations 3 and 5 are characterized by relatively low $max(\Pi)$ for both seasons and time scales of averaging, whereas $max(\Pi)$ in station 1 is relatively high in winter and low in summer. Both stations 1 and 3 are located in valleys. The directions of maximum/minimum predictability at station 1 are oriented to the directions of along/cross-valley for both seasons and timescales of averaging. These orientation features are not surprising given the local topography: the valley at station 1 is narrow, and the predictability features suggest that the connection between large-scale flow and surface flow is strongly determined by funneling of winds. However, for station 3, the directional relationships between the orientation of maximum/minimum predictability and along/cross-valley directions vary with season; while the direction of maximum predictability is along-valley in winter, this direction is substantially rotated away from the valley axis in summer. In other words, the direction of maximum predictability in summer is near the direction of across-valley, which is counter-intuitive of the role funneling may play in connection between large-scale flow and surface flow.

Although stations 1, 3, and 5 are all expected to be influenced by local wind systems associated with mountainous terrain, there are differences among the characteristics of predictability at the three stations. This suggests that there is no single universal explanation for the observed characteristics of predictability. Terrain features may be one factor, but it is not the only one. The overall predictability is high and anisotropy is weak at stations 2 and 4 located in relatively flat terrains as indicated by small I_{TP}. Characteristics of predictability at stations 2 and 4 also differ from each other; in particular, local wind systems associated with land-water contrast are likely to influence predictability at station 2.

By comparing the polar plot of $\Pi(\theta)$ with polar plots of $\sigma(\theta)$ and $kurt(\theta)$ of wind components in Fig. 3.6, one can see that, in general, predictability of wind components is higher when the variability is larger (larger $\sigma(\theta)$) and when the kurtosis is smaller. Moreover, sometimes the directions $\hat{e}_{\bar{U}}$ and Π_{max} are similar, while other times these directions are not well-aligned.

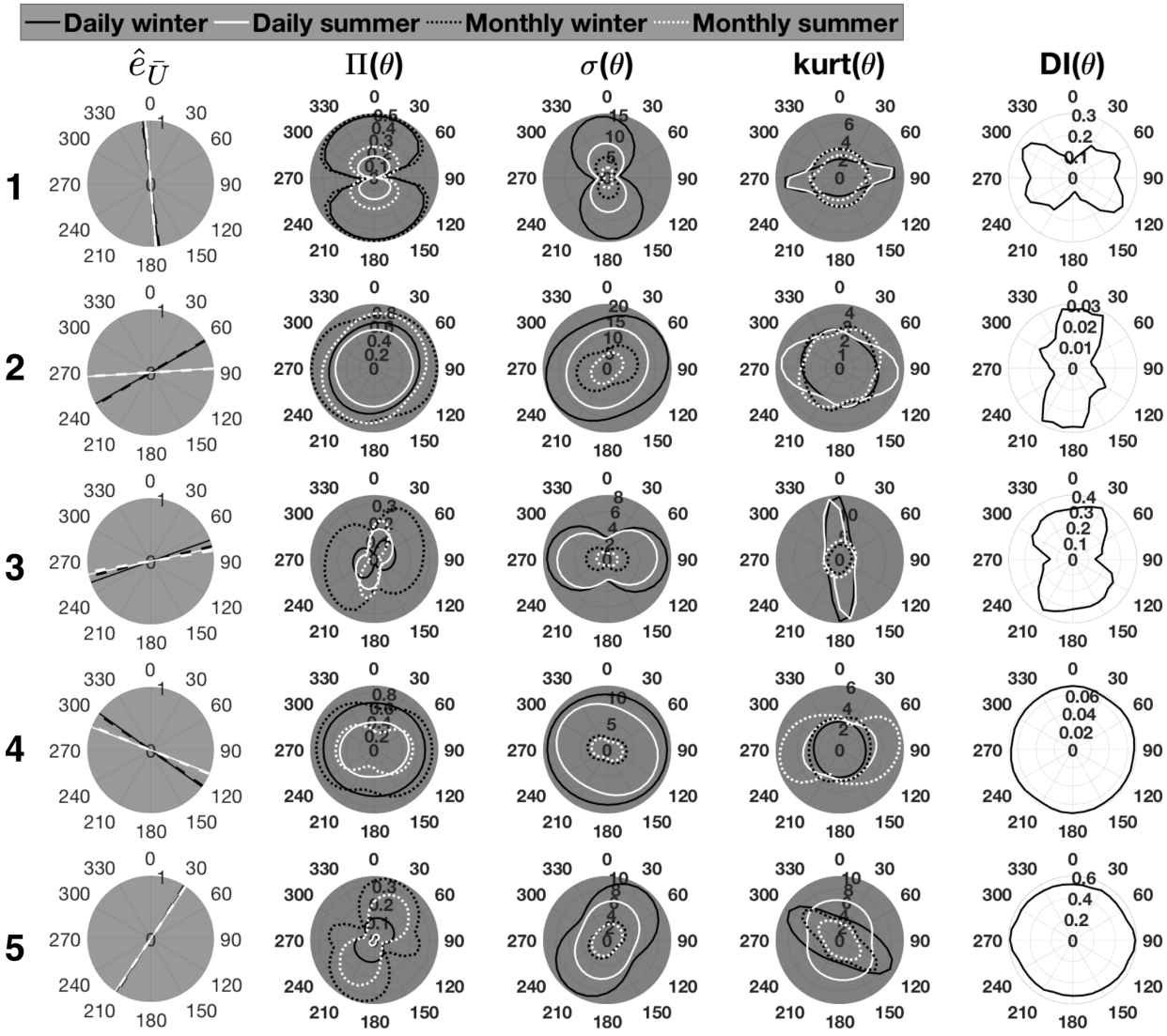

Figure 3.6: Predictability $\Pi(\theta)$ and potential explanatory factors: directions of mean wind vectors ($\hat{e}_{\bar{U}}$), statistical properties of surface wind components ($\sigma(\theta)$, $kurt(\theta)$) and topographic complexity $DI(\theta)$ for the numbered stations shown in Fig. 3.5

As discussed above, a clear relationship between the anisotropy of predictability and the orientation of local topographic complexity holds in some cases but not in others. Inspection of Fig. 3.6 demonstrates that the relationships between predictive characteristics and each explanatory factor are imperfect. For example, the (relatively weak) summer predictability at station 3 is such that the directions of *maximum*/minimum predictability are characterized with *small*/large standard deviation and *large*/small kurtosis, and $\hat{e}_{\bar{U}}$ is close to perpendicular to the direction of Π_{max}. One can see from these examples that the re-

lationships between $\Pi(\theta)$ and explanatory factors are not simple, and I need to consider more stations in a range of geographical settings. To this end, I will now consider these relationships across all stations in order to assess any broad relationships between explanatory factors and predictability in terms of both magnitude and direction.

3.3.3 Factors related to magnitude of predictability

The probability distributions of three measures of predictability ($max(\Pi)$, $min(\Pi)$, $\alpha(\Pi)$) conditioned on measures of the potential factors, calculated according to Eq. (3.4), are presented in Figs. 3.7-3.9.

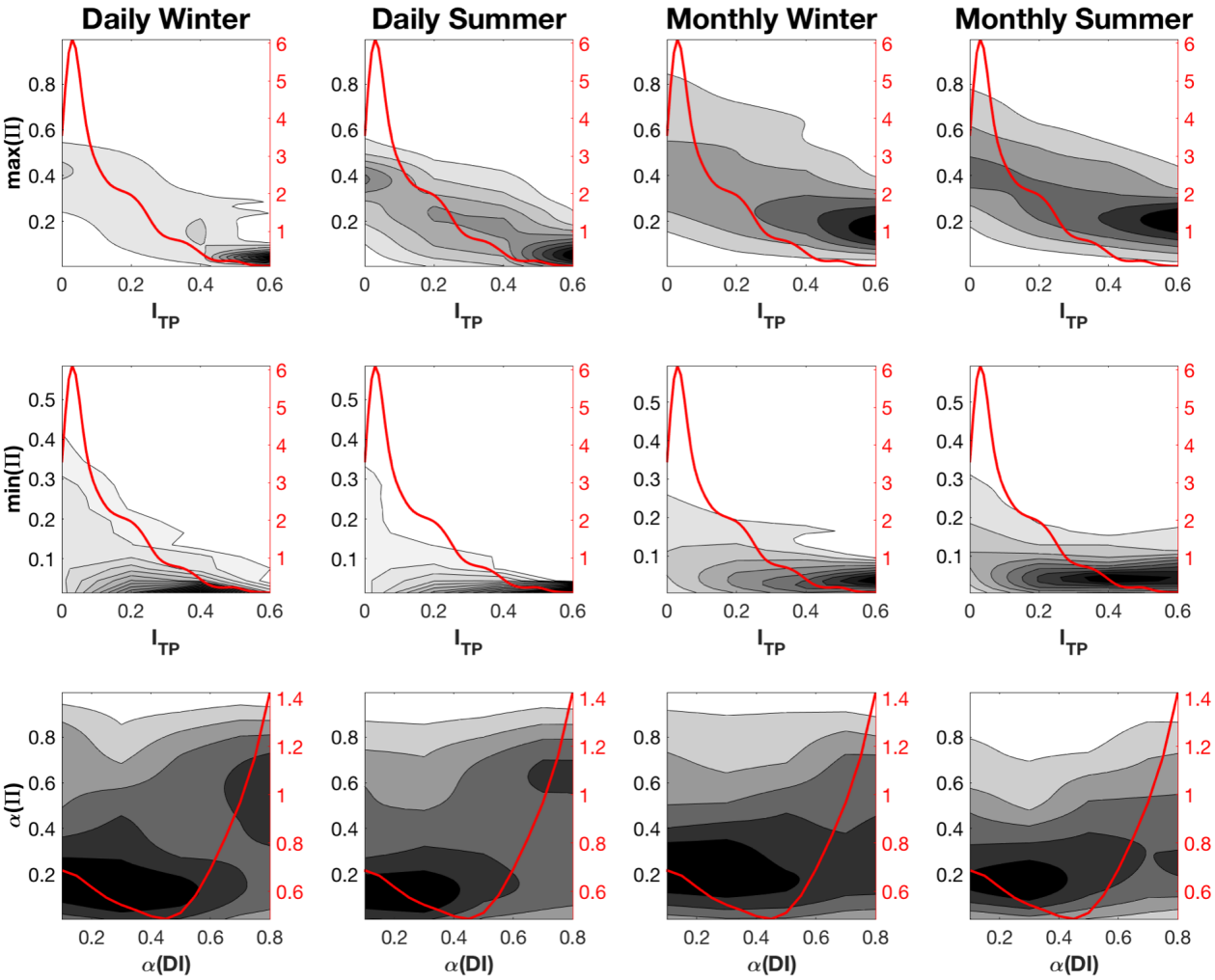

Figure 3.7: Probability density functions of $max(\Pi)$, $min(\Pi)$ and $\alpha(\Pi)$ conditioned on magnitude and anisotropy of topographic complexity (respectively I_{TP} and $\alpha(DI)$) for daily and monthly averaged winds in both winter and summer. Darker color indicates values of higher probability density. The red curve indicates a kernel density estimate of the marginal probability density of the explanatory factor related to topographic complexities: I_{TP} in rows 1 and 2, and $\alpha(DI)$ in row 3.

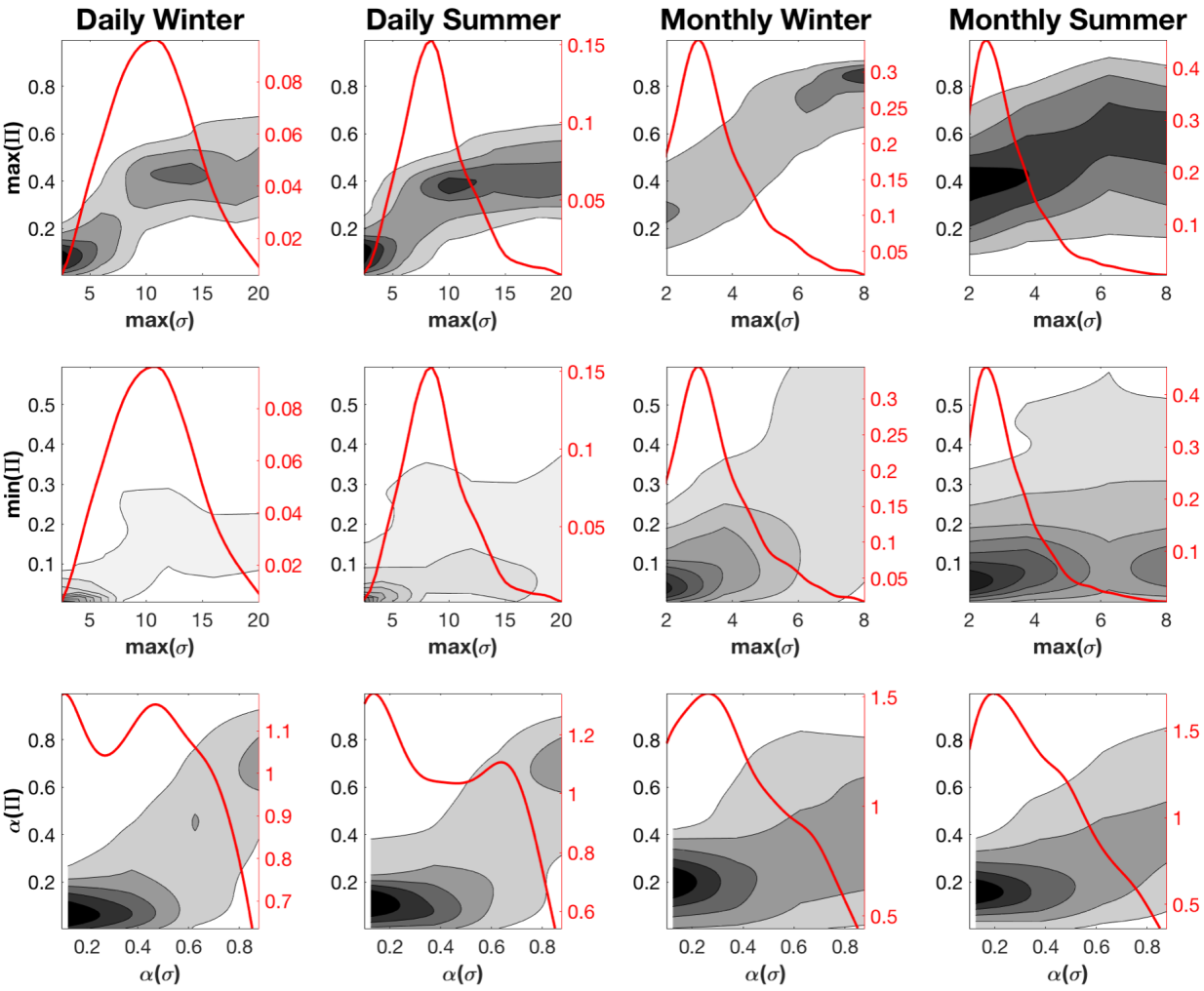

Figure 3.8: As in Fig. 3.7 for distributions of predictability conditioned on the maximum and anisotropy of standard deviation of wind components. The red curve indicates kernel estimate of probability density of the explanatory factor related to variability of wind components: $max(\sigma)$ in rows 1 and 2, and $\alpha(\sigma)$ in row 3.

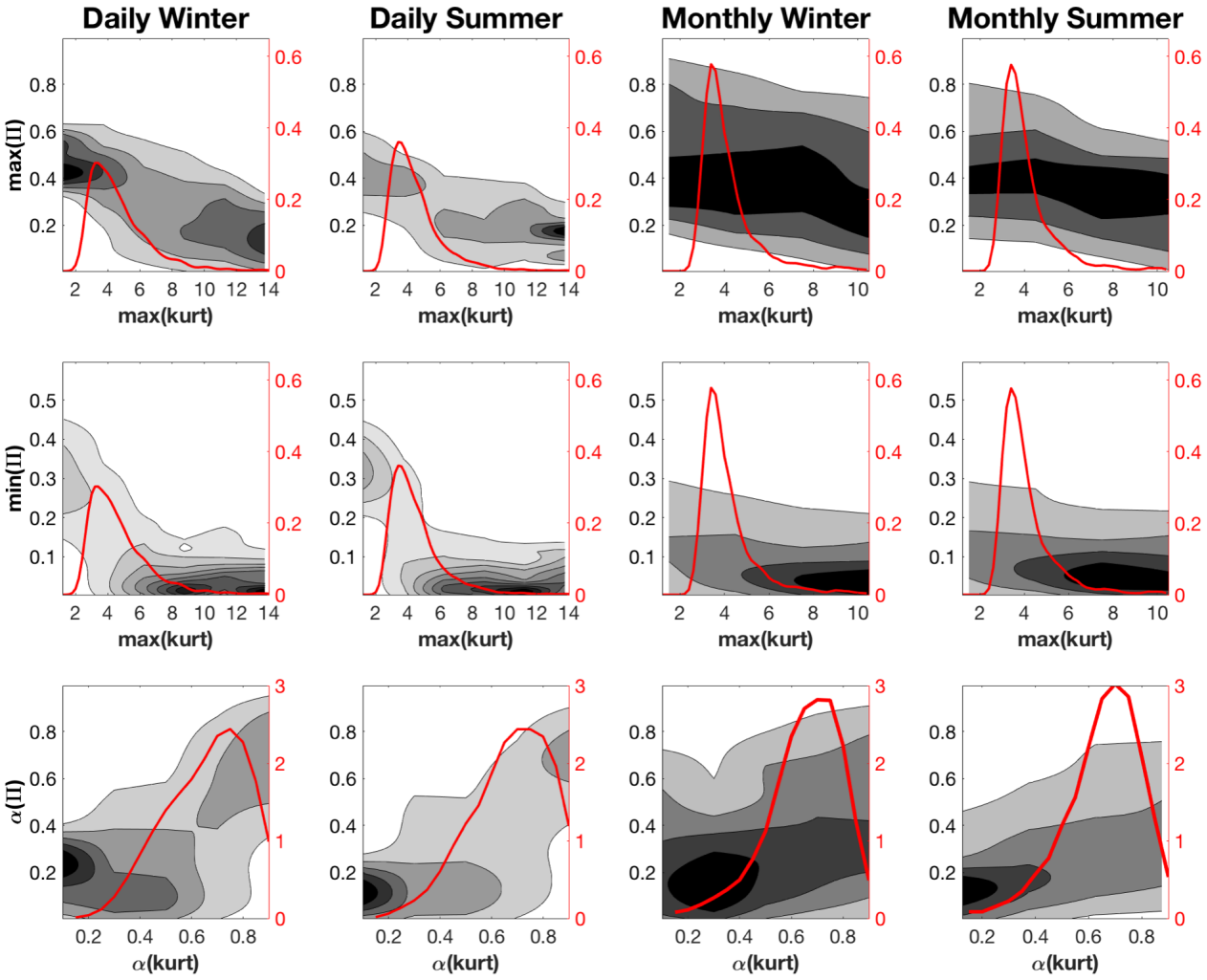

Figure 3.9: As in Fig. 3.7 for the distribution of predictability conditioned on the maximum and anisotropy of the kurtosis of wind components. The red curve indicates kernel estimate of probability density of the explanatory factor related to kurtosis of wind components: $max(kurt)$ in rows 1 and 2, and $\alpha(kurt)$ in row 3

Fig. 3.7 shows that lower predictability tends to be associated with topographically complex terrain. In particular, max(Π) decreases steadily with I_{TP}. Given a large value of I_{TP}, I find that both $max(\Pi)$ and $min(\Pi)$ are small. However, this statement cannot be reversed, especially for the relationship between $min(\Pi)$ and I_{TP}. In addition, the relationship between predictive anisotropy $\alpha(\Pi)$ and anisotropy of directional topographic complexity $\alpha(DI)$ is weak. Many stations are in places with similar directional distributions of topography but different predictive anisotropy.

Fig. 3.8 shows that higher predictability tends to be associated with more variable wind

components, and that the association between predictability and variability of wind components is more evident for $max(\Pi)$ than $min(\Pi)$. The increase of $max(\Pi)$ values saturates at intermediate values of $max(\sigma)$; beyond this point, increasing variability does not improve predictability. Furthermore, Fig. 3.8 also shows that there is a positive correlation between anisotropy of variability and anisotropy of predictability for both seasons and timescales of averaging, but this pattern is stronger for daily averaged data.

Fig. 3.9 shows that higher predictability tends to be associated with wind components characterized by data distributions with lighter tails (or flatter centers) as indicated by smaller values of kurtosis. In particular, highest predictability corresponds to kurtosis less than or near a value of 3. However, the relative frequency of $max(kurt)$ shown as the red curve in Fig. 3.9 indicates that the number of stations with $max(kurt)$ smaller than 3 is small. Out of 2109 stations, there are fewer than 280 stations with $max(kurt) < 3$ for daily averaged data, and fewer than 170 stations with $max(kurt) < 3$ for monthly averaged data. The most common kurtosis values are near 3 (corresponding to the Gaussian value). Predictability generally decreases as kurtosis of components increases away from this value. In this way, Fig. 3.9 shows that stations with wind components with non-Gaussian distribution characterized by heavier tails or more peaked centers are more likely to have lower predictability, and this pattern is stronger for daily averaged data. Moreover, while $max(\Pi)$ decreases steadily with $max(kurt)$, $min(\Pi)$ tends to rapidly decrease to 0 as $max(kurt)$ approaches 4. Lastly, $\alpha(\Pi)$ tends to increase with $\alpha(kurt)$, especially for daily averaged data.

Seasonal differences of statistical relationships shown in Figs. 3.7-3.9 are small. However, the results display a clear difference between daily and monthly timescales of averaging. In particular, the relationships between predictability and statistical properties of wind components are weaker for monthly averaged data than daily averaged data, especially for the relationships between predictability and kurtosis. A simple statistical model presented in Section 3.4.2 attempts to give a qualitative explanation for the relationships between predictability and statistical properties shown in this section.

3.3.4 Factors versus direction of predictability

Directions of predictability may be related to directional variability of topographic complexity $DI(\theta)$ and statistical properties of surface wind fluctuations, as well as directions of mean surface wind vectors $(\bar{u}, \bar{v})$. Histograms of rank correlation ρ between $\Pi(\theta)$ and $DI(\theta)$, $\sigma(\theta)$, $kurt(\theta)$ are considered for all stations (Fig. 3.10).

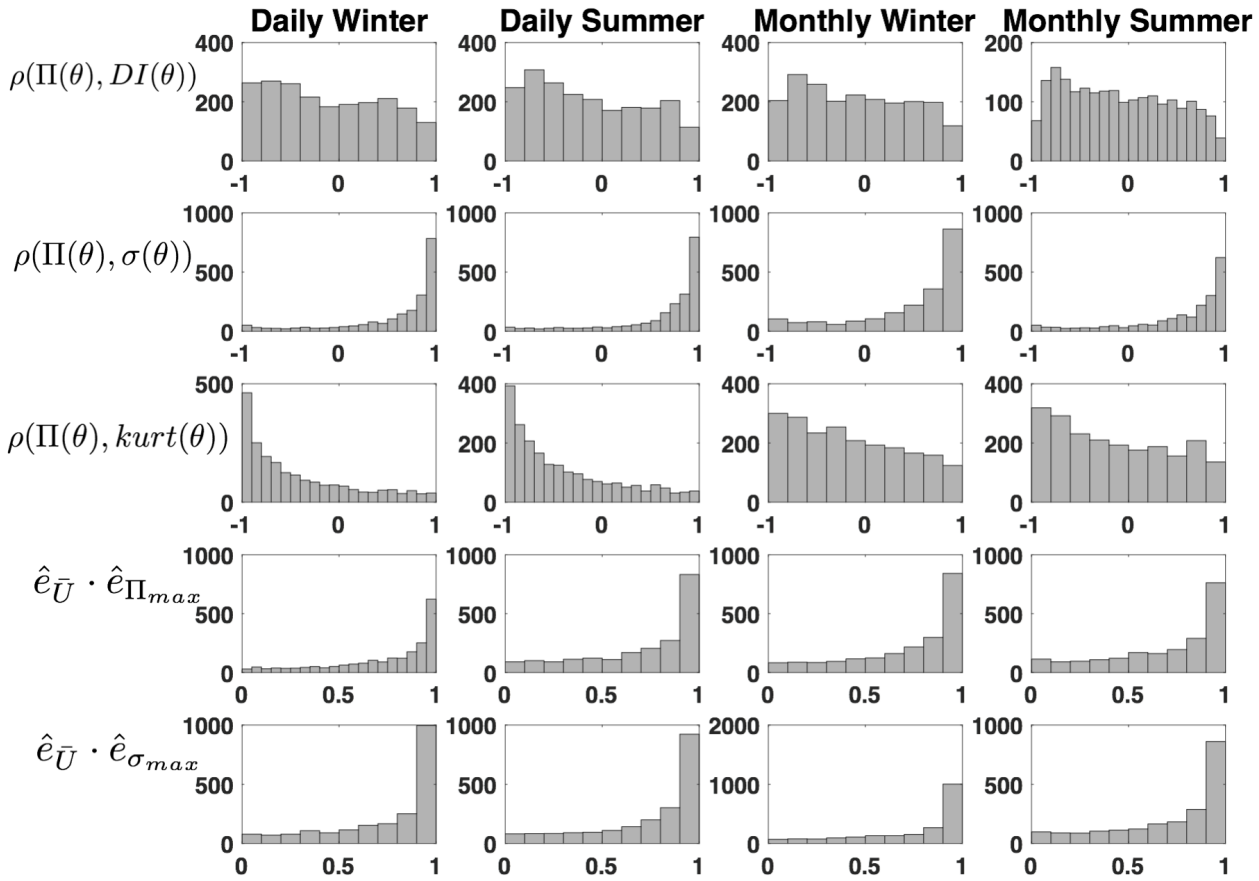

Figure 3.10: The first three rows show histograms of rank correlation coefficient between directional variation of predictability $\Pi(\theta)$ and the factors $DI(\theta)$, $\sigma(\theta)$ and $kurt(\theta)$. The last two rows show histograms of $\hat{e}_{\Pi_{max}} \cdot \hat{e}_{\bar{U}}$ and $\hat{e}_{\sigma_{max}} \cdot \hat{e}_{\bar{U}}$ to assess the directional relationships among mean wind vectors $< \bar{u}, \bar{v} >$, $max(\Pi)$ and $max(\sigma)$

For both monthly and daily timescales in winter and summer, values of $\rho(\Pi(\theta), \sigma(\theta))$ are dominated by strong positive directional correlation, concentrated near the value of 1. That is, the most variable wind components are generally the best predicted components. For daily averaged winds, values of $\rho(\Pi(\theta), kurt(\theta))$ are dominated by strong negative directional correlation, concentrated around values near -1. It follows that in general better predicted wind components are characterized with light tails or flatter distribution centers indicated by small kurtosis. On the other hand, wind components characterized by heavy tails or peaked centers as indicated by large kurtosis tend to be poorly predicted. While this pattern is substantially weakened for monthly averages, the general tendency remains. The relative frequency of $\rho(\Pi, DI)$ tends to decrease away from -1, but the overall directional relationship between predictability and terrain variability is weak. This indicates while

better predictability is somewhat more likely to be found in the directions of less variable terrain, this effect is weak. The directional relationship between predictability and mean wind vectors $< \bar{u}, \bar{v} >$ is shown by the dot product of unit vectors: $\hat{e}_{\Pi_{max}}$ and $\hat{e}_{\bar{U}}$. The histograms of $\hat{e}_{\Pi_{max}} \cdot \hat{e}_{\bar{U}}$ show that most stations tend to have mean wind vectors parallel to directions of $max(\Pi)$. Moreover, histograms of $\hat{e}_{\sigma_{max}} \cdot \hat{e}_{\bar{U}}$ show that directions of mean wind vectors are parallel to directions of $max(\sigma)$ for most stations. It is common for maximum predictability and maximum variability to both be aligned along the mean wind. Of these factors, only the directional relationship between kurtosis and predictability of wind components shows considerable difference between daily and monthly predictions. The directional relationships between predictability and all factors show negligible difference between winter and summer.

3.4 Discussion

3.4.1 Statistical analysis based on observations

The results presented in the previous section show that the relationships between predictability and explanatory factors display differences between daily and monthly averaging timescales. I propose two potential reasons for these differences. First, variability of large-scale tropospheric predictors is predominantly on synoptic timescales, while near-surface winds can also be influenced by mesoscale processes characterized by shorter timescales. Averaging the data over longer timescales will suppress the locally-driven variability more than that associated with large-scale processes, thereby strengthening the statistical relationship between predictor and predictand quantities. Second, according to the central limit theorem, the distribution of monthly-averaged atmospheric quantities is in general expected to be closer to Gaussian than daily-averaged quantities. As multivariate Gaussian distributions are characterized by linear relationships between variables, it follows that the nonlinearity of the relationship between two datasets of climate data is often diminished as the time scale of averaging becomes longer (Yuval and Hsieh, 2002). As a result, predictability by a linear regression model is likely to be higher for longer timescales of averaging as kurtosis approaches 3. This fact is also consistent with the observation that the relationship between predictability and kurtosis is weaker for the longer timescale of averaging as shown in Fig. 3.9 and Fig. 3.10, since there are fewer pairs of large kurtosis and small predictability as most data are concentrated in the region of small kurtosis values for monthly averaged data.

The relationships between predictability and potential explanatory factors indicate either that one causes the other, or that they have a common physical cause. For example, characteristics of local wind systems are influenced by local terrain, and local wind systems contribute to statistical properties of surface wind components as well as the direction of mean surface wind components. Anisotropy in standard deviation and kurtosis of wind components can both be created by anisotropy in surface topography. This study can neither make a general statement about how topographic complexity, magnitude of variability, and shape of data distribution (i.e. degree of non-Gaussianity) of surface wind components are related to each other nor establish the cause and effect between these factors and characteristics of predictability. What I have achieved in this chapter is the identification of general patterns related to predictability of surface wind components. Since the patterns shown in the relationships between predictability and statistical properties are stronger than topographic complexity, I will develop a descriptive model aiming at clarifying the relationship between wind predictability and statistical properties (i.e standard deviation and kurtosis) of wind component fluctuations in an idealized conceptual framework. Note that statistical predictability of wind components is not directly related to the mean wind vector, as the regression analysis considers fluctuations around the mean. It follows that the apparent relationship between the orientation of the mean wind and predictability must result because the variability characteristics and the mean state must share a common physical cause. In the future, a physically-based study of relationships between predictability and physical phenomena related to topographic complexity and atmospheric circulations is needed in order to clarify the physical sources of predictive anisotropy.

3.4.2 Statistical analysis based on a descriptive model

I now present an idealized statistical model used to characterize the relationship between predictability and the statistical properties of surface wind components. Let $U(\theta)$ be the surface wind component in the direction θ. I will assume that $U(\theta)$ can be partitioned into two parts:

$$U(\theta) = U_x + g(\theta)U_y, \tag{3.5}$$

where U_y and U_x are respectively variables perfectly correlated with and completely uncorrelated with the large-scale flow. I further assume that U_x and U_y are mutually uncorrelated. The variable U_x represents that part of surface winds which is entirely associated with small-scale local processes and/or is nonlinearly related to free tropospheric variability such that there is statistical dependence but zero correlation. By construction, U_x cannot

be modeled by linear regression with free tropospheric predictors. For simplicity, I assume that variability of U_x is isotropic, and that the distribution of U_y is Gaussian with a kurtosis of 3. The quantity $0 \leq g(\theta) \leq 1$ encodes the anisotropy of the part of $U(\theta)$ that is linearly predictable by free tropospheric variability. The structure of $g(\theta)$ accounts for predictive anisotropy inherent to the character of free tropospheric variability as well as the influence of surface inhomogeneities such as terrain. By construction, $\Pi(\theta) = corr^2(U(\theta), U_y)$ measures the predictability of $U(\theta)$ by a linear regression model with free-tropospheric predictors. Based on the above assumptions, the statistical properties of the surface wind components σ, *kurt* and statistical predictability Π can be calculated as follows:

$$\sigma(\theta) = \sqrt{V_x + V_y G(\theta)}, \tag{3.6}$$

$$kurt(\theta) = \frac{K_x \beta^2 + G(\theta)^2 K_y + 6G(\theta)\beta}{\beta^2 + 2G(\theta)\beta + G(\theta)^2}, \tag{3.7}$$

$$\Pi(\theta) = corr^2(U, U_y) = \frac{G(\theta)}{\beta + G(\theta)}, \tag{3.8}$$

where $G(\theta) = g^2(\theta)$, $V_x = \sigma^2(U_x)$, $V_y = \sigma^2(U_y)$, $K_x = kurt(U_x)$, $K_y = kurt(U_y)$, and $\beta = V_x/V_y$. Since G varies with the direction of wind projection θ, the quantities σ, *kurt* and Π are functions of θ. By calculating the first derivative of σ, *kurt* and Π with respect to G, I know that σ and Π increase with G, and *kurt* decreases with G when $K_x > 3$. For simplicity, I do not consider $K_x < 3$ as wind components with kurtosis larger than 3 are dominant in observations. It follows that:

$$min(\sigma) = \sqrt{V_x + V_y G_{min}}, \tag{3.9}$$

$$max(\sigma) = \sqrt{V_x + V_y}, \tag{3.10}$$

$$min(kurt) = \frac{K_x \beta^2 + K_y + 6\beta}{\beta^2 + 2\beta + 1}, \tag{3.11}$$

$$max(kurt) = \frac{K_x \beta^2 + G^2{}_{min} K_y + 6G_{min}\beta}{\beta^2 + 2G_{min}\beta + G^2{}_{min}}, \tag{3.12}$$

$$min(\Pi) = \frac{G_{min}}{\beta + G_{min}}, \tag{3.13}$$

$$max(\Pi) = \frac{1}{\beta + 1}; \tag{3.14}$$

where $G_{min} = min(g^2(\theta))$ and I have taken $max(g^2(\theta)) = 1$.

An ensemble of the quantities given by Eqs. (3.9) - (3.14) is obtained by sampling the parameter values randomly. I do not tune the parameters to match data characteristics of individual stations, but some restrictions are applied to the sampling. Specifically, in all samples, $K_y = 3$ and values of K_x are drawn randomly from a uniform distribution between 3 and 10. V_y and V_x are sampled so that most of the time V_y is larger than V_x (that is, I assume a generally moderate to large "signal-to-noise ratio"). In the present analysis, V_x and V_y are positive numbers drawn from a normal distribution with the following parameters: $\overline{V_x} = 5$ and $\sigma(V_x) = 2.5$, $\overline{V_y} = 5$ and $\sigma(V_y) = 10$. As illustrated in Fig. 3.11, the patterns of the relationships between simulated predictability and statistics of wind components (kurtosis and variability) qualitatively resemble those in observations (Figs. 3.7-3.9).

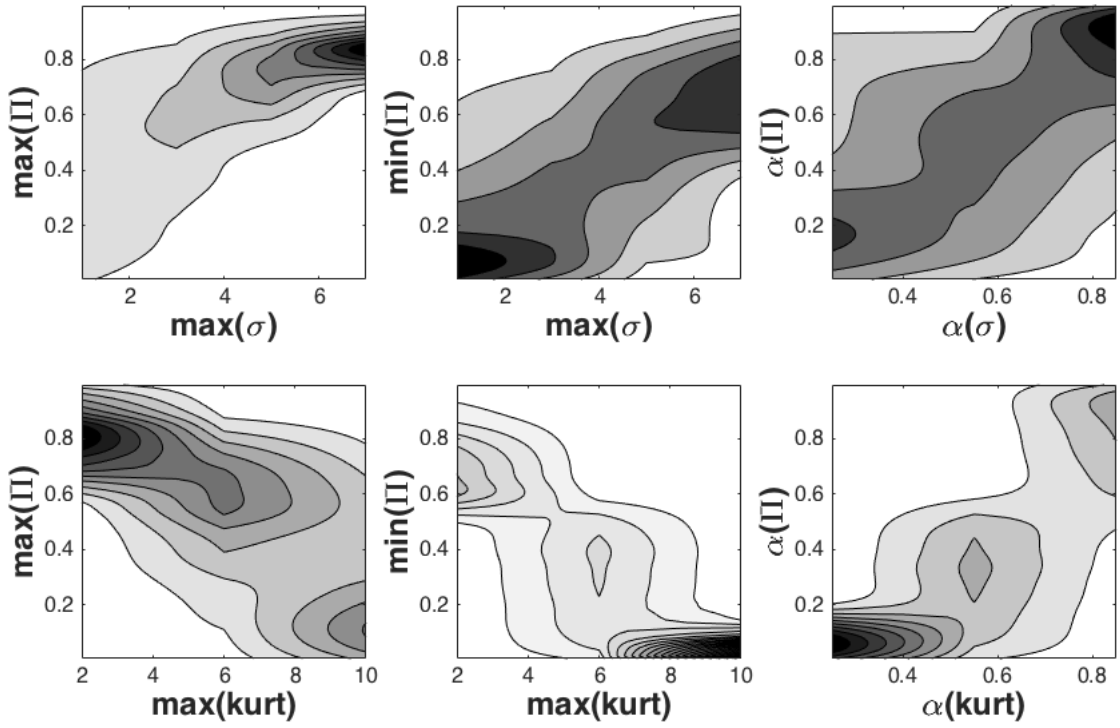

Figure 3.11: Probability density functions of the maximum, minimum and anisotropy of statistical predictability conditioned on the maximum and anisotropy of standard deviation and kurtosis for the idealized model Eqs. (3.9) - (3.14).

Specifically, the results of the simulation show that: (1) predictability (both $max(\Pi)$ and $min(\Pi)$) tends to increase given larger variability (i.e. larger $max(\sigma)$) and a lower level of non-Gaussianity ($max(kurt)$ approaching 3) of wind components; (2) the relationship

between the predictability and variability of wind components is stronger for $max(\sigma)$ and $max(\Pi)$ than for $max(\sigma)$ and $min(\Pi)$, (3) $max(\Pi)$ saturates for larger values of $max(\sigma)$, and (4) the relationship between predictability and kurtosis of wind components is stronger for $max(kurt)$ and $min(\Pi)$ than $max(kurt)$ and $max(\Pi)$. Moreover, the anisotropy of predictability tends to be stronger given stronger anisotropy of the wind statistics (i.e. $\alpha(\sigma)$ and $\alpha(kurt)$). All of these features can also be found in the results of observation as discussed in Section 3.3.3.

This simple descriptive model makes no assumption regarding the nature of the transfer function used for the statistical prediction (except to the extent that nonlinear dependence is included in U_x), and it shows that much of the association between the linear predictability of surface wind components and their statistical properties can be accounted for by a simple partitioning of surface wind variability into a portion that is linearly related to free-tropospheric variability and a portion that is not. Although this simple descriptive model cannot explain the causes of predictive anisotropy, it provides a simple characterization of it and relates predictive anisotropy to "signal-to-noise ratio" corresponding to the relative magnitudes of large-scale and local influences on the wind.

In particular, from Eqs. 3.13 and 3.14, I obtain that $\alpha \to G_{min}$ for $V_x \gg V_y$, and $\alpha \to 1$ in the opposite limit $V_y \gg V_x$. This indicates that if the influence of the unpredictable component becomes larger (i.e. increased noise indicated by large V_x), the prediction becomes more anisotropic. When $V_x = 0$, fluctuations in all directions are perfectly correlated with and predictable by the large-scale flow. As the variability in the linearly unpredictable part increases, the predictability of surface winds in the direction of smallest $g(\theta)$ (i.e. wind components with least predictive signal) is reduced more than in the direction of largest $g(\theta)$ (i.e. wind components with strongest predictive signal) and predictive anisotropy increases. Similarly, those directions with smaller $g(\theta)$ are more influenced by the non-Gaussian local variability in the sense of being heavier tailed or more peaked in the distribution of $U(\theta)$, resulting in the association of larger kurtosis with smaller predictability.

3.5 Conclusion

In this chapter, I have assessed the linear statistical relationship between large-scale atmospheric flow in the free troposphere and surface wind variability. The strength of this statistical relationship is important in determining the efficacy of statistical prediction, such as statistical downscaling. Particular attention has been paid to the anisotropy of predictability. I have demonstrated that predictive anisotropy is a common characteristic at

surface meteorological stations at a range of locations across the world. In regions away from complex topography, both the magnitude and direction of predictive anisotropy are spatially continuous, indicating possible large-scale organization by both the surface and the flow aloft.

Furthermore, I conducted a preliminary study investigating how different aspects of fluctuating surface winds are related to the predictive anisotropy. The results demonstrate that low predictability is often associated with complex terrain and that the best-predicted wind components generally lie in the direction of largest variability and smallest kurtosis, which generally correspond to the direction of the time-averaged wind. The results are effectively characterized by an idealized model in which surface wind variability is partitioned into a large-scale, linearly predictable part and a local-scale, linearly unpredictable part. The broad qualitative agreement of this model with the observed features of surface wind predictability provides evidence of the underlying hypothesis that predictive anisotropy can be characterized by the relative strength of large-scale "signal" and local-scale "noise" of the surface wind components. I cannot provide a physical mechanism for predictive anisotropy from the perspective of atmospheric dynamics related to large-scale and local-scale atmospheric circulations that can influence the strength of "signal" and "noise".

A subsequent study in Chapter 4 will investigate the extent to which the use of a linear transfer function limits statistical predictability: that is, if the observed predictive anisotropy is a consequence of the use of a simple statistical model. If the underlying relationship between surface wind components and larger-scale predictors in the free atmosphere is more nonlinear in some directions than others, linear transfer functions can result in low predictability of surface wind components in these directions, and linear predictive anisotropy will emerge. The association between directions of relatively poor prediction and high kurtosis suggests the possibility for improvement of prediction by nonlinear models.

On the other hand, predictive anisotropy may be inherent to physical phenomena at the surface and/or flow aloft. While the use of historical data cannot directly address the changes in predictability of surface winds in response to changes in large-scale circulation or weather patterns, the results in this chapter indicate that, in general, predictability should become larger and more isotropic as a result of any changes that increase the signal-to-noise ratio. I do not present any physical explanation for controls on the "signal-to-noise" ratio. Determination of the physical explanation requires a more detailed, physically-based analysis of the connection between near-surface flow and large-scale as well as local-scale

variability in atmospheric circulations. Such an analysis is an important direction for future study.

Chapter 4

Linear and nonlinear regression prediction of surface wind components

4.1 Introduction

Surface winds are a climatic field of interest because of the broad range of societal and economic sectors they affect, including agriculture, transport, and energy systems (Stull, 2000). However, the modeling skill of surface winds using global climate models (GCMs) is limited due to coarse horizontal and vertical resolution and difficulties simulating other boundary layer processes (Holtslag et al., 2013; He et al., 2010a). Driving finer-resolution dynamical models by GCM output is one approach to the simulation of surface winds. This approach has the advantage of modeling surface winds based on physics, but the drawback is that dynamical models are computationally expensive and also subject to resolution- and parameterization-dependent biases. An alternative approach is through computationally cheaper statistical prediction using well-resolved, large-scale predictors.

Given the importance of surface winds, it is beneficial to assess the predictive skill of surface winds by statistical prediction, which is a typical application of supervised learning (Hsieh, 2009). Specifically, statistical prediction requires a transfer function (TF) derived from the statistical relationship between predictors (e.g. large-scale climate fields in the atmosphere aloft) and predictands (local-scale surface winds) based on historical data, such that the TF can be then applied for new prediction. The skill of statistical prediction depends on how well the TF can model the relationship between predictors and predictands. Therefore, the predictability resulting from statistical predictions is strongly influenced by the characteristics of the predictor-predictand relationship, such as the functional form

(linear vs nonlinear) of the predictor-predictand relationship, and the signal-to-noise ratio (SNR). Statistical prediction can be classified into linear or nonlinear approaches according to whether the TF used to model the predictor-predictand relationship is implemented by linear or nonlinear regression methods.

Previous studies have showed that predictability of surface wind components projected onto different compass directions by linear TF is often characterized by anisotropy, such that different projections are predicted with different levels of skill (Salameh et al. 2009; van der Kamp et al. 2012; Culver and Monahan 2013; Sun and Monahan 2013; Chapter 3). Chapter 3 showed that predictive anisotropy by linear regression (LR) becomes strong when the overall SNR of predictor-predictand relationship (across all directions) is small, and directional predictability approaches isotropy when the overall SNR becomes large. The argument was based on a simple descriptive model partitioning the surface winds into two parts, perfectly correlated and uncorrelated with the large-scale atmospheric flow respectively. In this context, the "signal" was defined in terms of a linear statistical relationship.

A factor other than noise that can also lead to weak correlation between surface winds and large-scale flow (and therefore poor linear statistical predictability) is a nonlinear statistical relationship between predictands and predictors. If the predictive signal of the predictor-predictand relationship is nonlinear, the overall predictability by linear TF should be lower than using nonlinear TF. Chapter 3 showed that the directions of low linear predictability are often aligned with wind components characterized by large kurtosis. Linear predictive models are optimal when the joint distribution of the predictor and predictand are Gaussian. Therefore, it is possible that the anisotropy of linear-regression based prediction of surface wind components may result from variation of linearity of the predictor-predictand relationship in different directions. There is no a priori reason to expect that this relationship is linear, as atmospheric dynamics are themselves nonlinear.

While a nonlinear TF can result in higher predictability than linear TF because it can model a broader class of functional relationships, in practice, the cross-validated predictability by nonlinear TF may be lower than that of linear TF when the predictor-predictand relationship contains a large amount of noise. More complex algorithms are more likely to fit the noise as well as the predictive signal of the predictor-predictand relationships, which leads to the problem of overfitting. The problem of overfitting becomes more pronounced as the number of observations used in the statistical analysis is reduced.

This chapter aims at clarifying whether predictability of surface wind components resulting from linear TF is improved by allowing for nonlinear functional relationships. I

also assess if the predictive anisotropy found in previous studies (van der Kamp et al. 2012; Culver and Monahan 2013; Sun and Monahan 2013; Chapter 3) is an artifact of the use of a linear-regression based TF. Specifically, I compare the characteristics of predictability (i.e. magnitude and anisotropy) of surface wind components using nonlinear TF and linear TF for 2109 land stations across a wide range of locations across the globe (the same set of stations considered in Chapter 3).

I consider three common nonlinear machine learning methods as TFs for statistical prediction of surface wind components: neural network (NN), support vector machine (SVM) and random forest (RF), and compare the results with the statistical prediction by linear regression (LR). These three nonlinear methods have been applied to prediction of surface wind speed in a few previous studies. For example, Sailor et al. (2000) developed a methodology based on NN for downscaling GCM output to predict surface wind speed. Mohandes et al. (2004) applied SVM to predict daily averaged wind speed at Madina, Saudi Arabia, and compared the performance of SVM with NN. Their results indicate that SVM outperforms NN at this site. Davy et al. (2010) applied RF based statistical prediction to model wind variability at a coastal location in Victoria, Australia. They found that RF based statistical prediction outperforms linear regression, and the overall accuracy of RF was competitive with NN. These previous studies on statistical prediction only use a small number of stations in limited geographic regions, and the predictand in most of these studies is wind speed. The present study undertakes a systematic comparison of statistical prediction by linear and the three nonlinear methods (NN, SVM and RF) using data from meteorological stations at a wide range of locations across the world in order to get a more comprehensive picture of the statistical prediction of surface wind components.

This chapter is organized as follows. Section 4.2 presents the implementation of nonlinear models used in this chapter. Section 4.3 presents the results of comparing predictability of surface wind components using statistical predictions with different TFs. Section 4.4 presents a discussion comparing the skill of linear and nonlinear regression methods. The conclusions are given in Section 4.5. Brief descriptions of nonlinear regression methods are presented in Appendix A.

4.2 Implementation of nonlinear models

A detailed discussion of the nonlinear regression methods considered in this study is presented in Appendix A. Various software tools exist to implement nonlinear regression

analysis. The tools in this study are selected on the basis of robustness of results and flexibility of implementation. In this regard, MATLAB function 'fitrsvm' in the Statistics and Machine Learning Toolbox (MathWorks, 2017a) and the Python function 'RandomForestRegressor' in the Scikit-learn library of Python (Python, 2016) are used for implementing SVM and RF in this study. Since I choose different training procedures for daily-averaged and monthly-averaged data according to their different data properties, the R package 'nnet' (Ripley and Venables, 2016) and the MATLAB Neural Network toolbox (MathWorks, 2017b) are used to implement NN for daily- and monthly-averaged data. The details will be explained later in this section.

In all regression analyses, four predictor variables (X) are related to one predictand (Y). Each nonlinear regression method requires the specification of parameters related to model architecture. The challenge is to determine parameters which yield robust regression models for this study. To do this, I randomly select 100 stations from all available surface meteorological stations and apply the nonlinear methods with different configurations to predict daily and monthly averaged surface wind components. A controlled experiment approach is used, so that for each parameter being tested all other settings are the same. The parameters related to model architecture remain fixed for all stations considered once they have been determined by this analysis (done separately for each statistical method, averaging period, and season) .

The NN model used for this study has the following structure. The architecture of the NN model is denoted $4-N_h-1$ (see A.1), where 4 and 1 refer to the number of predictor and predictand variables respectively, and N_h is the number of hidden neurons. The number N_h is determined by controlled experiments as described above. Since the length of the time series of daily averaged data is about 30 times longer than that of the monthly averaged data, more hidden neurons are used to model the daily averaged surface wind components. I choose $N_h = 5$ for daily-averaged data and $N_h = 2$ for monthly averaged prediction of surface wind components based on consideration of a range of values from $N_h = 2$ to $N_h = 30$. Not only the model architecture, but the training procedures for predicting daily and monthly averaged surface wind components also differ.

Early stopping is a method used to avoid overfitting the statistical model, in which a subset of the data is reserved to assess out-of-sample model performance during the training process. For early stopping, a fraction of data must be set aside for validation. Early stopping is used for monthly averaged prediction by NN but not for daily averaged prediction. This choice is based on the results of the study of Amari et al. (1996). Their numerical experiments demonstrated that overfitting is not a problem when the number of

training samples (N observations) is larger than the number of model parameters (N_p) by more than a factor of thirty ($N > 30N_p$), and that in this situation reserving a fraction of data for validation in early stopping is not better than using the whole dataset for training until convergence of the numerical optimization algorithm. However, when $N < 30N_p$, it is necessary to use measures to prevent overfitting such as early stopping. In this study, each training set uses 32 out of 33 years (i.e. 1980-2012) of winter or summer data. This means there are approximately 90×32 records of daily averaged data and 3×32 records of monthly averaged data. The number of parameters for a $4-5-1$ and $4-2-1$ NN model is 31 and 13 respectively. Daily prediction meets the threshold of $N > 30N_p$ but the monthly prediction does not.

The study of Amari et al. (1996) provided an equation to estimate the optimal fraction of data for validation,

$$f_{opt} = \frac{\sqrt{2N_p - 1} - 1}{2(N_p - 1)}. \tag{4.1}$$

Based on Eq. (4.1), 16% of each segment of monthly averaged data for training (i.e. 32 years) are randomly chosen as validation data for early stopping in each training session. Because of the different NN training procedures for daily averaged and monthly averaged data, the R package 'nnet' and the MATLAB Neural Network toolbox are used respectively for these analyses. Early stopping is automatically incorporated into the functions provided by MATLAB Neural Network toolbox. Although early stopping can be turned off in the MATLAB implementation, empirical tests in this study show that the MATLAB Neural Network toolbox is generally slower than 'nnet' in R, especially for large datasets. Therefore, I apply 'nnet' of R in predicting daily averaged data in this study.

Predictability by SVM regression is influenced by the choice of two parameters, ε and C, and the choice of kernel functions. The parameter ε sets the limit of the deviation of the regression solution from training data in the input space. However, the limit set by ε is not strict since observations exceeding the limit will also be "tolerated" (i.e. included in the regression process) as long as they are within the limit of deviation set by the box constraint C. The radial basis function (rbf) and linear function are chosen as the kernel types for daily and monthly averaged predictions, respectively, based on empirical tests as described before.

Modeling by RF requires the specification of the maximum number of features used for data partition when looking for the best split in tree regressions and the number of tree regressions used for ensemble averaging. In this study, the maximum number of features is taken to be the number of variables of predictors (i.e. 4), and the number of tree regressions

is chosen to be 100. Empirical tests show that there is no significant improvement in predictability of either daily or monthly averaged surface wind components when the number of tree regressions in RF is larger than 100.

4.3 Results

I contrast two aspects of the predictability of surface wind components $\Pi(\theta)$ resulting from the linear regression model and the three nonlinear regression models: the overall magnitude and the directional variability of predictability respectively quantified by $\overline{\Pi(\theta)}$ and $\alpha(\Pi)$ (Eq. 2.7). Nonlinear methods (primarily SVM for daily-averaged and NN for monthly-averaged data), outperform LR in terms of $\overline{\Pi(\theta)}$ at most stations (Fig. 4.1).

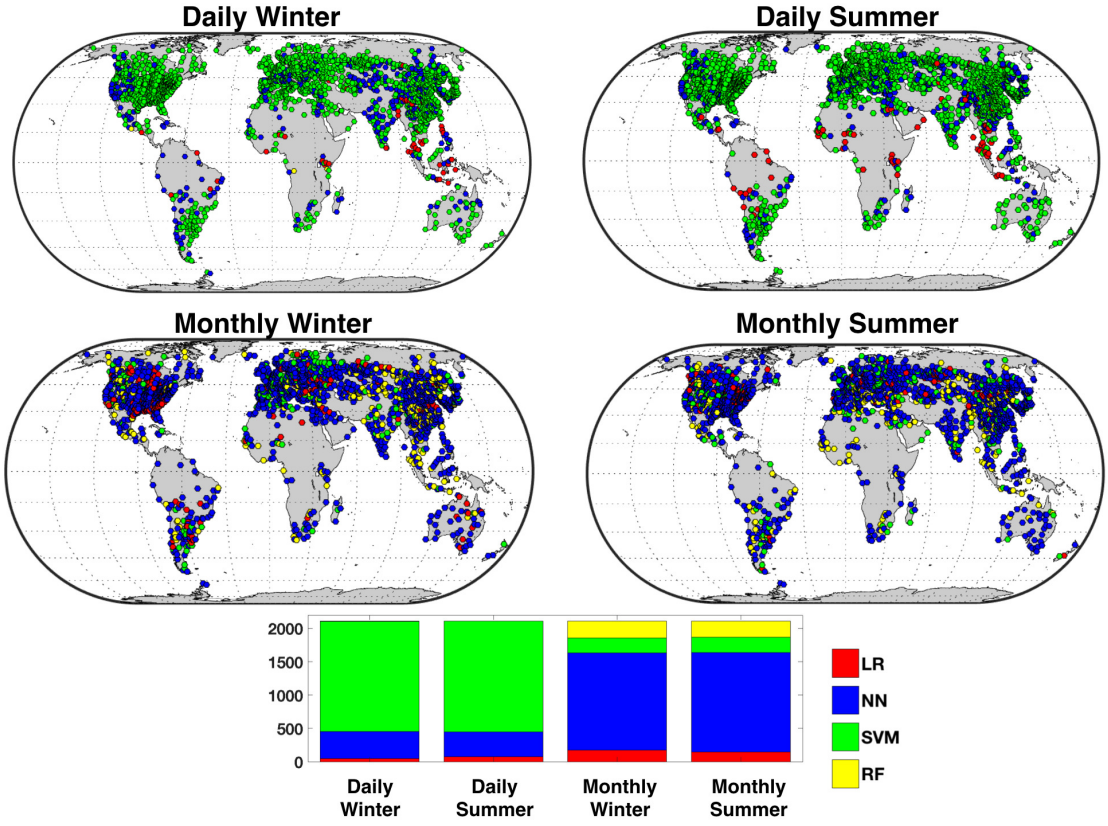

Figure 4.1: Spatial distribution of the regression method which gives highest $\overline{\Pi(\theta)}$ at each station for four cases of prediction: predicting daily and monthly averaged surface wind components in winter and summer. The number of stations belonging to each regression method that gives the highest $\overline{\Pi(\theta)}$ are shown for each class of prediction

The same method generally gives the highest $\overline{\Pi(\theta)}$ for both summer and winter prediction at most stations, as is evident in the minimal seasonal differences of the spatial patterns in Fig. 4.1. In contrast, there are distinct differences between the results of monthly and daily averaged predictions. For daily (monthly) averaged data, SVM (NN) outperforms other methods at the majority of stations. Those stations for which LR is the best method of predictability are generally distributed in the subtropics for daily timescales, and in the midlatitudes for monthly timescales.

Fig. 4.2 compares the values of $\overline{\Pi(\theta)}$ resulting from LR with those resulting from the three nonlinear methods: NN, SVM and RF.

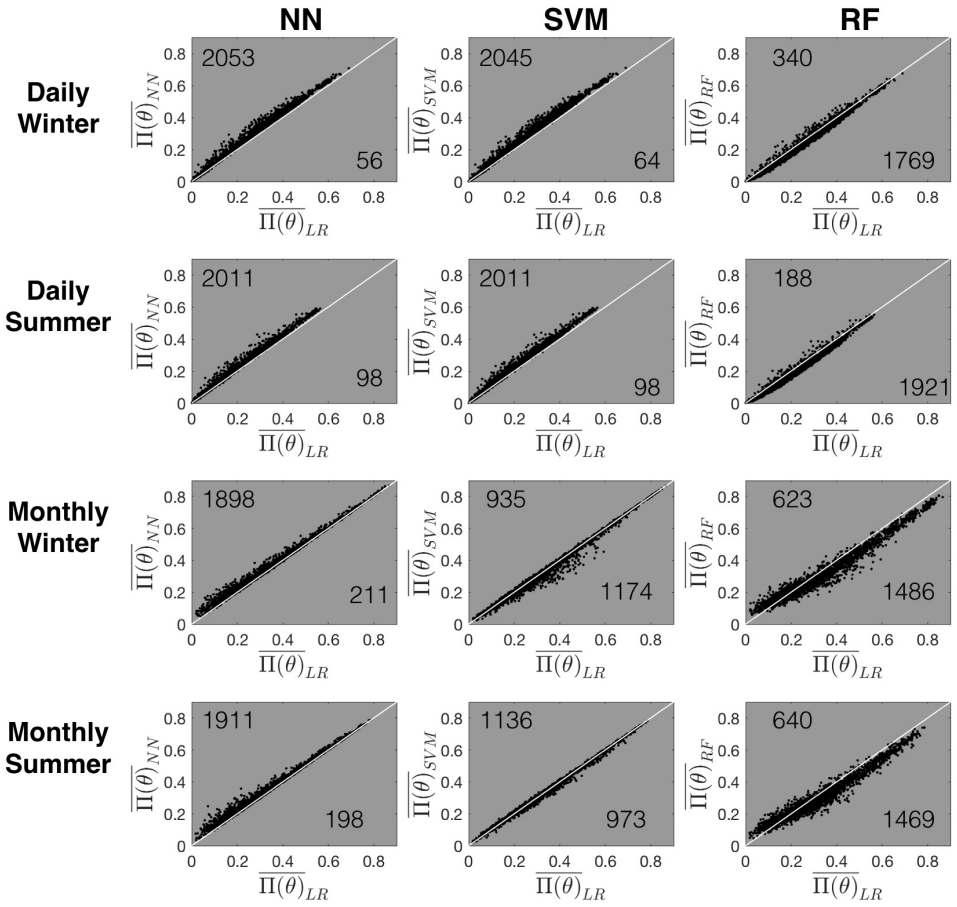

Figure 4.2: Scatter plots of $\overline{\Pi(\theta)}$ of predictability of surface wind components resulting from LR against those resulted from NN, SVM and RF respectively at the 2109 surface meteorological stations are shown with the 1:1 line. The number of stations with larger $\overline{\Pi(\theta)}$ by linear regression and nonlinear regression is labeled at the lower right and upper left respectively

The difference between $\overline{\Pi(\theta)}$ resulting from LR and any of the three nonlinear regression methods is not large for most stations: scatter of points falls close to the 1:1 line. In general, regions of relatively low predictability by linear regression remain poorly predicted using nonlinear regression methods. The similarity between linear and nonlinear regression methods is most pronounced for NN and SVM; prediction skills from RF show substantially more scatter. Nonlinear regression methods differ from each other by the complexity of the algorithm and may be more or less advantageous in detecting an underlying nonlinear signal in different situations. For the data considered in this study, the predictability by RF is generally lower than those of NN and SVM (Figs. 4.1 and 4.2).

Fig. 4.3 compares the anisotropy of $\Pi(\theta)$ resulting from LR and the three nonlinear regression methods.

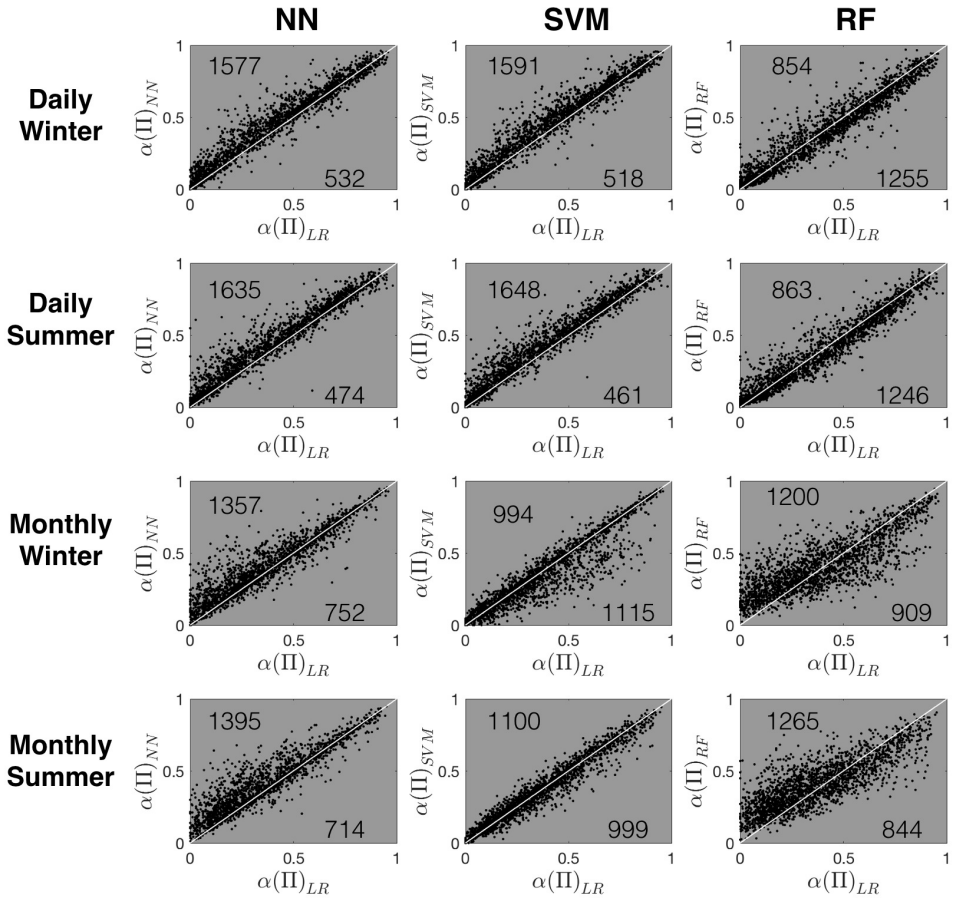

Figure 4.3: As in Fig. 4.2 for $\alpha(\Pi)$

The predictive anisotropy values generally fall around the 1:1 but with a larger scatter than the values of $\overline{\Pi(\theta)}$ in Fig. 4.2. Broadly speaking, regions of strong(weak) predictive

anisotropy resulting from linear regression are generally regions of strong(weak) anisotropy from nonlinear regression methods. However, many stations have slightly weaker nonlinear predictive anisotropy (as indicated by the larger number of points falling above the 1:1 line than below it). Among the various nonlinear regression methods, the variation between linear and nonlinear $\alpha(\Pi)$ is largest for RF.

The results of this study show that nonlinear regression methods (NN, SVM and RF) do not substantially improve predictability of surface wind components relative to LR. Although at many stations linear predictive anisotropy is stronger than nonlinear predictive anisotropy resulting from at least one of the nonlinear regression, the differences are small. The results suggest that model complexity is not a major factor in determining the overall magnitude and anisotropy of statistical predictability of surface wind components.

I now consider the spatial distribution of differences in predictive skill between linear and nonlinear regression models by comparing linear predictability, $\Pi(\theta)_{LR}$, with the highest predictability among the three nonlinear regression models along each direction of projection, denoted $\Pi(\theta)_{NL}$. Specifically, I consider the spatial distribution of $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ and $\frac{\alpha(\Pi)_{NL}}{\alpha(\Pi)_{LR}}$. Only wintertime results are shown as differences between seasons are relatively small. From a broad view, stations with large differences between linear and nonlinear prediction are usually located in regions characterized by surface heterogeneity, such as major mountain ranges or land/sea contrast (Figs. 4.4 and 4.5).

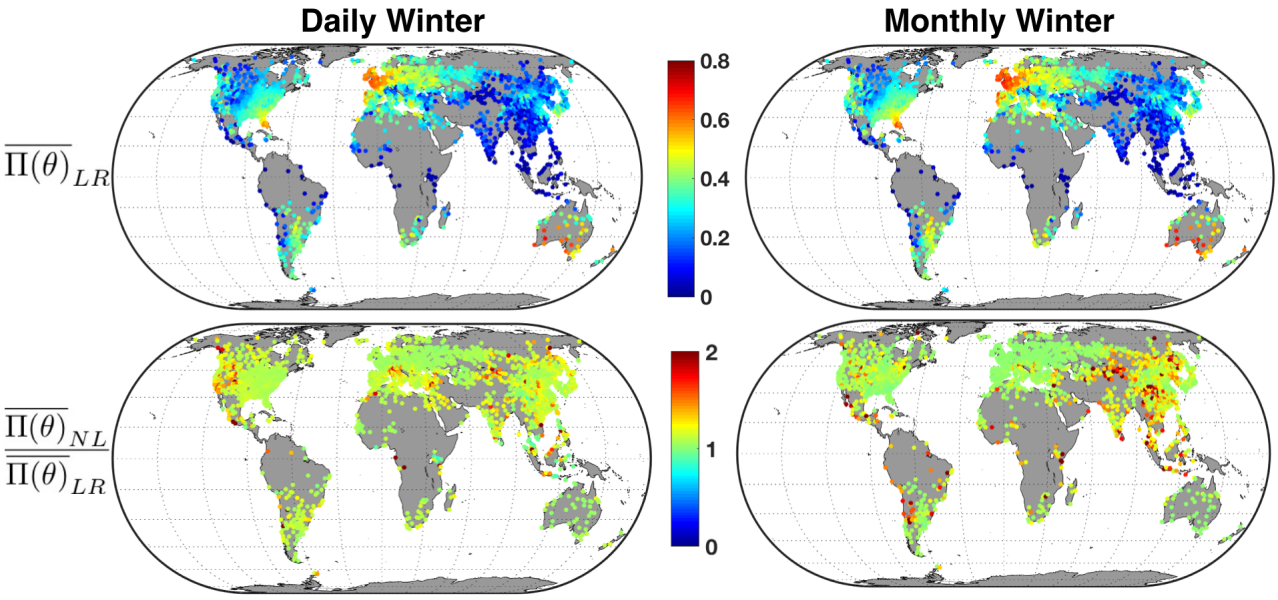

Figure 4.4: Spatial distribution of $\overline{\Pi(\theta)}_{LR}$ and $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ for daily and monthly averaged prediction of winter surface wind components

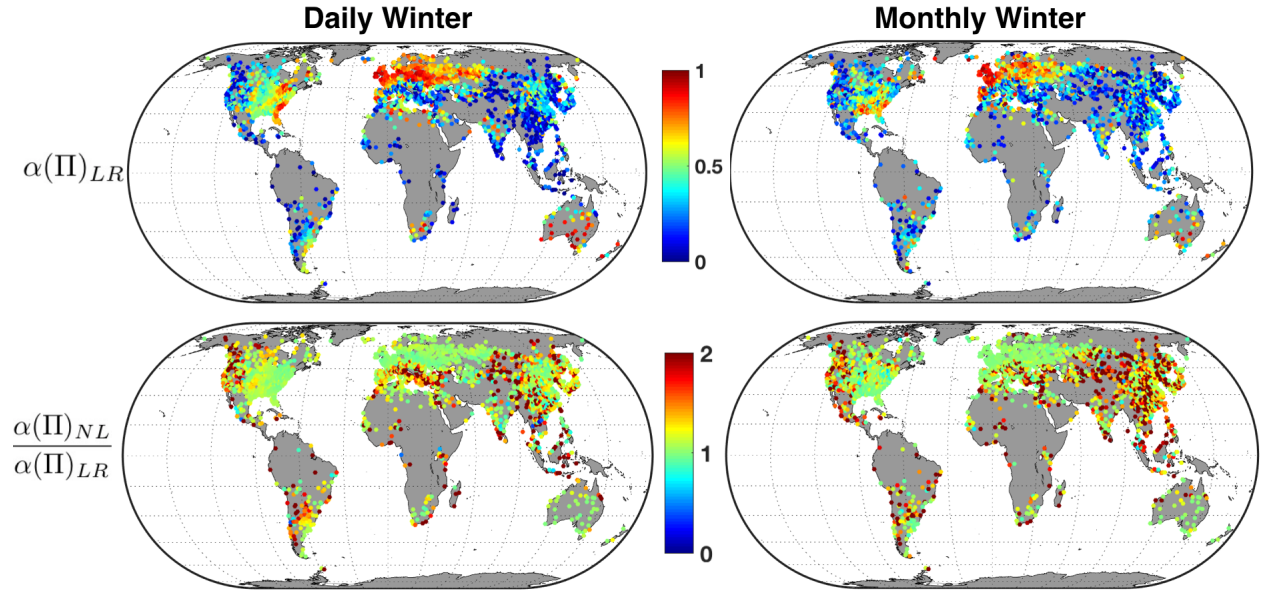

Figure 4.5: Spatial distribution of $\alpha(\Pi)_{LR}$ and $\frac{\alpha(\Pi)_{NL}}{\alpha(\Pi)_{LR}}$ for daily and monthly averaged prediction of winter surface wind components

The pattern varies with time scales and is different for $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ and $\frac{\alpha(\Pi)_{NL}}{\alpha(\Pi)_{LR}}$. Most stations with large values of $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ for daily prediction are located in the mountainous regions of

North America and Asia, with a few stations in coastal regions. For monthly prediction, large values of $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ are nearly absent in the North American Cordillera but are commonly found in the mountainous regions of Asia and South America as well as along the coast of South America, Africa and South East Asia. The association of large values of $\frac{\alpha(\Pi)_{NL}}{\alpha(\Pi)_{LR}}$ with mountainous and coastal regions is more obvious than that of $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ for both daily and monthly predictions. The locations with large values of $\frac{\overline{\Pi(\theta)}_{NL}}{\overline{\Pi(\theta)}_{LR}}$ and $\frac{\alpha(\Pi)_{NL}}{\alpha(\Pi)_{LR}}$ tend to correspond to locations of low linear predictability and strong linear predictive anisotropy. That is, the nonlinear models tend to outperform LR in areas of relatively low statistical predictability.

Representative stations are selected to illustrate typical locations where relatively large differences between $\Pi(\theta)_{LR}$ and $\Pi(\theta)_{NL}$ are observed (Fig. 4.6). Station information is listed in Table 4.1. Stations 1,2 and 5 are in mountainous regions, and stations 3 and 4 are characterized by land/sea contrast.

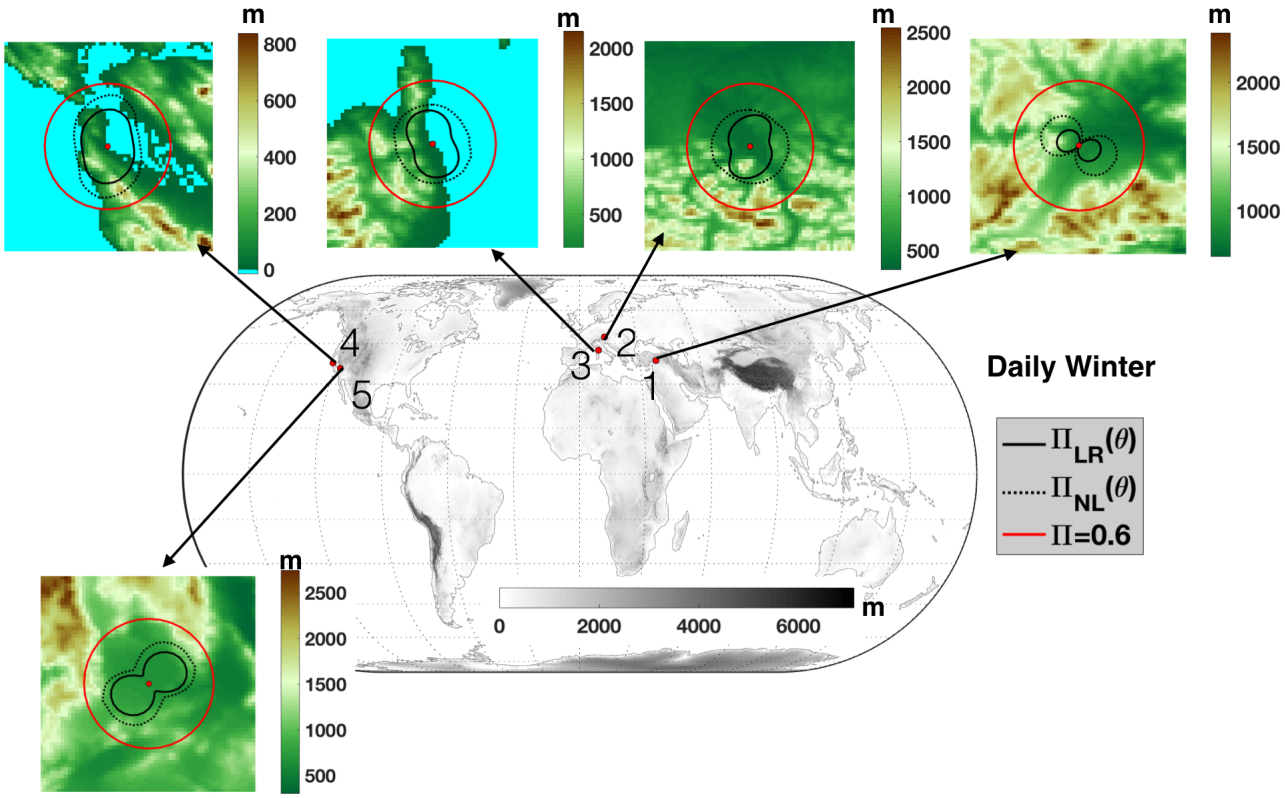

Figure 4.6: Comparison of $\Pi(\theta)_{LR}$ with $\Pi(\theta)_{NL}$ resulting from the best predictability among NN, SVM, RF in each direction of projection of surface winds at five representative stations for daily averaged prediction in winter

Table 4.1: Representative meteorological stations shown in Fig. 4.6. The name refers to the international ID of the stations

Station	Name	Location	Latitude(°)	Longitude(°)
1	LTAT	Malatya Erhac Airport, Turkey	38.4400	38.0900
2	LOWS	Salzburg Airport, Austria	47.7930	13.0040
3	LFKB	Bastia, France	42.5530	9.4840
4	KSFO	San Francisco Intl Airport	37.6190	-122.3750
5	KNID	China Lake, California, US	35.6860	-117.6920

At all stations considered, nonlinear predictability exceeds that of LR in all directions. For stations 2 and 3, predictive anisotropy resulting from linear regression is weakened greatly to almost isotropic predictability by nonlinear regression; in contrast, predictive anisotropy by nonlinear regression becomes stronger than that by linear regression at station 1. For other stations, the small increase in predictability by nonlinear regression does not substantially change the predictive anisotropy.

Chapter 3 showed that the strength and anisotropy of LR-based prediction of surface wind components are related to the magnitude and non-Gaussianity of their fluctuations (as measured respectively by standard deviation and kurtosis). Specifically, the best-predicted surface wind components tended to be along the direction of the most variable fluctuations closest to having a Gaussian distribution. I now investigate if there are relationships between the directional structure of statistical properties of surface wind components (i.e. standard deviation and kurtosis) and differences in predictability resulting from linear and nonlinear regression models. In this analysis, I exclude stations of very small predictability (those for which both $\overline{\Pi(\theta)}_{NL}$ and $\overline{\Pi(\theta)}_{LR}$ are smaller than 0.1). Small differences in predictability between linear and nonlinear methods at such stations are not meaningful. First, I define $\beta(\theta) = \frac{\Pi_{NL}(\theta)}{\Pi_{LR}(\theta)}$ along each direction of projection. Fig. 4.7 shows the directional relationship of $\beta(\theta)$ with $\sigma(\theta)$ and $kurt(\theta)$ across all valid stations as measured by the histogram of the rank correlation coefficients $\rho(\sigma,\beta)$ and $\rho(kurt,\beta)$.

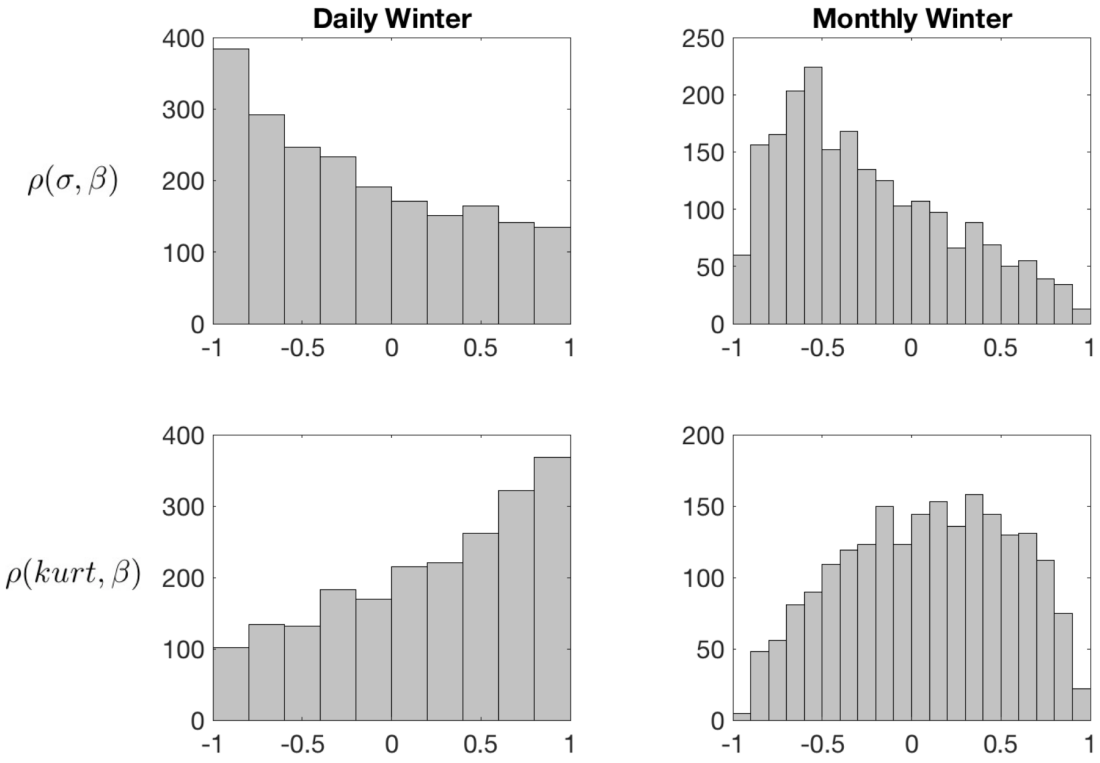

Figure 4.7: Histograms of rank correlation coefficients between directional variation of $\beta(\theta) = \frac{\Pi(\theta)_{NL}}{\Pi(\theta)_{LR}}$ and the statistical properties of surface wind components: directional standard deviation $\sigma(\theta)$ and kurtosis $kurt(\theta)$ of surface wind components

Values of the rank correlation coefficient near 1 indicate that maxima of these directional quantities are aligned, as are minima, while values near -1 indicate that maxima in one quantity are oriented with minima in the other. Correlations near zero correspond to no common orientations of directional structures.

In general, $\rho(\sigma,\beta)$ and $\rho(kurt,\beta)$ are dominated by strong negative values and strong positive values respectively. This pattern is stronger for daily predictions and weaker for monthly predictions. The result implies that the improvement of nonlinear predictability over linear predictability tends to be larger in the direction of surface wind components with small fluctuations and non-Guassian distribution with heavier tails or sharp central peaks as indicated by smaller standard deviation and larger kurtosis. On the other hand, nonlinear predictability tends to be equal to or smaller than linear predictability in the direction of surface wind components with larger fluctuations and smaller kurtosis. As kurtosis values less than 3 and β values less than 1 are uncommon in data considered in this study, LR tends to have comparable skill to nonlinear regression in the direction of surface wind

components characterized by larger fluctuations and data distribution closer to Gaussian.

To assess how enhancement of predictive skill by nonlinear regression models can be related to the strength of fluctuations and non-Gaussianity of surface wind components, I plot the maximum of $\beta(\theta)$ over all directions as functions of the directional maxima of σ and kurtosis (Fig. 4.8).

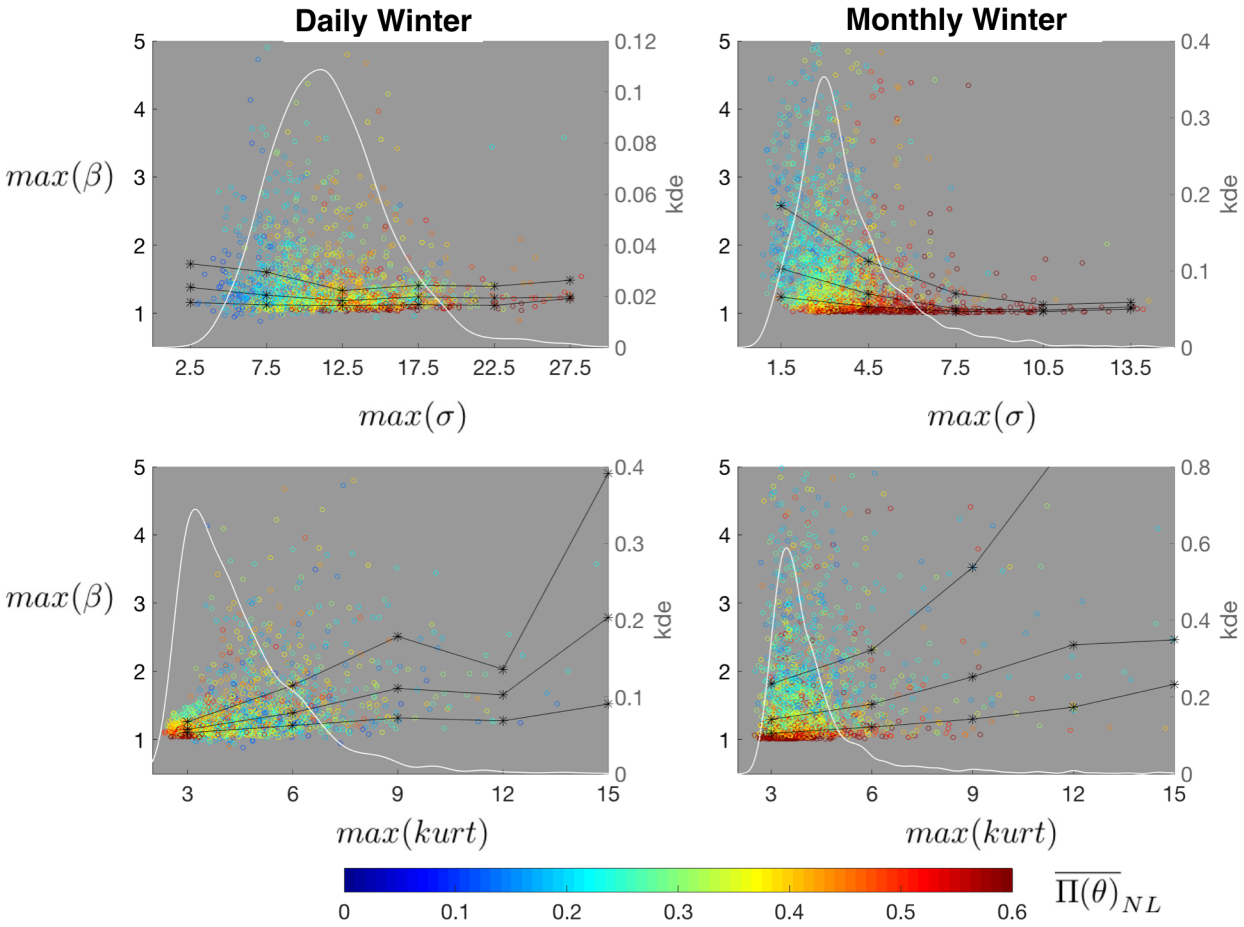

Figure 4.8: Scatter plots of $max(\beta)$ vs $max(\sigma)$ and max$(kurt)$ of surface wind components colored by $\overline{\Pi(\theta)}_{NL}$. The white curve indicates kernel density estimation (kde) of the statistics: $max(\sigma)$ and $max(kurt)$. The black lines show the 25th, 50th and 75th percentile of $max(\beta)$ estimated from even-sized bins of statistical values $max(\sigma)$ and $max(kurt)$ centered at the values corresponding to each plot mark

This figure also shows 25th, 50th, and 75th percentiles of $max(\beta)$ estimated from even-sized bins of $max(\sigma)$ and $max(kurt)$. Despite the presence of considerable scatter, the figure shows that relatively large increases of predictability by using nonlinear predictabil-

ity (i.e. large values of max(β)) are more likely to be found at large values of max($kurt$) and small values of max(σ). Notice that relatively large values of β are generally associated with low overall predictability (not many stations with relatively large $max(\beta)$ correspond to high $\overline{\Pi}_{NL}$). For the majority of stations with $max(kurt)$ close to 3, Π_{NL} differs little from Π_{LR}, and stations characterized by large fluctuations of surface wind components (i.e. large values of $max(\sigma)$) also tend to have similar linear and nonlinear predictability.

4.4 Discussion

The general form of a regression model is:

$$y = f(x;\gamma) + \text{noise}, \tag{4.2}$$

where $f(x;\gamma)$ is the model functional relationship between predictors and predictands, and γ is a vector of parameters. For a particular regression model applied to a particular data set, the "noise" includes contributions from variations in y not dependent on x (the intrinsic noise), and systematic biases between the class of functional forms that can be taken by $f(x;\gamma)$ and the true functional relationship between predictor and predictand. Poor predictability of y by model $f(x;\gamma)$ can therefore result from: (1) low SNR of the relationship between predictors and predictands, resulting from large variability of the intrinsic noise or a weak predictor-predictand relationship, or (2) the inadequate representation of the predictor-predictand relationship by the regression model $f(x,\gamma)$. By considering a collection of linear and nonlinear regression models with the same predictors and predictands, I include a broad range of possible representations of the relationship between free-tropospheric fields and surface wind components.

If the predictor-predictand relationship is characterized by strong noise (i.e. small values of SNR), the resulting statistical predictability will not be high regardless of the type of regression. With large intrinsic noise, the cross-validated predictability resulting from linear regression may exceed that of a nonlinear regression model because nonlinear regression models generally have a larger number of parameters and are more prone to fitting noise. In contrast, if the predictor-predictand relationship is characterized by relatively weak intrinsic noise, the predictability may depend systematically on the type of regression models used for statistical prediction. If the underlying predictive signal is strong (with a large SNR), both linear regression and nonlinear regression models will result in high predictability when the underlying relationship is close to linear, so long as the non-

linear regression considered can reduce to linear regression (as is the case for all of the nonlinear models considered in this study). In contrast, when the true relationship is nonlinear, prediction by LR should be inferior to that by the other regression methods.

Following the above discussion, there are two possible reasons that may explain why nonlinear regression does not substantially outperform linear regression. First, there is no evident nonlinear predictive signal, and second, the intrinsic noise associated with the predictor-predictand relationship is strong. The available information cannot unambiguously indicate which of these two possible reasons is more important in explaining why nonlinear regression cannot greatly outperform linear regression at most stations. However, the relationships between the statistical properties of surface wind components and β can provide some clues on this issue.

Chapter 3 showed that high linear predictability tends to be associated with surface wind components characterized by relatively large fluctuations with near-Gaussian distribution, and poor linear predictability is generally associated with wind components characterized by small fluctuations and non-Gaussian distribution. The association of high linear predictability and $\beta \approx 1$ with surface wind components characterized by large fluctuations and Gaussian distribution may suggest that the nonlinearity of predictive signals at these stations is weak.

On the other hand, the results show a systematic improvement in predictive skill resulting from nonlinear regression for surface wind components characterized by small fluctuations and non-Gaussian distribution with heavy tails or sharp central peaks. Although the trend of systematic improvement is not strong, it suggests that the nonlinear predictive signal is stronger than linear predictive signal at these stations. It should be emphasized that any such nonlinear predictive signals in general are weak at these stations as indicated by low predictability for both linear and nonlinear models. It is difficult to be certain that nonlinear predictive signals in the predictor-predictand relationship are the only cause of $\beta > 1$ at these stations since low predictability may be attributed to strong intrinsic noise in the predictor-predictand relationship. In other words, one cannot rule out the possibility that overfitting may also play a role in causing $\beta > 1$ for stations characterized by low predictability resulting from both linear and nonlinear regressions.

4.5 Conclusion

I have evaluated the predictability of surface wind components by statistical prediction using linear and a range of nonlinear regression methods as transfer functions at 2109 land stations across a wide range of locations. The results show that linear predictability of surface wind components at most stations is lower than predictability obtained by at least one of the three nonlinear regression methods, but the difference in predictability of surface wind components by linear regression and the best of the three nonlinear regression methods is generally small. Where there are improvements, these are generally found at stations where the overall level of predictability is low for all methods. Moreover, the anisotropy of predictability of surface wind components projected along different directions is a common feature for both linear and nonlinear regression methods, although stations with strong linear predictive anisotropy tend to have slightly weaker nonlinear predictive anisotropy. I found that many of the stations with relatively large changes in the magnitude and degree of anisotropy of predictability are in regions characterized by surface heterogeneity (e.g. major mountain ranges, land-water contrast). This statement cannot be reversed: not all stations in regions of heterogeneous landscape show relatively large differences in skill between linear and nonlinear regression predictions. Moreover, although the difference between linear and nonlinear predictability is generally small, at individual stations the nonlinear predictability tends to be higher than linear predictability in the direction of wind components with smaller fluctuations and non-Gaussian distributions characterized by heavier tails or higher peaks. This pattern is more obvious for daily averaged prediction than monthly averaged prediction. Averaging time series generally causes them to become more Gaussian and the statistical relationship between variables to be more linear (Yuval and Hsieh, 2002), which may explain the fact that improvements in nonlinear prediction relative to linear prediction are associated with surface wind components with more non-Gaussian distributions. However, this effect is not strong.

The relatively small change in skill of prediction by linear and nonlinear regression methods indicates that statistical predictability of surface wind components is generally not substantially influenced by the functional form but limited by the intrinsic noise. Future study is needed to assess what physical factors determine the magnitude of the intrinsic noise in the predictor-predictand relationship, which following Chapter 3 can be interpreted as being associated with local variability of atmospheric circulations on meso- and smaller scales. A subsequent study in Chapter 5 investigates how well RCMs and reanalysis products can simulate the physical processes responsible for influencing the intrinsic noise in

the predictor-predictand relationship. Finally, the results in this chapter suggest that since nonlinear predictive models generally do not substantially improve predictability of surface wind components using the predictors in this study, there is no advantage building the transfer function of statistical prediction using nonlinear regression methods that are more computationally expensive to apply than linear regression.

Chapter 5

Comparison of linear predictability of surface wind components from observations with simulations from RCMs and reanalysis

5.1 Introduction

Surface winds are a climatic field of importance in economic and societal sectors including air quality, agriculture, and transport. Global climate models (GCMs) can effectively model large-scale processes (e.g. synoptic to planetary scales). However, their coarse resolution (typically on the order of 100 km; Church et al. 2013) generally limits their skill in modeling surface wind components as they do not resolve the smaller microscale to mesoscale processes that also influence surface winds. One approach to predicting surface wind components is through statistical prediction in which a transfer function (TF) is built based on statistical relationships between station-based surface data and large-scale climate fields.

Previous studies have shown that the predictability of surface wind components by statistical prediction with linear TFs is often characterized by predictive anisotropy. That is, the predictability of surface wind components can vary with the direction of projection. Salameh et al. (2009) found that only one of the zonal and meridional wind components (u and v) at stations located in valleys of French Alps can be predicted well using statistical prediction with a generalized additive model as the transfer function. Other studies

used linear statistical prediction to predict surface wind components projected onto compass directions from 0° to 360° at 10° intervals. For example, van der Kamp et al. (2012) and Culver and Monahan (2013) applied linear statistical prediction to predict surface wind components in western and central Canada, and Monahan (2012) and Sun and Monahan (2013) studied linear prediction of sea surface winds. These studies found that predictability of surface wind components generally varies with directions of projection in the regions they considered. Chapter 3 further investigated the predictability of surface wind components at a large number of land stations across the globe, and found that linear predictive anisotropy is a common feature. Chapter 4 showed that predictive anisotropy is not an artifact of the use of a linear TF but is also found when statistical predictions are made using nonlinear regression models.

A question of interest is what limits predictability along certain directions of projections. Salameh et al. (2009) attributed the unequal predictability of zonal and meridional wind components at the location they considered to the orientation of the valley, as the across-valley wind component is characterized by local variability unexplained by large-scale climate fields. van der Kamp et al. (2012) and Culver and Monahan (2013) found similar topographically-oriented predictive anisotropy in the Western Cordillera of North America, and Chapter 3 showed that low predictability and strong predictive anisotropy tend to be associated with regions marked by surface heterogeneity, such as mountainous and coastal regions. These previous studies suggest that topography plays a role in limiting predictability of surface wind components along certain directions. However, topography does not seem to be the only factor determining predictive anisotropy of surface wind components as some locations are found with maximum predictability aligned across-valley rather than along-valley (van der Kamp et al., 2012; Culver and Monahan, 2013), and strong predictive anisotropy can also occur in regions with relatively flat terrain as shown in Chapter 3 and over the oceans (Monahan, 2012; Sun and Monahan, 2013). Chapter 3 showed that the best predicted surface wind components tend to be those of highest variability and with distributions closest to Gaussian, and that the poorest predictability is generally associated with wind components characterized by relatively weak variability and non-Gaussian distribution. It appears that no single factor determines predictive anisotropy.

I study a set of metrics of predictability in this chapter, with an emphasis on predictive anisotropy, by assessing how well these metrics can be represented by comprehensive physically-based models (i.e. different regional climate models and a global reanalysis) in three regions: North America (NAM), Europe-Mediterranean Basin (EMB) and East Asia (EAS). Regional climate models (RCMs) are a form of dynamical downscaling in which

physical processes are simulated at finer scale than by GCMs. Dynamical downscaling by RCMs is an alternative to statistical downscaling, and one of the motivations of this study is to assess how well RCMs can represent the observed statistical features of the relationship between free-tropospheric flow and surface winds. A systematic analysis of where the models represent this relationship well, and where they do not, also provides more insight regarding the physical controls on the statistical predictability of near-surface winds. The accuracy of RCMs depends on a number of factors, including the accurate representation of boundary conditions, the accuracy of driving data, the size of the domain, and the proper parameterization of physical processes (Rummukainen, 2010). I only consider simulations from RCMs driven by observationally-constrained reanalysis boundary conditions. Although local-scale dynamics near the land surface generally cannot be modeled with good skill by reanalysis (He et al., 2010b), large-scale features in the free troposphere are well represented by reanalysis products. Therefore, free-tropospheric climate fields output by reanalysis are generally considered as reliable boundary conditions. By using reanalysis-driven RCMs, I can focus the discussion of the simulation of predictive anisotropy on physical processes described by regional models rather than considering the potential systematic bias inherent in the driving GCMs. Although reanalysis products are generally closer to observation than simulations from both GCMs and RCMs, they also have limitations due to observational constraints and the accuracy of the assimilation models (Parker, 2016).

Neither RCMs nor reanalyses can be regarded as perfect representations of point observations. In general, the differences in statistical relationships between mid-tropospheric state and surface winds in models and observations can be attributed to (1) unresolved or poorly represented physical processes in the model and (2) the difference between point measurements (observation) and averages over the scale of a gridbox (models). While my analysis is not able to distinguish between these two sources of difference, irrespective of the source of difference, this characterization is useful from the perspective of determining the utility of the RCMs as tools for dynamical downscaling. As well, in either case, differences between observed and modelled statistical relationships provide insight regarding the relative importance to these relationships of large-scale, resolved processes and small-scale, unresolved processes.

I also further elaborate a mathematical model of directional predictability introduced in Chapter 3 based on an idealized partitioning of surface wind components into large-scale “signal” and local “noise”. The idealized model provides a conceptually organizing perspective on the controls on statistical predictability of surface winds and how well these

are represented in more comprehensive models like RCMs and reanalysis. I only consider linear statistical prediction in this chapter because the results of Chapter 4 show that the predictive skill resulting from nonlinear regression based TFs is not very different from linear TFs. Finally, while Chapter 4 considered statistical prediction of both daily and monthly averaged surface winds, I focus on daily averaged data quantities in this chapter as the time period of the RCM simulations considered in this study may not be long enough to give robust results of statistical prediction using monthly averaged data.

This chapter is organized as follows. Section 5.2 presents the data considered and methods used in the comparison of observed and modeled features of statistical prediction. Section 5.3 introduces the idealized mathematical model used as a conceptual framework for understanding controls on surface predictability. Section 5.4 compares the structures of statistical prediction in observation and comprehensive models. A discussion relating the results to the idealized model is presented in Section 5.5, and conclusions are in Section 5.6.

5.2 Data of RCMs and reanalysis

In this Chapter, I consider statistical predictions of surface wind components by linear prediction at a subset of the total 2109 land stations (Fig. 1.1) consisting of 557 stations in NAM, 595 stations in EMB, and 715 stations in EAS. These regions are chosen based on the availability of RCM simulations, and because they have higher station densities than other areas of the Northern Hemisphere (Fig. 5.1). Statistical prediction presented in this chapter follows the approach shown in Chapter 2.

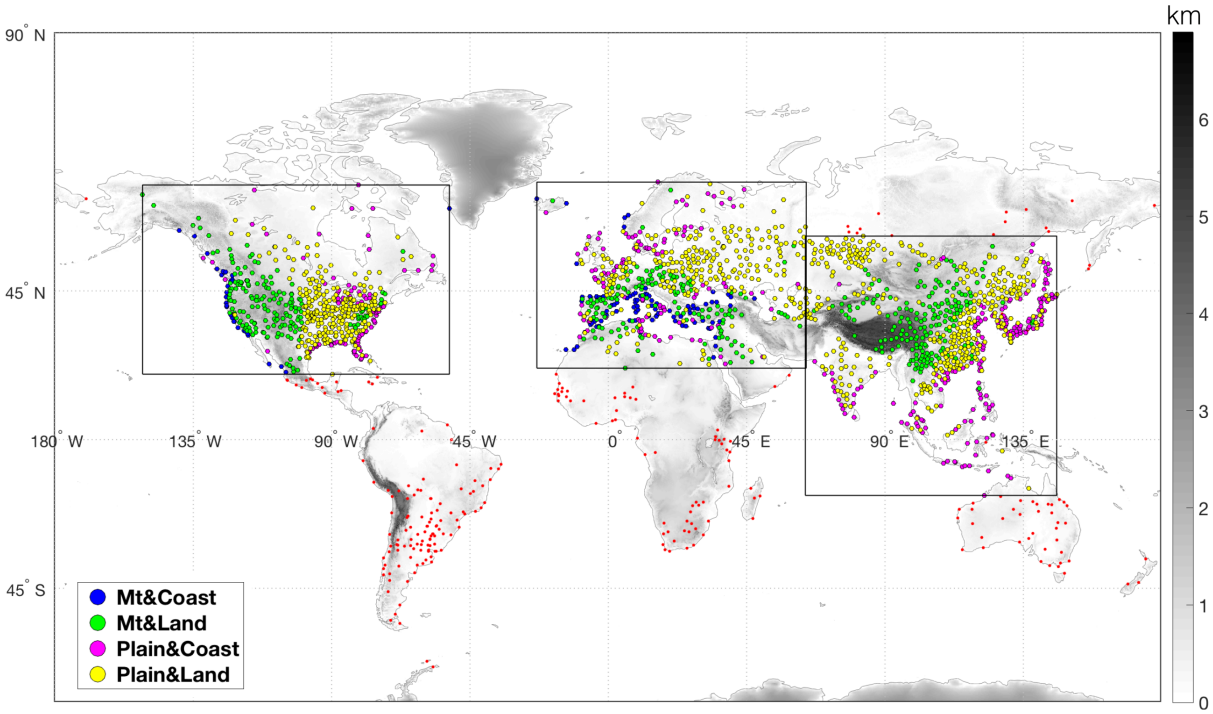

Figure 5.1: The locations of the 2109 land stations used for statistical prediction of surface winds. The domains of North America (NAM; 557 stations), Europe-Mediterranean Basin (EMB; 595 stations) and East Asia (EAS; 715 stations) are outlined. Stations in the three domains are classified into 4 groups according to local topography and proximity to water.

In order to assess the connection between surface heterogeneity and characteristics of predictability of surface wind components, I classify these stations according to two categories: (1) whether the station is in a mountainous region or in flat terrain (denoted Mt or plain) (2) whether the station is adjacent to water or inland (denoted coast and land). These two categories result in the 4 groups of stations illustrated in Fig. 5.1. The classification of station locations as mountain or plain is based on the maximum elevation within 0.2° from the station location. If the maximum elevation is larger than 1000 m, the station is classified as a mountain station; otherwise, it is a plain station. The radius of 0.2° is chosen to ensure that the classification is based on local terrain. The elevation data used for the classification is 1 arc-minute global relief data from the ETOPO1 Global Relief Model (Amante and Eakins 2009; downloaded from `https://www.ngdc.noaa.gov/mgg/global/global.html`). Coastal stations are classified using the coastline data provided by the Mapping Toolbox of MATLAB (MathWorks, 2016). If the location of a station is within the range of 30 km from the nearest coastal boundary, the station is classified as coastal. In such loca-

tions, the surface winds are likely to be influenced by land-water contrast since sea breezes commonly extend inland as far as 30 km (Oke, 2002).

The RCM simulations in this study are chosen according to available time coverage and resolution from existing runs from the Coordinated Regional Climate Downscaling Experiment (CORDEX) project as well as the North American Regional Climate Change Assessment Program (NARCCAP). One of the purposes of the CORDEX project is to provide a quality-controlled dataset of downscaled information as a model evaluation framework (Giorgi et al., 2009). The NARCCAP project is currently the most comprehensive regional climate-modeling project for climate change impact studies in North America (Mearns et al., 2009). Longer duration simulations can lead to more robust statistical results, and higher resolution can potentially contribute to modeling the local-scale physical processes with higher accuracy. All RCM simulations used in this study are at least 19 years long. The horizontal resolution for all RCMs in this study is at least 50 km or finer. Table 5.1 summarizes the basic information for all datasets used in this study. More detailed information specific to the RCMs is shown in Table 5.2.

Table 5.1: Summary information for each type of data considered in this study

Data type	Labels	Source	Region available	Spatial resolution	Temporal resolution	Period
Observation	OBS	WeatherData function of Mathematical 9.0	global land		hourly	1980/01/01-2012/12/31
Reanalysis	NCEP2	http://www.esrl.noaa.gov/psd/	global region	surface: $\sim$ 200km; 500hPa: $2.5^\circ \times 2.5^\circ$	daily	1980/01/01-2012/12/31
RCMs	NA1	http://climate-modelling.canada.ca/climatemodeldata/canrcm/CanRCM4/	North America	25km	daily	1989/01/01-2009/12/31
	NA2	http://www.narccap.ucar.edu	North America	50km	3-hourly	1980/01/01-2000/12/31
	EUR1	http://www.cordex.org/index.php/esgf-menu	Europe	12.5km	daily	1980/01/01-2010/12/31
	EUR2	http://climate-modelling.canada.ca/climatemodeldata/canrcm/CanRCM4/	Europe	25km	daily	1989/01/01-2009/12/31
	EAS1	https://cordex-ea.climate.go.kr/main/dataSurvey.do?menuId=download	East Asia	50km	daily	1989/01/01-2008/12/31
	EAS2	https://cordex-ea.climate.go.kr/main/dataSurvey.do?menuId=download	East Asia	50km	daily	1989/01/01-2008/12/31

For prediction using reanalysis surface winds, u and v in Eq. (2.3) for calculating the surface wind components (i.e. predictands) are the eastward wind and northward wind at 10 m from NCEP-2 Reanalysis data (Kanamitsu et al., 2002). To compare prediction results between observations and model products, I wish to obtain values of model fields at locations nearest to the observational stations considered in this study. Since the location of an observational station generally will not correspond exactly to a point in RCM or reanalysis grids, I need to estimate values of u and v at each station using the model output.

Table 5.2: Details of RCMs considered in this study

Labels	Driving reanalysis	Regional models	Modeling Group	Project	References
NA1	ECMWF-ERAINT	Canadian Regional Climate Model 4 (CanRCM4)	Canadian Centre for Climate Modelling and Analysis (CCCma)	CORDEX	Scinocca et al., 2016
NA2	NCEP2	Weather Research & Forecasting Model (WRF)	Pacific Northwest National Lab, US	NARCCAP	Mearns et al., 2014
EUR1	ECMWF-ERAINT	Rossby Centre regional atmospheric model (RCA4)	Swedish Meteorological and Hydrological Institute (SMHI)	CORDEX	Strandberg et al., 2015
EUR2	ECMWF-ERAINT	Canadian Regional Climate Model 4 (CanRCM4)	Canadian Centre for Climate Modelling and Analysis (CCCma)	CORDEX	Scinocca et al., 2016
EAS1	ECMWF-ERAINT	HadGEM3-RA	National Institute of Meteorological Research (NIMR), Korea	CORDEX	Davies et al., 2005
EAS2	NCEP2	Weather Research & Forecasting Model (WRF)	Seoul National University (SNU), Korea	CORDEX	Skamarock et al., 2005

Two methods to estimate u and v at station locations based on grid point values are considered. The first one simply takes u and v at the grid point nearest to the station. The second representation is based on inverse distance weighting. Specifically, a group of neighboring points surrounding a station is identified, and the representation of u and v at the station is the weighted average of u and v at these grid points:

$$x_{station} = \frac{\sum_{i=1}^{N} x_i w_i}{\sum_{i=1}^{N} w_i}, \tag{5.1}$$

where x stands for u or v, N is the number of neighboring grid points, and w_i is the inverse square of the distance between each grid point and the station location. In this study, I use weighted averages with $N = 8$. The number of neighboring grid points $N = 8$ is chosen based on empirical tests which show no significant difference in values of u and v by Eq. (5.1) for $N \geq 8$. Moreover, the difference in results between the first and second representation is small and does not change the final results (not shown). Note that both representations still potentially suffer from biases resulting from the fact that grid-point values represent spatial averages on the order of the grid resolution, while station data are point measurements. In particular, the model fields will not include local variability on scales smaller than the grid resolution.

The predictors in the study consist of free tropospheric meteorological fields: temperature T, geopotential height Z, zonal wind U, and meridional wind V at 500 hPa. Following Chapter 2, the four predictor fields are chosen from a domain of $40^\circ \times 40^\circ$ centered at each station. For prediction of observational data and reanalysis surface fields, the four predictor fields are from the NCEP-2 reanalysis product. For prediction of RCM surface fields, the four predictor fields are taken from the output of the corresponding RCMs in a domain of $40^\circ \times 40^\circ$ centered at the station. Since the resolution of RCMs is generally much finer than

that of NCEP-2 reanalysis, only RCM grid points closest to each grid point of reanalysis fields within the $40^\circ \times 40^\circ$ domain are kept.

The time period of the observations and the reanalysis product is 1980/01/01 to 2012/12/31. The available time period for RCMs is generally shorter; approximately 20 years for most of the RCMs used in this study (Table 5.1). I divide data into seasons of June, July, August (JJA) and December, January, February (DJF). The minimum sample size of multiple linear regression based on a comprehensive study by Green (1991) is $50+8m$ (where m is the number of predictors). Accordingly, I should have at least $N_{sample} = 82$ data points for each regression model as $m = 4$ in this study. As a result, I only consider daily averaged prediction for the analysis in this chapter since the data size of monthly-averaged RCM output for a given season with 20 years of data ($N_{sample} = 20\,\text{years} \times 3\,\text{months/year}$) may not be large enough to obtain a robust statistical fit.

5.3 Idealized model of predictability

I use an idealized model to provide a conceptual framework for the controls on characteristics of surface predictability. The idealized model, which builds on a similar model in Chapter 3 is based on the assumption that surface wind variability can be partitioned into distinct signal and noise contributions when using large-scale climate variables for statistical prediction. By definition, the signal refers to the part of the surface winds perfectly predicted by the large-scale circulations, and the noise refers to the part of surface winds that originates from local processes and cannot be explained by large-scale predictors. The decomposition implies that the signal and noise are uncorrelated. A wind vector can be expressed in terms of the components $\tilde{u}$ and $\tilde{v}$ in an orthogonal basis $(\hat{e}_{\tilde{u}}, \hat{e}_{\tilde{v}})$

$$\begin{cases} \tilde{u}(t) = \tilde{u}_s(t) + \tilde{u}_n(t) \\ \tilde{v}(t) = \tilde{v}_s(t) + \tilde{v}_n(t), \end{cases} \tag{5.2}$$

where the subscripts s and n respectively denote signal and noise. Note that it is free to choose the orientation of the basis vectors in this decomposition; they do not need to align with the zonal and meridional directions. The wind component projected onto direction θ (with θ measured clockwise away from $\hat{e}_{\tilde{v}}$) is

$$U(\theta) = U_s(\theta) + U_n(\theta), \tag{5.3}$$

where $U_s(\theta) = \tilde{u}_s \sin(\theta) + \tilde{v}_s \cos(\theta)$ and $U_n(\theta) = \tilde{u}_n \sin(\theta) + \tilde{v}_n \cos(\theta)$. The predictability of surface wind components $\Pi(\theta)$ is then measured by the squared correlation coefficient

$$corr^2(U, U_s) = \frac{cov^2(U, U_s)}{var(U)var(U_s)}. \tag{5.4}$$

It follows that
$\Pi(\theta) = \frac{\text{var}(\tilde{v}_s)\cos^2(\theta)+\text{var}(\tilde{u}_s)\sin^2(\theta)+\text{cov}(\tilde{u}_s,\tilde{v}_s)\sin(2\theta)}{[\text{var}(\tilde{v}_s)+\text{var}(\tilde{v}_n)]\cos^2(\theta)+[\text{var}(\tilde{u}_s)+\text{var}(\tilde{u}_n)]\sin^2(\theta)+[\text{cov}(\tilde{u}_s,\tilde{v}_s)+\text{cov}(\tilde{u}_n,\tilde{v}_n)]\sin(2\theta)}$.

I define

$$\eta = \frac{\text{var}(\tilde{v}_s)}{\text{var}(\tilde{v})}, \tag{5.5}$$

and

$$\zeta = \frac{\text{var}(\tilde{u}_s)}{\text{var}(\tilde{u})}. \tag{5.6}$$

The quantities η and ζ represent the fractions of surface wind components projected onto $\hat{e}_{\tilde{v}}$ and $\hat{e}_{\tilde{u}}$ that can be explained by the large-scale predictors; in other words, they represent the signal variance fraction of $\tilde{v}(t)$ and $\tilde{u}(t)$. As by construction, $\text{cov}(\tilde{u}, \tilde{v}) = \text{cov}(\tilde{u}_s, \tilde{v}_s) + \text{cov}(\tilde{u}_n, \tilde{v}_n)$, and using the definition of the correlation coefficient $\rho(x,y) = \frac{\text{cov}(x,y)}{\sqrt{\text{var}(x)\text{var}(y)}}$, I obtain

$$\rho(\tilde{u}, \tilde{v}) = \rho_s \sqrt{\eta\,\zeta} + \rho_n \sqrt{(1-\eta)(1-\zeta)}, \tag{5.7}$$

where $\rho_s = \rho(\tilde{u}_s, \tilde{v}_s)$ and $\rho_n = \rho(\tilde{u}_n, \tilde{v}_n)$. It follows that I can express $\Pi(\theta)$ in any direction in terms of the fractions of variance of surface wind components explained by the large-scale predictor along $\hat{e}_{\tilde{u}}$ and $\hat{e}_{\tilde{v}}$ (i.e., ζ,η), and the anisotropy of variability of surface wind components projected onto directions of the coordinates, $\gamma = \sqrt{\frac{\text{var}(\tilde{v})}{\text{var}(\tilde{u})}}$, correlations ρ, and ρ_s:

$$\Pi(\theta) = \frac{\eta\,\gamma\cos^2(\theta) + \zeta\,\frac{1}{\gamma}\sin^2(\theta) + \rho_s\sqrt{\eta\,\zeta}\,\sin(2\theta)}{\gamma\cos^2(\theta) + \frac{1}{\gamma}\sin^2(\theta) + \rho(\tilde{u},\tilde{v})\,\sin(2\theta)}. \tag{5.8}$$

The constraint that $0 \leq \Pi(\theta) \leq 1$ follows as each of the squared wind component correlations $\rho^2(\tilde{u}, \tilde{v})$, $\rho^2(\tilde{u}_s, \tilde{v}_s)$, $\rho^2(\tilde{u}_n, \tilde{v}_n)$ fall between zero and one. This model reduces to the form of that in Chapter 3 if $\rho(\tilde{u}, \tilde{v}) = 0$ and $\text{var}(\tilde{u}_n) = \text{var}(\tilde{v}_n)$. While the parameters η, ζ, γ, ρ and ρ_s are distinct in the model, their observed distributions are not necessarily independent. This point will be discussed further in Section 5.5.

5.4 Results

In this section, I present the comparisons of metrics of predictability derived from comprehensive models (i.e. RCMs and reanalysis) with those from observations in the NAM, EMB and EAS regions. Maps of the spatial distribution of metrics of predictability ($min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$) are shown in Figs. 5.2-5.4.

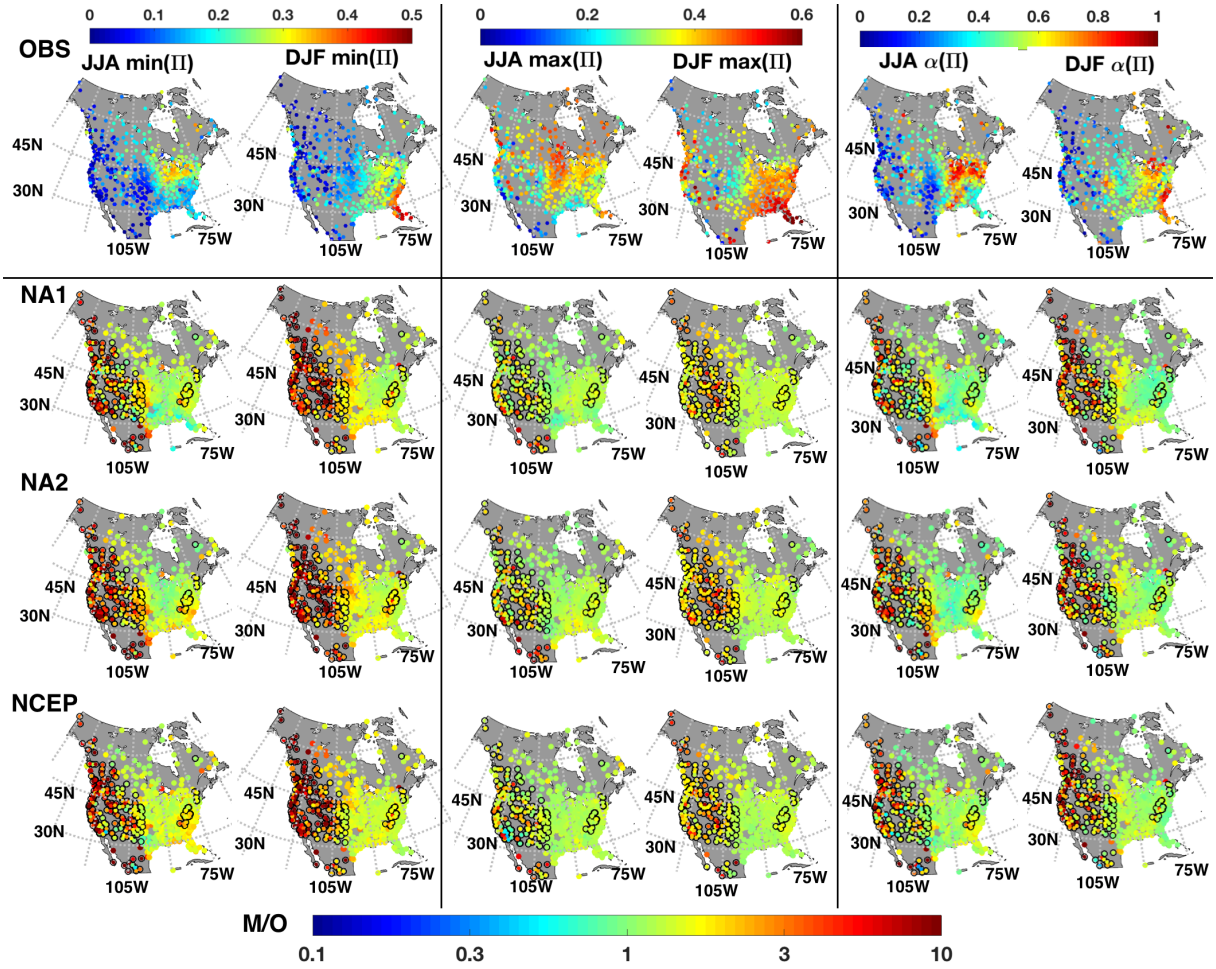

Figure 5.2: Top row: observed daily $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$ for JJA and DJF data for the NAM domain. The second, third and fourth rows show the comparison of predictability metrics derived from observations with those derived from the two RCM models and NCEP-2 reanalysis in terms of the ratio of the modelled value to the observed. The color scale is logarithmic. Stations with black outlines are in mountainous regions; those without are in plain regions. M/O stands for model/observation.

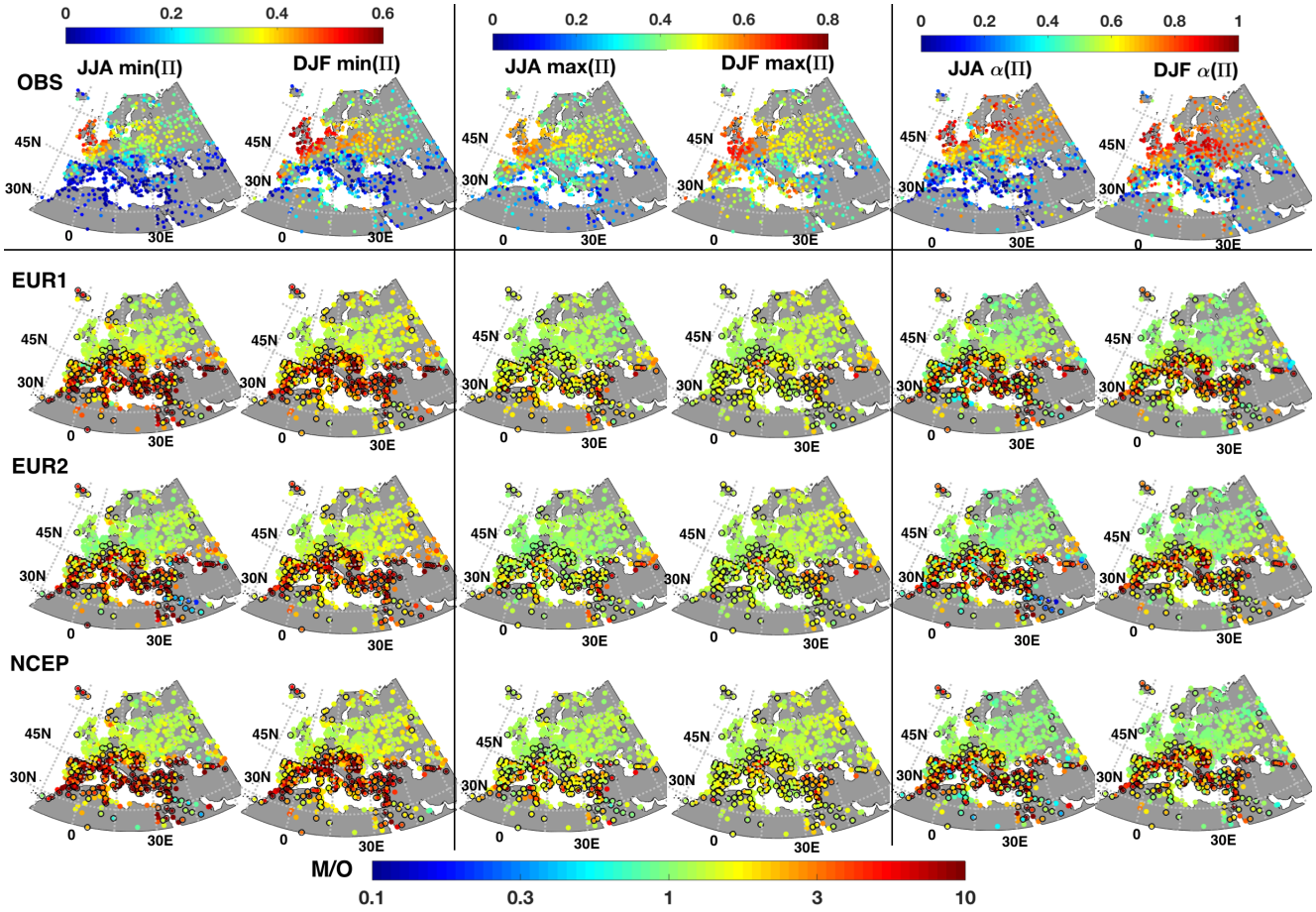

Figure 5.3: As in Fig. 5.2, but in the EMB domain.

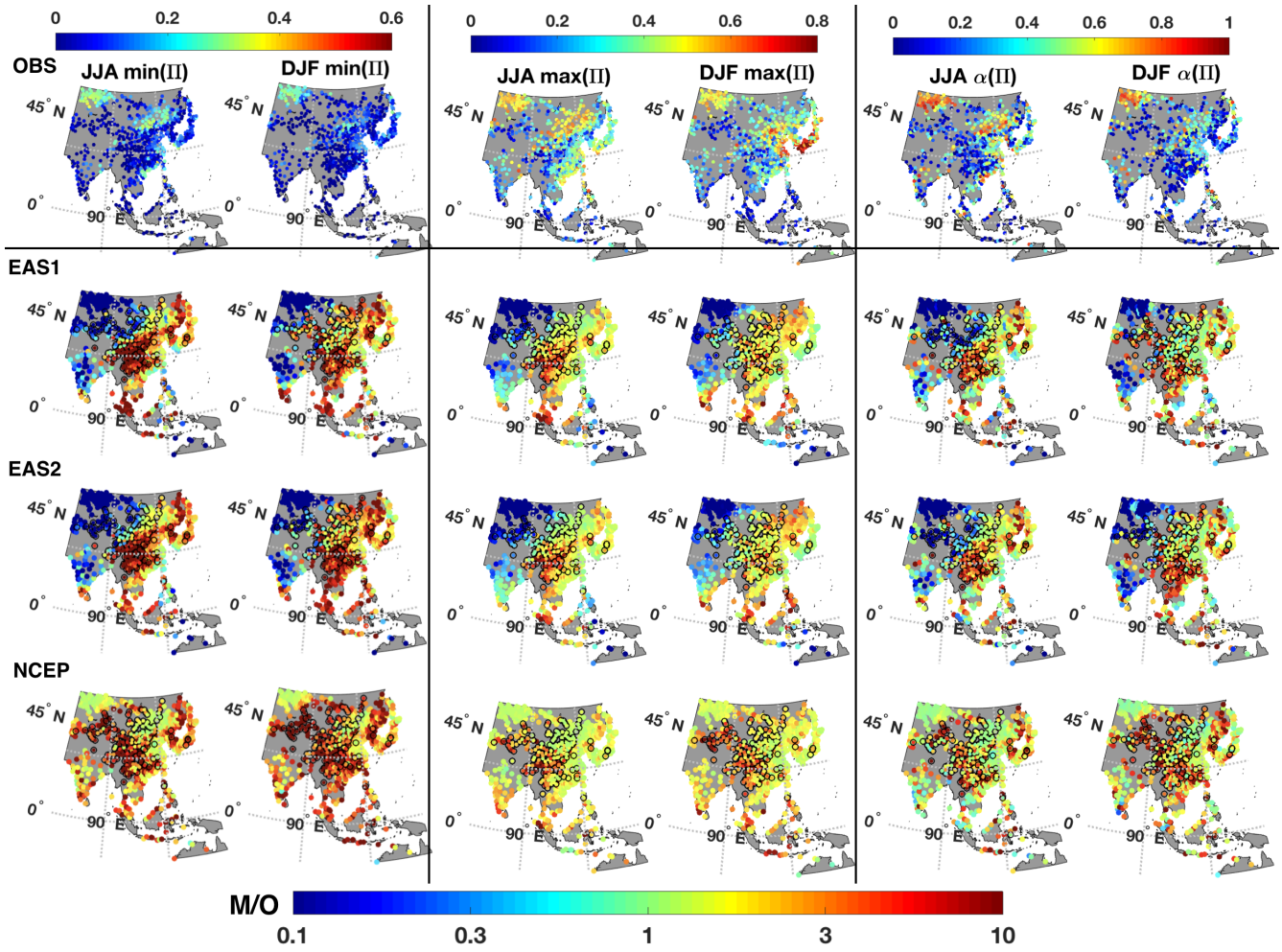

Figure 5.4: As in Fig. 5.2, but in the EAS domain

The comparison of predictability metrics resulting from comprehensive models and observations is quantified using the ratio $\frac{M}{O}$, where M refers to the values of $min(\Pi)$, $max(\Pi)$ or $\alpha(\Pi)$ derived from comprehensive models, and O refers to the corresponding values from observations. Observations show that mountainous Western NAM generally has lower predictability and stronger predictive anisotropy than the relatively flat Eastern NAM (Fig. 5.2). Southern EMB (characterized by both complex terrain and land/sea contrast) generally has lower observed predictability and stronger predictive anisotropy (Fig. 5.3) than North-Central Europe and western Russia. The EAS domain (Fig. 5.4) also shows some connection between terrain and observed metrics of predictability. For example, central Asia extending from Western China to the western boundary of the domain is characterized by both mountainous and relatively flat terrain. This region is generally characterized by low predictability and strong predictive anisotropy, except for the relatively flat region

in the area North of 45°N and west of 90°E where overall predictability is higher and predictive anisotropy is weaker. There are more stations with higher observed $max(\Pi)$ and weaker predictive anisotropy in northeast China, and most of Japan as well as in India. The terrain in these regions is relatively flat.

The relationships between metrics of predictability and topography are similar in NAM and EMB for each of $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$. The values of all three measures modeled by both RCMs and reanalysis are higher than in observations in regions characterized by mountainous terrain (i.e. western NAM and southern EMB), while modeled values of $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$ are generally in reasonable agreement with observations in relatively flat regions. In EAS, there is no systematic contrast in values of $\frac{M}{O}$ between mountainous and plain regions, but there is a distinct contrast between the region west of roughly 90°E where modeled values of $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$ by the two RCMs are significantly smaller than observations, and the region east of 90°E where large values of $\frac{M}{O} > 1$ are common.

Amongst all three domains considered, the region west of 90°E in EAS is the only region where extensive underestimation of metrics of predictability ($\min(\Pi)$, $max(\Pi)$ and α) from RCMs is observed. In addition, the area north of 45° and west of 90°E in EAS is particularly notable as this area is characterized by observations with relatively high predictability and nearly isotropic directional predictability, whereas the rest of the area of underestimation west of 90°E is characterized by relatively low predictability and strong predictive anisotropy in observations. The contrast of underestimation and overestimation for areas divided by 90°E may be due to some systematic bias of the regional dynamic downscaling models used in the two RCMs in this region. The reason for this bias is not clear, and one possible reason could be that upstream information on the background flow outside the model domain is not properly represented by the RCMs.

The relationship between metrics of predictability from comprehensive models and observations can be summarized in Taylor diagrams (Figs. 5.5-5.7) which is a graphical tool to assess how closely modeled results can match observation by quantifying the spatial correlation between model and observation (R), centered root-mean-square difference (E), and spatial deviation of observation- (σ_o) and model-based fields (σ_m) (Taylor, 2001). Mathematically, the relationships among these statistics can be summarized by:

$$E = {\sigma_m}^2 + {\sigma_o}^2 - 2\sigma_m\sigma_o R. \tag{5.9}$$

The construction of the Taylor diagram is based on the similarity between Eq. (5.9) and the

Law of Cosines (Taylor, 2001):

$$c^2 = a^2 + b^2 - 2abCos(\phi), \tag{5.10}$$

where ϕ is the angle between a and b. It is a common practice to normalize the statistics (i.e., E, σ_m, σ_o) by σ_o so that the reference point (i.e. observation) can be 1 unit away from the origin.

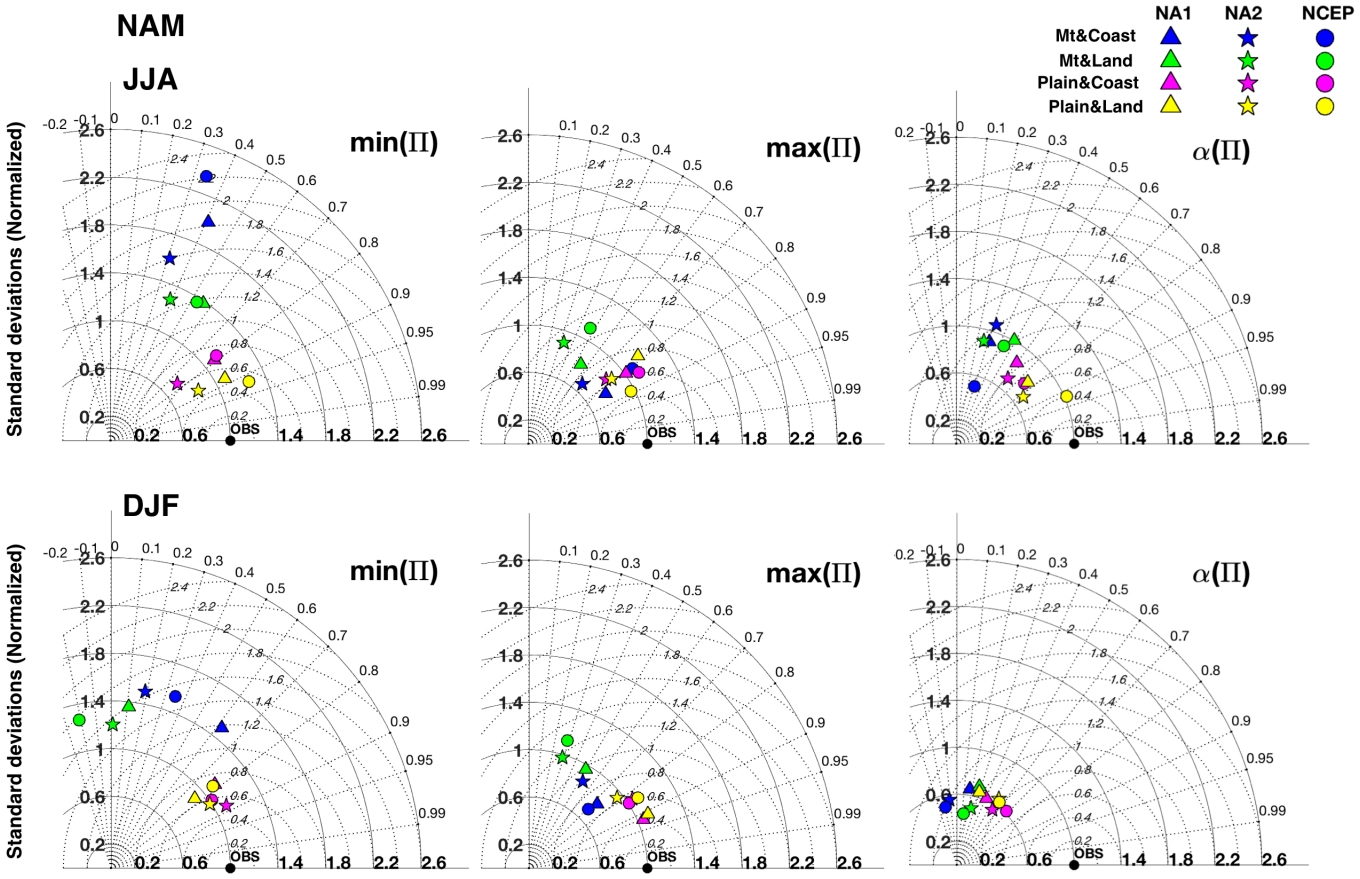

Figure 5.5: Taylor diagrams showing the comparison of predictability of surface wind components obtained from observations with those obtained from models (i.e. RCMs NA1, NA2 and NCEP-2 reanalysis) for groups of stations classified by their terrains in NAM. The angular coordinate corresponds to spatial correlation between modeled results and observation. The arcs of black solid circle indicate the spatial deviation of the modeled results, and the arcs of black dotted circle indicate the centered root-mean-square difference. The closer the distance of a mark (representing a model) to the OBS point (i.e. observation), the better the results from the model are.

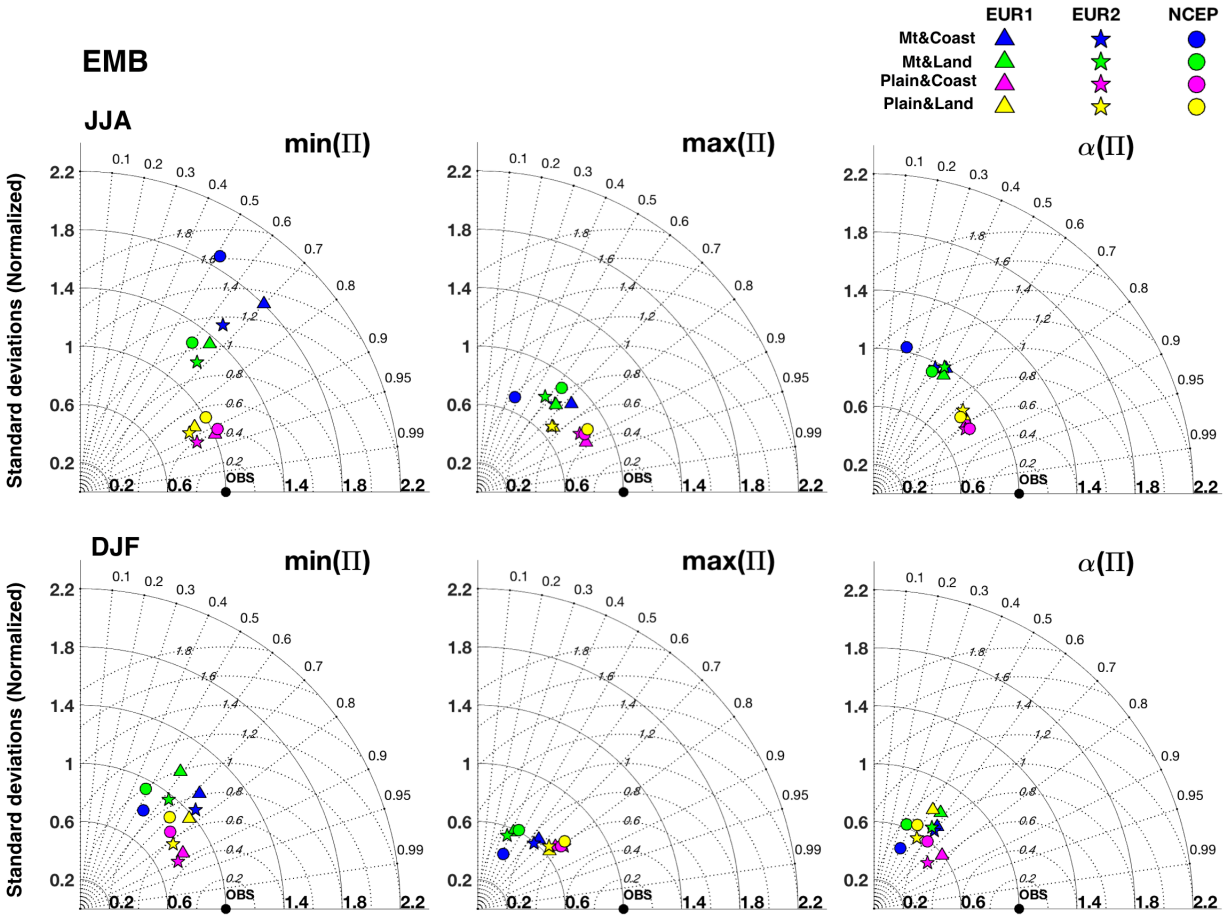

Figure 5.6: As in Fig. 5.5, but in EMB.

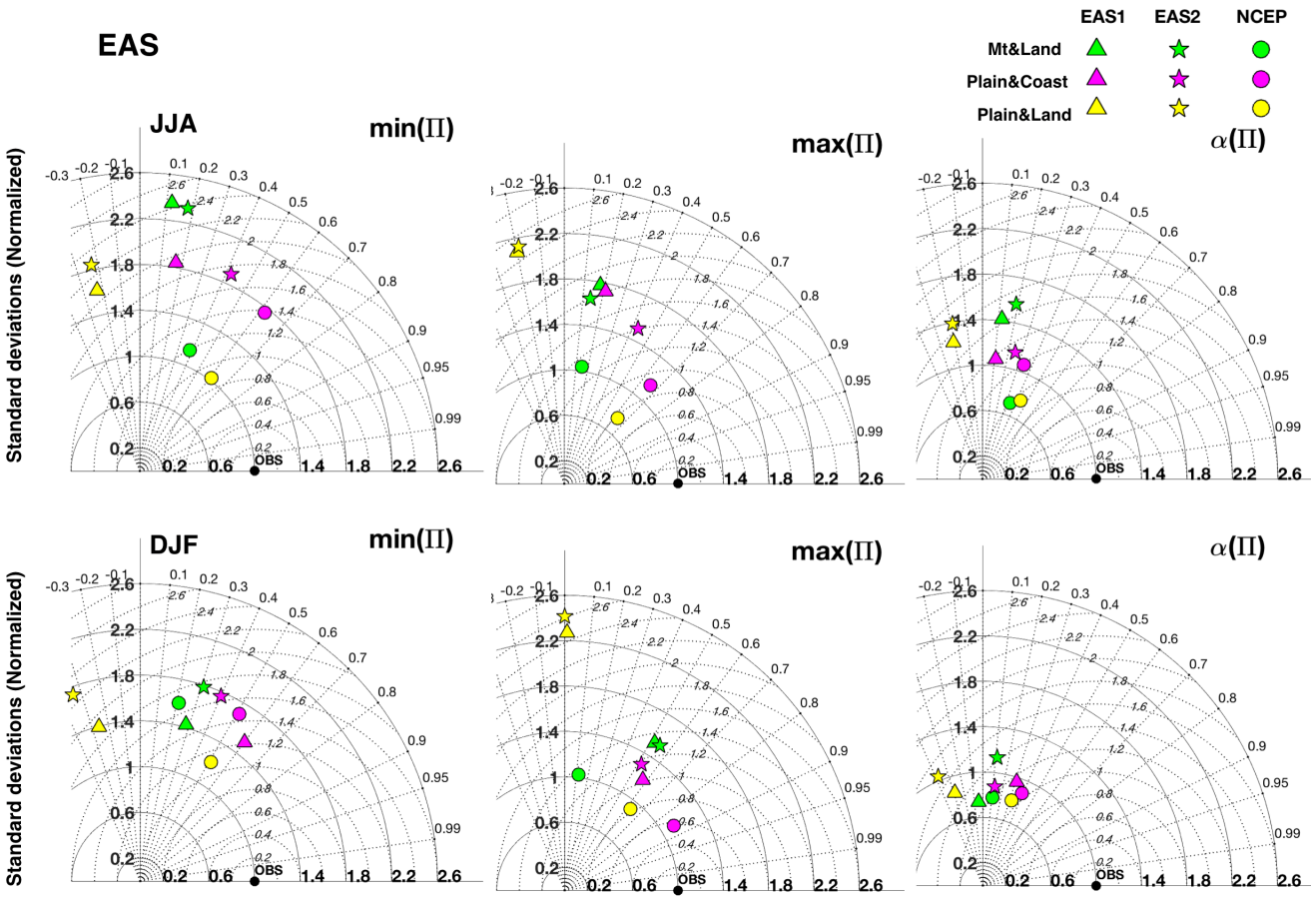

Figure 5.7: As in Fig. 5.5, but in EAS.

To facilitate comparison between Taylor diagrams for different regions, seasons, and terrain types, all fields are normalized by the spatial standard deviation of the corresponding observational fields. Several overall patterns are evident from Figs. 5.5-5.7. The differences in predictive structure resulting from various models for different terrain groups are larger for $min(\Pi)$ than for $max(\Pi)$ or $\alpha(\Pi)$. The difference between mountainous and plain terrain is more evident than the difference between coast and land. Seasonal differences are more pronounced for $min(\Pi)$ in mountainous regions, and more evident in NAM and EMB than in EAS. There is no systematic pattern of similarities or differences between inland and coastal stations in either season. Model differences between the two RCMs in each domain are small in general, and no one RCM considered is systematically better than the other in any of the domains considered.

As can also be seen by the maps in Figs. 5.2 and 5.3, in NAM and EMB, predictability measures from comprehensive models are generally closer to those from observations for

stations in plain regions relative to mountainous ones. This pattern is most evident for $min(\Pi)$. In EAS, modeled quantities related to predictability are generally not close to observations for any type of terrain (see also Fig. 5.4). The relatively coarse-resolution (but observationally informed) reanalysis is not systematically better or worse than RCMs in NAM and EBM and in the mountainous region in EAS. In contrast, for the plain region in EAS, both coastal and inland, reanalysis is generally closer to observations than both RCMs, and this is especially evident for inland stations in the plain regions.

5.5 Discussion

I will focus on DJF results as the seasonal differences between predictability measures simulated by comprehensive models and from observations are generally not large, and only the difference between mountainous and plain terrain is discussed as there is no clear contrast between coastal and inland stations.

Overestimation of the metrics of predictability: $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$ is commonly observed in NAM and EMB and east of 90°E in EAS. Panels (a),(b),(c) in Fig. 5.8 show the relationships between $\frac{M}{O}$ and the the observed values for each of $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$. Panel (d) in Fig. 5.8 compares the direction of $min(\Pi)$ and direction of the maximum value of $\frac{M}{O}$, denoted $max(\frac{M}{O}[\Pi(\theta)])$ through histograms of the dot product of unit vectors aligned along these directions, denoted $\hat{e}_{max(\frac{M}{O}[\Pi])} \cdot \hat{e}_{max(\Pi)}$. Since there is little difference between the the two RCMs used in each domain in terms of the relationships between $\frac{M}{O}$ and observed values, I only show the relationships resulting from NCEP-2 reanalysis and one of the RCMs in each domain. In addition, since predictability metrics from RCMs in west of 90°E of EAS is variant, I only consider east of 90°E in EAS in the following analysis.

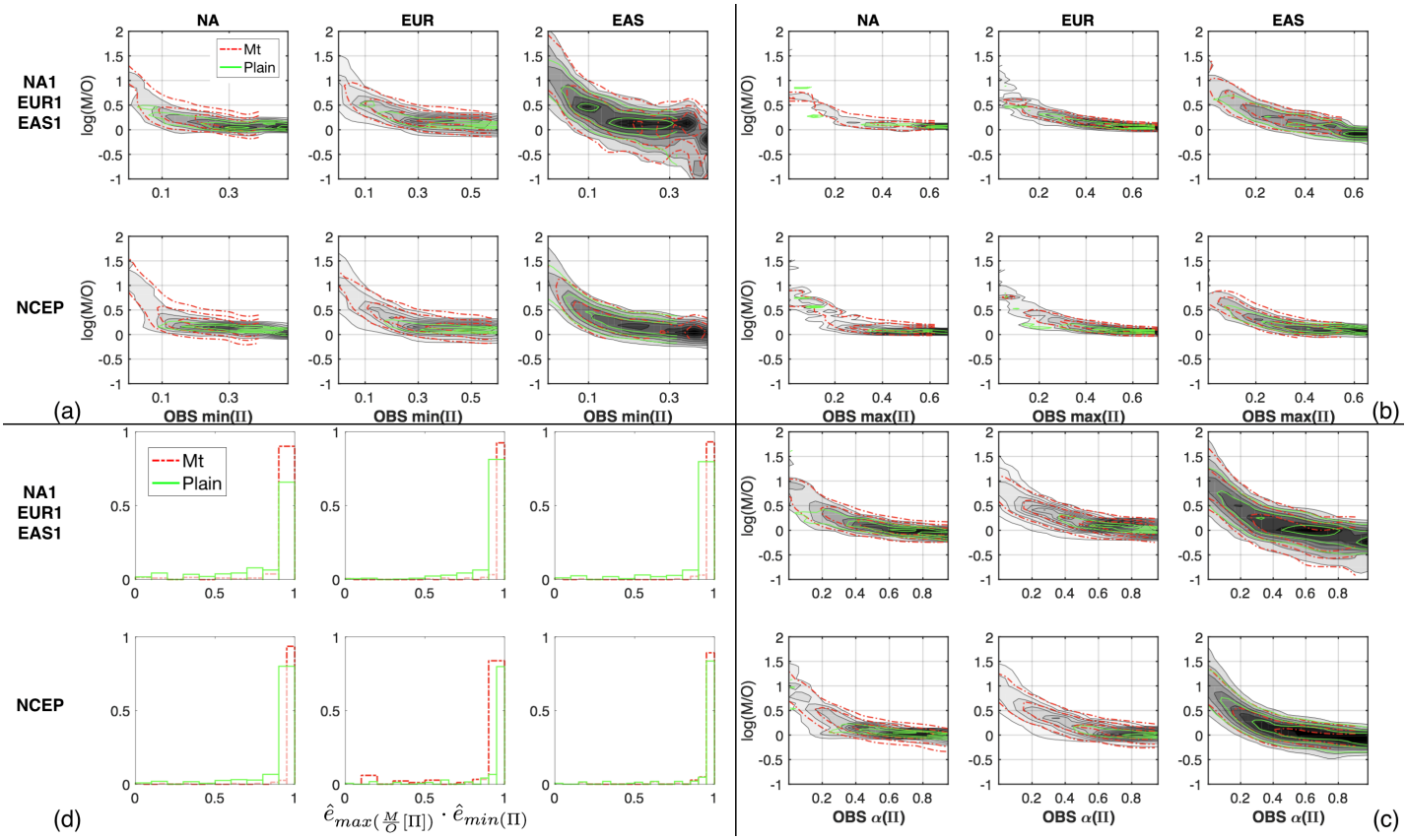

Figure 5.8: Panels (a)-(c): Estimated probability density functions of $log(M/O)$ conditioned on observed values of predictability metrics ($min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$). Filled contours are obtained by using data from all stations in the region; red and green contours are obtained by stations only from mountainous and plain regions respectively. Panel (d) shows the histogram of the dot product of the unit vectors in the directions of minimum predictability and maximum M/O for both mountainous and plain stations.

Panels (a),(b),(c) in Fig. 5.8 show that values of $\frac{M}{O}$ for $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$ are inversely correlated with the observed predictability metrics in all three domains considered. The ratio $\frac{M}{O}$ tends to exceed 1 for smaller values of observed predictability metrics, and $\frac{M}{O}$ tends to approach 1 as these predictability metrics increase. That is, in general overestimates of simulated predictability metrics are found in regions where they are small in observations. However, in EAS, $\frac{M}{O} < 1$ corresponds to relatively large observed predictability metrics, and the reason of this variant structure is that the division by 90°E is approximate, and there are still some variant stations with relatively large underestimates of predictability metrics in the east of 90°E. Panel (d) shows that the predominant value of

$\hat{e}_{max(\frac{M}{O}[\Pi])} \cdot \hat{e}_{max(\Pi)}$ is 1 in any domain and for both mountainous and plain terrain, which indicates that the largest overestimation of predictability by comprehensive models tends to occur along the direction of minimum observed predictability at most stations. Note also that while small values of observed predictability metrics and their overestimate by comprehensive models are more common in mountainous areas, they also occur in regions classified as plain.

Since linear predictability defines a measure of SNR of surface wind components, the general pattern shown in Fig. 5.8 indicates that the SNR of observed surface wind components in poorly-predicted directions tends to be inflated when using comprehensive models to model the surface winds, thereby artificially weakening predictive anisotropy. In contrast, relatively large values of SNR of observed surface wind components are approximately the same for modeled wind components output from comprehensive models. Overestimation of the predictability metrics is consistent with small-scale physical processes that contribute to predictive noise of surface wind components (i.e. resulting in small SNR) in reality not being captured by comprehensive models. Small values of observed $min(\Pi)$ are more likely to be over-predicted than small values of $max(\Pi)$ (the magnitude of $\frac{M}{O}$ is larger for $min(\Pi)$ than $max(\Pi)$ in the interval of small observed values) indicating that $min(\Pi)$ is more influenced by small-scale processes than $max(\Pi)$.

Using the conceptual framework derived from the idealized mathematical model, predictability metrics of surface wind components can be related to the model parameters (i.e. η, ζ, γ, ρ and ρ_s defined in Section 5.3). The physical processes unresolved by comprehensive models leading to over-prediction can result in comprehensive models having a different local/large-scale decomposition from observations and therefore occupying a different region of idealized model "parameter space" than observations. In order to study how the predictability metrics: $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$ from both observations and comprehensive models are related to the idealized model parameters, I focus on the orthogonal coordinates aligned along and across the direction of maximum predictability at each station, denoted respectively ($\hat{e}_{max(\Pi)}$,$\hat{e}_{max(\Pi)\perp}$). While the directions of $max(\Pi)$ and $min(\Pi)$ are rarely exactly perpendicular to each other, the direction of $min(\Pi)$ is often close to the direction perpendicular to $max(\Pi)$ as shown in Chapter 3.

Not all model parameters are independent of each other. In particular, there is a strong positive correlation between η and ζ. By definition of Eq. (5.8), ζ and η are respectively $max(\Pi)$ and $\Pi(\hat{e}_{max(\Pi)\perp})$ in the coordinate system defined by ($\hat{e}_{max(\Pi)}$,$\hat{e}_{max(\Pi)\perp}$), and η is often close to $min(\Pi)$ as $\hat{e}_{max(\Pi)\perp}$ is close to the direction of $min(\Pi)$ for most stations. In general, the maximum and minimum directional predictability tend to increase and de-

crease together (not shown). There is no evident relationship between any of the correlation parameters (i.e. ρ and ρ_s) and the variance parameters (i.e. η, ζ, γ).

Figs. 5.9-5.12 show the relationships between model parameters and the predictability metrics. The relationships are generally similar for both observations and comprehensive models, and those between ζ, η and the predictability metrics are stronger than other model parameters in all three domains considered. I only show the results in NAM (Fig. 5.9) and EAS (Fig. 5.10) as the relationships between ζ, η and predictability metrics in EBM (not shown) are similar to those in NAM and EAS.

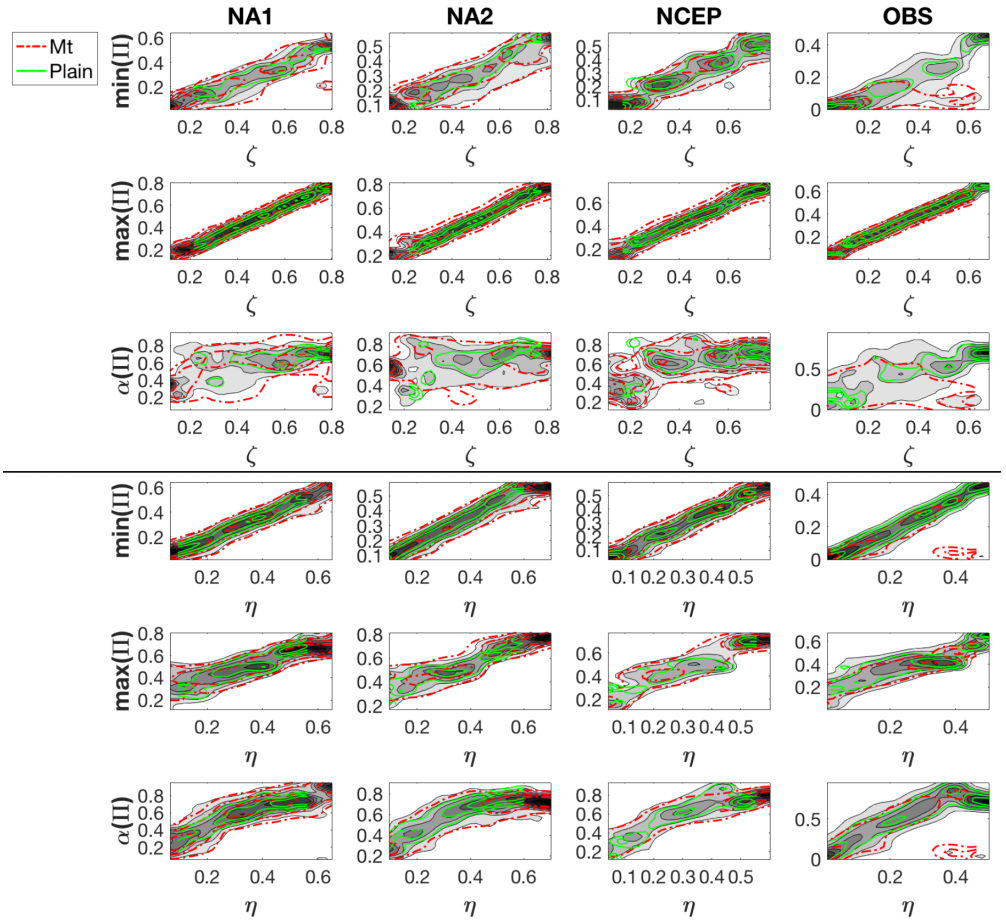

Figure 5.9: Estimated probability density functions of predictability metrics ($min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$) conditioned on ζ and η in the coordinate system defined by ($\hat{e}_{max(\Pi)}$,$\hat{e}_{max(\Pi)\perp}$) for observations and all comprehensive models in the NAM domain. Filled contours are obtained by using data from all stations in the region; red and green contours are obtained using only stations in mountainous and plain regions respectively.

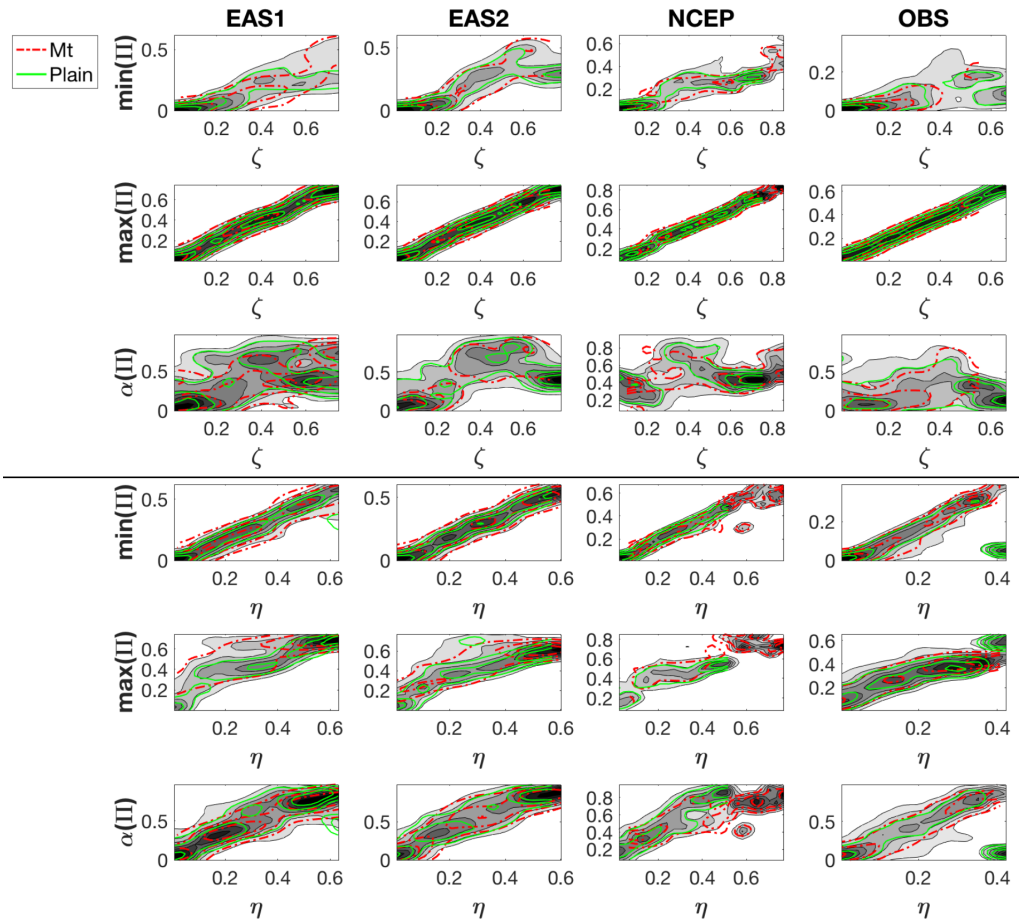

Figure 5.10: As in Fig. 5.9, but in EAS

The perfect linear correlation between ζ and $max(\Pi)$ is true by construction, while the correlation between between η and $min(\Pi)$ is also strong (consistent with the correlation between η and ζ). Evident variant structures in the observed relationships between η and $min(\Pi)$ are found in both NAM and EAS (1st row and 4th column in the lower panel of Figs. 5.9 and 5.10), resulting from large differences between the direction of $min(\Pi)$ and $\hat{e}_{max(\Pi)\perp}$ at the variant stations. In NAM, the locations of variant stations are located in mountainous regions consistent with the direction of $min(\Pi)$ being influenced by the orientation of topographic features. However, the variant stations in EAS are mostly found along the coastline of Japan (not shown), which suggests an influence from small-scale processes related to land-water contrast. The reasons for the differences in these variant observed behaviors in these two domains, which are not seen in the comprehensive models, are not clear.

Moreover, $\alpha(\Pi)$ generally increases with η but there is no evident relationship be-

tween α and ζ, indicating that predictive anisotropy is strongly influenced by the SNR of the surface wind component close to the direction of $min(\Pi)$, but the influence of SNR from the direction of $max(\Pi)$ is relatively weak. This supports the idea that anisotropy is generally controlled by local features that are more likely to influence $min(\Pi)$ by limiting the relationship between the predictors in the free troposphere and surface predictands, and many of these local features are not seen by the RCMs as manifested by overestimation of $min(\Pi)$.

Figure 5.11 shows that in NAM, values of $min(\Pi)$ and $\alpha(\Pi)$ tend to be lower for $\gamma = \frac{std(\tilde{v})}{std(\tilde{u})}$ much above or below the value of 1 and for stronger correlations $\rho(\tilde{u},\tilde{v})$ between surface wind components projected onto $\hat{e}_{max(\Pi)}$ and $\hat{e}_{max(\Pi)\perp}$. That is, both $min(\Pi)$ and $\alpha(\Pi)$ tend to be larger for more isotropic and less correlated vector fluctuations. The relationship between γ and predictability metrics is most evident in observations and least evident in reanalysis; the relationship between ρ and predictability metrics is stronger in observation and reanalysis than RCMs. Similar patterns described above can also be found in EMB (Fig. 5.12).

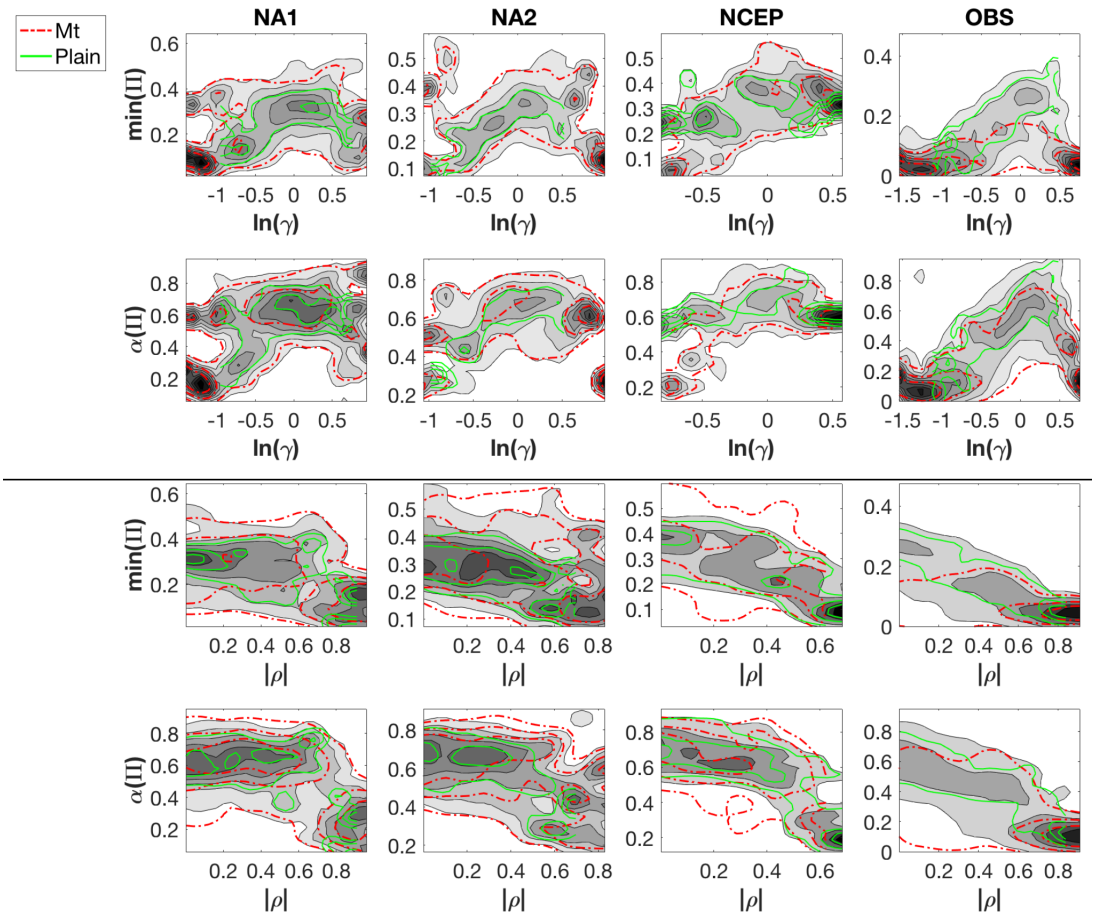

Figure 5.11: Estimated probability of predictability metrics given $ln(\gamma)$ and $|\rho(\tilde{u},\tilde{v})|$ in the coordinate system defined by ($\hat{e}_{max(\Pi)}$,$\hat{e}_{max(\Pi)^{\perp}}$) for observations and all comprehensive models used in NAM. Filled contours are obtained by using data from all stations in the region; red and green contours are obtained by stations only in mountainous and plain regions respectively.

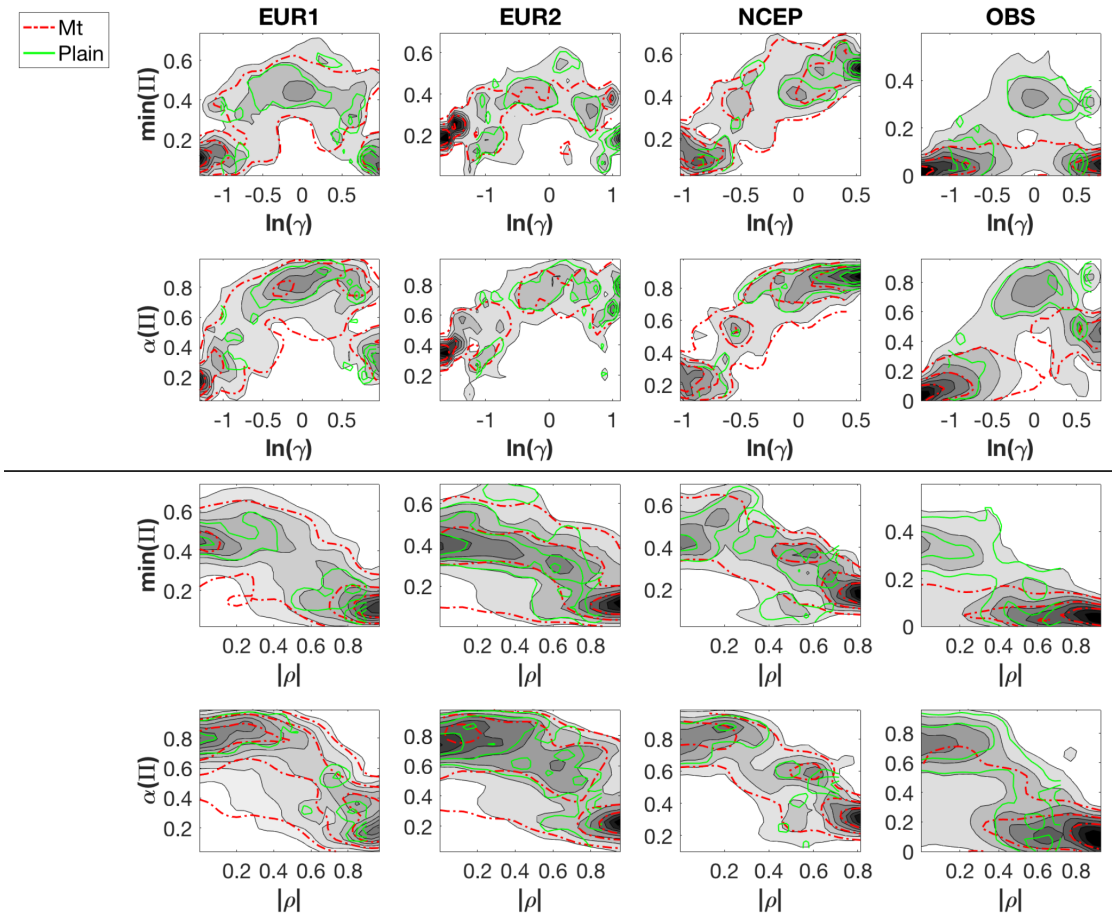

Figure 5.12: As in Fig. 5.11, but in EMB

The statistical relationships shown in Figs. 5.11 and 5.12 indicate that predictive anisotropy is related to the anisotropy of variability of surface wind components consistent with the results shown in Chapter 3, in which a positive correlation between $\alpha(\Pi)$ and the anisotropy of variability of surface wind components was found. Values of $max(\Pi)$ are generally not influenced by γ and ρ (not shown) in any of the three domains considered. This suggests that the same small-scale processes which contribute to lower predictability in some directions and stronger predictive anisotropy may induce anisotropy of variability of surface wind components and stronger correlation between surface wind components. For example, small scale processes related to surface inhomogeneities (such as linearly oriented topographic features, or coastlines) can result in correlations between vector wind components. It follows that lower predictability and stronger predictive anisotropy may have a common cause with anisotropy of variability of surface wind components, and strong correlation between surface wind components, and the common cause may be linked to

small-scale processes that are not resolved by RCMs.

For reasons that are unclear, EAS is generally variant in terms of having no systematic relationship between model parameters (γ and ρ) and metrics of predictability ($min(\Pi)$ and $\alpha(\Pi)$) in either observations or comprehensive models (not shown) as well as having underestimation of predictability metrics simulated by comprehensive models (Fig. 5.8).

5.6 Conclusion

I have compared characteristics of predictability of surface wind components (represented by the metrics: the directional maximum and minimum as well as the anisotropy) by linear statistical prediction using both station-based observational data and output from various comprehensive models (RCMs and NCEP-2 reanalysis) in three domains: North America (NAM; 557 stations), Europe-Mediterranean Basin (EMB; 595 stations) and East Asia (EAS; 715 stations). I divided data into four groups according to two categories of terrain: (1) stations adjacent to large water bodies (coastal) or (2) land stations and (3) in mountainous or (4) plain areas. In NAM and EMB, the characteristics of predictability from the use of comprehensive models in plain regions are generally close to those of observations while mountainous regions are dominated by overestimation of predictability metrics from the use of comprehensive models. In contrast, the difference between mountainous and plain terrains is not obvious in EAS where overestimation of predictability is commonly observed east of 90°E regardless of terrain, and the area west of 90°E is dominated by under-prediction from the use of RCMs (a pattern that is not observed in the reanalysis). There is no systematic pattern in characteristics of predictability associated with inland and coastal land stations.

In all three domains considered, the smallest SNR of observed surface wind components tends to be inflated when using comprehensive models thereby increasing overall predictability and artificially weakening predictive anisotropy. Regions of large SNR of observed surface wind components are generally in good agreement with comprehensive models with the exception of EAS where the largest values of SNR of observed surface wind components tend to correspond to underestimation of predictability and artificially strengthened predictive anisotropy. The general pattern of overestimation of predictability metrics suggests that the inability of the comprehensive models to capture small-scale physical processes (which contribute to predictive noise in observations) causes artificially inflated SNR of surface wind components biasing $min(\Pi)$ more than $max(\Pi)$. It follows that surface wind components from the use of comprehensive models show the same local/large-

scale decomposition as in observation but exist in a different region of "parameter space". This interpretation is reinforced through consideration of the relationships between metrics of predictability and parameters of an idealized model that decomposes surface wind variations into uncorrelated large-scale and local contributions. This model also suggests that lower predictability and stronger predictive anisotropy may have a common cause (related to local-scale processes) with anisotropy of variance of surface wind components and strong correlations of surface wind components. The reason for the underestimation of relatively large values of predictability in EAS by the RCMs (but not the observationally-constrained reanalysis product) is unclear, and is an interesting direction for future study.

An important overall conclusion following the results in this chapter is that variability in $min(\Pi)$ is most strongly influenced by small-scale processes and is therefore the major contributor to predictive anisotropy. Comprehensive models on the scale of RCMs and the reanalysis do not capture these small-scale processes. One area of future study is to more precisely identify the scale of the missing small-scale physical processes in RCMs, which may enhance the utility of the RCMs as tools for dynamic downscaling. In particular, predictability from mesoscale models on a small domain can be used to investigate just how small the domain needs to be in order for the models to capture the physical processes related to the local features. Special attention is evidently needed to study the physical processes related to local wind systems in EAS and their representations in RCMs.

Chapter 6

Conclusions

As outlined in Chapter 1, the objective of this study is to provide a global characterization of statistical predictability of surface wind components and build a general framework to explain characteristics of statistical predictability with an emphasis on predictive anisotropy (variation of predictability of surface wind components with directions of projection). This thesis also considers the factors that can influence characteristics of statistical predictability of surface wind components. In particular, it addresses the question of whether the factors are physical or are related to the types of regression methods used as transfer functions in the statistical prediction.

6.1 Summary of results

By assessing the linear statistical predictability of surface wind components at 2109 land stations across the globe, I found that predictive anisotropy is a common characteristic across the globe. Strong predictive anisotropy and low predictability are generally associated with terrain characterized by surface heterogeneity, such as mountainous regions and to a lesser degree, land/water contrast; however, marked predictive anisotropy and low predictability are also found in regions away from evident surface heterogeneity. The comparison of monthly and daily averaged predictions shows that there are more stations with higher overall monthly predictability than daily predictability in both summer and winter, and the seasonal difference of predictability is generally small.

In order to test whether the limitation of linear predictability along some directions resulting in low minimum directional predictability and strong predictive anisotropy is due to inadequate representation of the predictor-predictand relationship by the linear regression

model, linear statistical predictability is compared with that by three nonlinear regression methods (i.e. NN, SVM and RF) at the same 2109 stations. The results show that although there is a modest trend of systematic improvement in nonlinear predictability for surface wind components with fluctuations of relatively small magnitude or large kurtosis, more complex nonlinear regression methods do not in general substantially outperform linear regression in the cross-validated statistical prediction of surface wind components. Furthermore, predictive anisotropy cannot be eliminated or substantially weakened by using nonlinear regression methods. Improvement in predictability and weakening of predictive anisotropy by nonlinear regression generally occur in locations characterized by surface heterogeneity. Overall, the type of regression methods (either linear or nonlinear) is not a major factor influencing characteristics of predictability of surface wind components such as predictive anisotropy.

Potential factors that can influence linear statistical predictability have been identified as topographic complexity, statistical properties of surface wind components (i.e., standard deviation and kurtosis of wind components), and directions of mean surface wind vectors. I found that poor predictability of surface wind components is generally associated with components characterized by relatively weak and non-Gaussian variability and more likely to be in areas characterized by complex topography, and that predictive anisotropy is positively correlated with anisotropy of variability and anisotropy of kurtosis. Moreover, the directions of the best predicted surface wind components tend to correspond to the directions of most variable surface wind components, the directions of surface wind components characterized by Gaussian variability (i.e. $kurt \approx 3$, since kurtosis of surface wind components less than 3 is not common), and the directions of mean wind vectors.

Surface wind components simulated by RCMs and reanalysis were also studied using linear statistical prediction in three major domains in the Northern Hemisphere, namely North America (NAM), Europe-Mediterranean Basin (EMB) and East Asia (EAS), and the resulting predictability metrics (i.e. $min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$) were compared with those from observations. The comparison shows that in NAM and EMB, values of predictability metrics tend to be higher than the corresponding observed values in regions of low predictive skill (particularly mountainous regions), whereas predictability metrics from the simulations of comprehensive models are in good agreement with those from observation regions of relatively high predictive skill (particularly plain regions). In contrast, EAS is characterized by regions in which under-estimations of predictability metrics occur for relatively large metrics in observation. The reasons for the variant behaviour of EAS are not clear. Small values of observed $min(\Pi)$ are more likely to be over-predicted than small val-

ues of $max(\Pi)$. Over-estimates of predictability suggest that small-scale physical processes that contribute to the predictive noise of surface wind components (i.e., resulting in small values of signal-to-noise ratio) in reality are not resolved by the physical parameterization of comprehensive models, and the SNR by simulations of comprehensive models is artificially inflated. Therefore, small-scale processes that cannot be resolved by comprehensive models (e.g., RCMs) can contribute to low predictability and strong predictive anisotropy, and $min(\Pi)$ is more influenced by small-scale processes than $max(\Pi)$.

Two idealized mathematical models based on the assumption that surface wind variability can be partitioned into distinct signal and noise contributions are developed to aid the explanation of the characteristics of predictability of surface wind components. The first mathematical model also makes two additional assumptions, that the variability of predictive noise is isotropic, and there is no correlation between surface wind components. The first one is used to provide a conceptual organizing of results related to the relationship between the predictability metrics ($min(\Pi)$, $max(\Pi)$ and $\alpha(\Pi)$) and the statistical properties of surface wind components (variability and kurtosis). Simulation of the relationships by this model shows that predictive anisotropy can be related to the SNR corresponding to the relative magnitudes of large-scale and local influences on the wind. In particular, the model shows that if the influence of the unpredictable component becomes larger (increased noise), the prediction becomes more anisotropic, and those directions that are more influenced by unpredictable noise are more likely to be found with non-Gaussian local variability, resulting in the association of larger kurtosis with smaller predictability.

The second idealized mathematical model is an extension of the first one without the two additional assumptions in the first one. The main utility of the second model is to characterize the predictability of arbitrary wind components in terms of the statistical properties of components along two specified orthogonal directions. The relationships between predictability metrics and parameters derived from this model for both observations and comprehensive models indicate that the strongest influence on characteristics of predictability comes from the SNR of surface wind components; predictive anisotropy is more influenced by the variability of $min(\Pi)$ than $max(\Pi)$. Moreover, in NAM and EMB but not in EAS, lower $min(\Pi)$ and stronger $\alpha(\Pi)$ are associated with more anisotropic variability of surface wind components and stronger correlations between surface wind components, which suggests that lower predictability and stronger predictive anisotropy have a common cause with the anisotropy of variability of surface wind components and strong correlation between surface wind components. Since the results of this study indicate that $min(\Pi)$ is influenced by small-scale processes, I again interpret this cause to be related to small-scale

processes that cannot be resolved by RCMs.

6.2 Achievements and future studies

This study enhances our understanding of statistical prediction of surface winds by characterizing the statistical predictability and exploring the relationships between predictability metrics and potentially influential factors at an unprecedentedly large number of stations. The main achievements of the study are (1) the exclusion of differences in the form of regression methods of transfer functions as a major factor in influencing statistical predictability, (2) characterization of relationships between predictability metrics and influential factors on a global scale, and the use of such relationships to argue that physically-caused changes in predictability vary with SNR. The comparison of simulated predictability metrics from comprehensive models and those from observations provides evidence that those physical phenomena are likely to be small-scale processes that cannot be captured well by RCMs and reanalysis. Beyond this I cannot provide any detailed physical explanation for the controls on SNR responsible for limiting of statistical predictability.

Physically based studies are needed to identify specific physical processes responsible for the observed characteristics of predictability (predictive anisotropy), through a more detailed analysis of the connection between near-surface flow and large/local-scale variability in atmospheric circulations possibly with the aid of mesoscale models. Such an analysis would need to consider a range of different regions, as the characteristics of predictability based on statistical analysis in this study show substantial regional variations.

Appendix A

Nonlinear regression methods

There are three nonlinear regression methods used in this study: neural network (NN), support vector machines (SVM) and random forests (RF). These three methods are supervised learning methods in which a function is inferred from a set of training data consisting of pairs of predictors and predictands in a process referred to as training. After training, the estimated function can be used for new predictions. Supervised learning methods differ from each other by the algorithms they use to infer the regression function from training data. A brief introduction to the NN, SVM and RF algorithms will now be presented.

A.0.1 Neural network

The method of NN in nonlinear regression is inspired from the structure of neurons in biology. There are many types of neural network. In this study I use feed-forward neural networks, which is the most widely-applied NN. In feed-forward NN, the predictand is related to the predictor by a sequence of linked computational elements known as hidden layers (Hsieh, 2009). Fig. A.1 demonstrates the structure of a feed-forward NN used to model a regression with P predictors and K predictands by a single layer of M hidden neurons. This structure is similar to that of a two-stage regression model (Hastie et al., 2011).

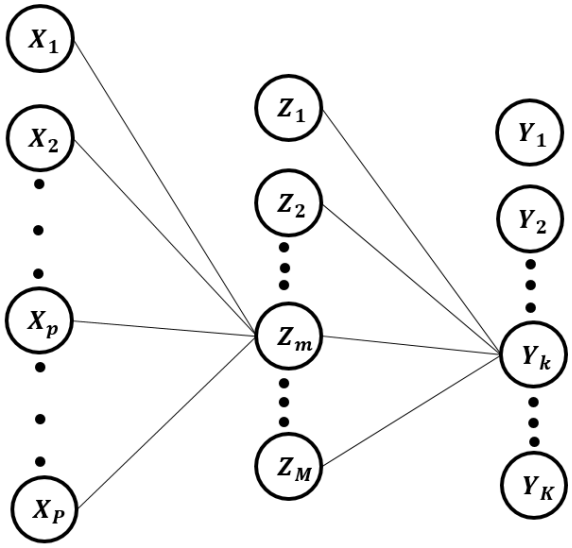

Figure A.1: Schematic of a single hidden layer, feed-forward neural network, where X_i and Y_i are predictors and predictands, and Z_i are the hidden neurons

In the first stage of the NN regression, the values of the hidden neurons $Z = [Z_1, ..., Z_m, ..., Z_M]$ are computed as linear combinations of the inputs $X = [X_1, ...X_p, ..., X_P]$,

$$Z_m = s(\alpha_{0m} + \alpha_m^T X), \quad m = 1, ..., M, \tag{A.1}$$

where α_{0m} are the offset scalars and α_m are the weight vectors of predictors X. The function $s(v)$ is usually chosen to be the sigmoid function $s(v) = 1/(1+e^{-v})$ which asymptotically saturates for large positive and negative values of v: $s(v) \to 0$ as $v \to -\infty$ and $s(v) \to 1$ as $v \to +\infty$. In the second stage of the regression models, each predictand Y_k is computed as the linear combination of hidden neurons $Z = [Z_1, ..., Z_m, ..., Z_M]$,

$$\widehat{Y_k} = \beta_{0k} + \beta_k^T Z, \quad k = 1, ..., K, \tag{A.2}$$

where $\widehat{Y_k}$ is the modeled value of target Y_k, β_{0k} are the offset scalars, and β_k are the weight vectors. Together, the parameters α_{0m}, α_m and β_{0m}, β_k are also called the weights of the NN model.

Training of the NN is done using an optimization algorithm to seek weights of the model that minimize the objective function $J = \sum_{k=1}^{K} \sum_{i=1}^{N} (Y_{ik} - \widehat{Y}_{ik})^2$ where N is the number of observations. Parameters of the initial state can be chosen randomly. The minimization of

the objective function J is done iteratively until some convergence criterion is met. The primary issue in training neural network is overfitting. Measures to prevent overfitting should be considered in both model architecture (i.e. number of hidden neurons) and model training.

Hidden neurons The complexity of a neural network is increased by increasing the number of hidden neurons. A neural network with one hidden layer of a finite (but sufficiently large) number of neurons can approximate any continuous function to arbitrary accuracy (Csáji, 2001). The modeling challenge is choosing the right number of hidden neurons. Models with too few hidden neurons might not have the flexibility to capture the nonlinear signal of the data. Models with too many hidden neurons might be so flexible that they will fit the noise of the data. The appropriate number of hidden neurons can be chosen from empirical testing (as described in Section 4.2).

Training methods Early stopping is a common method of preventing overfitting, in which the training process stops well before J reaches the global minimum. In this method, the training data are divided into two subsets. All data in the first subset undergo the training process as described above to update the weights of the model. Each iteration in the training is referred to as an epoch. The training process is repeated for many epochs. The second dataset is used to evaluate the objective function at each epoch. As the number of training epochs increase, the value of J over the validation data generally decreases at first before increasing. Training beyond the point at which J starts to increase will only contribute to model overfitting and is therefore stopped. The model parameters are those obtained at the minimum of J over the validation data. Note that the separation into training and validation sets is done as part of the parameter estimation process, and is distinct from the data subsetting associated with cross-validation.

A.0.2 Support Vector Machine

Support vector machines are characterized by the use of kernel functions which represent the mapping of observations in the input space into a high-dimensional feature space where linear regression can be used. Therefore, a nonlinear model in the input space can be learned from a subset of observations (support vectors) by linear regression in the high-dimensional feature space. The following mathematical formulation is based on the documentation of the Statistics and Machine Learning Toolbox in MATLAB (MathWorks,

2017c).

Suppose x_n represents one case of training data in the input space with observed response value y_n, and $g(x_n)$ represents a mapping function which maps x_n in the input space to the feature space. The function, $f(x_n)$, which is used to model y_n, can be constructed as

$$f(x_n) = g(x_n)^T \beta + b. \tag{A.3}$$

The goal of SVM regression is to find a function $f(x)$ that deviates from y by a value no larger than ε for each observation of training data, and at the same time is as flat as possible. Flatness requires that the weight vector of $f(x)$ should be as small as possible. The problem is formulated as minimizing the objective function

$$J(\beta) = \frac{1}{2}\beta^T \beta, \tag{A.4}$$

subject to the constraint that all residuals less than ε: that is for all n, $|y_n - (g(x_n)'\beta + b)| \leq \varepsilon$. The set of x values satisfying $|y - f(x)| \leq \varepsilon$ is known as the ε-tube. However, for a given ε, not all observations may satisfy the constraint of falling in the ε-tube; therefore, slack variables ζ_n and ζ_n^* are used to allow the regression models to tolerate errors up to the value of $\varepsilon + \zeta_n$ or $\varepsilon + \zeta_n^*$, where ζ_n and ζ_n^* represent the upper and lower limit of the extension of the ε-tube respectively (Fig. A.2).

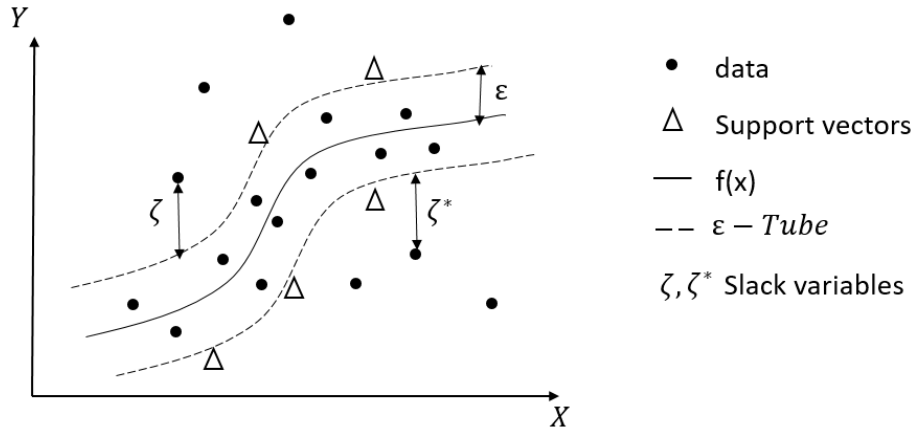

Figure A.2: Schematic of regression by support vector machine

By including slack variables, the minimization of the objective function becomes

$$J(\beta) = \frac{1}{2}\beta^T\beta + C\sum_{i=1}^{N}(\zeta + \zeta^*), \tag{A.5}$$

subject to: for all n
$y_n - (x_n'\beta + b) \leq \varepsilon + \zeta_n$,
$(x_n'\beta + b) - y_n \leq \varepsilon + \zeta_n^*$,
$\zeta_n \geq 0$
$\zeta_n^* \geq 0$.
The constant $C > 0$ determines the largest deviations from ε that can be tolerated. The formulation of $J(\beta)$ in Eq. (A.5) is also known as the primal formula (Vapnik, 2013).

Estimating the weights of $f(x)$ of SVM regression in Eq. (A.5) corresponds to minimizing the ε-insensitive loss function defined as

$$L_\varepsilon = \begin{cases} 0 & \text{if } |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & \text{if } |y - f(x)| > \varepsilon \end{cases} \tag{A.6}$$

for which the errors associated with observations within the ε-tube are ignored (Vapnik, 2013). This optimization problem of minimizing Eq. (A.6) is computationally simpler to solve in its Lagrange dual formulation (MathWorks, 2017c). Constructing a Lagrangian function for Eq. (A.6) requires nonnegative multipliers α_n and α_n^* for each observation x_n. The dual formulation of minimizing Eq .(A.6) involves minimizing

$$L(\alpha) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) < g(x_i), g(x_j) > + \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) - \varepsilon\sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*), \tag{A.7}$$

where $< g(x_i), g(x_j) >$ is the inner product of the predictors after mapping, and Eq. (A.7) is subject to
$\sum_{n=1}^{N}(\alpha_n - \alpha_n^*) = 0$,
and that for all n:
$0 \leq \alpha_n \leq C$,
$0 \leq \alpha_n^* \leq C$,
$\alpha_n(\varepsilon + \zeta_n - y_n + f(x_n)) = 0$,
$\alpha_n^*(\varepsilon + \zeta_n^* - y_n + f(x_n)) = 0$,
$\zeta_n(C - \alpha_n) = 0$,

$\zeta_n^*(C-\alpha_n^*)=0$.
These conditions indicate that Lagrange multipliers $\alpha_n=0$ and $\alpha_n^*=0$ when observations are inside the ε-tube. The dual formulation Eq. (A.7) is solved by using quadratic programming techniques, the details of which are beyond the scope of this paper but can be found in Platt (1998). The solution has the form:

$$f(x)=\sum_{n=1}^{N}(\alpha_n-\alpha_n^*)<g(x_n),g(x)>+b. \tag{A.8}$$

In other words, the solution $f(x)$ only depends on those x_n which satisfy $(\alpha_n-\alpha_n^*)\neq 0$, and this subset of x_n from training data is denoted support vectors which fall within a distance ζ or ζ^* from the boundary of ε-tube as shown in Fig. A.2. Since by construction of SVM regression models, most cases of training data are inside the ε tube, the number of support vectors is small compared to the number of observations in training data. The transformation function $g(x)$ that maps each x_n of training data in the input space to the feature space is unknown, but $<g(x_n),g(x)>$ in Eq. (A.9) can be approximated by kernel functions $K(x_n,x)=<g(x_n),g(x)>$. There are many admissible kernel functions, and some common kernel types are listed in the following:

- linear kernel: $K(x_n,x)=<x_n,x>$
- polynomial kernel: $K(x_n,x)=(1+x'_nx)^p$, where $p=2,3,4...$
- radial basis function kernel: $K(x_n,x)=\exp(-\gamma\|x_n-x\|^2),\gamma>0$
- sigmoid kernel: $K(x_n,x)=\tanh(\gamma<x_n,x>+\tau),\gamma>0$.

The class of functions that can be approximated are determined by the chosen kernel form.

Generally, there are three factors that can influence the accuracy of SVM regression: the values of C, ε and the chosen kernel form. The parameter C determines the trade-off between model complexity (i.e. the flatness of $f(X)$) and the degree to which deviations larger than ε are tolerated in the optimization of the loss function. Larger C indicates that more data are used in the process of training but the resulting model can be complex and overfit the data; on the other hand, smaller C might make the regression model prone to underfitting because the training process might not have enough training data to characterize the underlying structure. By controlling the width of the ε-tube in the training data,

the parameter ε influences the number of support vectors used in the training process. Finally, depending on the properties of training data, some kernel forms may work better than others.

A.0.3 Random Forests

Random forests, first proposed by Breiman (2001), belong to the category of ensemble learning methods that generate many regression models and aggregate their results (Hastie et al., 2011). As the name, random forest, suggests, the individual model is a tree-based regression model. The basic idea of a tree-based regression is partitioning the input space into a set of subspaces, and then fitting a simple model (usually a constant) in each subspace. The partition is generally done by binary recursive partition in which the input space is first split into two regions, and each sub-region is fit with a simple model. This splitting process is repeated in the resulting sub-regions until some stopping rule is applied as illustrated in Fig. A.3.

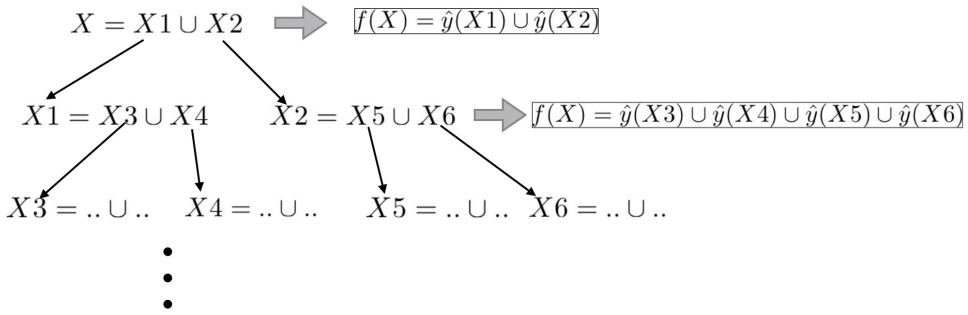

Figure A.3: Schematic of a tree regression by binary recursive partitioning. X represents the training data in the input space. $\hat{y}$ refers to a simple model used to fit the training data in each sub-space after the split. f stands for the resultant tree model at each level of partitioning.

The following formulation is based on Hastie et al. (2011). Suppose there are p input variables and one response for each of N observations in a dataset: that is, (x_i, y_i) for $i = 1, 2, ...N$ with $x_i = (X_{i1}, X_{i2}, ...X_{ip})$. Starting from the entire input space, the two planes after each split can be expressed as $R_1(j,s) = X|X_j \leq s$ and $R_2(j,s) = X|X_j > s$, where X_j and s represent the splitting variable and splitting point respectively. The best fit is achieved

by seeking the variable X_j and split point s which solve

$$\min[\min \sum_{x \in R_1(j,s)} (y_i - c_1)^2 + \min \sum_{x \in R_2(j,s)} (y_i - c_2)^2], \tag{A.9}$$

where c_1 and c_2 are the constants used to model values in R_1 and R_2 respectively. The split variable X_j and split point s which lead to best fit can be determined by scanning through input variables x_i, and depending on the objective of a regression problem, it is not necessary to use all p input variables in determining the best fit. The response y_i can be modeled by the tree regression

$$T(x) = \sum_{m=1}^{M} c_m I(x \in R_m), \tag{A.10}$$

where R_1, R_2, ..., R_M are regions resulting from partition in the tree regression, and c_1, c_2, ..., c_m are the constants used to model responses in the corresponding region.

Tree regression is generally prone to overfitting as the training processes use all training data including noise. The essential idea is to average many noisy but approximately unbiased models, thereby reducing the variance (Hastie et al., 2011). Averaging many regression trees constructed randomly from bootstrapped samples, the resulting model can be expressed as

$$f(X) = \frac{1}{B} \sum_{b=1}^{B} T(X_b), \tag{A.11}$$

where B is the total number of trees, and b characterizes the bth tree-based regression $T(X)$ from a bootstrap sampling. In general, there is no specific rule to determine the fraction of data used for bootstrap model construction. The Python function 'RandomForestRegressor' used in this study adopts the strategy of building each tree from drawing a sample of equal size of the input training data with replacement. The estimate of error is obtained by prediction over data not in the bootstrap sample (out-of-bag or OOB data).

One of the biggest merits of random forest analysis is its simplicity. There are only two parameters for RF, and the solution is not very sensitive to their values (Liaw and Wiener, 2002). The first parameter is the number of input variables needed to consider when seeking the best split during tree regression. The second parameter is the number B of individual tree regressions used in RF.

Bibliography

Amante, C. and B. W. Eakins, 2009: 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24 19 pp.

Amari, S.-i., N. Murata, K.-R. Müller, M. Finke, and H. H. Yang, 1996: Statistical theory of overtraining-is cross-validation asymptotically effective? In *Advances in neural information processing systems*, 176–182.

Breiman, L., 2001: Random forests. *Machine learning*, **45**(1), 5–32.

Cheng, C. S., G. Li, Q. Li, H. Auld, and C. Fu, 2012: Possible impacts of climate change on wind gusts under downscaled future climate conditions over Ontario, Canada. *Journal of Climate*, **25**(9), 3390–3408.

Church, J. A., P. U. Clark, A. Cazenave, J. M. Gregory, S. Jevrejeva, A. Levermann, M. A. Merrifield, G. A. Milne, R. S. Nerem, P. D. Nunn, et al., 2013: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *Sea level change* 1137–1216.

Csáji, B. C., 2001: Approximation with artificial neural networks. *Faculty of Sciences, Eötvös Loránd University, Hungary*, **24**, 48.

Culver, A. M. and A. H. Monahan, 2013: The statistical predictability of surface winds over western and central Canada. *Journal of Climate*, **26**(21), 8305–8322.

Curry, C. L., D. van der Kamp, and A. H. Monahan, 2012: Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. I. Predicting wind speed. *Climate dynamics*, **38**(7-8), 1281–1299.

Davies, T., M. J. Cullen, A. J. Malcolm, M. Mawson, A. Staniforth, A. White, and N. Wood, 2005: A new dynamical core for the Met office's global and regional modelling of the

atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **131**(608), 1759–1782.

Davy, R. J., M. J. Woods, C. J. Russell, and P. A. Coppin, 2010: Statistical downscaling of wind variability from meteorological fields. *Boundary-layer meteorology*, **135**(1), 161–175.

Giorgi, F., C. Jones, G. R. Asrar, et al., 2009: Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin*, **58**(3), 175.

Green, S. B., 1991: How many subjects does it take to do a regression analysis. *Multivariate behavioral research*, **26**(3), 499–510.

Hastie, T. J., R. J. Tibshirani, and J. H. Friedman, 2011: *The elements of statistical learning: data mining, inference, and prediction*. Springer.

He, Y., A. Monahan, C. Jones, A. Dai, S. Biner, D. Caya, and K. Winger, 2010a: Land surface wind speed probability distributions in North America: observations, theory, and regional climate model simulations. *Journal of Geophysical Research*, **115**(D04), 103.

He, Y., A. H. Monahan, C. G. Jones, A. Dai, S. Biner, D. Caya, and K. Winger, 2010b: Probability distributions of land surface wind speeds over North America. *Journal of Geophysical Research: Atmospheres*, **115**(D4).

Holtslag, A., G. Svensson, P. Baas, S. Basu, B. Beare, A. Beljaars, F. Bosveld, J. Cuxart, J. Lindvall, G. Steeneveld, et al., 2013: Stable atmospheric boundary layers and diurnal cycles: challenges for weather and climate models. *Bulletin of the American Meteorological Society*, **94**(11), 1691–1706.

Hsieh, W. W., 2009: *Machine learning methods in the environmental sciences: Neural networks and kernels*. Cambridge university press.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, Y. Shi-Keng, et al., 2002: Ncep-doe amip-ii reanalysis (r-2). *Bulletin of the American Meteorological Society*, **83**(11), 1631.

Liaw, A. and M. Wiener, 2002: Classification and regression by randomforest. *R news*, **2**(3), 18–22.

Lu, H., 2008: Modelling terrain complexity. In *Advances in Digital Terrain Analysis*. Q. Zhou, B. Lees, and G. Tang, Eds., Springer, 159–176.

Mao, Y. and A. Monahan, 2017: Predictive anisotropy of surface winds by linear statistical prediction. *Journal of Climate*, **30**(16), 6183–6201.

Maraun, D., F. Wetterhall, A. Ireson, R. Chandler, E. Kendon, M. Widmann, S. Brienen, H. Rust, T. Sauter, M. Themeßl, et al., 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, **48**(3).

MathWorks, 2016: Mapping Toolbox. `https://www.mathworks.com/help/map/`.

MathWorks, 2017a: fitrsvm. `https://www.mathworks.com/help/stats/fitrsvm.html`.

MathWorks, 2017b: Getting Started with Neural Network Toolbox. `https://www.mathworks.com/help/nnet/getting-started-with-neural-network-toolbox.html`.

MathWorks, 2017c: Understanding Support Vector Machine Regression. `https://www.mathworks.com/help/stats/understanding-support-vector-machine-regression.html`.

Mearns, L., S. McGinnis, R. Arritt, S. Biner, P. Duffy, W. Gutowski, I. Held, R. Jones, R. Leung, A. Nunes, et al., 2007,updated 2014: The North American regional climate change assessment program dataset. National Center for Atmospheric Research Earth System Grid data portal, Boulder, CO.

Mearns, L. O., W. Gutowski, R. Jones, R. Leung, S. McGinnis, A. Nunes, and Y. Qian, 2009: A regional climate change assessment program for North America. *Eos, Transactions American Geophysical Union*, **90**(36), 311–311.

Mohandes, M., T. Halawani, S. Rehman, and A. A. Hussain, 2004: Support vector machines for wind speed prediction. *Renewable Energy*, **29**(6), 939–947.

Monahan, A. H., 2012: Can we see the wind? Statistical downscaling of historical sea surface winds in the subarctic northeast pacific. *Journal of Climate*, **25**(5), 1511–1528.

Oke, T. R., 2002: *Boundary layer climates*. Routledge.

Parker, W. S., 2016: Reanalyses and Observations: Whats the Difference? *Bulletin of the American Meteorological Society*, **97**(9), 1565–1572.

Platt, J., 1998: Sequential minimal optimization: A fast algorithm for training support vector machines. Technical report, Microsoft Research.

Pryor, S., J. T. Schoof, and R. Barthelmie, 2005: Empirical downscaling of wind speed probability distributions. *Journal of Geophysical Research: Atmospheres*, **110**(D19).

Python, 2016: RandomForestRegressor-scikit-learn 0.18.1 documentation. `http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html`.

Ripley, B. and W. Venables, 2016: Package 'nnet'. `https://cran.r-project.org/web/packages/nnet/nnet.pdf`.

Rummukainen, M., 2010: State-of-the-art with Regional Climate Models. *Wiley Interdisciplinary Reviews: Climate Change*, **1**(1), 82–96.

Sailor, D., T. Hu, X. Li, and J. Rosen, 2000: A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. *Renewable Energy*, **19**(3), 359–378.

Salameh, T., P. Drobinski, M. Vrac, and P. Naveau, 2009: Statistical downscaling of near-surface wind over complex terrain in southern France. *Meteorology and Atmospheric Physics*, **103**(1-4), 253–265.

Schoof, J., 2013: Statistical downscaling in climatology. *Geography Compass*, **249-265**(4).

Scinocca, J., V. Kharin, Y. Jiao, M. Qian, M. Lazare, L. Solheim, G. Flato, S. Biner, M. Desgagne, and B. Dugas, 2016: Coordinated global and regional climate modeling. *Journal of Climate*, **29**(1), 17–35.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF version 2. Technical report, National Center For Atmospheric Research Boulder Co Mesoscale and Microscale Meteorology Div.

Strandberg, G., L. Bärring, U. Hansson, C. Jansson, C. Jones, E. Kjellström, M. Kupiainen, G. Nikulin, P. Samuelsson, and A. Ullerstig, 2015: *CORDEX scenarios for Europe from the Rossby Centre regional climate model RCA4*. SMHI.

Stull, R. B., 2000: *Meteorology for scientists and engineers: a technical companion book with Ahrens' Meteorology Today*. Brooks/Cole.

Sun, C. and A. Monahan, 2013: Statistical downscaling prediction of sea surface winds over the global ocean. *Journal of Climate*, **7938–7956**.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, **106**(D7), 7183–7192.

van der Kamp, D., C. L. Curry, and A. H. Monahan, 2012: Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. ii. Predicting wind components. *Climate dynamics*, **38**(7-8), 1301–1311.

Vapnik, V., 2013: *The nature of statistical learning theory*. Springer Science & Business Media.

Wolfram, 2016: WeatherData source information. `http://reference.wolfram.com/language/note/WeatherDataSourceInformation.html`, Accessed: 2016-01-01.

Yuval and W. Hsieh, 2002: The impact of time-averaging on the detectability of nonlinear empirical relations. *Quarterly Journal of the Royal Meteoro- logical Society*, **583**(1609-1622).

Zhao, P. and B. Yu, 2006: On model selection consistency of Lasso. *Journal of Machine Learning Research*, **7**(Nov), 2541–2563.