Projecting Climate Change Impact on the Regional Scale:
Future Changes of Wind and Wind Energy in California

Ph.D. Dissertation

University of California, Davis

Meina Wang

April 2019

Projecting Climate Change Impact on the Regional Scale:
Future Changes of Wind and Wind Energy in California

By

MEINA WANG

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ATMOSPHERIC SCIENCE

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Professor Paul Ullrich, Chair

_____

Professor Ian Faloona

_____

Professor Shu-hua Chen

Committee in Charge

2019

Dedication

This thesis is dedicated to my parents.

Acknowledgement

My deepest gratitude goes to my advisor, Professor Paul Ullrich. He always encourages us to go on paths that spark true passion. It is his encouragement that leads me to work on the very beginning of this dissertation, and he has been giving me unfailing support ever since, both academically and mentally. What he taught me throughout the years is so much more beyond atmospheric science, and I feel extremely fortunate to have him as my mentor.

I would also like to thank Dr. Dev Millstein at LBNL. It was a wonderful experience working with him on the wind energy project.

I would also like to thank Professor Travis O'Brien. I truly appreciate the opportunity he offered me to work on the coastal fog project. It was always a pleasure meeting with him.

I would like to thank Xue Meng Chen, Alan Rhoades, Dustin Grogan, who are senior members of our grad group, and our department IT members Bill Broadley and Terri Knight. They have given me tremendous help throughout the years.

I would like to extend my gratitude towards my dissertation committee, Professors Ian Faloona, Shu-hua Chen, Yufang Jin, and Sara Pryor, who provided invaluable comments and suggestions for this dissertation.

Last but not least, I would like to thank my parents, for supporting and encouraging me to take on this adventure. Special thanks to Luman Qu.

Without their support and encouragement, none of this work would have happened.

Abstract

Unprecedented climate change is expected in this century. Although its impact on the regional scale is less well-understood. This thesis aims to bridge the gap between the relatively well-studied future change of synoptic-scale weather systems, and the localized change under future climate.

First, we focused on one of the most common coastal phenomena within California's San Francisco Bay Delta and Central Valley - marine air penetration (MAP) events, which are a broad category of meteorological features that include the well-known Delta Breeze (DB). Summertime MAP episodes, identified by an inland cooling pattern and onshore wind, are selected in both observational and reanalysis dataset using a newly developed objective criteria. Correlations between MAP occurrence and synoptic-scale meteorology are then examined to understand the drivers of MAP. Historical data are first used to determine if any significant trends have been observed in the frequency and character of MAP events. Future trends in MAP are then predicted based on an investigation of how synoptic-scale meteorology will be modified in the 21st century under CMIP5 climatological projections. Based on our analysis of historical and future MAP frequency (under RCP8.5), a weak positive trend ($\sim$0.2 days/century) is identified in the occurrence of sufficiently strong summertime MAP days, although further study is needed.

Second, we analyzed the California wind resource variability under the future climate. Shifting wind patterns are an expected consequence of global climate change, with direct implications for wind energy production. However, wind is notoriously

difficult to predict, and significant uncertainty remains in our understanding of climate change impacts on existing wind generation capacity. In this study, historical and future wind climatology and associated capacity factors at five wind turbine sites in California are examined. Historical (1980-2000) and mid-century (2030-2050) simulations were produced using the Variable-Resolution Community Earth System Model (VR-CESM) to understand how these wind generation sites are expected to be impacted by climate change. A high-resolution statistically downscaled WRF product provided by DNV GL, reanalysis datasets MERRA-2, CFSR, NARR, and observational data were used for model validation and comparison. These projections suggest that wind power generation capacity throughout the state is expected to increase during the summer, and decrease during fall and winter, based on significant changes at several wind farm sites. This study improves the characterization of uncertainty around the magnitude and variability in space and time of California's wind resources in the near future, and also enhances our understanding of the physical mechanisms related to the trends in wind resource variability.

Third, we expanded our analysis to the end-of-century, and developed a statistical model for wind patterns classification. Wind energy production is expected to be affected by shifts in wind patterns that will accompany climate change. However, many questions remain on the magnitude and character of this impact, especially on regional scales. In this study, clustering is used to group and analyze wind patterns in California using model simulations from the Variable-Resolution Community Earth System Model (VR-CESM). Specifically, simulations have been produced that cover both historical (1980-2000) and end-of-century (2080-2100) time periods. Once clustered, observed changes to wind patterns can be analyzed in terms of both the change in frequency of those clusters and changes to winds within clusters. A further examination of the synoptic-scale fields associated with each cluster then provides a better understanding of how changes to large-scale meteorological fields are important for driving changes in localized wind speeds.

All three studies greatly advance our understanding of the climate change impact on the regional scale. Future projections for wind and wind energy resources in California provide valuable and actionable insight for stakeholders, and such information is crucial in aiding future climate mitigation and adaption.

Contents

x

Chapter 1

Introduction

The global climate is changing, and climate signals have already been detected across a wide range of observations [1]. Such changes will have inevitable impact on human and natural systems [2]. Much previous research has been focused on analyzing climate change on the global scale. However, there still remains great uncertainty in future climate changes at the regional level. The latter will have direct impact on health and economics of local communities. Climate projections on the regional level can also provide actionable insight to local stakeholders for future climate mitigation and adaptation.

The main focus of this thesis is to analyze the regional climate change impact within California, and link the localized changes to synoptic-scale meteorological conditions. In particular, we choose to detect the regional climate change through its impact on wind and wind energy resources in California. This thesis also serves the purpose of better understanding the relationship between the future behavior of synoptic-scale weather systems and predicted behavior on a regional-scale.

The structure of this thesis is as follows: Chapter 1 provides an introduction and review of literature on the topics covered in this thesis. Chapter 2 focuses on identifying and analyzing the future change of marine air penetration events, which is the prominent wind pattern during summertime around the San Francisco Delta region.

The associated synoptic-scale meteorological background is also discussed. Chapter 3 reports our study on the future projection of wind energy resources across five major wind farm sites across California under mid-century. A state-of-the-art climate model, the variable-resolution CESM, is used to model near-surface wind speed under the future climate. Synoptic-scale drivers behind the regional-scale wind changes are then identified. Chapter 4 expands on both Chapter 2 and 3 by extending the time frame to end-of-century, and incorporating a full suite of wind patterns into the analysis. An unsupervised machine learning algorithm, agglomerative clustering, is applied to group different wind patterns from the unlabeled modeling data. The impact on wind energy resources is assessed by analyzing the future change of capacity factors at each wind farm site. Chapter 5 includes summaries of the main findings from previous chapters, and discussions of future work.

1.1 Marine Air Penetration in California's Central Valley

This study focuses on a particularly important coastal phenomenon around the San Francisco Bay Delta region - marine air penetration (MAP) events [3, 4, 5], also commonly referred to as the Delta Breeze (DB). However, MAP events are more general, and are composed of two phenomena that are quite different in scale - the daily sea breeze circulation and the Pacific coast monsoon. The sea breeze is a very well studied phenomenon with a vast amount of literature on this topic. It is produced by the differential heating between the land and the adjacent water mass. Whereas the most noticeable characteristic of a sea breeze event is the frontal-like feature at the leading edge of its circulation [6], the Pacific coast monsoon is described as a slow and steady transport of the marine air inland from the Pacific anticyclone [7]. MAP events occur whenever cool and moist marine air penetrates inland, whereas DB events must be associated with a sea breeze front. This study

focuses on analyzing the more general phenomenon MAP.

MAP mostly occurs from the late spring through early autumn months, and usually peaks in intensity during the summertime, when the temperature and pressure gradients between the land and the adjacent water mass are the greatest [7]. This phenomenon is more closely controlled by synoptic-scale meteorology. It can act as a meso-scale "amplifier," which can influence the synoptic-scale "signal" [8]. During the summertime, MAP events can be intensified by DB events, as the latter is superimposed on the monsoon flow [5].

MAP events occur in many coastal locations throughout the world, and are strongly correlated with local topography, making prediction of such events particularly difficult and localized. Literature on this topic has widely varying nomenclature for these events, including marine air penetration (MAP) [3, 4], marine air intrusion [9, 10, 11], marine air invasion [7], or onshore surge of marine air [8], but essentially all describe the onshore flow component that causes the cool and moist marine air to penetrate inland. Most of the literature on this topic may refer to MAP events and the sea breeze interchangeably. However, the dominant component for MAP events is its associated synoptic background, and has been discussed in Fosberg and Schroeder [5].

There are many local studies of MAP. While most studies use direct measurement from observations to identify MAP events [3, 8, 5], some other studies also use numerical model simulations [12, 10]. However, detailed studies focusing on the MAP events around the San Francisco Bay Delta region is still lacking.

MAP events can have impact on people's lives in many ways. It can be a relief to people by transporting cool moist marine air inland, which can alter local temperature and decrease the health risks due to prolonged heat exposure. Accurate forecasting of MAP events is of great economical value due to its large impact on the local electricity demand by altering of the temperature [13]. Also, MAP can influence local air quality via inland ventilation [14, 15], and this wind pattern is im-

portant for determining the placement of wind turbines for wind power generation. Literature analyzing the future change in the temperature and wind fields implying stronger sea breeze events under the climate change [16]. However, detailed analysis on the future change of MAP is still lacking. Therefore, a better understanding of MAP is called for to improve our knowledge of the impacts from climate change. The complete and detailed report of this study is described in Chapter 2.

1.2 The future of wind energy in California

Renewable energy is collected from renewable resources, such as solar, wind, and the earth. In the case of wind energy, turbines in wind power systems collect and convert the kinetic energy of the wind into electricity. By 2016, renewable energy has taken up to 12.2% of the total energy production in the United States. Among all the different renewable electricity sources, wind power contributes to 37% of the total renewable electricity generation, ranking it as the second to largest contributor, only after hydro power [17]. Wind energy has also been projected to surpass hydro power in 2019 for the first time, making it the leading renewable power source in the United States [18]. Within the United States, California is the first state to build large wind farms back in the 1980s [19]. Over the last few years, the amount of wind-generated electricity in California has grown rapidly, to where it met almost 7% of the state's total system power by 2017 [20]. At present, California remains among the states that have the largest amount of installed wind power capacity. A study by Barthelmie and Pryor [21] also indicates that a moderate wind energy deployment plan could help delay 1–6 years from crossing the 2 °C warming threshold.

Looking back on the history of wind energy development, wind energy has been proven to be a promising and reliable energy source. However, wind varies significantly across space and time, which impacts the distribution of wind energy. Wind energy, like many other renewable energy resources, is influenced by the global cli-

mate change through its impact on global energy balance and the associated atmospheric circulation [22, 23]. There are many studies that have looked at the climate change impact on wind energy in different regions around the globe [24, 25, 26, 27]. The study by Schaeffer et al. [24] provides an overview of the impacts from climate change on the renewable energy chain. It also identifies the current knowledge gaps and challenges, and states that the detailed future projections of wind speed variability at the hub height level still remain largely unknown. Also, due to the nature of the wind energy, it can not be stored in a similar fashion as hydropower. Thus, the temporal variation of wind speed has a significant impact on the wind energy output, which is also subjected to be influenced by the global climate change [25]. Wind farms typically have shorter lifespan than hydropower dams, which make it more adaptable to the future climate in the long-term. All of the above-mentioned imply the importance of an accurate projection of hub height level wind energy variability under the future climate. Such information will be essential for stakeholders with regards to wind project development and financing [28].

This study aims to advance our understanding of the California hub height wind resource variability till the mid-century. This time frame is chosen as it is within the typical lifespan of 20-25 years for existing wind farms. Results from this study can also provide valuable insight into future wind project planning to mitigate risks from climate change. In this study, we utilized a state-of-the-art climate model with high spatial and temporal resolutions to simulate the future wind resources across California. Using this model, we are able to resolve the complex Californian topography at a high resolution, and analyze the synoptic-scale features that are driving local changes to wind resources. The high temporal resolution of this model allows us to have a more accurate representation of the wind profile. Five major wind farm sites across California, Shiloh, Altamont Pass, Alta, San Gorgonio, Ocotillo, are chosen to analyze the wind variability at each location. This selection is a combination of existing wind farms, and locations with growth potential for developing future wind

5

projects. For the complete description of this study, refer to Chapter 3.

1.3 Clustering analysis of wind patterns in California

This study is a continuous study from Chapter 3. In this study, we extend the future projections of California wind resource variability to the end-of-century, and include a full suite of wind patterns in California. Many studies have been focused on analyzing the climate change impact on wind energy resources. Karnauskas, Lundquist, and Zhang [29] find robust reductions in wind resources in northern mid-latitudes based on simulations from various climate models. Scaling down the analysis to the regional level yields a variety of results. Yu et al. [30] find positive trend in wind speed over the Midwest, the Great Plains, and the U.S. Northeast, while negative trends in the east and in some areas in California. They also detect impact from climate modes on seasonal wind speed change across the U.S.. Within California, Duffy et al. [31] use high resolution climate model, and detect a decreasing trend of potential wind power production in Tehachipi during fall season. These studies show the importance of improving our understanding of the wind resources variability in space and time, which can help in many aspects of future wind project development.

In order to include a complete suite of wind patterns into our analysis, we use a clustering algorithm to group days with similar wind patterns. Clustering is a commonly used method to classify patterns from unlabeled data. There have been studies using such method to find wind patterns. Berg et al. [32] uses a clustering analysis to identify the impact from ENSO on near surface wind in Southern California. Conil and Hall [33] apply a mixture clustering model to examine the primary wind regimes of local atmospheric variability in Southern California. Clustering has also been used to identify the wind patterns associated with localized air pollution

[34, 35].

In this study, we provide a statistical method to analyze the hub-height wind profile without requiring prior knowledge of various wind types. Though we focus on the five wind farm sites across California from Chapter 3, the methodology is generally applicable to any other locations as well. Results from this study link changes in wind clusters with its synoptic-scale meteorological backgrounds, and improve our understanding of California wind resources variability by the end-of-the-century. The full description of this study is presented in Chapter 4.

Bibliography

[1] Dennis L Hartmann et al. "Observations: atmosphere and surface". In: *Climate change 2013 the physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, 2013.

[2] Myles R Allen et al. "IPCC fifth assessment synthesis report-climate change 2014 synthesis report". In: (2014).

[3] Lars E Olsson, William P Elliott, and Sheng-I Hsu. "Marine air penetration in western Oregon: An observational study". In: *Monthly Weather Review* 101.4 (1973), pp. 356–362.

[4] Robert G Read. "Marine air penetration of the Monterey Bay coastal strip and Salinas Valley, California". In: (1971).

[5] Michael A Fosberg and Mark J Schroeder. "Marine air penetration in central California". In: *Journal of Applied Meteorology* 5.5 (1966), pp. 573–589.

[6] STK Miller et al. "Sea breeze: Structure, forecasting, and impacts". In: *Reviews of geophysics* 41.3 (2003).

[7] Mark J Schroeder et al. "marine air invasion of the Pacific Coast a problem analysis". In: *Bulletin of the American Meteorological Society* 48.11 (1967), pp. 802–808.

[8] Clifford F Mass, Mark D Albright, and Daniel J Brees. "The onshore surge of marine air into the Pacific Northwest: A coastal region of complex terrain". In: *Monthly weather review* 114.12 (1986), pp. 2602–2627.

[9] Naoyuki Kurita et al. "Influence of large-scale atmospheric circulation on marine air intrusion toward the East Antarctic coast". In: *Geophysical Research Letters* 43.17 (2016), pp. 9298–9305.

[10] Julien P Nicolas and David H Bromwich. "Climate of West Antarctica and influence of marine air intrusions". In: *Journal of Climate* 24.1 (2011), pp. 49–67.

[11] Jaya Naithani, Hubert Gallée, and Guy Schayes. "Marine air intrusion into the Adelie Land sector of East Antarctica: A study using the regional climate model (MAR)". In: *Journal of Geophysical Research: Atmospheres* 107.D11 (2002).

[12] Charles Chemel and Ranjeet S Sokhi. "Response of London's urban heat island to a marine air intrusion in an easterly wind regime". In: *Boundary-layer meteorology* 144.1 (2012), pp. 65–81.

[13] TD Davis et al. "Guessing Mother Nature's Next Move". In: *PUBLIC UTILITIES FORTNIGHTLY* 143.8 (2005), p. 48.

[14] Scott Beaver and Ahmet Palazoglu. "Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality". In: *Journal of Applied Meteorology and Climatology* 45.12 (2006), pp. 1710–1726.

[15] John E Simpson. *Sea breeze and local winds*. Cambridge University Press, 1994.

[16] Bereket Lebassi et al. "Observed 1970–2005 cooling of summer daytime temperatures in coastal California". In: *Journal of Climate* 22.13 (2009), pp. 3558–3573.

[17] *Renewable energy in the United States*. https://en.wikipedia.org/wiki/Renewable_energy_in_the_United_States. Accessed: 2019-3-21.

[18] US EIA. "Energy Information Administration, Short-Term Energy Outlook, 2010". In: $http://www.eia.gov/forecasts/steo/pdf/steo\_full.pdf$ (2019).

[19] *History of U.S. Wind Energy.* https://www.energy.gov/eere/wind/history-us-wind-energy. Accessed: 2019-3-21.

[20] California Energy Commission. *California Energy Almanac, Total Electricity System Power.* 2016.

[21] RJ Barthelmie and Sara C Pryor. "Potential contribution of wind energy to climate change mitigation". In: *Nature Climate Change* 4.8 (2014), p. 684.

[22] Volker V Quaschning. *Renewable energy and climate change.* Wiley, 2019.

[23] M King Hubbert. "The energy resources of the earth". In: *Scientific American* 225.3 (1971), pp. 60–73.

[24] Roberto Schaeffer et al. "Energy sector vulnerability to climate change: a review". In: *Energy* 38.1 (2012), pp. 1–12.

[25] Sara C Pryor and RJ Barthelmie. "Climate change impacts on wind energy: A review". In: *Renewable and sustainable energy reviews* 14.1 (2010), pp. 430–437.

[26] André Frossard Pereira de Lucena et al. "The vulnerability of wind power to climate change in Brazil". In: *Renewable Energy* 35.5 (2010), pp. 904–912.

[27] SC Pryor, RJ Barthelmie, and Erik Kjellström. "Potential climate change impact on wind energy resources in northern Europe: analyses using a regional climate model". In: *Climate Dynamics* 25.7-8 (2005), pp. 815–835.

[28] A Tindal. "Financing wind farms and the impacts of P90 and P50 yields". In: *EWEA Wind Resource Assessment Workshop.* Vol. 11. 2011.

[29] Kristopher B Karnauskas, Julie K Lundquist, and Lei Zhang. "Southward shift of the global wind energy resource under high carbon dioxide emissions". In: *Nature Geoscience* 11.1 (2018), p. 38.

[30] Lejiang Yu et al. "Temporal and spatial variability of wind resources in the United States as derived from the Climate Forecast System Reanalysis". In: *Journal of Climate* 28.3 (2015), pp. 1166–1183.

[31] Phillip B Duffy et al. "Climate Change Impacts on Generation of Wind, Solar, and Hydropower in California". In: *California Energy Commission. Publication Number: CEC-500-2014-111.* (2014).

[32] Neil Berg et al. "El Niño-Southern oscillation impacts on winter winds over Southern California". In: *Climate dynamics* 40.1-2 (2013), pp. 109–121.

[33] Sebastien Conil and Alex Hall. "Local regimes of atmospheric variability: A case study of Southern California". In: *Journal of Climate* 19.17 (2006), pp. 4308–4325.

[34] Scott Beaver and Ahmet Palazoglu. "Influence of synoptic and mesoscale meteorology on ozone pollution potential for San Joaquin Valley of California". In: *Atmospheric Environment* 43.10 (2009), pp. 1779–1788.

[35] Ling Jin, Robert A Harley, and Nancy J Brown. "Ozone pollution regimes modeled for a summer season in California's San Joaquin Valley: A cluster analysis". In: *Atmospheric environment* 45.27 (2011), pp. 4707–4718.

Chapter 2

Marine Air Penetration in California's Central Valley: Meteorological Drivers and the Impact of Climate Change

2.1 Introduction

Unprecedented climate change is expected in the next century, with inevitable and significant impacts anticipated on the health, welfare, and economies worldwide. Although the large-scale impacts of climate change are well understood, there remains substantial uncertainty with regard to how it will impact local-scale meteorology, which is of greater relevance for local stakeholders. Since projections of climate change over the coming century are typically only available at resolutions that are insufficient to represent local-scale meteorological features, there is a pressing need for the development of new statistical techniques to connect these unresolved features to well-resolved large-scale and synoptic-scale meteorological patterns [1]. To this end, the present study investigates the development and application of one

such technique to a particular coastal phenomenon – marine air penetration (MAP) events [2, 3, 4, 5]. Literature on this topic has widely varying nomenclature for these events, including marine air penetration (MAP), marine air intrusion (MAI), marine air invasion, or onshore surge of marine air. However, these terms essentially all describe the onshore flow component that causes the cool and moist marine air to penetrate far inland. The cooling that comes with the MAP events is the most socially influential impact since it can alter the local energy consumption, as well as bring cooling relief during summertime to densely populated area such as the California Delta region. MAP exemplifies many of the issues associated with projecting climate change at fine spatial scales: Namely, the character of MAP events strongly depends on coastal topography, but restrictions on model resolution (driven by computational constraints) requires topography to be smoothed out, leading to a poor representation of MAP in global models.

The goal of this study is to understand the connections between MAP events of sufficient strength and large-scale meteorology, and use that insight to understand potential trends in MAP events in the future. Our approach proceeds as follows: First, an objective criteria for identifying MAP events is developed based on available meteorological observations. Next, five synoptic-scale meteorological indicators are identified that are strongly associated with MAP events. Finally, these indicators are applied to historical reanalysis and future Coupled Model Intercomparison Project-Phase 5 (CMIP5) projection to identify any trend in MAP frequency.

The remainder of this paper is as follows. First, the character and impacts of MAP events in California are described in section 2.2. Section 2.3 outlines the datasets that have been used in this analysis, and section 2.4 describes our criteria for objective detection of MAP days. Using this criteria for MAP occurrence, section 2.5 then describes the anomalous large-scale meteorological fields that we have found to be associated with MAP events. These fields are then used in section 2.6 to identify five objective indicators associated with MAP occurrence, and a statistical model is

developed for predicting MAP days based on these indicators. This statistical model is then applied in section 2.7 to understand how MAP frequency may be changing in response to a warming climate. Discussion and conclusions follow in section 2.8.

2.2   Marine Air Penetration

MAP events occur in many near-coastal locations throughout the world, driven by rough coastal topography that blocks marine air from penetrating inland everywhere except through one or more narrow regions. In this study, we focus specifically on MAP events that occur in the California San Francisco Bay Delta and Central Valley regions (see Figure 2.1). In this context, MAP is a general meteorological feature composed of both a mesoscale and a synoptic scale phenomena: the daily sea breeze circulation, which naturally arises in response to a land/ocean temperature contrast, a product of a differing heat capacity of the land versus the ocean [3, 6]; and the Pacific coast monsoon [3], a slow and steady transport of the marine air inland from the North Pacific anticyclone [7]. This latter component enhances the on-shore sea breeze flow and allows for marine air to penetrate farther inland than the sea breeze would normally allow. Note that MAP should not be considered interchangeably with the well-known California Delta Breeze, which must be associated with a sea breeze front [8, 9].

MAP events in California most often occur from the late spring through early autumn months, and usually peak in intensity and frequency in the summertime, when the temperature and pressure gradients are the greatest between land and ocean [7]. The coastal topography of California is necessary for MAP events, and acts as a meso-scale "amplifier," which can influence the synoptic-scale "signal" [10]. Further, during summertime, MAP events can be intensified by the Delta Breeze, as the latter is superimposed on the monsoon flow [2].

MAP episodes are a dominant feature of summertime meteorology in California's

14

near-Delta Central Valley. A MAP event can be a relief from summertime heat by transporting cool moist marine air inland, which can alter local temperature and decrease health risks due to prolonged heat exposure. In turn, a MAP event can drive down electricity demand, which is typically dominated by indoor environmental conditioning during summer months [11], or alter local air quality via inland ventilation [12, 13]. However, MAP events exist along a spectrum, ranging from weak MAP episodes that rarely carry marine air beyond the San Francisco Bay Delta through strong events that effectively ventilate a large fraction of the California Central Valley. The strongest MAP events, and the ones that are the focus of this study, are associated with the combination of a sea breeze and a large-scale flow pattern that enhances the wind field. These events are of particular interest, as it is events of this nature that the potential to counteract climate change induced warming in the Central Valley.

Past studies of MAP have typically used observational datasets to identify individual events [2, 4, 10]. A study by Zaremba and Carroll [14] analyzed the more general summer wind flow regimes over the Sacramento Valley in 1991. They classified the wind regime by observing the wind direction in Davis, which sits just north of the San Francisco Bay Delta. Over their study period (May - September 1991), 72% of the days were classified as having some of the features of marine air intrusion (southerly wind at Davis).

Sources that have analyzed projected temperature and wind fields suggest stronger sea breeze events will occur in coastal California in response to climate change [15]. However, it is unclear if their assessment, which is motivated by a projected increase in land/ocean temperature contrast, is applicable to MAP, which is more strongly connected with synoptic-scale meteorological fields [2]. Therefore, our study aims to understand if a similar trend is also anticipated for MAP events.

Table 2.1: Selected CMIP5 models for assessing future trends in MAP days.

Model name	Grid resolution (lat x lon)	Institution
CCSM4	0.9424°x 1.25°	National Center for Atmospheric Research, USA
EC-ERATH	1.1215°x 1.125°	EC-EARTH consortium, European Centre
MRI-CGCM3	1.12148°x 1.125°	Meteorological Research Institute, Japan
HadGEM2-ES	1.25°x 1.875°	Met Office Hadley Centre, UK
MPI-ESM-MR	1.8653°x 1.875°	Max Planck Institute for Meteorology, Germany

2.3   Datasets

This study uses a number of datasets for the purposes of identifying MAP events. For understanding the local character of MAP events, observational data from California Irrigation Management Information System (CIMIS) stations [16] (`http://www.cimis.water.ca.gov/`), National Weather Service (NWS) stations [17], and radiosonde observation data from Oakland airport (OAK station number 72493) (University of Wyoming, Department of Atmospheric Science `http://weather.uwyo.edu/upperair/sounding.html`) were used. The CIMIS data is hourly observation, and sounding data is 12-hourly, only the afternoon hours were used for both datasets. For understanding the large-scale meteorological patterns associated with MAP events, the National Center for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) is used [18]. Among the available CFSR datasets, we chose its analysis subset, which provides 0.5-degree spatial resolution with outputs every 6 hours between 1979-2010. For understanding future trends in MAP frequency, we use data from the 21st century Coupled Model Intercomparison Project Phase 5 (CMIP5) [19] and a 21st century AMIP-style simulations produced using the Variable Resolution Community Earth System Model (VR-CESM) [20, 21].

The five CMIP5 datasets we have used are given in Table 2.1, along with their nominal grid resolution. This selection of datasets represents relatively even coverage from research groups around the world.

2.4 An Objective Criteria for MAP Events

Since MAP events are not a binary phenomena, no singular objective criteria is available for the identification of these features. Namely, MAP events are a general feature that includes any instance of a "sufficiently strong" onshore wind capable of pushing marine air into the Central Valley. Among the summer wind flow regimes in California, as categorized by Zaremba and Carroll [14] (hereafter ZC), we are interested in events of category I-III, which are associated with onshore flow through the San Francisco Bay Delta. These wind regimes, when compared with the mean summertime climatology, are associated with a local onshore wind of particular strength in the late afternoon, cooler inland temperatures, and higher specific humidity. Figure 2.1 shows a map of this region, along with the typical path of MAP winds.
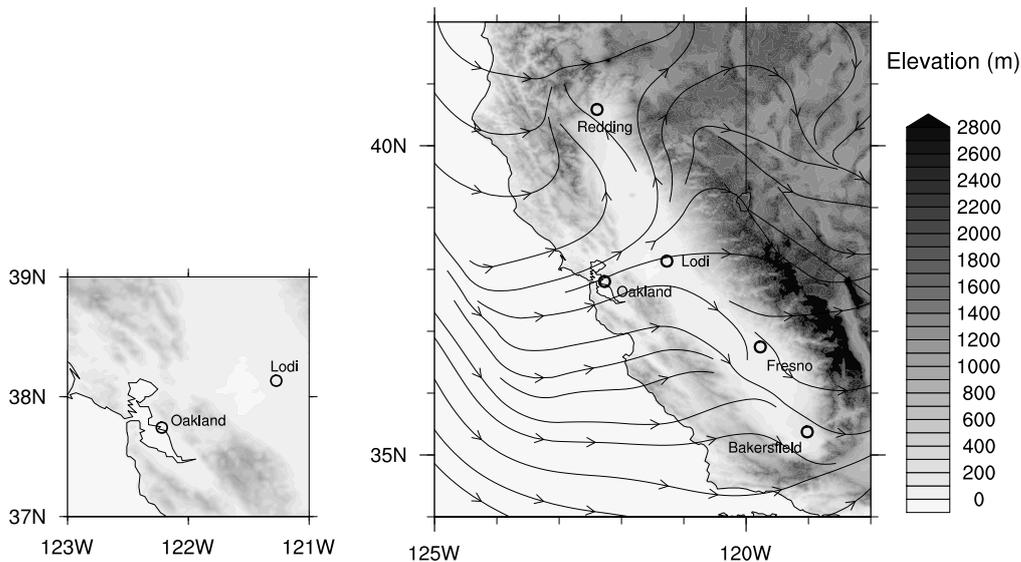
Figure 2.1: (Left) A topographical map of the San Francisco Bay area and Carquinez Strait showing the position of Oakland and Lodi. (Right) A topographical map of California depicting the California Central Valley and the locations of Oakland, Lodi, and Fresno. Wind streamlines for a typical MAP episode (obtained from CFSR data) are superimposed.

In developing an objective criteria for MAP events, it is important to consider any potential issues that may arise due to the observed fields, such as observational

uncertainty and signal-to-noise ratio. With this in mind, a temperature and wind direction criteria is proposed: First, a MAP event must drive a localized cooling through the Bay Delta, which we identify from a sufficiently large surface temperature difference between an inland location in the south of the Central Valley that is mostly insulated from on-shore flows (Fresno, 36.8°N, 119.7°W) and the eastern tip of the San Francisco Bay Delta (Lodi, 38.1°N, 121.4°W). These locations are also chosen, in part, because of the fidelity of the CIMIS temperature measurements at each location. Here, the spatial temperature difference is advantageous since it effectively removes the historical climate change signal (no observable trend was found in the temperature difference from CIMIS observations over the 1980-2010 period), and only triggers for events which deliver a sufficiently strong cooling along the San Francisco Bay. Second, a MAP event must be associated with an onshore wind at the Oakland radiosonde. To develop this criteria, local wind direction at 900hPa was assessed using both radiosonde measurements from Oakland at 00Z (5pm in local time) and inland meteorological station measurements. Among these two, the Oakland radiosonde was found to be the most reliable indicator of a MAP-like wind regime (i.e., ZC category I-III), as inland and near-surface wind observations exhibited more noise due to local surface effects. A criterion for the specific humidity field at Lodi was also considered, but was found to be poorly indicative of strong on-shore flow through the Delta and ventilation of the Central Valley. The weak signal was likely due to the cooler temperatures and relatively low specific humidity of the North Pacific.

The strength of the temperature and wind criteria has been tuned to provide a representative sample of "sufficiently strong" MAP events, while avoiding false positives that were clearly not associated with a MAP-like wind regime, and weak MAP days associated with the diurnal sea breeze but not with a large-scale meteorological forcing. Specifically, the temperature criteria was chosen to equal the mean temperature difference between Fresno and Lodi minus one standard devia-

tion. Consequently, the proposed observational criteria is given by:

Observational MAP Criteria

(a) The surface temperature difference between Fresno and Lodi must be greater than or equal to 7°C

(b) The 900hPa across-shore wind speed must be greater or equal to 3 m/s. The across-shore direction is defined as perpendicular to the west coast shoreline at 150 degree SE, and positive across-shore wind speed indicates onshore wind.

Here the 7°C difference was selected by first calculating the mean temperature difference between Fresno and Lodi (9.0°C), minus one standard deviation (2.1°C). Our motivation was to include the 'stronger' cooling days (at Lodi) while eliminate the 'weaker' ones, then further select the MAP events using the on-shore wind criteria. These two criteria added together were to select the on-shore flow with sufficient cooling effect around the Delta region. The on-shore wind direction was defined as wind direction ranging from 150° to 330°, as an approximation of on-shore direction with respect to the US west coastline. The across-shore wind speed was then defined as the wind component from the on-shore wind that was perpendicular to the coastline. From the wind direction histogram (not shown), the on-shore wind was found to be the most frequent.

As discussed earlier, MAP events exist along a spectrum and so there remains some latitude in defining thresholds to ensure only the most significant events are captured. The thresholds chosen above allow us to eliminate the weakest MAP events while also maintaining a sufficient statistical sample of these events. Simply using the two observational MAP criteria above, we were able to detect a lower surface temperature, higher surface water vapor mixing ratio, and higher surface wind speed during the afternoon hours on MAP days (blue lines in Figure 2.2) compared to non-MAP day (black lines in Figure 2.2) from hourly CIMIS data.

Surface temperature at Fresno (dash lines in Figure 2.2) didn't show much influence under MAP days, due to Fresno sits further from the Delta region.
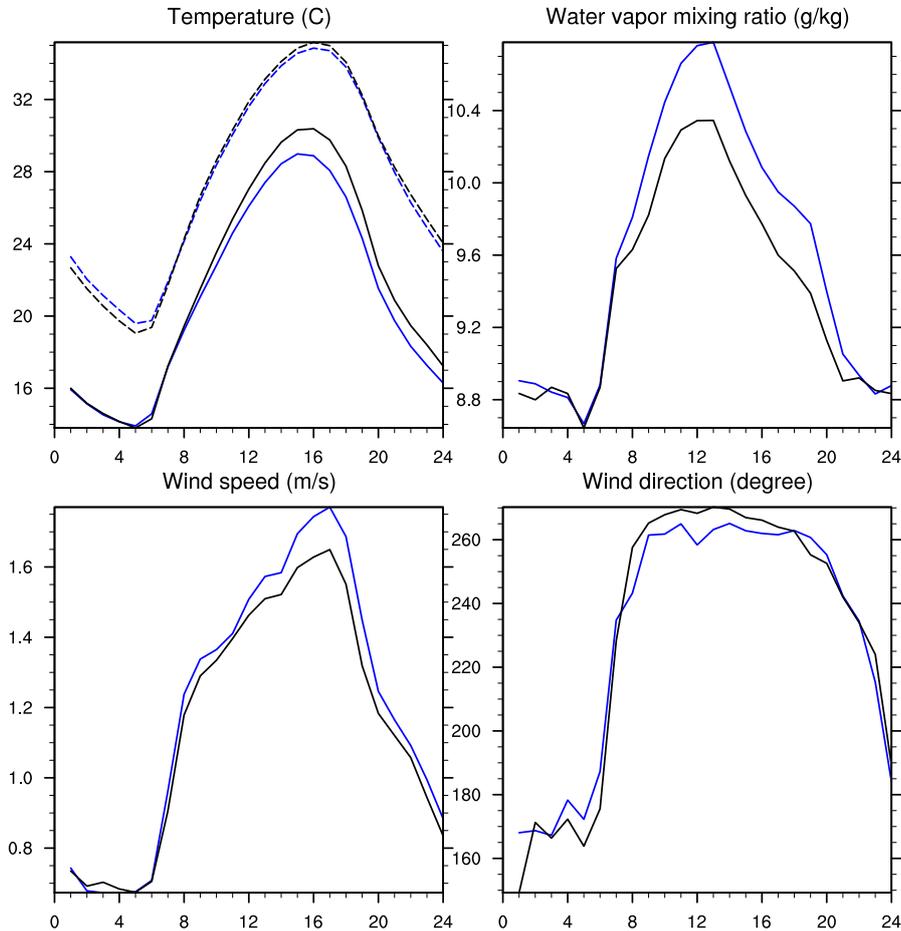
Figure 2.2: Hourly mean values of surface temperature, water vapor mixing ratio, wind speed and wind direction at Lodi (solid lines), California on days meeting the observational MAP criteria (blue) and days not meeting the criteria (black) from CIMIS observation. Dash lines shown are surface temperature at Fresno on MAP days (blue) and non-MAP days (black).

CFSR data from 2001 to 2010 were initially used in this study in order to assess the performance of the observational criteria. The number of MAP days that match the criteria for each month over the 10-year period is plotted in Figure 2.3. Consistent with the known climatology of MAP events [2], most MAP days occurred during the summer months (JJA). In our analysis, approximately 46% of all JJA days over the 2001-2010 period are classified as MAP days. Although this is much fewer than the 72% of summer days identified as MAI days by Zaremba and Car-
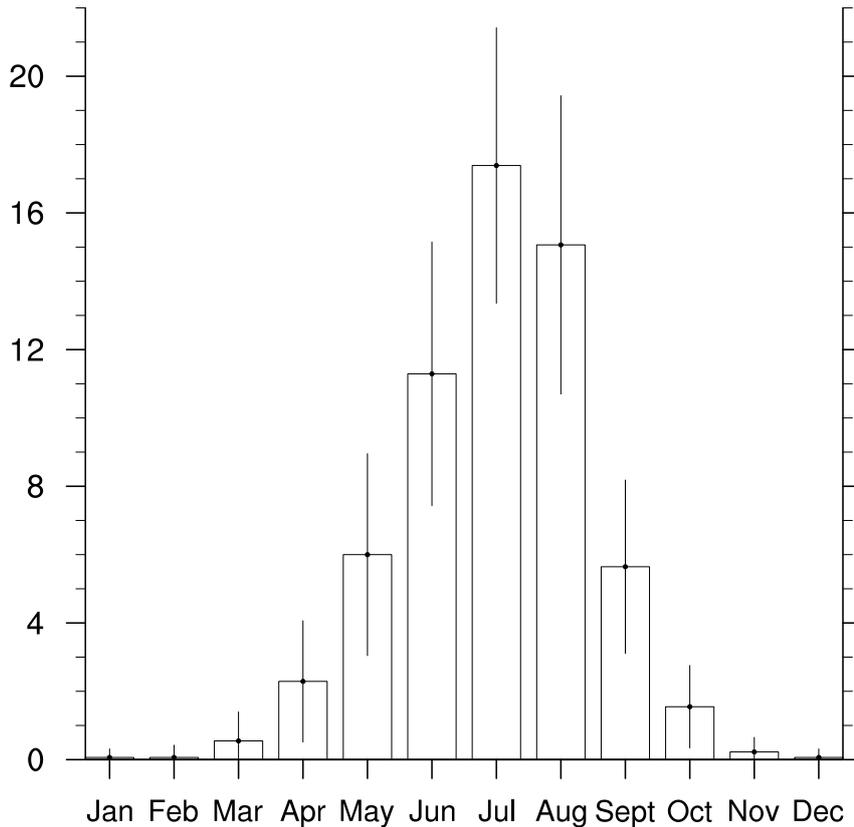
Figure 2.3: Average number of MAP days per month selected from CIMIS observation between 2001 and 2010. Vertical bars depict standard deviations.

roll [14], this difference again represents our freedom to isolate only MAP days of sufficient strength.

Since this study requires the large-scale climatological fields associated with MAP days, the CFSR reanalysis dataset was adopted. However, although CFSR is closely constrained to meteorological station data in this region, there are slight differences in the mean climatology of CFSR and CIMIS that needed to be addressed. Grid points in CFSR are not coincident with meteorological stations, and so the grid point data at (38°N, 121.5°W) was used for "Lodi" and data at (36.5°N, 120°W) was used for "Fresno". As a consequence, the mean temperature difference from CIMIS was observed to be 9.0°C and 10.1°C from CFSR. Further, CFSR exhibited a 1.7°C trend over the 30 year period 1980-2010 that was not present in CIMIS data, and there was no trend in the temperature difference between Fresno and Lodi from

CIMIS data. Finally, mean 900hPa wind speed at Oakland was measured as 3.1 m/s in CIMIS and 3.8 m/s in CFSR. Consequently, our thresholds for MAP events have been adjusted to account for the best match between the CFSR and CIMIS data. This led to the CFSR criteria for MAP days, as follows:

CFSR MAP Criteria

(a) The surface temperature difference between Fresno and Lodi must be greater than or equal to

$$8°C + 1.7°C \times \frac{\langle YEAR \rangle - 2005}{30}$$

(b) The 900hPa across-shore wind speed must be greater or equal to 1.87 m/s.

Using these adjusted criteria, the percentage of JJA MAP days in CFSR is found to be approximately 51% between 2001 and 2010 with a 73% overlap between CIMIS and CFSR datasets. Expanding our study period to 1979-2010 leads to a selection of approximately 47% of MAP days in CFSR, with a 69% overlap with CIMIS. Surface temperature, specific humidity and wind speed fields from the two sets of MAP events plotted in CFSR appear very similar.

2.5   Synoptic-scale Meteorology During MAP Events

For MAP days selected using the CFSR criteria, synoptic- and large-scale meteorological fields describing 700hPa geopotential height anomaly (a), 700hPa geopotential height (b), wind anomaly streamlines (c), wind streamlines (d), temperature anomalies (e) and specific humidity anomalies (f) are plotted in Figure 2.4.

On MAP days, the mean synoptic-scale meteorology is dominated by a high-pressure ridge in the Gulf of Alaska accompanied by a coastal trough through the
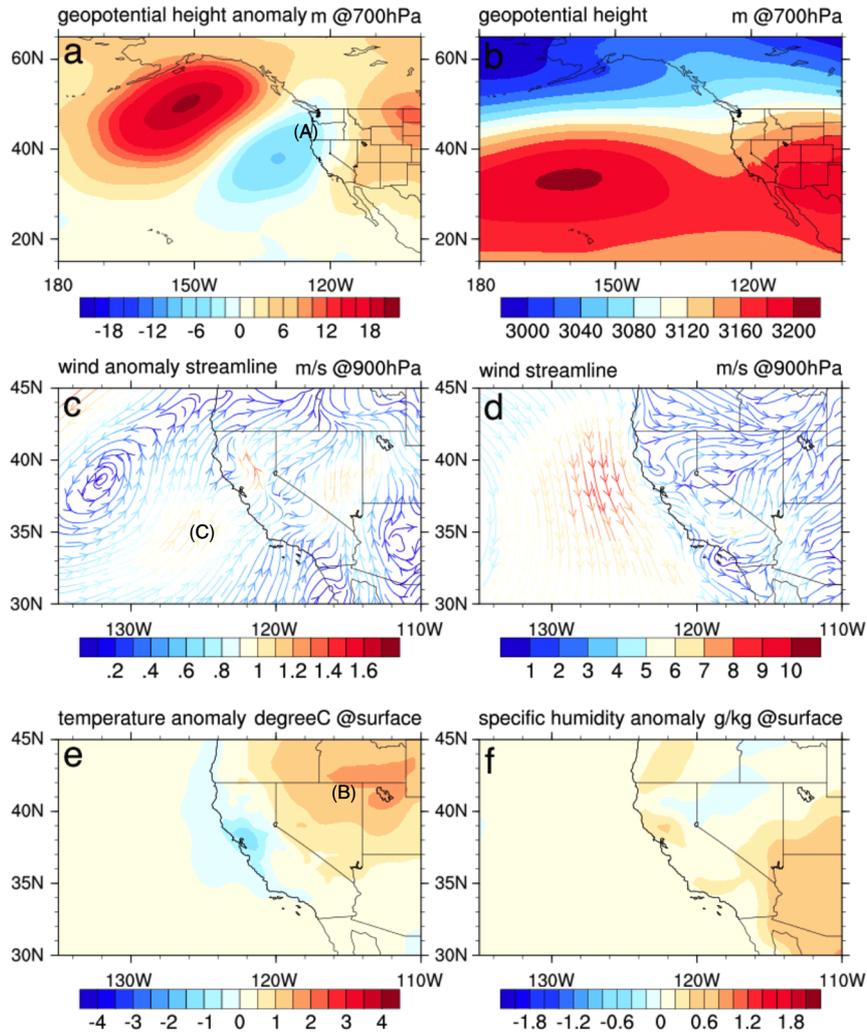
Figure 2.4: Geopotential height and height anomaly at 700 hPa, wind streamlines and anomaly streamlines at 900hPa, surface temperature anomaly and surface specific humidity anomaly on CFSR MAP days over the 2001-2010 "training" period from CFSR.

Pacific Northwest. The presence of this trough is effective at modifying the contours of the 700hPa geostrophic height (Fig. 2.4b) field to be approximately perpendicular to the coast, which in turn leads to a 700hPa geostrophic wind directed into the San Francisco Bay Delta. The coastal trough also leads to southerwesterly winds throughout the Western U.S. that are then responsible for a large positive temperature anomaly that extends throughout this region. Off-shore 900hPa winds, which are typically northerly, are weakened during MAP days and are shifted to be more closely aligned with the coast. In the San Joaquin Valley, 900hPa winds are observed to diverge away from the Delta, with the strongest winds directed northward. The near-surface onshore winds are then responsible for a drop in surface temperatures through the Bay Area and a corresponding increase in specific humidity as cool, moist oceanic air is forced through the Delta.

Days that only satisfied the on-shore wind criteria were similarly assessed, and were observed to be associated with a geopotential height anomaly that maintained a similar Rossby-wave pattern to Fig. 2.4a but was far less significant. This is indicative that the temperature criteria is effective at isolating days when the large-scale circulation enhances the on-shore flow.

2.6 Synoptic-scale Meteorological Indicators for MAP Events

The meteorological fields associated with MAP events that emerges from Figure 2.4 suggests that MAP days are associated with several potential large-scale meteorological indicators. For the purposes of this study, we have proposed the following five criteria as potential binary indicators of MAP days (all the anomalies are daily average values):

(I) Positive 700hPa geopotential height anomaly within the region bounded by 30°N - 60°N and 165°W - 120°W: As observed in Figure 2.4, MAP days are

associated with a dipole in the 700hPa geopotential anomaly field, with a ridge over the Gulf of Alaska and a trough over the Pacific Northwest. Although the position of the dipole is not consistent year-to-year, the average anomaly in the Gulf of Alaska still tends to be positive during MAP events.

(II) Negative 700hPa geopotential tendency at 44°N, 125°W (Fig. 2.4a, point (A)): In the days leading up to a MAP event, it is frequently the case that geopotential height anomaly dipole shifts eastward (geopotential height anomaly in the five days leading up to the MAP event is plotted in Figure 2.5), driving a negative geopotential tendency off of the Oregon Coast (Figure 2.6). As predicted by the quasi-geostrophic omega equation, a negative geopotential anomaly is associated with upward motion that is potentially associated with frontal activity.

(III) Inland daily average surface temperature anomaly at (41.5°N, 115.5°W) (Fig. 2.4e, point (B))greater than 1°C: As described earlier, and as apparent in Figure 2.4, MAP days are associated with a positive inland temperature anomaly related to the southwesterly wind field. This indicator appears to be maximal in the vicinity of 41.5°N, 115.5°W.

(IV) Along-shore 700hPa geopotential height difference between (42°N, 127°W) to (32.5°N, 121°W) less than or equal to 5m: Geostrophic onshore flow at the 700hPa level occurs when geopotential contours are perpendicular to the shore – or, equivalently, when there is a geopotential height gradient parallel to the shoreline. In Figure 2.4 the presence of geopotential contours perpendicular to the shoreline is indicative that this pattern is prominent during MAP days.

(V) Off-shore wind speed anomaly at (35°N, 125°W) (Fig. 2.4c, point (C)) less than or equal to 2 m/s: Large off-shore wind speeds (particularly northerly winds) tend to lead to suppressed MAP activity, as marine air is pushed along the shore and away from the Bay Delta. On the other hand, MAP days are associated with suppressed off-shore wind speeds that can encourage on-shore flow. The position (35°N, 125°W) was chosen because it is at the offshore region where the wind speed anomaly differs most between MAP and non-MAP days.

Using these five indicators, a multivariate logistic regression model was developed. Using the R generalized linear model (`glm`), it was then determined whether or not each indicator was a statistically significant indicator of a MAP event. The regression model indicated high significance for all of the five synoptic indicators for 2001-2010 (the p-values for each of the indicators were (I) 1.22×10^{-11}, (II) 2×10^{-16}, (III) 1.89×10^{-4}, (IV) 2×10^{-16}, and (V) 9.8×10^{-4}). The strongest indicators for MAP were purely associated with the 700hPa geopotential field and its corresponding tendency, indicative of the role played by synoptic-scale meteorology in driving MAP events. The receiver operating characteristic (ROC) curve (supplement Figure S1) illustrates the performance of this binary classifier system. The maximum AUC = 1 means the perfect differentiation between MAP and non-MAP, and AUC = 0.5 means the classifier is no better than a random guess. For our model, the area under curve (AUC) value was ∼0.74 for 2001-2010. Applying this procedure to 1979-1989 and 1990-2000 yielded AUC values of ∼0.69 and ∼0.71, respectively. This suggests that the model is consistent across all three decades, and indicates that the large-scale indicators described above will correctly discriminate MAP events from non-MAP events approximately 70% of the time.

To further understand the relationship between these indicators and the objective MAP criteria, we also assessed the performance of the model as the temperature difference criterion was varied between 1°C and 12°C. Indicators (II), (III) and (IV)
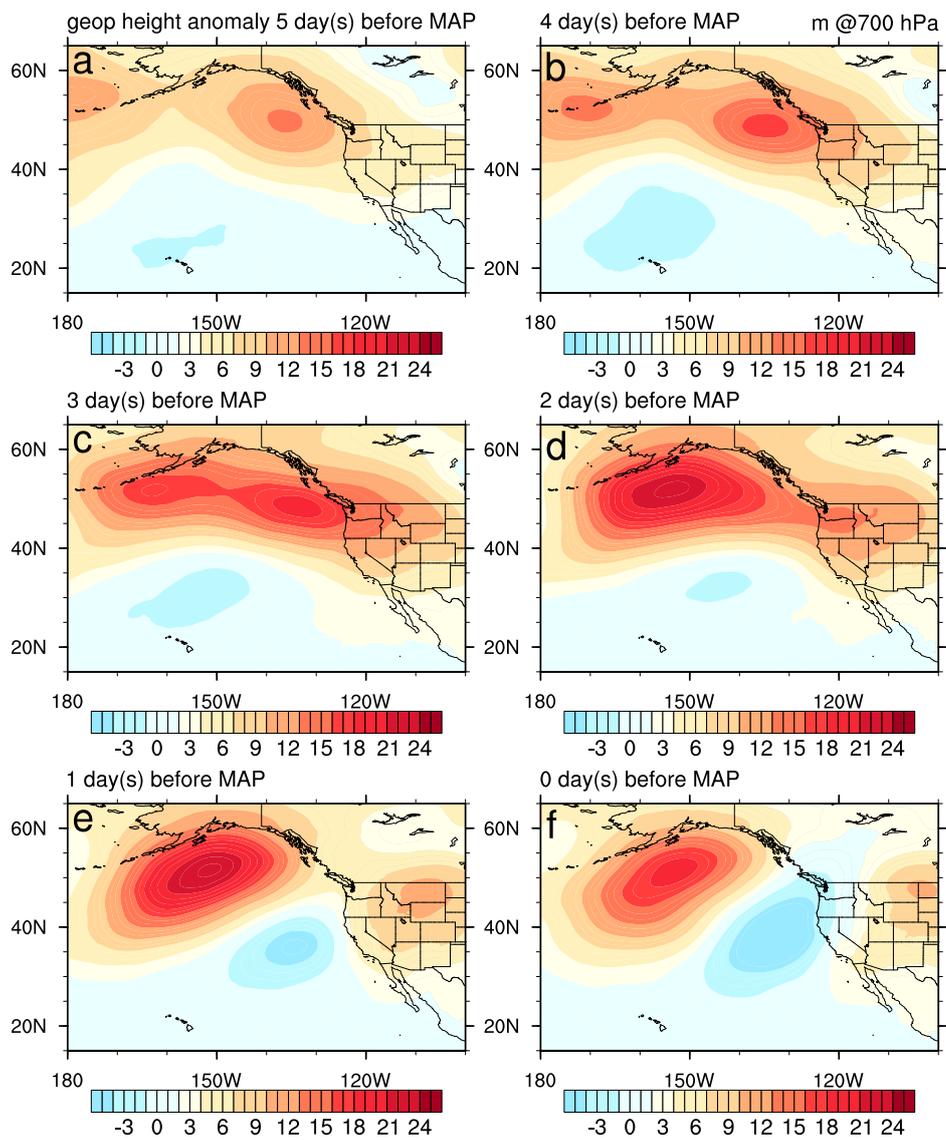
Figure 2.5: 700hPa geopotential height anomaly field for 5-0 day(s) before MAP events from CFSR.
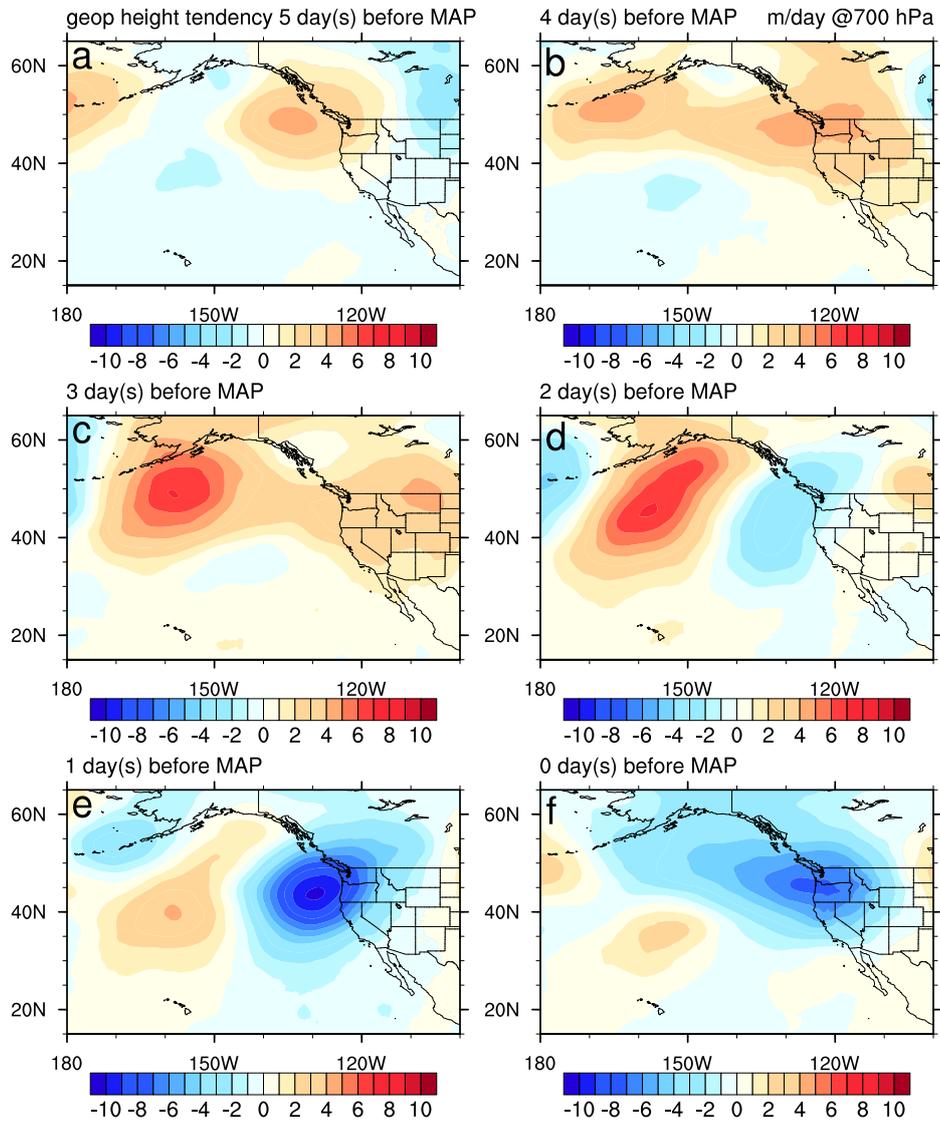
Figure 2.6: 700hPa geopotential tendency field for 5-0 day(s) before MAP events from CFSR.

were found to be strongly predictive across all temperatures, and so were primarily associated with the on-shore wind criterion. Indicators (I) and (V) were both sensitive to the temperature difference, with maximum predictability attained with a CFSR temperature difference of 8°C. Other values of the temperature difference led to indicators that were either not significantly correlated or produced a lower AUC value.

Using the five binary indicators, 32 daily categories are developed based on all possible configurations of indicators and the resulting classification given in Table 2.2. For each of these categories, 1 indicates the number of MAP and non-MAP events for each category, with 1 indicating the criterion holds, and 0 is false, for all JJA days between 1979-2010. As expected, categories that satisfy a larger fraction of the five indicators also tend to have a larger ratio of MAP to non-MAP days. However, days that simultaneously satisfy all five criteria (category 1) are fairly rare: Although this category includes more than four times as many MAP days as non-MAP days, only 46 such days occurred over the 32 year study period. Weighing categories by absolute frequency, category 2 (which does not trigger indicator (V)), category 4 (which does not trigger indicator (IV) or (V)), and category 18 (which does not trigger indicator (I) or (V)) exhibit the greatest occurrence of MAP days.

Figure 2.7 depicts the synoptic meteorology associated with all days in category 2. It captures 181 MAP days out of total 1326 MAP days, or 13.7% of the total number of MAP days between 1979-2010. Category 2 days have only a weak off-shore wind speed anomaly. However, they do feature a large positive geopotential height anomaly (Figure 2.7a) that sits in the Gulf of Alaska, accompanied by a weak negative geopotential height anomaly sitting at off the Oregon coast. This suggests a Rossby wave train directed eastward and slightly southward. The negative anomaly acts to redirect the strong northerly flow off of the California coast, leading to onshore winds directed through the San Francisco Bay Delta. The positive specific humidity anomaly (Figure 2.7f) extends south along the Baja peninsula, caused by

weakened off-shore winds through Southern California.

Figure 2.8 depicts category 4, the second-most frequent MAP event category by total count (135 MAP days, or 10.2% of total MAP days). This category is similar to category 2 in that it captures days with a positive geopotential anomaly (Figure 2.8a) in the north Pacific, but unlike category 2 it is not associated with an offshore trough. Although this configuration is fairly common among summer days, there is only a slightly better than even ($\sim 54\%$) chance for this category to produce a MAP event. In this case, the occurrence of MAP will be primarily driven by the local meteorology.

Figure 2.9 depicts category 18, the third most frequent MAP event category, which accounts for 124 MAP days, or 9.4% of the total number of MAP days. Figure 2.9a shows a much different large-scale pattern compared to category 2, particularly in the north Pacific, where there is a strong negative geopotential height anomaly centered at (45°N, 135°W) associated with a Rossby wave train that is inclined to the north. This large negative geopotential height anomaly triggers a substantial weakening of the offshore northerly winds, in turn permitting air to "leak" into the Delta. The strength of the geopotential anomaly dipole across the US West also enhances southerly flow in this region and drives a large inland positive temperature anomaly (satisfying criteria (III)).

For these three categories, lines of constant negative geopotential height anomaly lie perpendicular to the California coastline, driving the wind anomaly in the onshore direction. Categories 2 and 18 also feature an offshore trough, in agreement with Beaver and Palazoglu [12]. Although it may seem that the offshore trough would be a better indicator of MAP, it turns out that replacing the positive geopotential anomaly indicator (I) with a negative geopotential anomaly indicator off of the Oregon coast does not actually improve our ability to detect of MAP days.

In contrast, the most frequent non-MAP events fall into category 32 (depicted in Figure 2.10), and account for 223 (14.4%) non-MAP days out of 1554 total. In
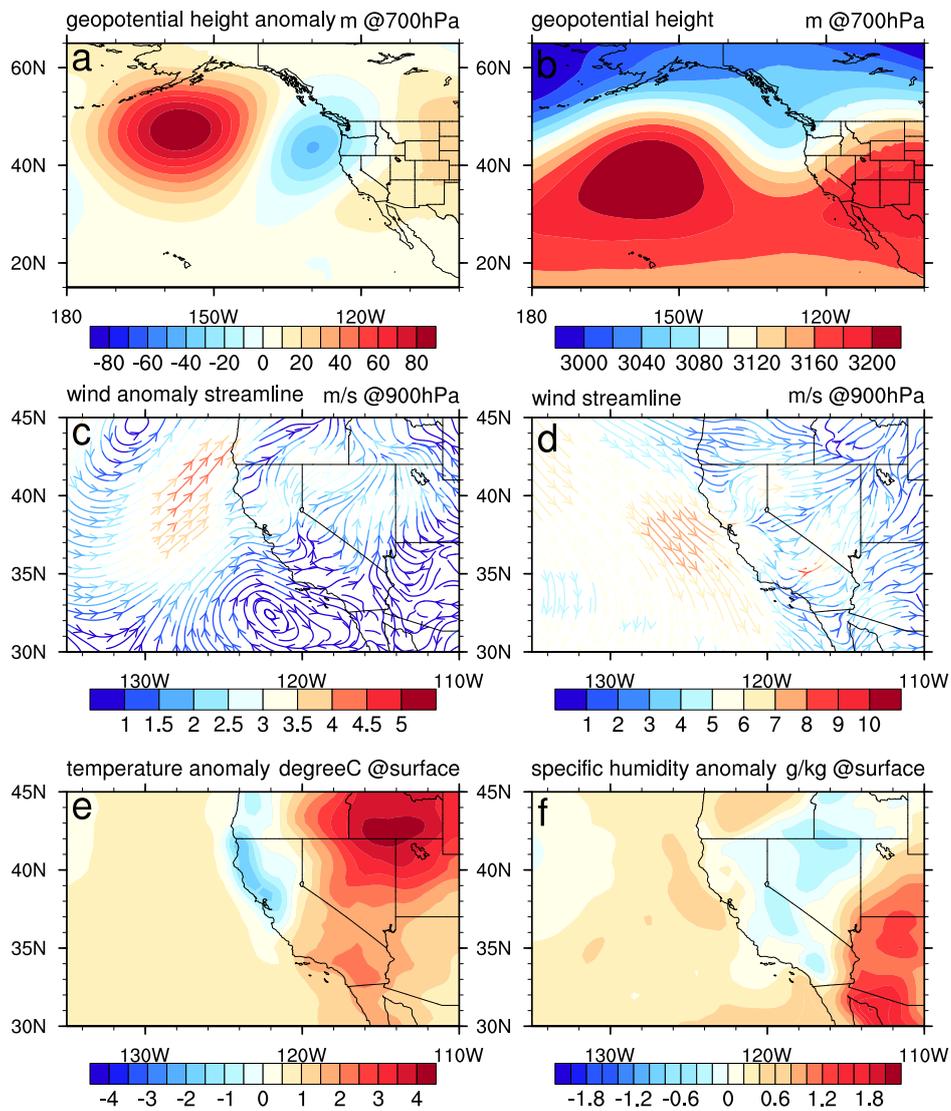
Figure 2.7: The synoptic meteorology over the period 1979-2010 of category 2 days (the most frequent MAP category) from CFSR.
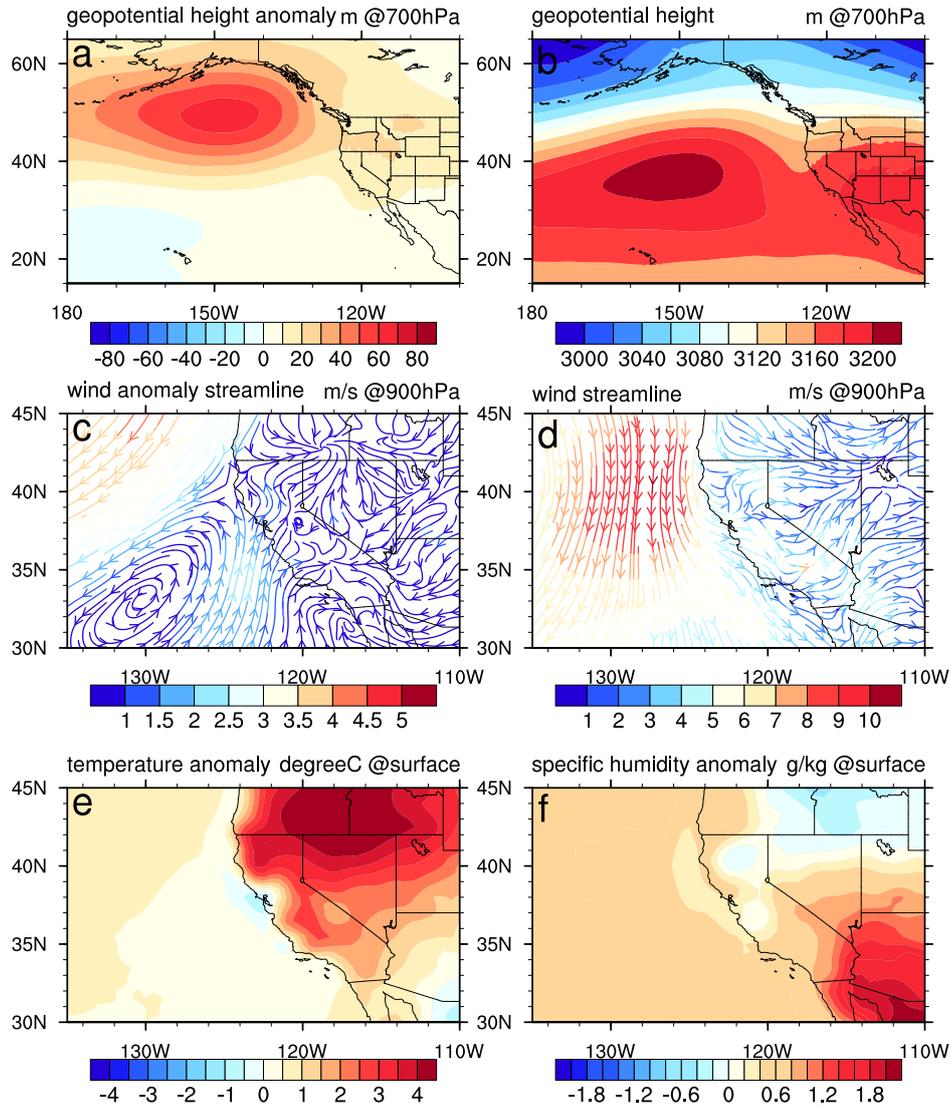
Figure 2.8: The synoptic meteorology over the period 1979-2010 of category 4 days (the second most frequent MAP category) from CFSR.
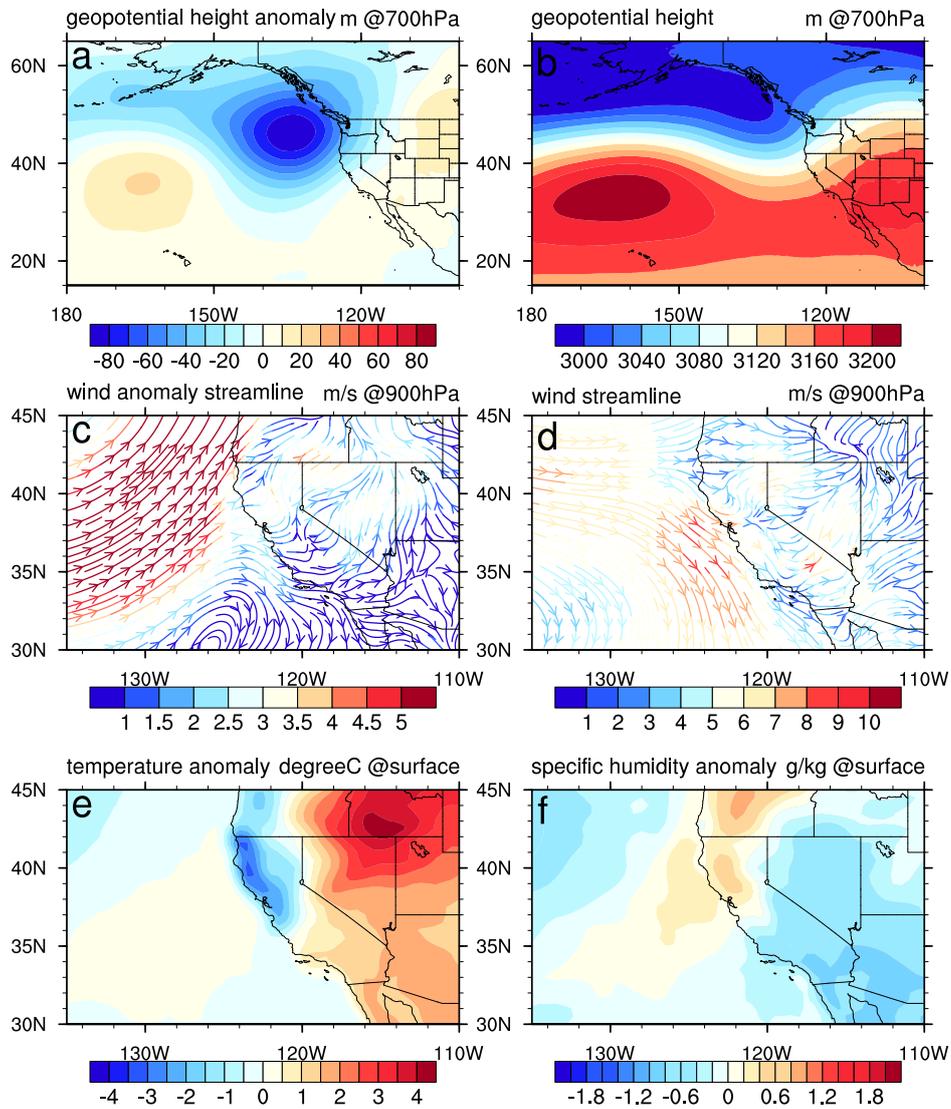
Figure 2.9: The synoptic meteorology over the period 1979-2010 of category 18 days (the third most frequent MAP category) from CFSR.

this case, the synoptic-scale pattern (figure 2.10a) is in opposition to the pattern in MAP category 2, with a major negative geopotential height anomaly sitting in the Gulf of Alaska, and a weak positive anomaly located near the Oregon coast. The intensity of the negative geopotential anomaly in this category partially explains why the mean synoptic-scale meteorology indicates a strong positive geopotential anomaly in the north Pacific associated with MAP days (Figure 2.4), even though geopotential anomaly in this region takes on opposing signs for the top two MAP categories – namely, a negative geopotential anomaly in this region is frequently associated with a blocking pattern that prevents MAP from occurring.

The second to most frequent non-MAP category is category 16 (Figure 2.11), which captures 11.1% of the non-MAP days, and has a dominant positive geopotential height anomaly (Figure 2.11a) centering at (45°N, 140°W). This pattern is again an off-shore blocking pattern, here exhibiting characteristics in opposition to category 18 (Figure 2.9).

Notably, the top two non-MAP categories both have positive geopotential height anomaly contour lines perpendicular to the coastline, enhancing the northerly flow off of the California coast, and blocking on-shore flow.

2.7   Future Trends in MAP Events

In this section, the categorization strategy from section 2.6 is used to project future trends in MAP days. Since MAP events are fairly localized and driven by quick variations in local topography, they are not actually resolved in the coarse resolution data typically produced by global climate models (GCMs). However, the indicators identified in section 2.6 are well-resolved, even in coarse-resolution GCM data. To estimate the number of MAP days in a given year, the following procedure was performed: For each JJA day over the period 1980-2100, the five binary MAP indicators are computed for each CMIP5 dataset. Cross-referencing the associated
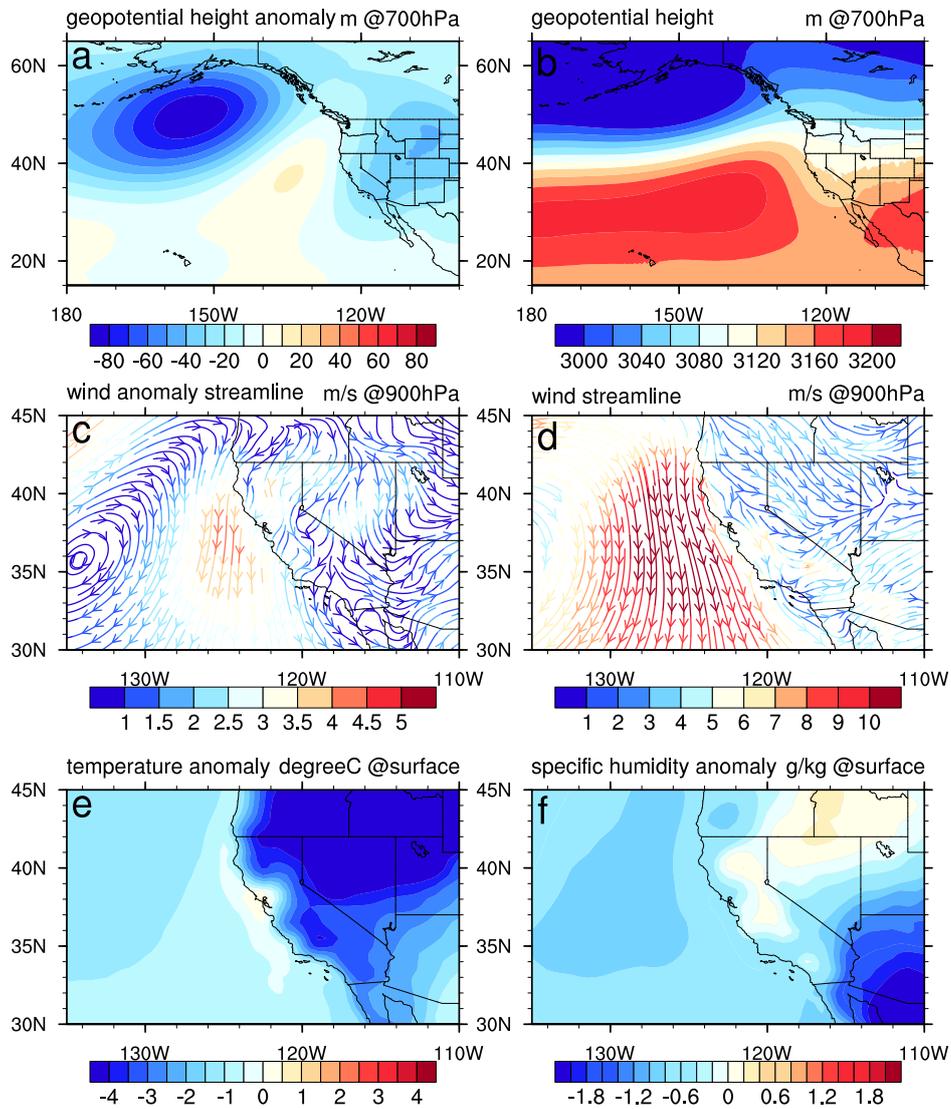
Figure 2.10: The synoptic meteorology over the period 1979-2010 of category 32 days (the most frequent non-MAP category) from CFSR.
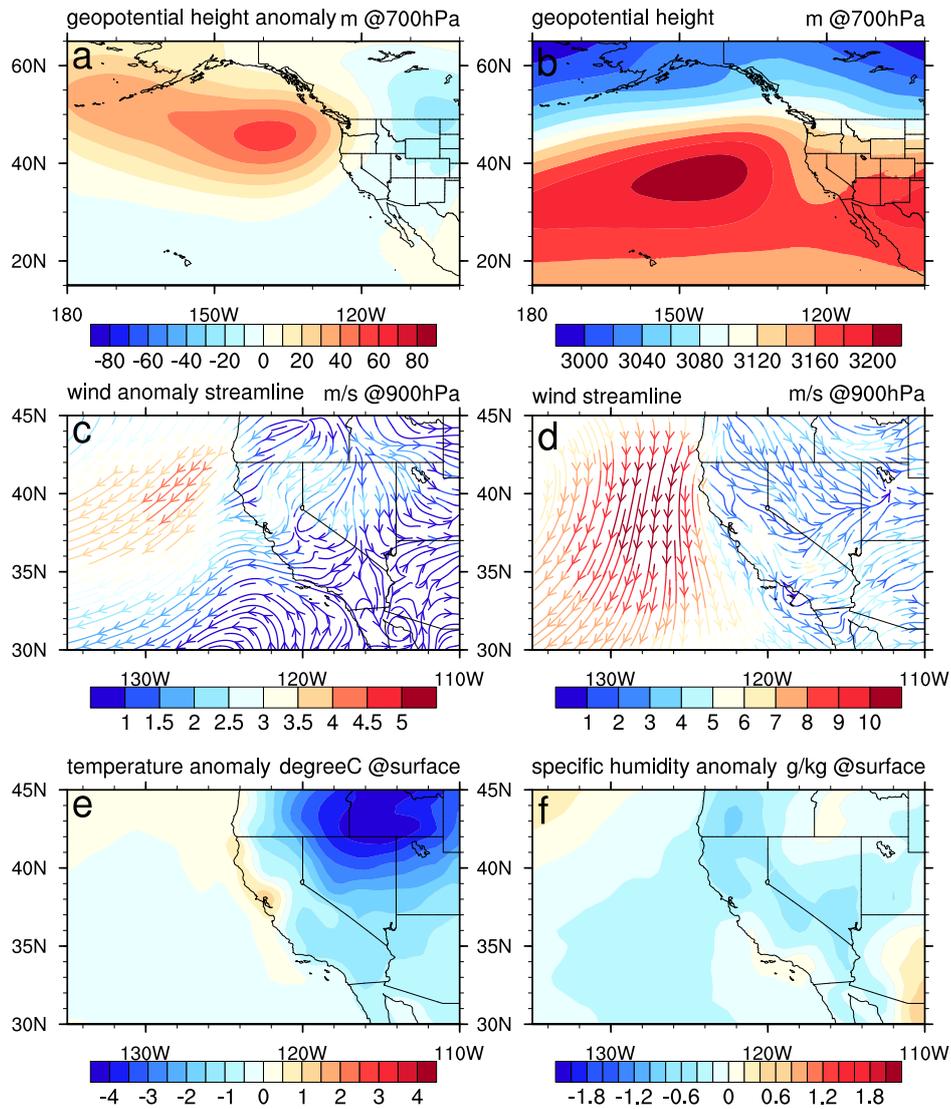
Figure 2.11: The synoptic meteorology over the period 1979-2010 of category 16 days (the second most frequent non-MAP category) from CFSR.

category (1-32) with the number of 1979-2010 MAP days then yields a corresponding probability of the day being a MAP day (i.e. category 1 is a MAP day $46/57 = 80.7\%$ of the time and not a MAP day $11/57 = 19.3\%$ of the time). Summing the MAP probability over all days then gives the mean number of MAP days expected over that year. This procedure then allows us to understand how changes in the frequency of the 32 MAP categories will impact the frequency of MAP days.

Historical MAP days are selected in the CMIP5 datasets and VR-CESM dataset of Rhoades et al. [21] using the five synoptic indicators, and compared this selection with historical MAP days from CFSR. Figure 2.12 shows that the number of MAP days was found to vary around 42 ± 4 out of the 92 days (JJA). In fact, there is surprisingly good agreement on the number of MAP days among all datasets (even though there was little agreement for a particular year), suggesting no significant biases in the representation of the synoptic-scale meteorology of this region.

Future trends in the five synoptic indicators in CMIP5 were then assessed under different climate scenarios, by analyzing experiments arising from the "business as usual" Representative Concentration Pathway 8.5 (RCP8.5). Indicators (I), (III) and (V) were all anomalous values calculated from each year mean, so this could include the annual variations during the study period. Therefore, building a linear regression line based on the corresponding climatological field was necessary so that the climatological trends were better represented, and then the anomaly values for each indicators (I), (III) and (V) were calculated against the values from regression lines. As expected, all models exhibited enhanced inland temperatures and increases in geopotential height anomaly through the 21st century, consistent with the well-known consequences of climate change, but exhibited mixed behavior for the individual indicator fields. Nonetheless, as shown in Figure 2.13, all datasets examined produced a small upward trend ($\sim$0.2 days/century on average). The primary driver for enhanced MAP came from reduced along-shore wind speeds at (35°N, 125°W) in all models, leading to more days that triggered indicator (V),
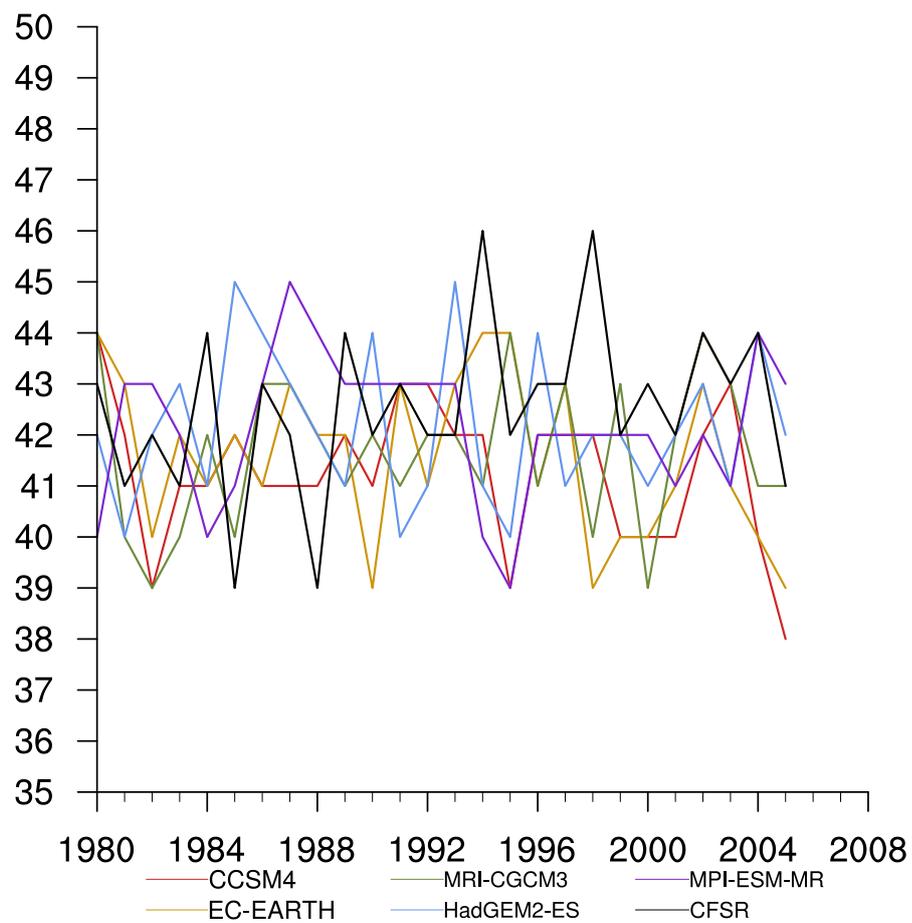
Figure 2.12: Number of MAP days selected over JJA from CFSR and historical CMIP5 data.

and this trend is consistent across all five CMIP5 models, though the magnitude varies. A slight reduction in geopotential height difference also emerged, which led to fewer days that triggered indicator (IV), although it was not enough to counteract the change in indicator (V). There is no clear agreement between models for the geopotential ridging indicator (I), the geopotential tendency (II) or the temperature indicator (III). In summary, this result suggests that climate change may be weakly conducive to the emergence of synoptic-scale fields that are favorable to MAP. Nonetheless, we note that a key limitation of this study is that it cannot be used to draw conclusions on weaker MAP events, or the effect of a cooler sea breeze (potentially associated with increased upwelling), both of which could also mitigate warming near the Delta region.

2.8   Conclusions

The goals of this project have been threefold: First, to develop a simple, objective criteria for the identification of MAP episodes in the California Central Valley from observational and reanalysis data; second, to characterize the synoptic- and large-scale meteorological fields that are correlated with these MAP episodes; and third, to bridge the gap between the relatively well-understood future behavior of synoptic- and large-scale weather systems and a local-scale meteorological phenomenon. In this study, it was observed that MAP episodes could be identified as the conjunction of a temperature difference along the central valley and an on-shore wind speed criteria. Using this criteria, MAP episodes were then associated with five large-scale indicators: (I) a positive geopotential anomaly in the Gulf of Alaska, (II) a negative geopotential tendency off of the Oregon coast, (III) an inland surface temperature anomaly in the north-west of the continental US, (IV) an off-shore geopotential height difference, and (V) an off-shore wind speed anomaly. These indicators then formed the foundation for a model for probabilitistic prediction of MAP episodes.
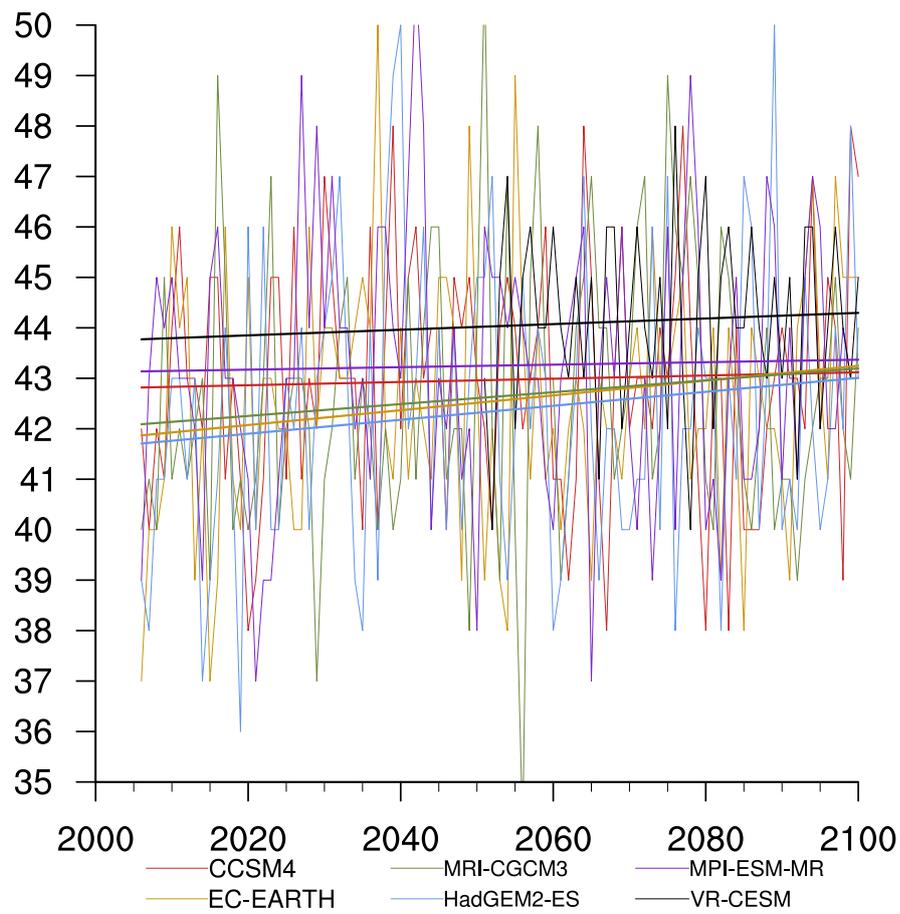
Figure 2.13: Number of MAP days selected over each summer season from CMIP5 and VR-CESM under RCP8.5, with horizontal lines indicate the linear trends from each CMIP5 dataset.

Finally, by analyzing MAP events through the lens of synoptic-scale indicators, this study finds that climate change may favor synoptic-scale conditions that are conducive to MAP events.

We note that the connections across scales that emerge from this study can be further used for improving forecasts of MAP events using available synoptic-scale fields. We also argue that a similar methodology can likewise be employed to develop statistical relationships between large-scale fields and local meteorological features.

Many other factors play an important role in the emergence of MAP that were not investigated here, including local meteorology, large-scale teleconnections and coastal upwelling. In particular, coastal upwelling is theorized to intensify under climate change [22], and has a strong influence on coastal temperatures and subsequently the strength of the cooling effect associated with MAP episodes. It also remains unclear how MAP interacts with other large-scale systems, such as the North Pacific Oscillation (NPO), Pacific Decadal Oscillation (PDO), and El Niño-Southern Oscillation (ENSO). Further, there is an incomplete understanding of how MAP interacts with other meteorological features, including coastal fog and Tule fog. These topics remain for the consideration of a future study.

Table 2.2: Number of MAP and non-MAP days in each of the 32 MAP categories for 1979-2010. Categories 2, 4, and 18 (bold) are the top three most frequent MAP categories, and categories 32 and 16 (bold) are the top two most frequent non-MAP categories.

	Indicators					1979-1989		1990-2000		2001-2010		1979-2010	
Category	I	II	III	IV	V	MAP	non-MAP	MAP	non-MAP	MAP	non-MAP	MAP	non-MAP
1	1	1	1	1	1	16	3	15	8	15	0	46	11
2	1	1	1	1	0	**69**	22	**68**	26	**44**	17	**181**	65
3	1	1	1	0	1	10	6	3	1	9	5	22	12
4	1	1	1	0	0	**44**	30	31	50	**60**	36	**135**	116
5	1	1	0	1	1	3	3	5	2	6	2	14	7
6	1	1	0	1	0	17	15	14	16	17	7	48	38
7	1	1	0	0	1	3	3	3	8	4	0	10	11
8	1	1	0	0	0	18	28	9	31	18	29	45	88
9	1	0	1	1	1	13	7	6	3	5	2	24	12
10	1	0	1	1	0	30	13	27	25	28	10	85	48
11	1	0	1	0	1	6	9	3	7	9	8	18	24
12	1	0	1	0	0	22	45	23	43	39	28	84	116
13	1	0	0	1	1	7	5	5	10	5	2	17	17
14	1	0	0	1	0	11	13	13	18	5	14	29	45
15	1	0	0	0	1	7	6	3	16	3	10	13	32
16	1	0	0	0	0	20	**64**	14	**50**	18	**58**	52	**172**
17	0	1	1	1	1	4	0	11	3	10	1	25	4
18	0	1	1	1	0	**42**	15	**41**	26	**41**	15	**124**	56
19	0	1	1	0	1	0	1	0	1	1	0	1	2
20	0	1	1	0	0	11	14	9	17	6	5	26	36
21	0	1	0	1	1	6	4	3	4	5	4	14	12
22	0	1	0	1	0	34	35	18	37	34	34	86	106
23	0	1	0	0	1	2	0	1	4	3	0	6	4
24	0	1	0	0	0	12	24	4	25	9	22	25	71
25	0	0	1	1	1	8	2	3	1	6	2	17	5
26	0	0	1	1	0	20	7	14	14	12	15	46	36
27	0	0	1	0	1	2	0	1	6	2	1	5	7
28	0	0	1	0	0	7	21	13	21	7	12	27	54
29	0	0	0	1	1	4	3	3	7	5	4	12	14
30	0	0	0	1	0	28	31	14	31	19	35	61	97
31	0	0	0	0	1	2	7	1	3	0	3	3	13
32	0	0	0	0	0	7	**69**	6	**92**	12	**62**	25	**223**

Bibliography

[1]   Richard Grotjahn et al. "North American extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends". In: *Climate Dynamics* 46.3-4 (2016), pp. 1151–1184.

[2]   Michael A Fosberg and Mark J Schroeder. "Marine air penetration in central California". In: *Journal of Applied Meteorology* 5.5 (1966), pp. 573–589.

[3]   Robert G Read. *Marine Air Penetration of the Monteray Bay Coastal Strip and Salinas Valley, California*. Vol. 71. 2. Moss Landing Marine Laboratories and Department of Meteorology, San Jose State College, 1971.

[4]   Lars E Olsson, William P Elliott, and Sheng-I Hsu. "Marine air penetration in western Oregon: An observational study". In: *Monthly Weather Review* 101.4 (1973), pp. 356–362.

[5]   Shih-Ang Hsu. *Coastal meteorology*. Elsevier, 2013.

[6]   STK Miller et al. "Sea breeze: Structure, forecasting, and impacts". In: *Reviews of geophysics* 41.3 (2003).

[7]   MJ Schroeder et al. "Marine Air Invasion of Pcific Coast: A Problem Analysis". In: *Bulletin of the American Meteorological Society* 48.11 (1967), pp. 802–+.

[8]   Annemarie Lynne Bellinder. *The California Delta Breeze*. University of California, Davis, 2012.

[9] Shane D Mayor. "Observations of seven atmospheric density current fronts in Dixon, California". In: *Monthly Weather Review* 139.5 (2011), pp. 1338–1351.

[10] Clifford F Mass, Mark D Albright, and Daniel J Brees. "The onshore surge of marine air into the Pacific Northwest: A coastal region of complex terrain". In: *Monthly weather review* 114.12 (1986), pp. 2602–2627.

[11] TD Davis et al. "Guessing Mother Nature's Next Move". In: *Public Utilities Fortnightly* 143.8 (2005), p. 48.

[12] Scott Beaver and Ahmet Palazoglu. "Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality". In: *Journal of Applied Meteorology and Climatology* 45.12 (2006), pp. 1710–1726.

[13] John E Simpson. *Sea breeze and local winds.* Cambridge University Press, 1994.

[14] Laura L Zaremba and John J Carroll. "Summer wind flow regimes over the Sacramento Valley". In: *Journal of Applied Meteorology* 38.10 (1999), pp. 1463–1473.

[15] Bereket Lebassi et al. "Observed 1970-2005 cooling of summer daytime temperatures in coastal California". In: *Journal of Climate* 22.13 (2009), pp. 3558–3573.

[16] Richard L Snyder. "California Irrigation Management Information System". In: *American Journal of Potato Research* 61.4 (1984), pp. 229–234.

[17] National Climate Data Center. *Quality controlled local climatological data.* 2011.

[18] S Saha et al. "NCEP Climate Forecast System Reanalysis (CFSR) 6-hourly Products, January 1979 to December 2010". In: *Research Data, The National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo, USA* (2010).

[19] Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. "An overview of CMIP5 and the experiment design". In: *Bulletin of the American Meteorological Society* 93.4 (2012), p. 485.

[20] Xingying Huang et al. "An evaluation of the variable resolution-CESM for modeling California's climate". In: *Journal of Advances in Modeling Earth Systems* 8 (2016), pp. 345–369. DOI: `10.1002/2015MS000559`.

[21] Alan M Rhoades et al. "Characterizing Sierra Nevada Snowpack Using Variable-Resolution CESM". In: *Journal of Applied Meteorology and Climatology* 55.1 (2016), pp. 173–196. DOI: `10.1175/JAMC-D-15-0156.1`.

[22] Mark A Snyder et al. "Future climate change and upwelling in the California Current". In: *Geophysical Research Letters* 30.15 (2003).

Chapter 3

The Future of Wind Energy in California:
Future Projections with the Variable-Resolution CESM

3.1 Introduction

Renewable energy installations, particularly wind and solar, have been rapidly deployed in recent years in an effort to displace existing fossil fuel-based energy sources [1]. Within the U.S., California was the first state to undertake development of large wind farms starting in the early 1980s. In terms of absolute capacity, California's wind-generated electricity has roughly doubled during the past five years, to meet 6.81% of the state's total system power (as of 2016). Research has also indicated that, globally, a moderate wind energy deployment plan in which wind displaces coal (i.e., 14% wind-derived electricity generation by 2050) would help delay by 1-6 years crossing the 2°C warming threshold, often considered a lower threshold for dangerous climate change [2]. The growing adoption of wind power

emphasizes that wind is a proven, reliable, and cost-effective source of low-emission power that can grow at scale. However, wind power is also dependent on sufficiently high wind speeds, which can significantly vary by location and time period. Several past studies have demonstrated historical decreases in near-surface wind speeds over many regions of the Northern Hemisphere, including the United States [3, 4]. Consequently, an understanding of present and future wind climatology is very important when determining where investments in the construction of new wind farms should be made. In particular, given that the lifespan of wind farms is typically around 20-25 years, climate change over the coming decades has the potential to significantly affect the wind farm productivity [5].

Like many other renewable energy technologies, wind energy is influenced by climate change through changes in global energy balance and resulting shfits in atmospheric circulation patterns [6]. The few studies that have examined the impact of climate change on wind resources over California using global and/or regional climate models [7] have been largely inconclusive. These prior studies have shown sensitivity to model setup, including choice of physics scheme, downscaling method, and number of models used [8, 9, 10, 11, 12, 13]. Furthermore, the spatial variability of wind energy resources and its sensitivity to model settings emphasizes the benefit of higher resolution models and multiple model inter-comparisons [7].

In order to better understand how climate change will impact wind energy resources in California, this study has utilized a state-of-the-art global climate modeling system with support for regional refinement, the Variable-Resolution Community Earth System Model (VR-CESM). The goal of this study is twofold: First, to validate, analyze, and understand the biases in the historical hub-height wind field as produced by VR-CESM, and second to use VR-CESM to understand how climatological trends will impact wind power. Seasonal synoptic-scale patterns were investigated as part of this work to better understand how shifts in large-scale systems can impact local-scale changes in wind energy. For this study we have divided

California into two primary sub-domains: Northern California (NC) sub-domain, which includes Shiloh and Altamont Pass sites, and Southern California (SC) sub-domain, which includes Alta, Tehachapi, San Gorgonio, and Ocotillo sites. These five wind farm locations constitutes a selection of both wind farm sites currently at service, and wind project sites are slated for new development. Note the Tehachapi wind farm (35°06'08" 118°16'58"W) is very close to the Alta Wind Energy Center (35°1'16"N 118°19'14"W), so only the Alta site was used for assessing the wind field in that area. Figure 3.1 depicts this region, along with the six wind farms and three atmospheric sounding locations.

Previous studies [14, 15] utilizing VR-CESM have demonstrated its competitiveness in studying high-resolution regional climatology when compared to other regional climate models, especially when non-local processes have significant influence on the local climatology. VR-CESM has demonstrated a much better representation of climatology within regions of complex topography, due to the relatively fine regional resolution compared with conventional GCM simulations [16, 17, 18].

The remainder of the paper is as follows. Section 3.2 describes the VR-CESM model setup and the datasets used in this study. In section 3.3, historical wind speeds are compared across all datasets, including the available sounding observational sites and surface observations. Future projections from the mid-century VR-CESM simulation are discussed in section 3.4. Changes to the synoptic-scale climatological background fields are also analyzed and described in this section. Discussion and conclusions follow in section 3.5.

3.2 Datasets

Two model simulations, three reanalysis products, and two observational datasets are used for model validation and inter-comparison of wind speed at hub height (summarized in Table 3.1). In this section we provide an overview of these prod-

48
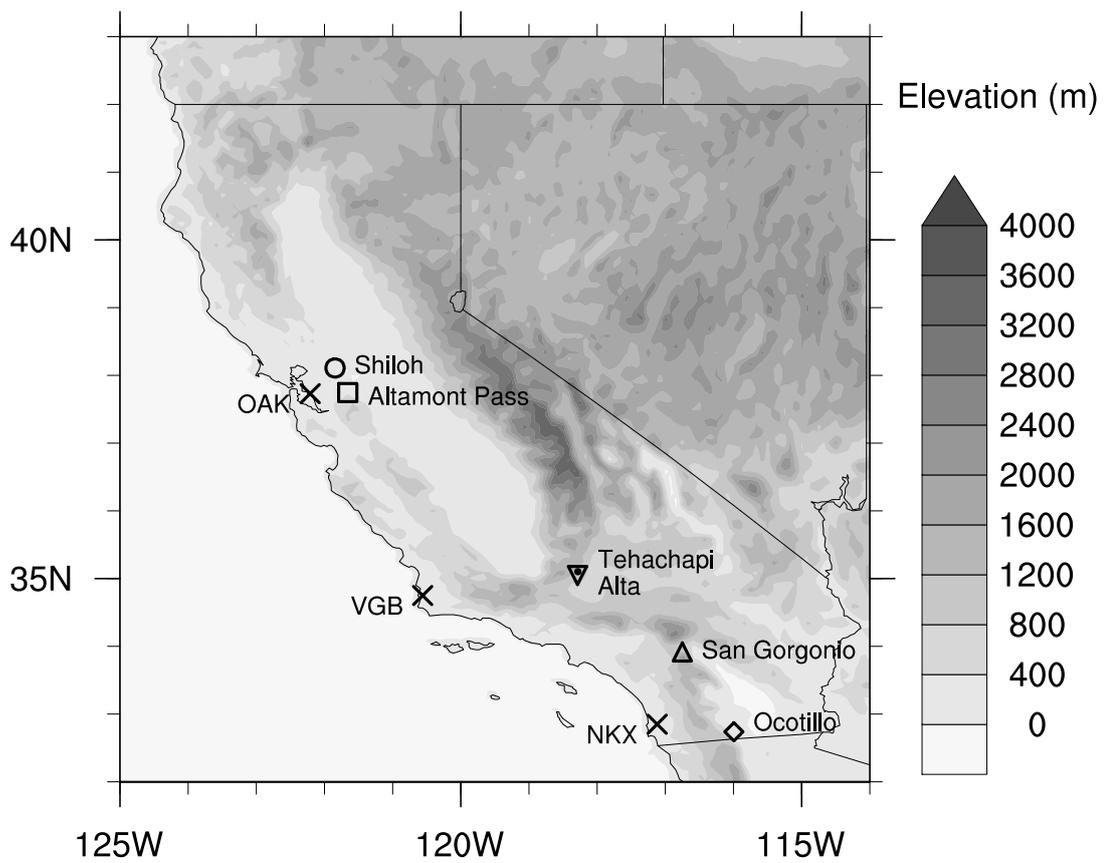
Figure 3.1: Six wind farm sites, and three sounding locations assessed in this study. Markers shown for each location are consistent as in the rest of the figures.

Table 3.1: Model, reanalysis, and observational datasets used in this study

Dataset	Spatial resolution	Temporal resolution	Time period	Source
VR-CESM	$\sim$14km(0.125°)	3-hourly	1980-2000; 2030-2050	Available upon request
Virtual Met	4km	1-hourly	1980-2000	N/A
MERRA-2	$\sim$55km(0.5°)	3-hourly	1980-2000	https://disc.gsfc.nasa.gov/datasets/M2I1NXASM$_V$5.12.4/$summary$
CFSR	$\sim$55km(0.5°)	6-hourly	1980-2000	https://rda.ucar.edu/datasets/ds093.0/
NARR	32km	3-hourly	1980-2000	https://www.esrl.noaa.gov/psd/data/gridded/data.narr.html
ISD	Point stations	1-hourly	1980-2000	https://www.ncdc.noaa.gov/isd
Soundings	OAK, VGB	12-hourly	1980-2000	http://weather.uwyo.edu/upperair/sounding.html
	NKX	12-hourly	1990-2000	

ucts. More details on these datasets can also be found in a separate data article [19].

3.2.1 Summary of datasets

VR-CESM (Global climate model product) CESM version 1.5.5, a fully coupled atmospheric, land, ocean, and sea ice model, was utilized for this study. All simulations used the F-component set (FAMPIC5), which prescribes sea-surface temperatures and sea ice but dynamically evolves the atmosphere and land surface component models. The atmospheric component mode is the Community Atmosphere Model, version 5.3 (CAM5) [20] with the spectral-element (SE) dynamical core [21] in its variable-resolution (VR) configuration [22]. The VR model grid used for this study, depicted in Figure 4.2, was generated for use in CAM and CLM with the open-source software package SQuadGen [23, 24]. On this grid the finest horizontal resolution is 0.125°($\sim$14km), with a quasi-uniform 1°mesh over the remainder of the globe. VR-CESM model used in this study is a hydrostatic model, therefore the vertical motion is not fully included. However, using the hydrostatic solver in our case allows the model to run much faster comparing to a non-hydrostatic solver, while the results still give a reasonable representation at 14km. Two simulations were conducted using this grid structure: First, the historical run covers the period from October 1st, 1979 to December 31st, 2000, with first three months discarded as the spin-up period, for a total of 21-years outputted every three hourly. This historical time period was chosen to provide an adequate sampling of the inter-annual variability, and to coincide with the time period from the
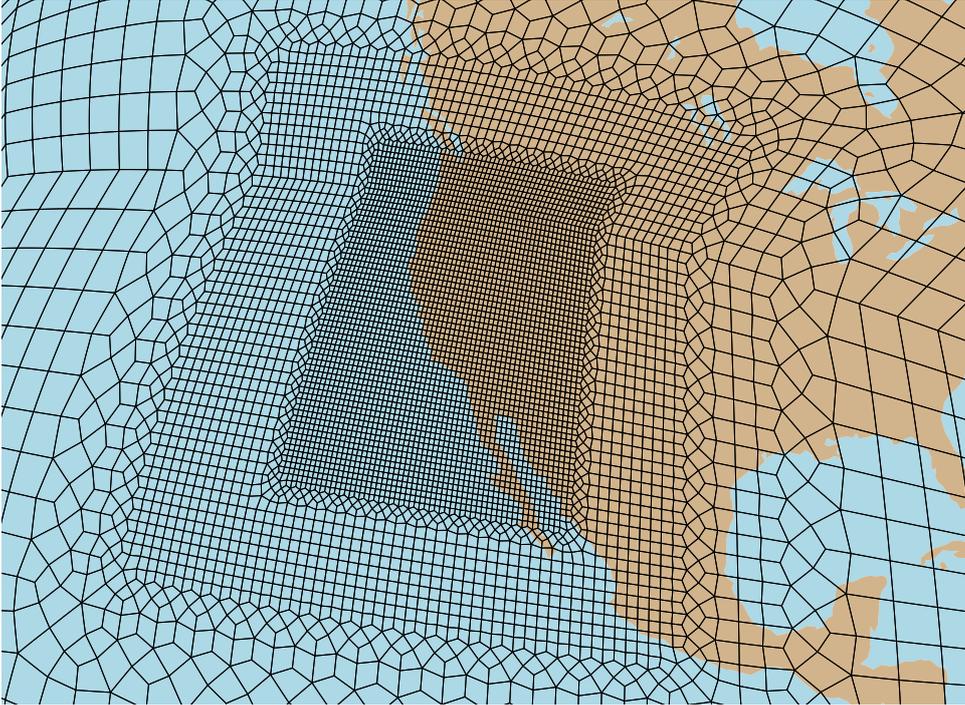
Figure 3.2: The VR-CESM grid used in this study, constructed by first successively refining a cubed-sphere grid with a 1°(111km) quasi-uniform resolution to a resolution of 0.125°($\sim$14km) over the western USA.

rest of the modeling and reanalysis datasets. For projections of future wind energy change, our mid-century simulation ran with the "business as usual" Representative Concentration Pathway 8.5 (RCP8.5) [25] from October 1st, 2029 to December 31st, 2050, again discarding the first three months for a total of 21-years. The future time period was chosen to emphasize the mid-century focus of this study and avoid divergence in the predicted impacts among different RCPs. Greenhouse gas (GHG) and aerosol forcings are prescribed based on historical or RCP8.5 concentrations for each simulation. More details on VR-CESM can be found in Rhoades et al. [14]. The complete namelist for CAM used in this study can be found at $https://github.com/meinaw/VR-CESM/blob/master/user\_nl\_cam$.

DNV GL Virtual Met (Dynamically-downscaled regional model product)
The Det Norske Veritas Germanischer Lloyd (DNV GL) Virtual Met product is derived from a hybrid dynamical-statistical downscaling system based upon the

Weather Research and Forecasting (WRF) model and an analog-based ensemble downscaling method (denoted as Virtual Met in table 3.1). The predictor consists of a coarse resolution WRF simulation that is run for the entire simulation period. To provide training data for the statistical model, a nested version of the same model is run at high resolution. The period over which the coarse and high-resolution runs overlap is called the training period, while the remaining portion is termed downscaling period. To downscale the predictor data outside of the training period, the best matching coarse estimates (termed "analogs") over the training period are found. The downscaled solution is then constructed from the set of high-resolution values that correspond to the best matching coarse analogs. This method is based upon Delle Monache et al.[26, 27].

The WRF simulation uses telescoping computational grids with one-way interaction. For this study the respective horizontal grid increments are 20 km and 4 km, with the 4 km grid centered over California. The initial and lateral boundary conditions are specified using MERRA-2, which is widely accepted in the wind energy community as a high-quality (albeit coarse resolution) wind product. The coarse model was run for the entire 01 Jan 1980 - 31 Dec 2015 period, and generated output every hourly, whereas the nested 4 km grid was run only during the last year of the full simulation (01 Jan 2015 to 31 Dec 2015). The high resolution downscaled dataset is then reconstructed for the entire 36-year period using the 4 km resolution training data and the 20 km simulation (both from the same WRF model configuration). The result is an hourly time series at each 4 km grid point for January 1st 1980 to December 31st 2015. Wind speed and direction at hub heights, including 50m, 80m, 140m, are predicted and output. This study purely utilized the 80m wind speed output, as the 80m hub height is typical for most present-day industrial wind turbines. DNV GL served solely as a data provider for this project, and is not responsible for any results from this study.

MERRA-2 (Reanalysis product) The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) is a reanalysis product for the satellite era using the Goddard Earth Observing System Data Assimilation System Version 5 (GEOS-50) produced by Global Modeling and Assimilation Office (GMAO) at NASA [28]. MERRA-2 integrates several improvements over the first version MERRA product, as described in Rienecker et al. [29]. For the fields used in this study, the spatial resolution is ∼55km with 3-hourly output frequency from 1980 to present. Vertical interpolation of MERRA-2 data, as described in the following section 3.3, was performed to calculate hub height wind speed at 80m. Variables used in vertical interpolation were extracted from two subsets: 3-hourly instantaneous pressure level assimilation [30], and hourly instantaneous single level assimilation [31] (extracted at 3-hourly frequency).

CFSR (Reanalysis product) The Climate Forecast System Reanalysis (CFSR) from NCEP (National Centers for Environmental Prediction) is a global, coupled reanalysis that spans from 1979 to present, with ∼55km spatial resolution and 6-hourly temporal resolution of relevant wind fields [32]. Notably, this temporal resolution is the lowest out of the five dataset used. The analysis subset was used in this study, and vertical interpolation was performed at 6-hourly frequency.

NARR (Reanalysis product) The North American Regional Reanalysis (NARR), another NCEP reanalysis product, features a slightly higher spatial resolution of ∼32km. It is a dynamically-downscaled data product with spatial coverage over North America, with 3-hourly temporal resolution from 1979 through present [33]. Hub height wind speeds from NARR were also calculated at this frequency.

ISD (In-situ observations) The Integrated Surface Database (ISD) from NOAA's National Centers for Environmental Information (NCEI) were used for assessment of hourly 10m wind speed from model and reanalysis. The ISD observational sta-

tions are distributed globally, with the highest concentration of stations found in North America. Stations across California that provide full year data were selected. As not all stations had continuous temporal coverage between 1980 to 2000, each year was examined separately so as to maximize the number of available stations. To compare 10m wind speeds from model and reanalysis datasets to ISD, the nearest grid point values to each of the ISD stations was used. Coastal stations were neglected in the analysis of 10m winds, due to coastal biases that tend to occur in near-surface coarse-resolution reanalysis. These biases tend to emerge because similarity theory is typically employed to extract 10m wind speeds, which produces distinctly different results over the ocean and land surface.

Upper air soundings (In-situ observations) Upper air soundings (vertical wind profiles) from all the available locations across California are incorporated into the comparison (University of Wyoming, Department of Atmospheric Science `http://weather.uwyo.edu/upperair/sounding.html`). The three sounding locations used in this study are OAK at Oakland airport (station number 72493), VBG at Vandenberg Air Force Base (72393), and NKX at San Diego (72293) (see Figure 3.1). The time period used in this study from the first two stations spans 1980 to 2000. NKX only has data available starting from September 1989, so only the full years 1990-2000 were assessed. Soundings were collected every 12 hourly at 00Z and 12Z, and logarithmic vertical interpolation was performed to calculate 80m wind at each sounding location. However, this logarithmic interpolation from sparsely sampled profile data could introduce uncertainties into the calculation.

3.2.2 Representation of topography

Local topography is particularly important in representing the wind field, particularly in the regions of significant topographic variability that tend to be well-suited for wind power generation. Consequently, the importance of model resolution can-
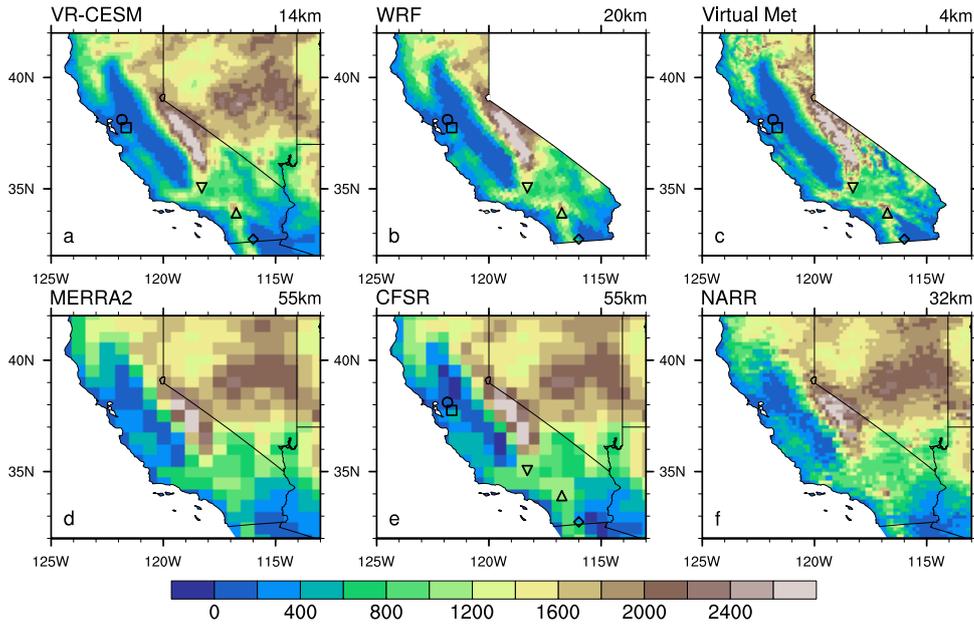
Figure 3.3: Topographical representation of California and surrounding regions from model (top row) and reanalysis (bottom row) datasets.

not be understated. Topographic profiles from each of the models and reanalysis datasets are plotted in Figure 4.1. As can be seen here, DNV GL WRF model ran at 20km resolution (b), which captures the dynamical wind field at this resolution, and then statistically downscaled to 4km resolution (c). VR-CESM uses a relatively smooth topography by comparison, due to its slightly lower spatial resolution of 14km (a). MERRA2, CFSR, and NARR (d-f) all have much more poorly refined topography, with a poor representation of the coastal ranges that are important for shaping the wind field. Note that these differences also imply that each model has a different altitude for the wind farms and sounding stations used in this study.

3.2.3 Wind speed interpolation method

The wind speed at each wind farm location 3.1 was determined using nearest grid point values to each wind farm site. To obtain hub-height wind vectors, vertical interpolation was performed on 3-hourly VR-CESM, 3-hourly MERRA-2, 6-hourly CFSR, and 3-hourly NARR products from 1980 to 2000. As mentioned above,

hub-height wind output is available directly from the DNV GL Virtual Met data product. Vertical interpolation of VR-CESM data uses the 3D wind field on hybrid surfaces and 10m altitude wind speed, which is computed from similarity theory. For VR-CESM data, the interpolation procedure is as follows: (1) the CAM5 hybrid coordinates are first converted to pressure coordinates within the column being analyzed, (2) the height of each pressure surface above ground level (AGL) is computed by subtracting the surface geopotential height from the geopotential height of the model level, (3) two model levels that bound the desired interpolation altitude are selected or, if the interpolation altitude is below the lowest model level, the lowest model level and 10m wind speed field are used, and (4) logarithmic interpolation is applied to obtain the wind speed at the desired interpolation altitude. The interpolation was done by fitting a log equation with the two levels bounding the altitude to be calculated, then with the log profile, interpolating the wind at desired altitude [34]. Vertically interpolated wind speed from MERRA-2, CFSR, NARR, and sounding observations all followed a similar procedure, and were calculated at three hub heights (50m, 80m, and 140m). Figure 3.4 to 3.7 show the interpolated hub-height wind speed at 50m and 140m, respectively, at northern and southern California. For wind speed at 80m, and detailed wind speed analysis, please refer to the next section. Different atmospheric stability will also lead to different vertical wind profile based on the wind speed power law. This will have impact on the vertical wind speed interpolation, though we didn't take this into consideration here.

Wind turbines can contribute to energy via the electric power system. This contribution is the total amount of usable energy supplied by the turbine per year [35]. The capacity factor (CF) is often defined as actual power output divided by the max amount of wind power that can be generated through the system. This wind speed and CF relationship is not continuous since there is a discontinuous minimum and maximum wind speed required to begin and cease wind power production (the
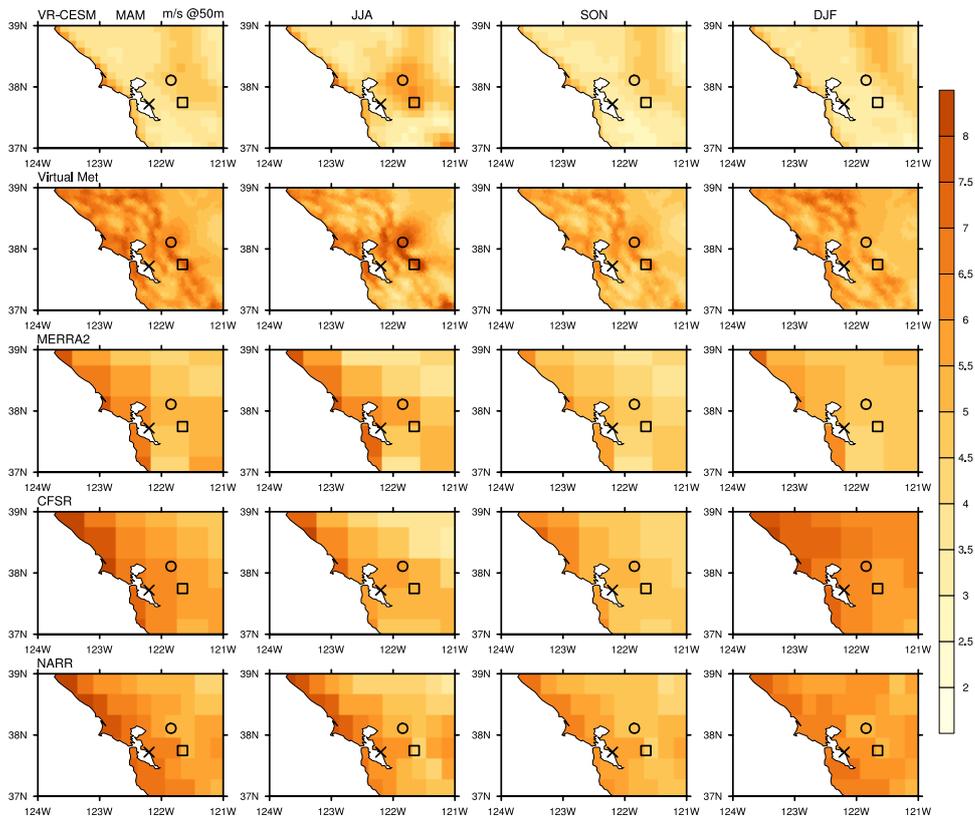
Figure 3.4: Seasonal average of interpolated 50m wind speed from each datasets for historical time period 1980-2000 in northern California domain.
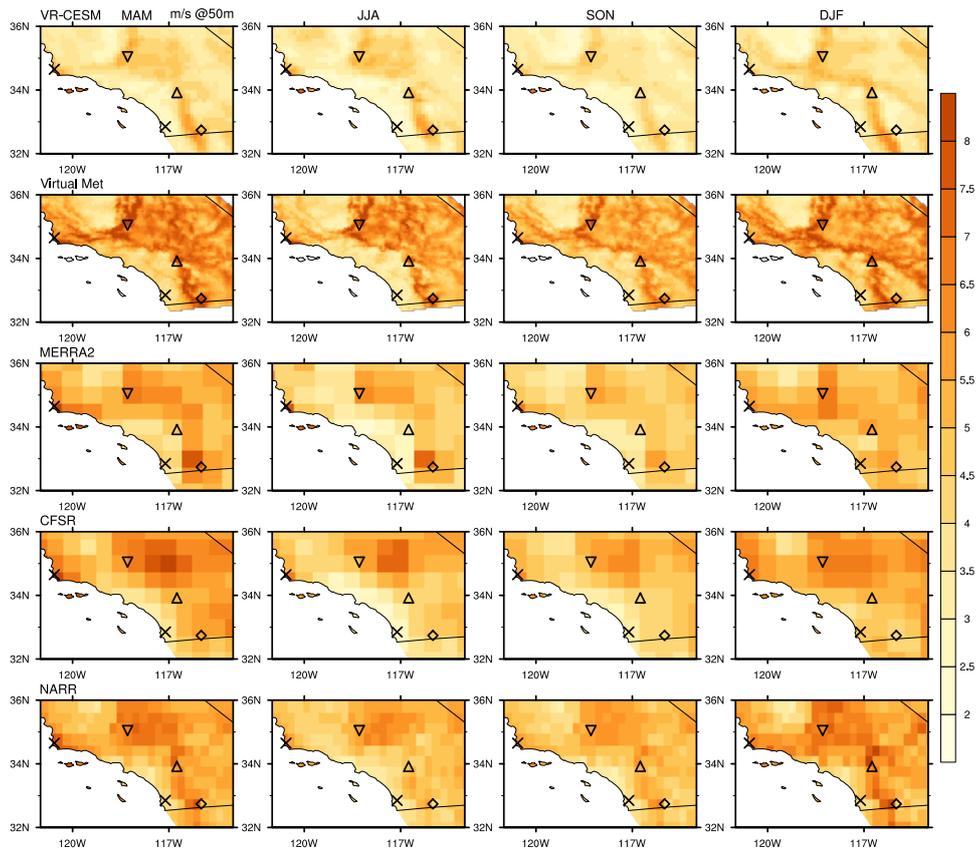
Figure 3.5: Seasonal average of interpolated 50m wind speed from each datasets for historical time period 1980-2000 in southern California domain.
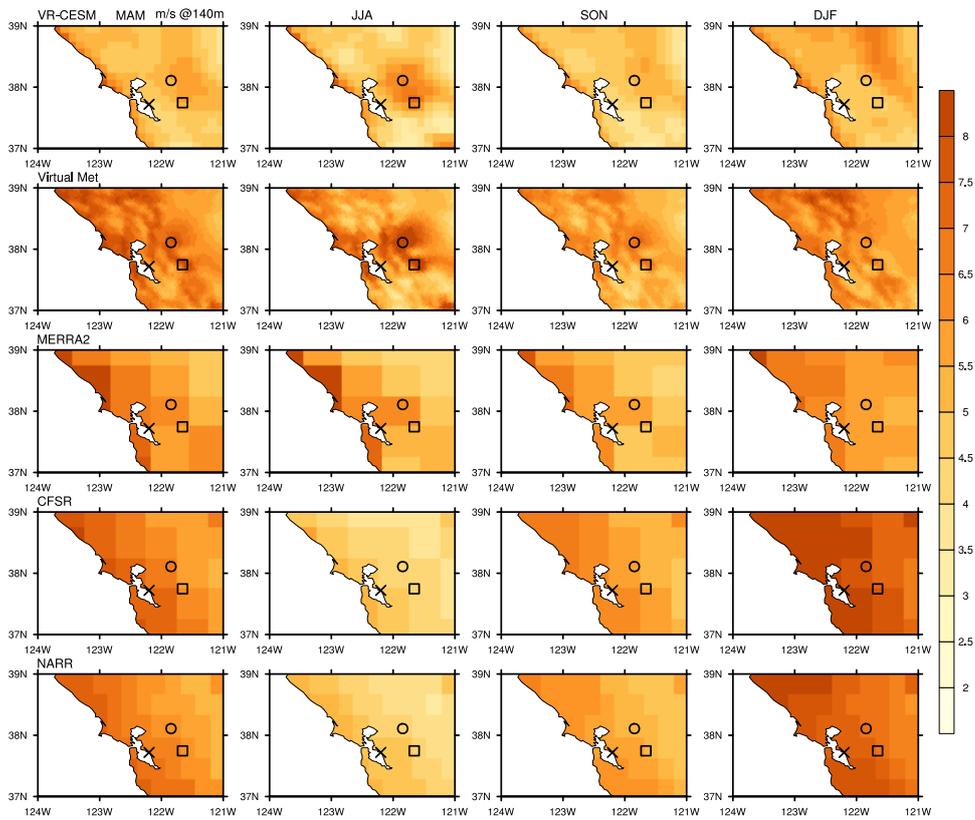
Figure 3.6: Seasonal average of interpolated 140m wind speed from each datasets for historical time period 1980-2000 in northern California domain.
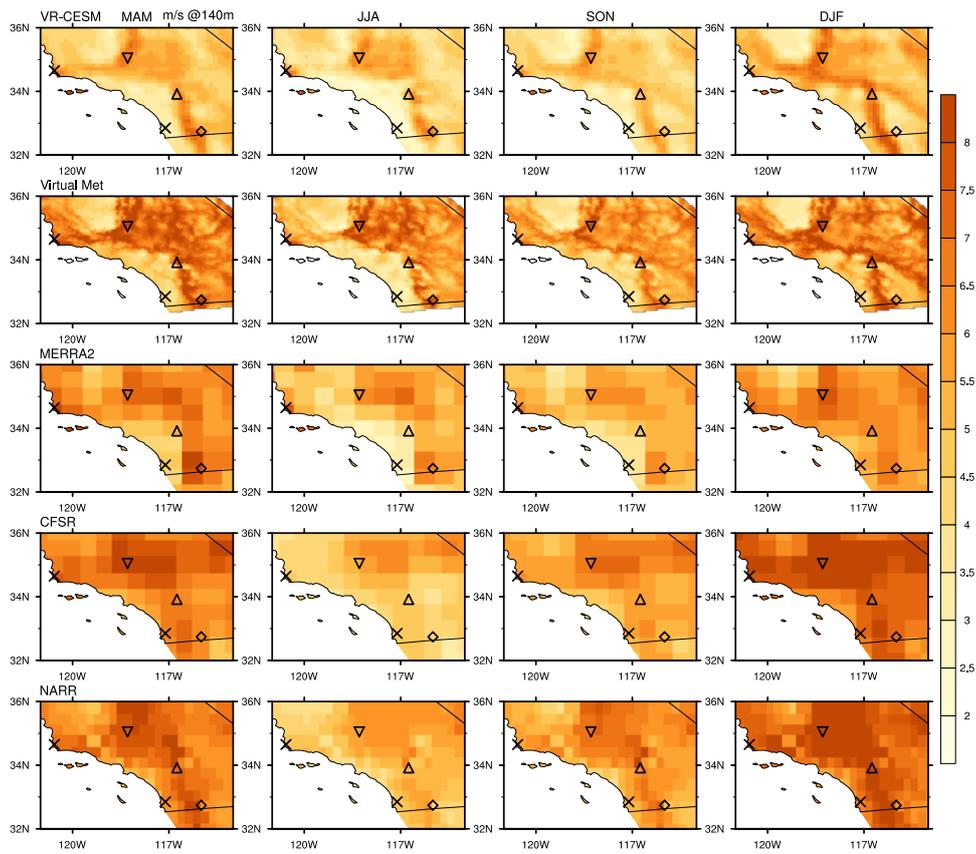
Figure 3.7: Seasonal average of interpolated 140m wind speed from each datasets for historical time period 1980-2000 in southern California domain.

Table 3.2: Power curves for wind farms across California. Each value corresponds to a 1m/s wind speed bin increment starting from 0m/s.

Wind farm	Power curve
San Gorgonio	IECclass1 = [0, 0, 0, 0.0043, 0.0323, 0.0771, 0.1426, 0.2329, 0.3528, 0.5024, 0.6732, 0.8287, 0.9264, 0.9774, 0.9946, 0.999, 0.9999, 1, 1, 1, 1, 1, 1, 1, 1]
Altamont Pass, Ocotillo	IECclass2 = [0, 0, 0, 0.0052, 0.0423, 0.1031, 0.1909, 0.3127, 0.4731, 0.6693, 0.8554, 0.9641, 0.9942, 0.9994, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
Alta, Shiloh	IECclass3 = [0, 0, 0, 0.0054, 0.053, 0.1351, 0.2508, 0.4033, 0.5952, 0.7849, 0.9178, 0.9796, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]

latter to avoid damage to the wind turbine under extreme wind conditions), and this is represented with different power curves associated with each of the wind farm sites. The calculated CF at each wind farm site is based on different characteristic power curves at that site, and do not include electrical losses during the power generation process. The normalized power curves at each wind farm sites, with each value corresponding to a 1m/s wind speed bin increment starting from 0m/s, are listed in Table 3.2. To calculate the CF, wind speed is multiplied with the corresponding power curve value from the corresponding wind speed bin, and then times 100 to convert the percentage values.

3.3 Model comparison and wind resources characterization

3.3.1 Methodology

The wind speed at each wind farm location was determined using nearest grid point values to each wind farm site. To obtain 80m wind vectors for this study, vertical interpolation was performed on 3-hourly VR-CESM, 3-hourly MERRA-2, 6-hourly CFSR, and 3-hourly NARR products from 1980 to 2000. As mentioned above, 80m wind output is available directly from the DNV GL Virtual Met data product used in this study, so values are extracted directly from the output from 1980 to 2000. Vertical interpolation of VR-CESM data uses the 3D wind field on

hybrid surfaces and 10m-altitude wind speed, which is computed from similarity theory. For VR-CESM data, the interpolation procedure is as follows: (1) the CAM5 hybrid coordinates are first converted to pressure coordinates within the column being analyzed, (2) the height of each pressure surface above ground level (AGL) is computed by subtracting the surface geopotential height from the geopotential height of the model level, (3) two model levels that bound the desired interpolation altitude are selected or, if the interpolation altitude is below the lowest model level, the lowest model level and 10m wind speed field are used, and (4) logarithmic interpolation is applied to obtain the wind speed at the desired interpolation altitude. Specifically, the interpolation was performed by fitting a log equation with the two levels bounding the altitude to be calculated, then interpolating the wind at desired altitude [34]. Vertically interpolated wind speeds from MERRA-2, CFSR, NARR, and sounding observations were all obtained a similar procedure, and were calculated at three hub heights (50m, 80m, and 140m). Further, wind speed at 80m was logarithmically interpolated for all three sounding profile locations, and compared with interpolated 80m wind speed at each sounding locations from all five model/reanalysis datasets.

The wind field enters into the maximum potential wind power $P(W)$ via the expression $P = \frac{1}{2}\rho A U^3$, where ρ is air density (kg/m^3), A is the cross section area of the turbine rotor (m^2), and U is wind speed at hub height (m/s). Given the cubic relationship between wind speed and wind energy potential, even a small change in wind speeds can lead to a substantial change to wind energy production. The energy contribution of wind turbines to the electric power system is then computed as the total amount of usable energy supplied by the turbine per year [35]. The capacity factor (CF) is often thus defined as actual power output divided by the maximum wind power output that can be generated through the system. This wind speed and CF relationship is not continuous, since there is a discontinuous minimum and maximum wind speed required to begin and cease wind power production (the
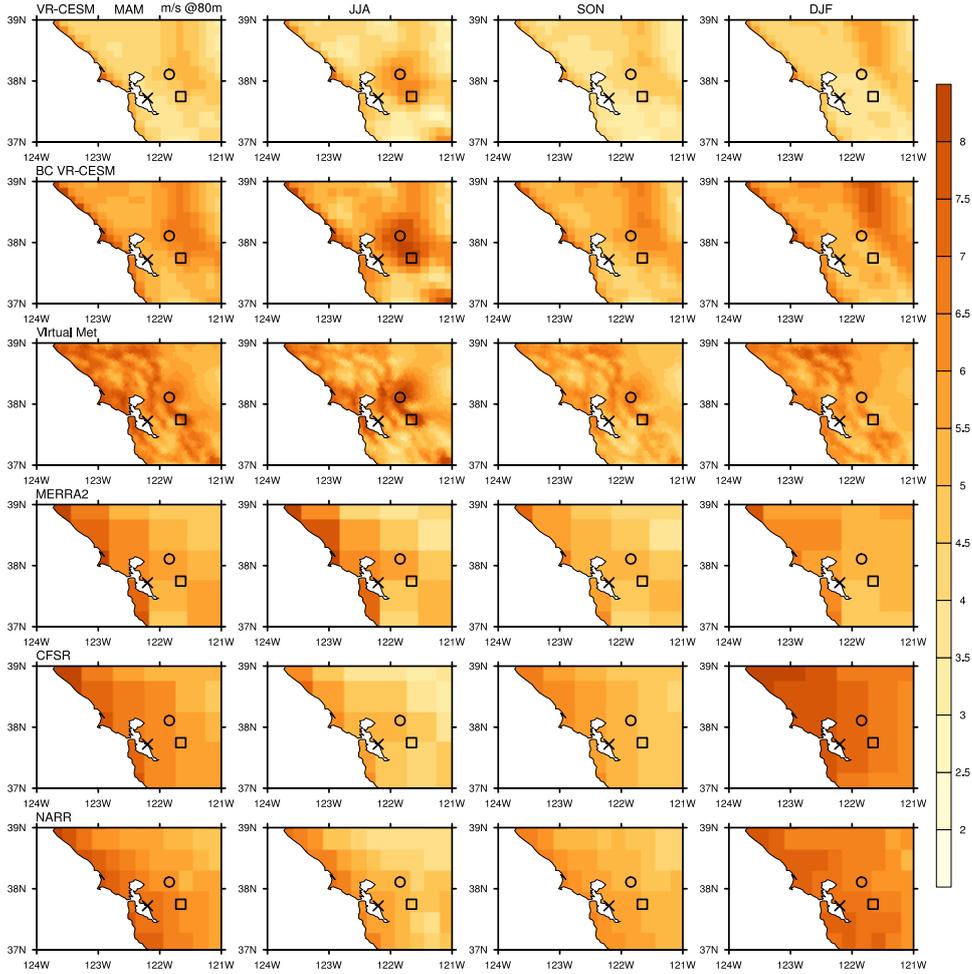
Figure 3.8: Seasonal average of interpolated 80m wind speed from each datasets for historical time period 1980-2000 in northern California domain.

latter to avoid damage to the wind turbine under extreme wind conditions), and this is represented with different power curves associated with each of the wind farm sites. For this study, the calculated CF at each wind farm site is based on different characteristic power curves specific to each site (see the data in brief accompanying this paper), and do not include electrical losses during the power generation process.

3.3.2 80m wind speed climatology

The remainder of the text focuses on the NC domain and SC domains. Figure 3.8 depicts the 80m wind speed fields (vertically interpolated values except for Virtual Met) from each of the datasets in the NC domain. Wind fields shown are

seasonal mean values for all March-April-May (MAM), June-July-August (JJA), September-October-November (SON), December-January-February (DJF) seasons between historical time period 1980-2000. Because of Virtual Met's high spatial resolution (4km), more topographic features are apparent in the wind field, whereas the MERRA-2, CFSR, and NARR wind fields are blockier due to their relatively coarse resolution. Comparing VR-CESM to Virtual Met, the overall pattern is very similar, although VR-CESM exhibits lower mean wind speeds overall. This difference will be further assessed as part of the wind farm site comparisons in section 3.3.2. Figure 3.9 depicts mean winds for the SC domain. Again, the patterns remain similar between VR-CESM and Virtual Met, but with a reduced wind magnitude.

Quantitatively, the VR-CESM and Virtual Met product outputs are highly correlated (~ 0.69), which suggests that the underlying physical mechanisms responsible for determining wind speed are similar between these two products. The slow wind speeds in VR-CESM are likely a consequence of excessive diffusion in the lowest model levels, and further hypothesized to be connected to a boundary layer parameterization in CESM that is not tuned for the high resolutions employed in this study (we anticipate addressing this issue in future work). To better match the wind speeds predicted in the virtual met product, we applied a multiplier of 1.30 to the VR-CESM results to produce a bias-corrected VR-CESM (BC VR-CESM) prediction. The value of this multiplier is determined by the mean wind speed difference between VR-CESM and the Virtual Met. As can be seen in Figure 3.8 and Figure 3.9, the wind magnitudes are more comparable to Virtual Met, the latter still produces more spatial variation as compared to BC VR-CESM. This difference in spatial variation can be attributed to the representation of topography in the model – as apparent in Figure 4.1, Virtual Met captures the rough rolling terrain of this region, whereas VR-CESM represents the coastal ranges as a single "mound." As a result, Virtual Met captures a detailed pattern of wind speed variation, whereas VR-CESM only captures a large-scale downslope winds off of this range. In Figure
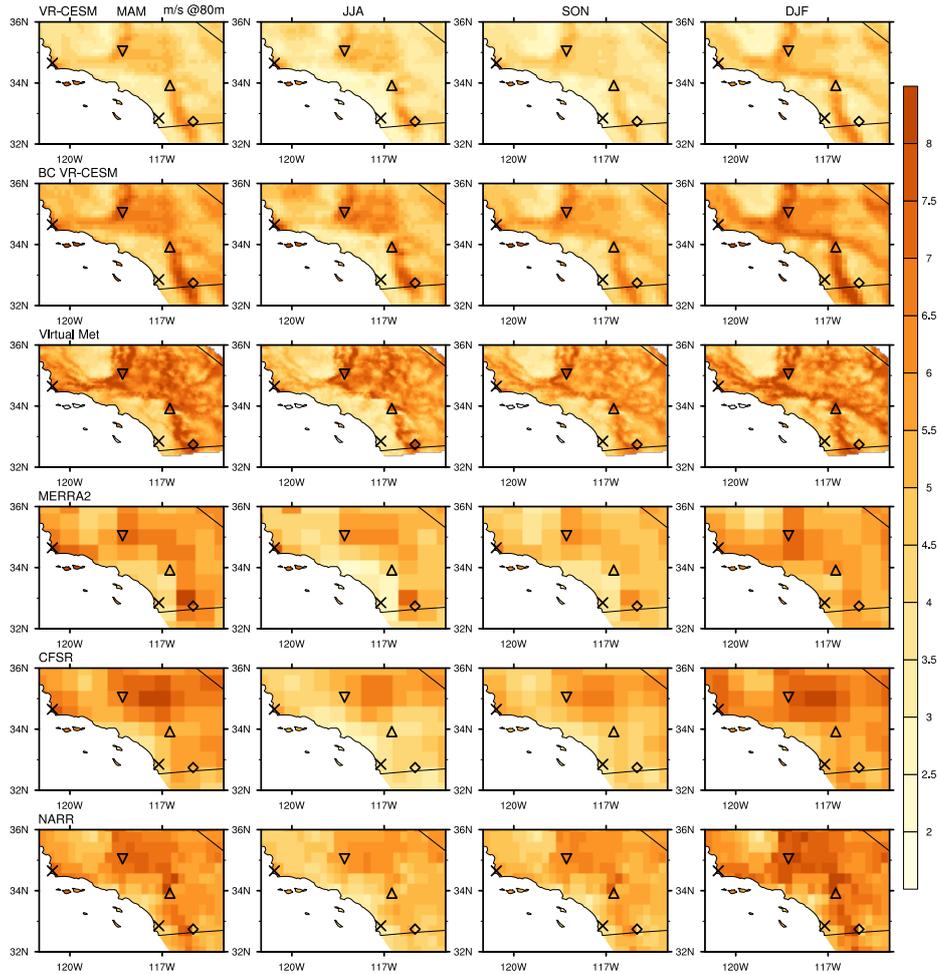
Figure 3.9: Seasonal average of interpolated 80m wind speed from each datasets for historical time period 1980-2000 in southern California domain.

3.12 we observe that the histograms of wind speed from BC VR-CESM are closer to WRF 20km, although the futher downscaled Virtual Met results exhibit much higher frequencies over the highest wind speed bins at all locations except San Gorgonio. For wind speed fields at the other two analyzed hub heights (50m and 140m), please refer to the data in brief. In general, higher altitudes tend to produce larger wind speeds, although the patterns remain quite similar.

Monthly climatological mean wind speeds at each wind farm site are depicted in Figure 3.10. As observed in Figures 3.8 and 3.9, Virtual Met tends to produce the highest overall wind speeds. Whereas VR-CESM exhibits a lower wind speed magnitude than Virtual Met, both datasets produce similar spatial patterns that

Table 3.3: Averaged Pearson pattern correlations between each pair of datasets as obtained from the seasonal mean 80m wind speed from 1980 to 2000.

VR-CESM				
Virtual Met	0.69			
MERRA-2	0.61	0.58		
CFSR	0.45	0.53	0.58	
NARR	0.45	0.52	0.51	0.77
Model name	VR-CESM	Virtual Met	MERRA-2	CFSR	NARR

are distinctly different than the other three reanalysis datasets. In particular, the coarser resolution reanalysis data tends to exhibit a weak seasonal cycle. Computing the correlation across monthly mean wind speeds between each dataset with Virtual Met, VR-CESM has the highest correlation (on average $\sim$0.87 over all five wind farm sites), followed by MERRA-2 ($\sim$0.55), and CFSR ($\sim$0.37). NARR ($\sim$0.17) exhibited the weakest correlation. To further quantify the spatial correlations between datasets, the centered Pearson pattern correlation (Table 3.3) was calculated for seasonal mean 80m wind speeds from all the datasets, with the domains masked to only include California, matching the domain from Virtual Met. As observed in Table 3.3, VR-CESM produces the highest pattern correlation ($\sim$0.69) with Virtual Met, followed by MERRA-2 ($\sim$0.58). Therefore, both temporal and spatial correlation comparisons suggest VR-CESM produces the most similar wind speed climatology (both temporally and spatially) to Virtual Met, followed by MERRA-2. NARR produces the lowest correlation in space and time – in fact, discrepancies in the spatial structure of NARR's wind climatology are likely indicative of potentially significant errors in its representation of wind speeds [David Pierce, personal communication]. At several sites (particularly San Gorgonio), the seasonality from the three reanalysis datasets is distinctly different from both VR-CESM and Virtual Met. This is again likely a direct result of the resolution discrepancy between the models and reanalysis – for instance, the San Gorgonio wind farm site sits along a narrow pass ($\sim$3km) between mountains, which is not resolved in the reanalysis datasets.

66
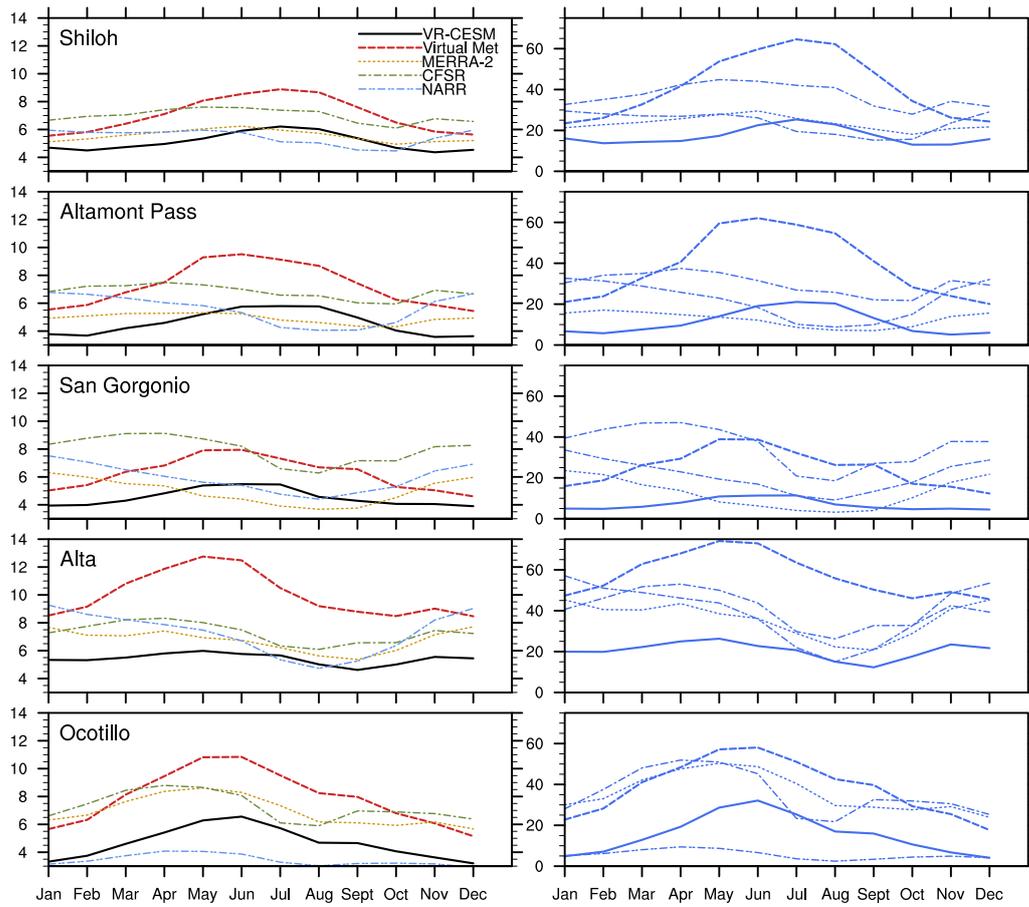
Figure 3.10: Monthly mean 80m wind speed (color-coded lines on left) and mean CF (blue lines on right) at each wind farm site from all datasets during historical time period 1980-2000.

The frequencies of instantaneous 80m wind speeds from each dataset in Figure 3.11. Wind speeds in almost all locations appear to follow a Weibull distribution, as is typical for wind speeds where the velocity in each coordinate direction is normally distributed [5, 36]. However, the Virtual Met data diverges from the Weibull distribution at several locations, which may be indicative of physical processes that are uniquely captured by this product at high spatial resolution. Specifically, Virtual Met produces higher wind speeds at a higher frequency than other datasets in many cases, leading to a greater spread among the wind speed bins. Frequencies from BC VR-CESM are closer to Virtual Met compared to VR-CESM due to increased wind speed, although there remains a mismatch in the shape of the distribution. The behavior of the Virtual Met data might be related to the analogous method used on WRF model, and further investigation is needed to analyze its impact on the hub-height wind speed. Unfortunately, the authors are presently unaware of any publicly available hub-height wind speed datasets that would allow direct validation of these results against observations.

3.3.3   10m wind speed climatology

The performance of VR-CESM is now assessed against the 10m hourly Integrated Surface Database (ISD). Although ISD incorporates hundreds of observation stations across California, many of these stations do not provide consistent observations over the relevant historical time period (1980-2000). In order to maximize the number of available stations each year, and ensure sure each year has complete data coverage, validation metrics (Table 3.4) were calculated separately for each year between 1980 and 2000. Also, to avoid issues with near-surface coastal flow, only inland observation stations were selected for comparison. After imposing these restrictions, an average of 100 inland stations were used from each year.

Table 3.4 provides the averaged seasonal bias and root-mean-square error (RMSE) at 10m altitude from our five datasets against ISD observations from 1980 to 2000.
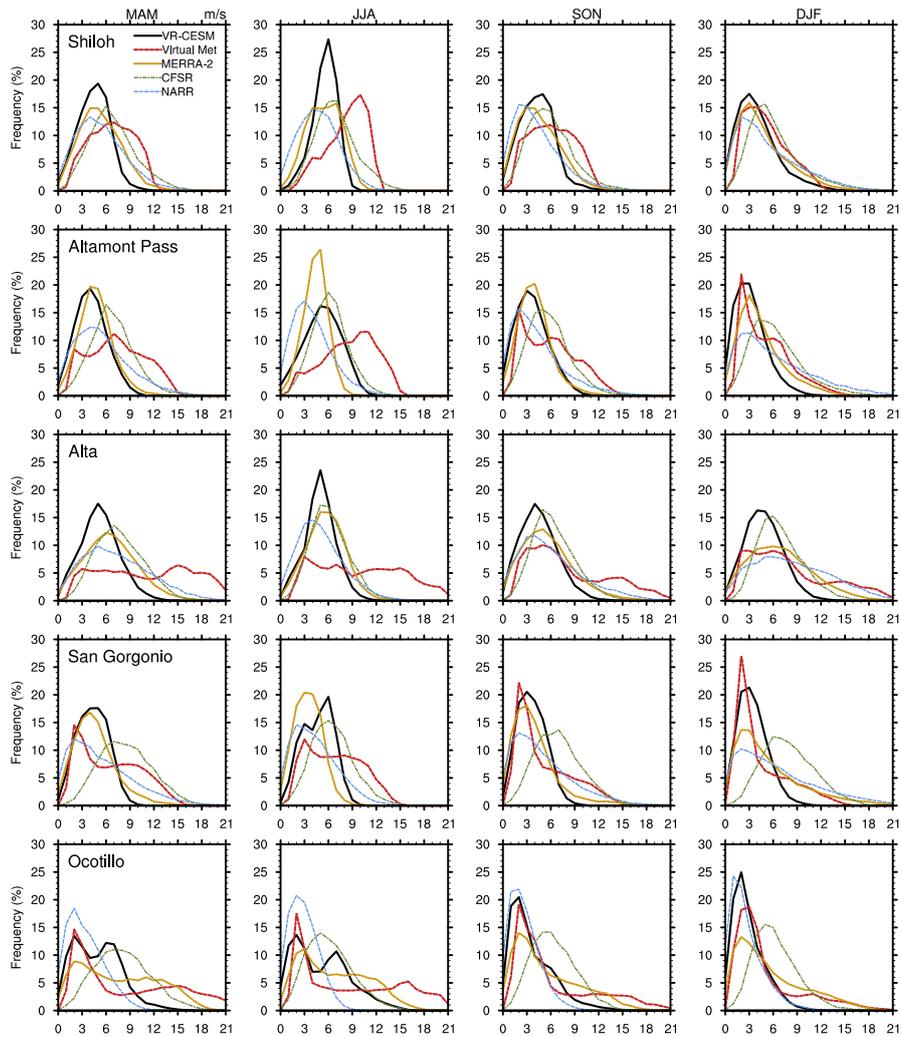
Figure 3.11: Frequencies for instantaneous 80m wind speeds from all datasets at each wind farm location for the historical time period 1980-2000 by season. The bin width is 1m/s and covers the range from 0m/s to 21m/s.
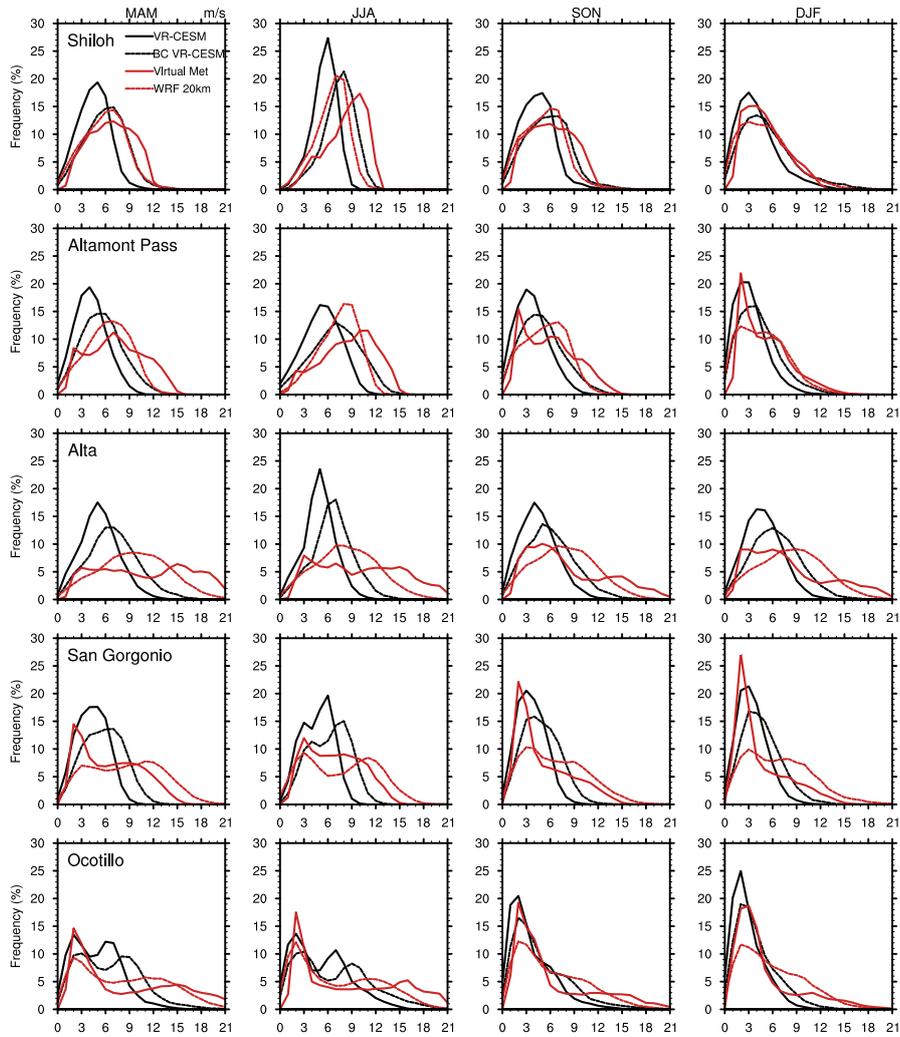
Figure 3.12: Frequencies for instantaneous 80m wind speed from bias-corrected VR-CESM (BC VR-CESM) and 20km WRF compared to VR-CESM and Virtual Met at each wind farm location for the historical time period 1980-2000. The bin width is 1m/s and covers the range from 0m/s to 21m/s.

Here, a negative (positive) bias indicates that the wind speed is lower (higher) than observations. As observed previously, VR-CESM tends to produce lower wind speeds than observation, whereas the Virtual Met produces overall higher wind speeds. MERRA-2 and Virtual Met exhibit similar differences, as MERRA-2 provides the boundary conditions for the WRF model; nonetheless, Virtual Met does produce higher mean wind speeds than MERRA-2, likely due to a positive wind bias that appears fairly consistently in the WRF model [37, 38]. Note that the values listed for Virtual Met in table 3.4 are dependent upon the specific WRF model configuration and initialization used in Virtual Met. Further investigation is required to understand biases in the WRF model. CFSR exhibits lower wind speeds for most of the year except the DJF season, whereas NARR produces higher wind speeds in all seasons. For MAM and JJA seasons, Virtual Met is very close to observations – namely, it shows a relatively small bias, whereas VR-CESM has strong negative biases in both seasons. In SON and DJF seasons, VR-CESM is closer to observations compared to Virtual Met, particularly during the DJF season (and closer to observations than all other datasets). As VR-CESM also obtains 10m wind using the lowest model level wind plus similarity theory, the biases in 10m wind have the potential to be conveyed to higher elevations during the calculation. So this 10m wind speed comparison with observation also provides us some insight into the possible biases for wind speed at 80m.

3.3.4 Comparison with soundings

Hub-height wind data in California is often produced through private investment and hence a closely guarded trade secret confidential to project owners. Consequently, for validation of our modeled hub-height wind speed data against observation, our assessment is limited to a select number of vertical sounding sites across California (listed in Section 3.2) for comparison of higher level wind speeds, and all of the three soundings are located near the coast (denoted by "X" in Figure 3.1)

71

Table 3.4: Bias and RMSE for 10m wind speed from all five datasets to inland ISD observational stations from 1980 to 2000. Bias and RMSE both have units of m/s.

Model name	Stats	MAM	JJA	SON	DJF	Annual average bias
VR-CESM	Bias	-0.80	-0.52	-0.32	-0.16	-0.45
	RMSE	1.23	1.06	0.88	0.85	
BC VR-CESM	Bias	-0.04	0.21	0.28	0.52	0.24
	RMSE	1.10	1.10	1.00	1.17	
Virtual Met	Bias	0.02	-0.03	0.40	0.56	0.24
	RMSE	0.97	1.02	0.94	1.02	
MERRA-2	Bias	-0.14	-0.13	0.23	0.52	0.12
	RMSE	0.87	0.92	0.78	0.91	
CFSR	Bias	-0.48	-0.50	-0.14	0.23	-0.22
	RMSE	1.11	1.11	0.83	0.88	
NARR	Bias	0.11	0.16	0.52	0.67	0.37
	RMSE	1.34	1.17	1.25	1.49	

with complex local topographies. The coarse resolution of these models requires them to average inland and offshore wind speeds, leading to skewed results. Also, the sounding observations are only measured twice daily. Both these factors take into account when doing interpolation to calculate 80m wind from sounding observations, and from model and reanalysis dataset at these sounding locations. In comparison, the three lower resolutions reanalysis datasets all project higher than observation wind speeds. At the OAK site, wind speed projected from VR-CESM is the closest (bias = 0.95m/s) to observations in terms of wind magnitude , though Virtual Met captures monthly variation better (correlation = 0.62). However, at VBG and NKX, none of the model datasets could be said to capture the values and seasonal variation particularly well, even though VR-CESM and Virtual Met are the closest among all.

Figure 3.13 shows the 80m wind speed from each model and reanalysis dataset, as well as the interpolated 80m wind speed from sounding observations. At the OAK site, wind speed projected from VR-CESM is the closest (bias = 0.95m/s) to observations in terms of wind magnitude , though DNV GL captures monthly variation better (correlation = 0.62).

The three reanalysis datasets all project higher wind speeds in comparison to

observations; the coarse resolution of these models requires them to average inland and offshore wind speeds, leading to skewed results. At VBG, all datasets projected higher wind speeds than observed. This might due to the location of VBG sounding sitting very close to coast ($\sim$5km), so that bias could be incorporated from the higher wind speeds from the adjacent ocean. Nonetheless, wind magnitudes from VR-CESM and DNV GL are very close to each other. Similarly at the NKX site, VR-CESM and DNV GL are again the closest among all, but at this sounding site, none of the datasets captured the monthly variation from observation particularly well. This might be because the NKX sounding site locates in the mountainous region where local topography is complicated, leading to the monthly variations not being captured. Overall, at these coastal locations, VR-CESM produced the closest results to observations at OAK site, whereas none of the model datasets could be said to capture the values and seasonal variation particularly well at the other two sites, which might due to biases coming from the higher wind speed over the ocean, or the complexity of the local topography.

3.3.5   Comparison between VR-CESM and Virtual Met

To further investigate the difference in wind field between VR-CESM and Virtual Met, the Virtual Met product was regridded to the VR-CESM grid and the difference taken. Figure 3.14 shows 1980-2000 seasonly mean wind speed difference from Virtual Met minus VR-CESM, with positive values indicates Virtual Met has higher wind speeds than VR-CESM. The difference is not spatially uniform – in particular, when comparing Figure 3.14 alongside Figure 4.1, Virtual Met projected higher wind speed over higher altitudes, and lower wind speed at lower altitudes. The five wind farm sites all sit at relatively high topography regions, and consequently Virtual Met projects higher values at all five locations from Figure 3.14, consistent with Figure 3.10.
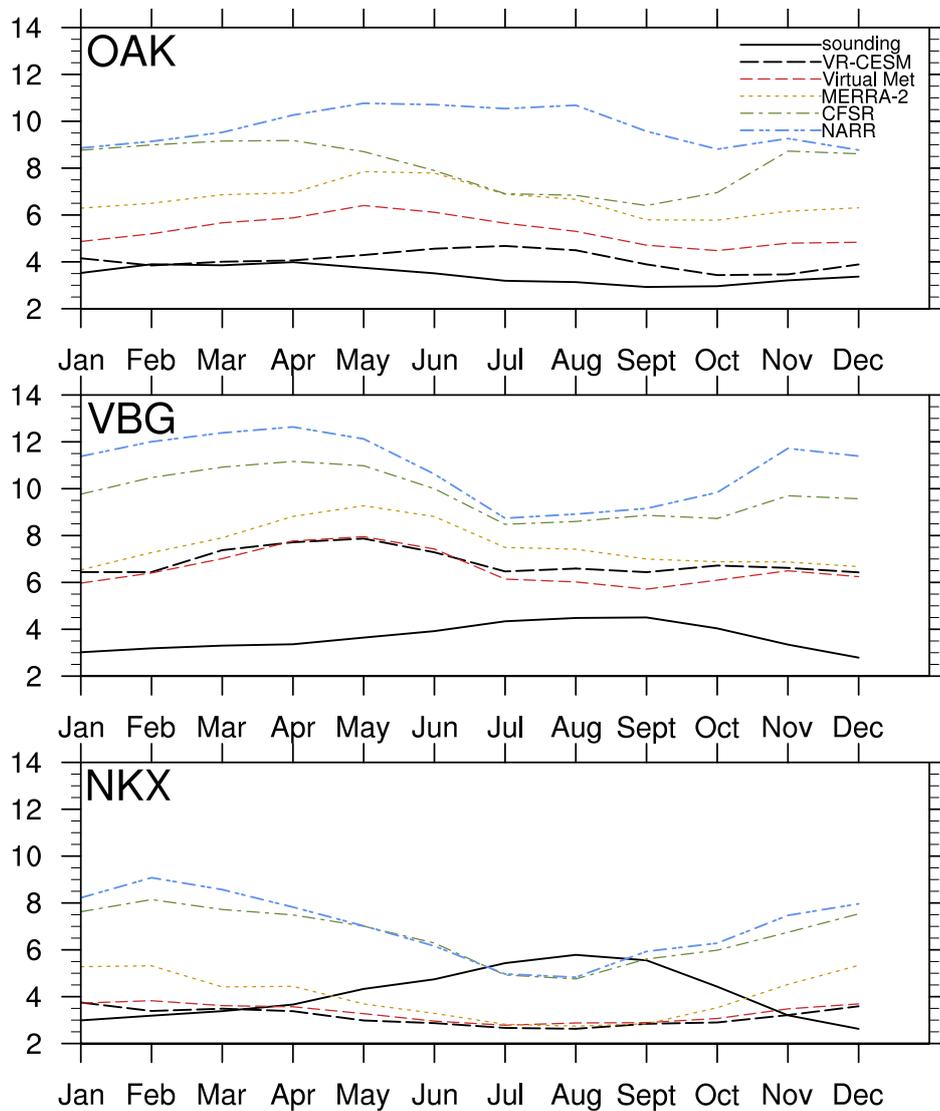
Figure 3.13: Comparison of monthly mean interpolated 80m wind speed from sounding observations with interpolated 80m wind speed from all datasets. The available time period at OAK and VBG is 1980-2000, and 1990-2000 at NKX.
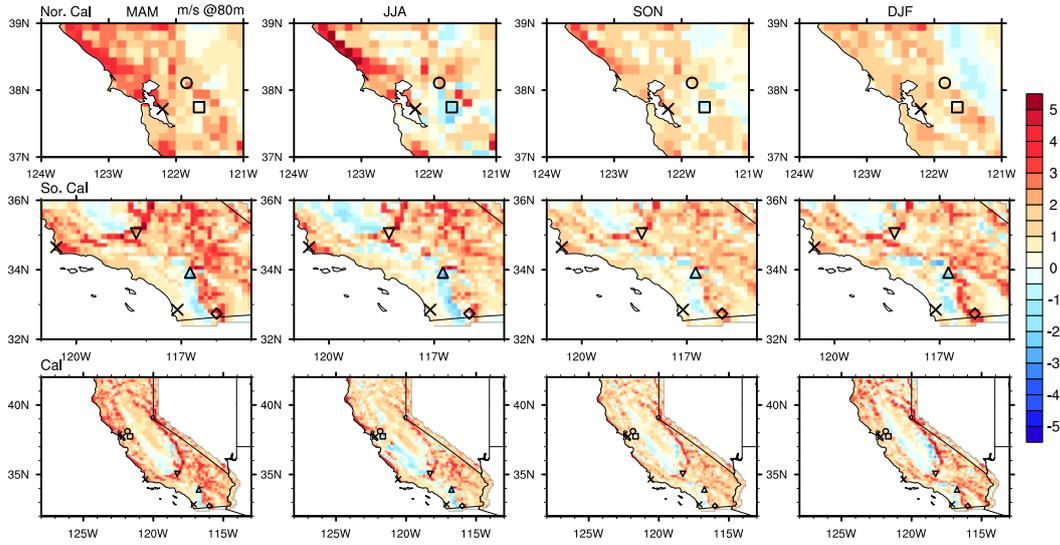
Figure 3.14: Comparison between DNV GL Virtual Met 4km and VR-CESM (Virtual Met minus VR-CESM) of interpolated 80m wind speed between 1980-2000 for northern, southern, and whole states of California domains.

3.4 Future projection

We now turn our attention to future projections of wind energy from VR-CESM mid-century simulation under the RCP8.5 "business as usual" scenario. In this section, seasonal wind power changes are first quantified from the mid-century projection, then understood in terms of the synoptic-scale meteorological shifts associated with these changes at each wind farm site.

3.4.1 Projected changes

Figure 3.15 compares the seasonal 80m wind speed change between mid-century and historical time periods (2030-2050 minus 1980-2000). These results indicate the SON, DJF, and MAM seasons exhibit decreases in wind speed for all seasons across most areas except for parts of the Central Valley (CV). However, JJA winds were projected to increase in magnitude throughout most of California, particularly through the SC domain.
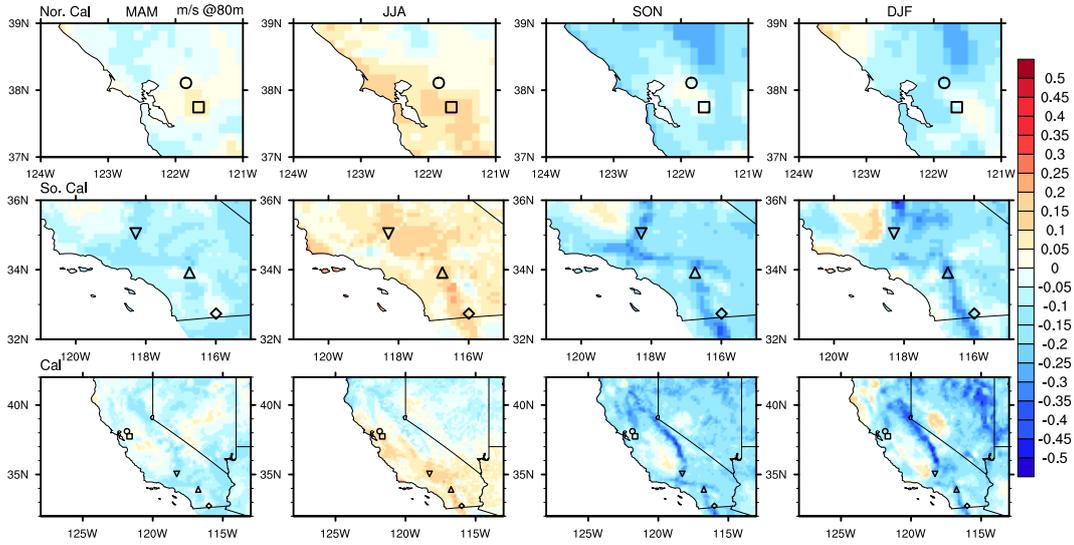
Figure 3.15: Comparison of VR-CESM seasonal averaged 80m wind speed between historical 1980-2000 and mid-century 2030-2050 (mid-century minus historical) for NC, SC, and California domains.

Comparing historical and future simulations, the seasonal pattern of CF and wind speed at each site was similar, with overall higher wind speeds during summer months, and lower wind speeds during winter months (Figure 3.16). All wind farm sites exhibit a net increase in both wind speed and CF during summer months (JJA), and decrease during winter months (DJF). Annual wind energy production decreases at all sites except Altamont Pass (Table 3.5). Consistent with Figure 3.16, JJA at all wind farm sites is associated with an increase in CF, while SON and DJF seasons lead to a decrease in CF. The SON CF decrease is consistent with results from Duffy et al. [39], which analyzed possible future trends at the Tehachapi wind farm site (denoted as · in Figure 3.1), and projected a significant decrease in wind speed throughout mid-century Fall months, and little change in Spring-Summer.

An increase in the frequency of lower wind speeds during SON and DJF seasons is indicative of the decreasing trend in wind speed through these two seasons. A decrease in the frequency of lower wind speeds during JJA, and increased frequency of higher wind speeds, is indicative of the increasing trend in wind speed during this
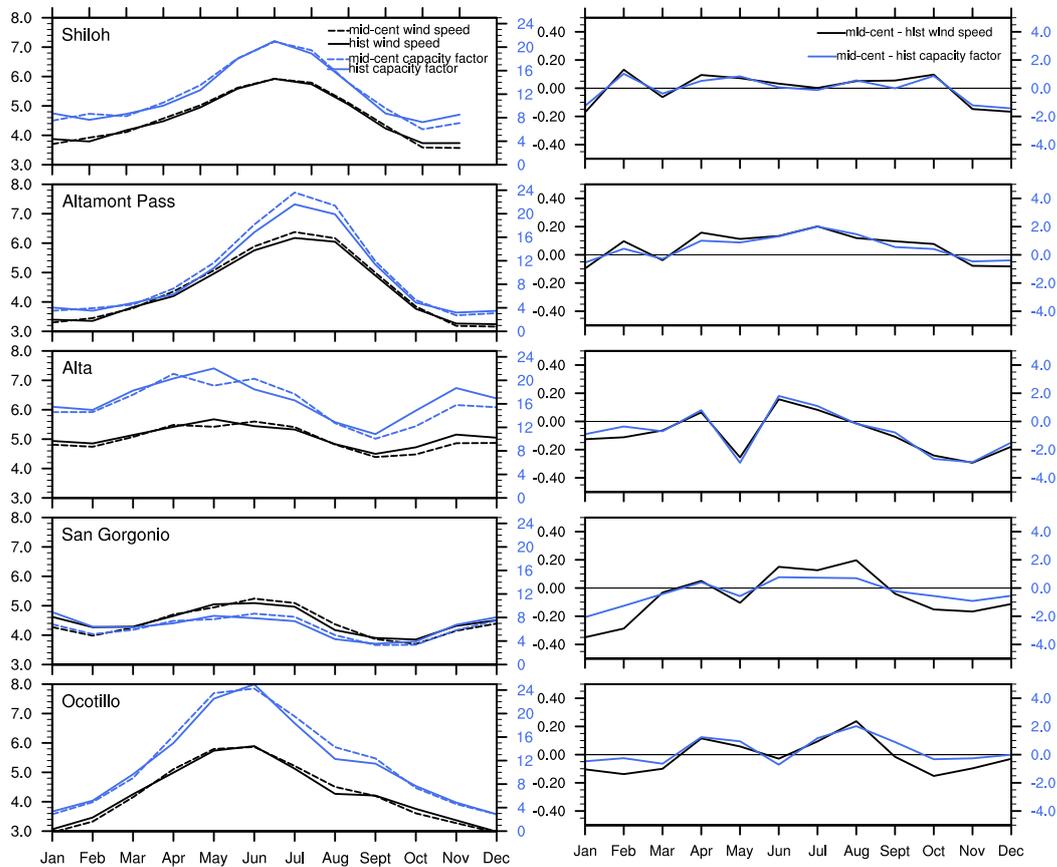
Figure 3.16: Comparison of 80m wind speed and capacity factor between historical and mid-century at each wind farm site.

Table 3.5: Seasonal and annual capacity factor changes (mid-century CF minus historical CF, divided by historical CF, and written as a percentage) at each wind farm site under mid-century 2030-2050 compared to historical 1980-2000. Boldface indicates a percent change above the 95% significance level.

wind farm	MAM	JJA	SON	DJF	annual
Shiloh	+ 0.2%	+ 0.4%	- 7.7%	- 5.8%	- 3.2%
Altamont Pass	+ 4.2%	**+ 7.5%**	- 4.5%	- 0.9%	+ 1.6%
Alta	- 5.1%	+ 8.3%	**- 13.3%**	- 7.3%	- 4.4%
San Gorgonio	- 2.4%	**+ 9.7%**	**- 10.9%**	**- 16.9%**	- 5.1%
Ocotillo	+ 1.6%	+ 5.6%	- 2.0%	- 9.0%	- 1.0%

season. Figure 3.17 depicts the differences in frequency between seasonal 80m wind speeds over the historical and mid-century periods from VR-CESM. The bold lines in Figure 3.17 correspond to the seasons with significant CF changes from Table 3.5.

3.4.2   Synoptic-scale drivers

In meteorology, synoptic-scale fields are associated with horizontal scales on the order of 1000km or more. Mean meteorological fields have been analyzed for seasons with significant CF changes to identify the synoptic-scale drivers that could influence the historical and mid-century wind climatology (JJA in Figure 3.18, SON in Figure 3.19, DJF in Figure 3.20). In particular, our analysis focuses on the 700hPa geopotential height field, which is defined as the height of 700hPa isobar surfaces above mean sea level, as well as surface pressure, surface temperature, and hub height wind speed at 80m overlaid with wind direction at the same height. The 700hPa geopotential height field was analyzed as it is reflective of the general circulation, with wind flow at this level largely geostrophic and hence following constant geopotential contours. The surface pressure field also impacts local wind speeds, and is closely associated with surface temperature changes. Synoptic-scale fields during the MAM season were not investigated, as there was no significant CF change detected over this period (see Table 3.5).

Through JJA (Figure 3.18), the 700hPa geopotential height field features an off-shore trough and geopotential height contour lines perpendicular to coast. This pattern is indicative of a typical summertime marine air penetration condition [40, 41, 42] and is driven by the off-shore trough modifying the geopotential height contour lines to be perpendicular to the coastline, allowing cool and moist marine air to penetrate inland. The location of the off-shore trough is directly responsible for driving marine air through the San Francisco Bay Delta. Relative to the historical period, the magnitude of the 700hPa geopotential height field under the mid-century increases (as a direct consequence of low-level warming). However, this increase is
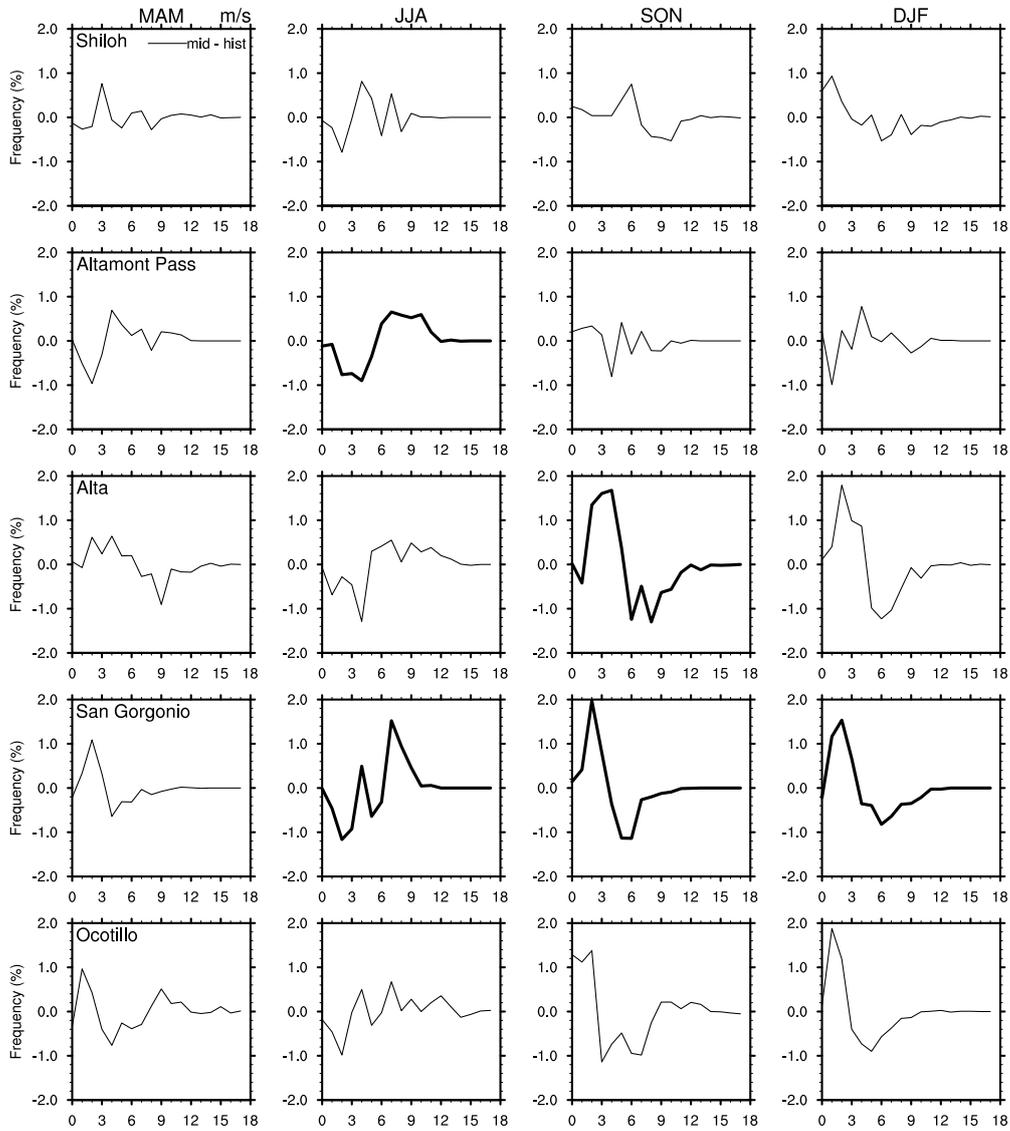
Figure 3.17: Differences in frequencies between mid-century 2030-2050 and historical 1980-2000 (mid-century minus historical) for seasonal averaged 80m wind speed from VR-CESM at each wind farm location. Bold lines correspond to significant changes from Table 3.5.

less pronounced over the Northern Pacific, which drives a weakening of the typically northerly wind pattern that traces the coastline in Northern California, and an increase in the on-shore flow pattern driven by the general circulation. This in turn leads to an increase in wind speeds through the San Francisco Delta region (Shiloh and Altamont Pass in NC domain). A shift in this synoptic-scale pattern also drives increased ventilation in the SC domain.

Surface pressure in JJA is also observed to increase more rapidly at higher altitudes – consequently the surface pressure in the Mojave desert increases more rapidly than the Central Valley, and leads to a weaker pressure gradient between the CV and Mojave. A similar observation was made by Miller and Schlegel [43] to explain a projected decrease in Santa Ana wind events in this region during the Fall season. Although this is a potential driver for wind speed decrease at Alta in SC, the impact of a reduced pressure gradient is counterbalanced by the changes to the large-scale geopotential height field, which enhances westerly wind throughout California.

Across both time periods, SON wind speeds are generally reduced in comparison to JJA, partly due to the decrease in land-sea temperature contrast, and associated reduction to marine air penetration. Comparing the 700hPa gepotential height field between historical and mid-century during SON, the entirety of the California coast is under the influence of the weakening of wind flow parallel to the coast, driven by the negative geopotential anomaly south of Alaska, and accompanied by a positive geopotential height anomaly over the continent. Through the SC domain, a weakening pressure gradient drives a decrease in the wind speed at Alta and San Gorgonio. This observation is in agreement with the observations of Duffy et al. [39], and leads to a projected 10-15% power potential decrease during Fall season in mid-century in the immediate vicinity of Tehachapi.

Through DJF (Figure 3.20), increased geopotential height over the sub-tropical western Pacific and the North American continent lead to a weaker northerly flow
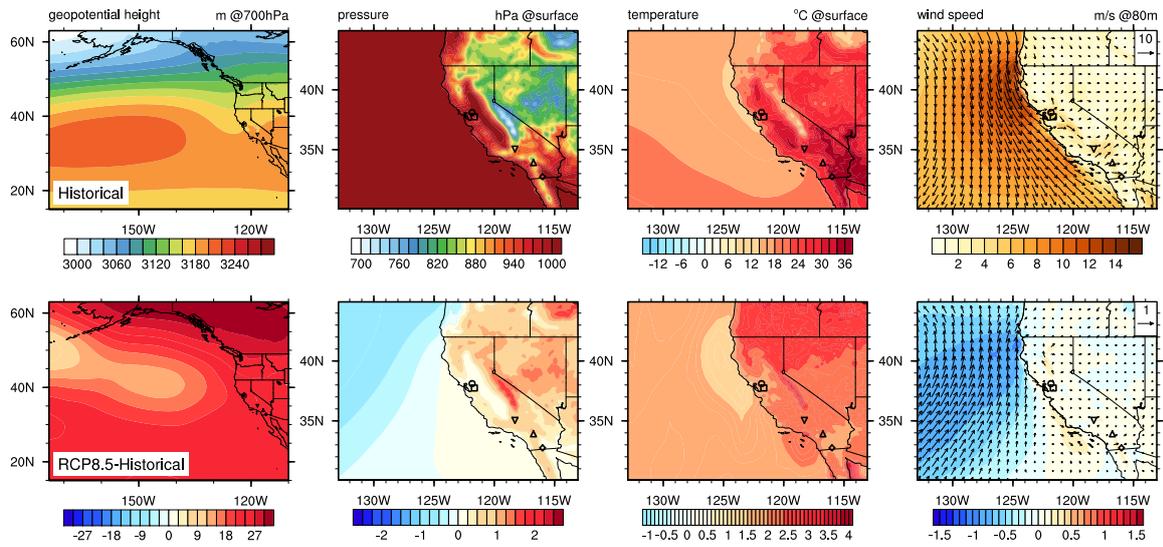
Figure 3.18: Seasonal mean 700hPa geopotential height, surface pressure, surface temperature, and 80m wind fields on historical 1980-2000 (top row), and the corresponding anomaly fields on mid-century 2030-2050 (bottom row) during JJA season. Anomaly values (bottom row) were calculated from subtracting mean historical fields (top row) from mean mid-century fields.
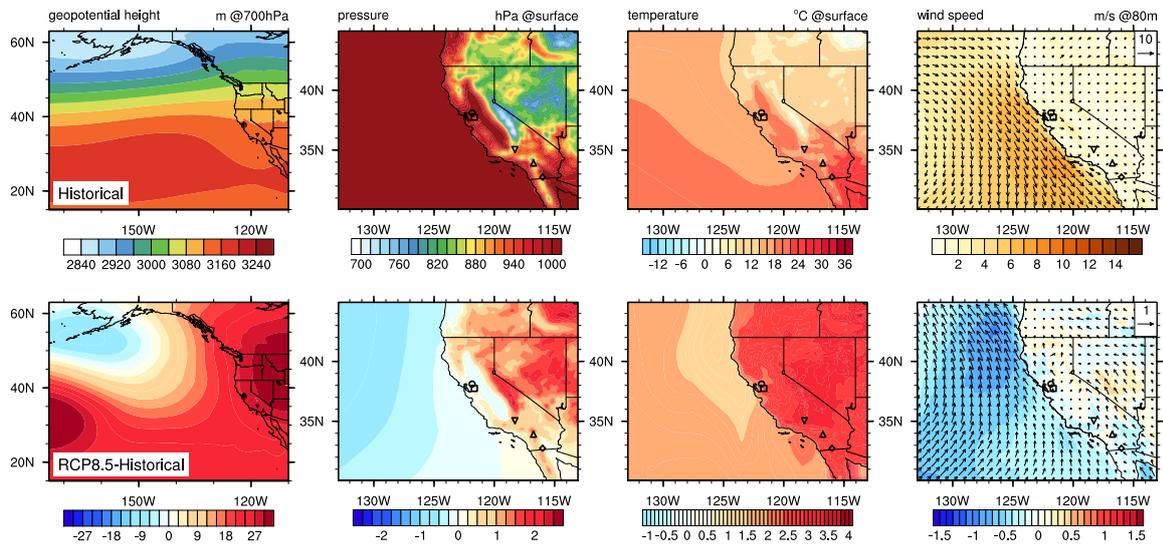
Figure 3.19: Seasonal mean 700hPa geopotential height, surface pressure, surface temperature, and 80m wind fields on historical 1980-2000 (top row), and the corresponding anomaly fields on mid-century 2030-2050 (bottom row) during SON season.
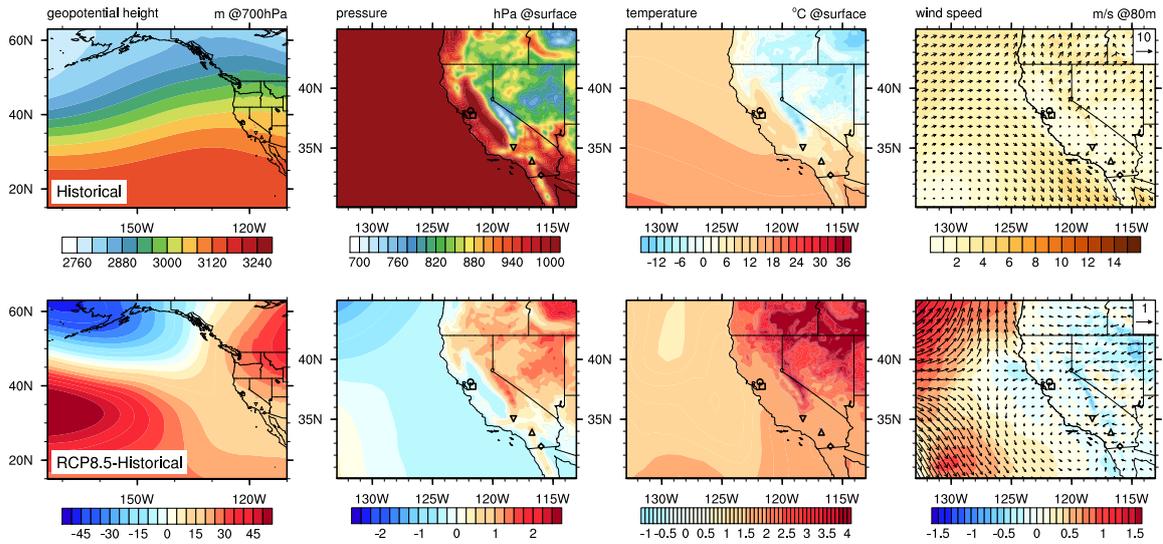
Figure 3.20: Seasonal mean 700hPa geopotential height, surface pressure, surface temperature, and 80m wind fields on historical 1980-2000 (top row), and the corresponding anomaly fields on mid-century 2030-2050 (bottom row) during DJF season.

parallel to the coast and a reduced on-shore flow. Further, with surface pressure decreases in the CV, the surface-level pressure gradient between the CV and the Mojave desert decreases, which would in turn be expected to drive lower wind speeds at the Alta wind farm site. The surface pressure gradient also decreases between the inland and the adjacent ocean near San Gorgonio wind farm site, which further enhances the wind speed decrease.

The seasonal meteorological patterns under the mid-century RCP8.5 scenario provide further evidence that future changes of wind energy in California will be influenced by both the synoptic-scale and local changes. Overall, the synoptic analysis suggests that the climate through mid-century will be conducive to higher wind speed across the whole state of California during JJA (5-10% at four of the five sites examined), and lower during SON (particularly at Alta and San Gorgonio which each exhibited a > 10% decrease) and DJF (with a 17% decrease at San Gorgonio). The changes to the surface pressure gradient between the Central Valley and the Mojave Desert appears robust across seasons and is a primary driver of wind speed decreases

in the SC domain. To ensure the synoptic-scale climatology of VR-CESM was not an outlier, synoptic-scale geopotential height fields were also examined across CMIP5 models over the same time period and similar trends were observed. Ensemble runs with VR-CESM could potentially add confidence to this study, and are a topic for future exploration once the identified biases in VR-CESM are addressed. Besides the mid-century time frame (2030-2050) that was studied in this paper, another VR-CESM simulation over the end-of-century time frame (2080-2100) was also conducted using the same model configuration. Wind speed change at each wind farm site from the end-of-century run had the same sign as the mid-century run, and relatively greater magnitude. The results from the end-of-century run adds confidence to our current analysis. However, due to the end-of-century time frame is outside the typical lifetime of a wind farm ($\sim$20-25 years), the analysis from end-of-century was not included in this paper.

3.5 Discussion and conclusions

The goal of this paper is twofold: First, to validate and assess the performance of VR-CESM as a tool for modeling near-surface wind speeds and, second, to leverage VR-CESM to assess the drivers of future wind speed change in California. The main conclusions of this paper follow.

The capacity of the VR-CESM variable-resolution global climate modeling system was assessed at correctly representing the historical character of wind field in California (1980-2000) against a high-resolution WRF statistically-downscaled wind data product, multiple reanalysis products, and publicly available observational data. Our results suggest that although VR-CESM generally exhibited a bias towards slower wind speeds inland, the monthly climatology and spatial pattern associated with the wind field was approximately consistent with observations. Although the wind climatology was greatly improved over coarse resolution reanal-

ysis products, we believe that the local model resolution (14km) is still too coarse for regions of rapid topographic variation. Nonetheless, rough agreement between simulated and observed wind fields led us to conclude that VR-CESM is correctly representing the key regional and synoptic-scale processes that are relevant for wind speed forecasts. Further work is needed to determine the source of the slow bias in near-surface wind speeds from CESM.

Second, this study aimed to project and understand hub-height wind speed changes at each wind farm site, using a VR-CESM mid-century (2030-2050) simulation under RCP8.5. To better understand the regional and synoptic-scale drivers that are responsible for these changes, our analysis targeted the meteorological patterns associated with large-scale shifts in wind character. The five major wind farm sites considered in this study spanned California. At almost all wind farm sites, significant seasonal changes were observed in the capacity factor, with an increase in summertime (JJA) resources and a decrease in fall (SON) and winter (DJF) under RCP8.5 at all five sites (Table 3.5). Synoptic-scale and localized drivers behind season wind energy change were also identified, and suggested climate change may favor synoptic patterns that lead to higher wind speed during JJA, and lower wind speed during SON and DJF.

Overall, this study improves the characterization of uncertainty around the magnitude and variability in space and time of California's wind resources in the near future, and also enhances our understanding of the physical mechanisms related to the trends in wind resource variability.

There are many climatological factors that impact on the wind energy in California, including correlations of wind speed with climate modes such as El Niño-Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), and North Atlantic Oscillation (NAO). Because of the long temporal frequency of these climate modes, there is some difficulty in disentangling how these climate modes have historically impacted wind resources. In this context, ensemble simulations with VR-

84

CESM may be valuable at modeling these connections. Future work could also address alternative statistical strategies for identifying change in wind fields: for instance, the use of a clustering method to analyze and group relevant wind patterns in California. Such a method could be used to investigate the potential historical and future trends from different wind patterns. Possible future study will also focus on analyzing the capacity of models to capture, and the climate change impact on intense and extreme winds.

Bibliography

[1] Ottmar Edenhofer et al. *Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change.* Cambridge University Press, 2011.

[2] RJ Barthelmie and SC Pryor. "Potential contribution of wind energy to climate change mitigation". In: *Nature Climate Change* 4.8 (2014), pp. 684–688.

[3] Robert Vautard et al. "Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness". In: *Nature Geoscience* 3.11 (2010), p. 756.

[4] SC Pryor, Rebecca Jane Barthelmie, and JT Schoof. "Inter-annual variability of wind indices across Europe". In: *Wind Energy* 9.1-2 (2006), pp. 27–38.

[5] SC Pryor and RJ Barthelmie. "Climate change impacts on wind energy: A review". In: *Renewable and sustainable energy reviews* 14.1 (2010), pp. 430–437.

[6] M King Hubbert. "Energy Resources of the Earth". In: *Scientific American* 224.3 (1971).

[7] DJ Rasmussen, T Holloway, and GF Nemet. "Opportunities and challenges in assessing climate change impacts on wind energy – a critical comparison of wind speed projections in California". In: *Environmental Research Letters* 6.2 (2011), p. 024008.

[8] Lejiang Yu et al. "Temporal and spatial variability of wind resources in the United States as derived from the Climate Forecast System Reanalysis". In: *Journal of Climate* 28.3 (2015), pp. 1166–1183.

[9] SC Pryor and RJ Barthelmie. "Assessing climate change impacts on the near-term stability of the wind energy resource over the United States". In: *Proceedings of the National Academy of Sciences* 108.20 (2011), pp. 8167–8171.

[10] David J Sailor, Michael Smith, and Melissa Hart. "Climate change implications for wind power resources in the Northwest United States". In: *Renewable Energy* 33.11 (2008), pp. 2393–2406.

[11] SC Pryor, RJ Barthelmie, and Erik Kjellström. "Potential climate change impact on wind energy resources in northern Europe: analyses using a regional climate model". In: *Climate Dynamics* 25.7-8 (2005), pp. 815–835.

[12] Cristina L Archer and Mark Z Jacobson. "Spatial and temporal distributions of US winds and wind power at 80 m derived from measurements". In: *Journal of Geophysical Research: Atmospheres* 108.D9) (2003 (2003).

[13] Moti Segal et al. "On the potential change in wind power over the US due to increases of atmospheric greenhouse gases". In: *Renewable Energy* 24.2 (2001), pp. 235–243.

[14] Alan M Rhoades et al. "Characterizing Sierra Nevada snowpack using variable-resolution CESM". In: *Journal of Applied Meteorology and Climatology* 55.1 (2016), pp. 173–196.

[15] Xingying Huang et al. "An evaluation of the variable-resolution CESM for modeling California's climate". In: *Journal of Advances in Modeling Earth Systems* 8.1 (2016), pp. 345–369.

[16] Alan M Rhoades, Paul A Ullrich, and Colin M Zarzycki. "Projecting 21st century snowpack trends in western USA mountains using variable-resolution CESM". In: *Climate Dynamics* (2017), pp. 1–28.

[17] Xingying Huang and Paul A Ullrich. "Irrigation impacts on California's climate with the variable-resolution CESM". In: *Journal of Advances in Modeling Earth Systems* 8.3 (2016), pp. 1151–1163.

[18] Colin M Zarzycki et al. "Effects of localized grid refinement on the general circulation and climatology in the Community Atmosphere Model". In: *Journal of Climate* 28.7 (2015), pp. 2777–2803.

[19] Meina Wang, Paul Ullrich, and Dev Millstein. "Datasets on hub-height wind speed comparisons for wind farms in California". In: *Data in Brief* 19 (2018), pp. 214–221.

[20] Richard B Neale et al. "Description of the NCAR community atmosphere model (CAM 5.0)". In: *NCAR Tech. Note NCAR/TN-486+ STR* (2010).

[21] John M Dennis et al. "CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model". In: *The International Journal of High Performance Computing Applications* 26.1 (2012), pp. 74–89.

[22] Colin M Zarzycki et al. "Aquaplanet experiments using CAM's variable-resolution dynamical core". In: *Journal of Climate* 27.14 (2014), pp. 5481–5503.

[23] P. A. Ullrich. "SQuadGen: Spherical quadrilateral grid generator". In: *University of California, Davis, Climate and Global Change Group software.[Available online at `http://climate.ucdavis.edu/squadgen.php`.] (2014)* (2014).

[24] O Guba et al. "The spectral element method on variable resolution grids: Evaluating grid sensitivity and resolution-aware numerical viscosity". In: *Geosci. Model Dev. Discuss* 7 (2014), pp. 4081–4117.

[25] Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. "An overview of CMIP5 and the experiment design". In: *Bulletin of the American Meteorological Society* 93.4 (2012), p. 485.

[26] Luca Delle Monache et al. "Probabilistic weather prediction with an analog ensemble". In: *Monthly Weather Review* 141.10 (2013), pp. 3498–3516.

[27]   Luca Delle Monache et al. "Kalman filter and analog schemes to postprocess numerical weather predictions". In: *Monthly Weather Review* 139.11 (2011), pp. 3554–3570.

[28]   Ronald Gelaro et al. "The modern-era retrospective analysis for research and applications, version 2 (MERRA-2)". In: *Journal of Climate* 30.14 (2017), pp. 5419–5454.

[29]   Michele M Rienecker et al. "MERRA: NASA's modern-era retrospective analysis for research and applications". In: *Journal of climate* 24.14 (2011), pp. 3624–3648.

[30]   "Goddard Earth Sciences Data and Information Services Center (GES DISC), MERRA-2 `inst3_3d_asm_Np`: 3d,3-Hourly,Instantaneous,Pressure-Level,Assimilation,Assimilated Meteorological Fields V5.12.4". In: (). DOI: `10.5067/QBZ6MG944HW0`.

[31]   "Goddard Earth Sciences Data and Information Services Center (GES DISC), MERRA-2 `inst1_2d_asm_Nx`: 2d,1-Hourly,Instantaneous,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4". In: (). DOI: `10.5067/3Z173KIE2TPD`.

[32]   S Saha et al. "NCEP Climate Forecast System Reanalysis (CFSR) 6-hourly Products, January 1979 to December 2010". In: *Research Data, The National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo, USA* (2010).

[33]   Fedor Mesinger et al. "North American regional reanalysis". In: *Bulletin of the American Meteorological Society* 87.3 (2006), pp. 343–360.

[34]   CG Justus and Amir Mikhail. "Height variation of wind speed and wind distributions statistics". In: *Geophysical Research Letters* 3.5 (1976), pp. 261–264.

[35]   Matthias Fripp and Ryan H Wiser. "Effects of temporal wind patterns on the value of wind-generated electricity in California and the Northwest". In: *IEEE Transactions on Power Systems* 23.2 (2008), pp. 477–485.

[36] Stanton E Tuller and Arthur C Brett. "The goodness of fit of the Weibull and Rayleigh distributions to the distributions of observed wind speeds in a topographically diverse area". In: *International Journal of Climatology* 5.1 (1985), pp. 79–94.

[37] Susumu Shimada et al. "Accuracy of the wind speed profile in the lower PBL as simulated by the WRF model". In: *Sola* 7 (2011), pp. 109–112.

[38] D Carvalho et al. "WRF wind simulation and wind energy production estimates forced by different reanalyses: comparison with observed data for Portugal". In: *Applied Energy* 117 (2014), pp. 116–126.

[39] Phillip B Duffy et al. "Climate Change Impacts on Generation of Wind, Solar, and Hydropower in California". In: *California Energy Commission (2014) CEC-500-2014-111.* (2014).

[40] Meina Wang and Paul Ullrich. "Marine Air Penetration in California's Central Valley: Meteorological Drivers and the Impact of Climate Change". In: *Journal of Applied Meteorology and Climatology* 57.1 (2018), pp. 137–154.

[41] Scott Beaver and Ahmet Palazoglu. "Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality". In: *Journal of Applied Meteorology and Climatology* 45.12 (2006), pp. 1710–1726.

[42] Michael A Fosberg and Mark J Schroeder. "Marine air penetration in central California". In: *Journal of Applied Meteorology* 5.5 (1966), pp. 573–589.

[43] Norman L Miller and Nicole J Schlegel. "Climate change projected fire weather sensitivity: California Santa Ana wind occurrence". In: *Geophysical Research Letters* 33.15 (2006).

Chapter 4

Future Projections of Wind Patterns in California with the Variable-Resolution CESM: A Clustering Analysis Approach

4.1   Introduction

It is expected that wind energy production, as with many other environmentally-sourced renewable energy technologies, will be directly impacted by climate change. However, the highly localized character of wind fields, driven by a strong sensitivity to local topography, makes it difficult to model and project wind fields at the scales needed for stakeholders. Nonetheless, a better understanding of the variability of localized wind fields is essential to future wind energy resources planning and could help reduce the risk of selecting future wind project locations.

Even with the known difficulties with modeling wind, some progress has been made in better understanding this important resource. Past studies have focused on analyzing the climate change impact on localized wind fields, and the associated

change in wind energy generation potential [1, 2, 3, 4]. Wang, Ullrich, and Millstein [1] assessed the climate change impact through mid-century on California wind energy resources, and found that wind speed (and hence wind energy production) is likely to increase in summer, and diminish during fall and winter. Another study by Duffy et al. [5] also concluded that available wind energy in California will decrease in fall and winter. Yu et al. [6] detected upward trends in wind speeds across areas of the US Great Plains and Intermountain West, but downward trends in the east and in some parts of California. However, these past studies have only assessed overall trends of wind patterns on seasonal scales, or focused only on one specific type of wind pattern.

In this study, we present a new approach that leverages an unsupervised machine learning algorithm, agglomerative clustering, to group wind patterns from unlabeled data into wind clusters. The unlabeled input data for the clustering algorithm is produced using the Community Earth System Model (CESM), a global climate modeling system that has some demonstrable skill with modeling wind [1]. More details about the model can be found in Section 4.2. This clustering technique is leveraged to provide insight into the drivers and variability of different wind patterns. Once clusters have been identified, changes in wind fields between historical and end-of-century are decomposed into change in the cluster frequency and the change within each cluster. The insights gained from this decomposition then serve as our starting point for explaining significant trends that should be expected in the future. We investigate the cause of within-cluster wind speeds change by analyzing synoptic-scale fields associated with each cluster. However, we do not investigate the drivers of future change to the frequency of clusters, as these changes depend on global meteorological patterns that are beyond the scope of this study. Finally, seasonal changes of wind energy are assessed, along with the local impact of observed changes from wind clusters. Given appropriate regional climate data, this technique has the potential to be adapted to essentially any geographic region.

This work builds on a previous study by Millstein et al. [7], who used clustering to identify the characteristics of ten selected clusters over the historical time period. Their study then investigated the wind regime changes over the period of 1980-2015 in California, and further analyzed the impact on local wind energy resources. The present study works to expand the time scope of Millstein et al. [7] to the end of the 21st century, and detect any significant trends associated with the most relevant wind clusters.

For the purposes of this study, we have divided California into two sub-domains: the Northern California (NC) domain, which includes Shiloh and Altamont Pass wind plant sites, and the Southern California (SC) domain, which includes Alta, San Gorgonio, and Ocotillo sites (Figure 4.1). These five wind plant locations include both wind plant sites currently in service, and wind project sites targeted for future development. Due to differences in wind patterns that emerge between NC and SC domains, the clustering algorithm was applied to the two domains separately.

The remainder of this paper is as follows: In section 4.2 we describe the VR-CESM model setup and the clustering algorithm used in this study. Results are presented in section 4.3, followed by discussion and conclusions in section 4.4.

4.2 Methods

This study uses model output from the Community Earth System Model (CESM), a widely-used global climate model [8, 9]. Three time periods were separately simulated, including historical (1980-2000), mid-century (2030-2050), and end-of-century (2080-2100). However, the mid-century period that was the focus of Wang, Ullrich, and Millstein [1] is not considered in this study, and is only used to provide additional input for the clustering procedure. All simulations used the same model setup, enabling us to compare across time frames, with differences only in prescribed sea-surface temperatures and greenhouse-gas forcing. Details on model validation,
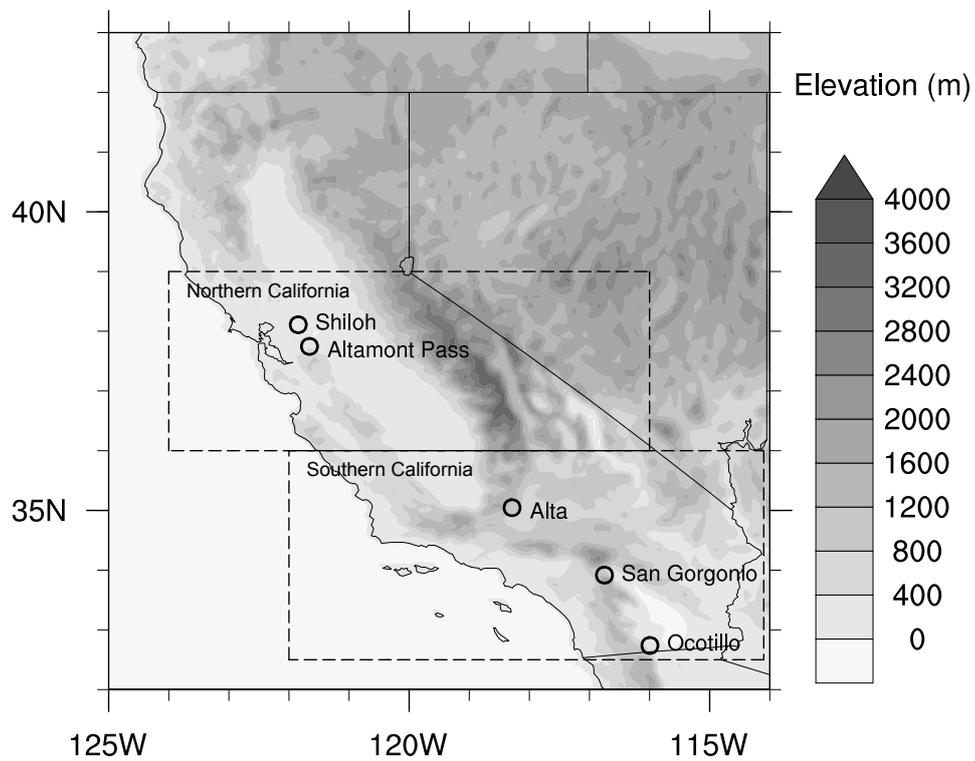
Figure 4.1: The Northern California (NC) and Southern California (SC) domains with dash line bounding boxes, along with the five wind plant locations. This figure is a reproduction of Figure 1 from Millstein et al. [7].

including comparison with observational stations, reanalysis datasets, and other modeling products, can be found in Wang, Ullrich, and Millstein [1].

4.2.1 Description of VR-CESM (global climate model product)

CESM version 1.5.5 was used for this study with the F-component set (FAMPIC5), which prescribes sea-surface temperatures and sea ice but dynamically evolves the atmosphere and land surface component models. The atmospheric component model is the Community Atmosphere Model, version 5.3 (CAM5) [8] with the spectral-element (SE) dynamical core Dennis et al. [10] in its variable-resolution (VR) configuration [11]. More details on VR-CESM can be found in Rhoades et al. [12, 13], and Huang et al. [14]. The VR model grid used for this study, depicted in Figure 4.2, was generated for use in CAM and CLM with the open-source software package SQuadGen [15, 16]. This grid has a finest horizontal resolution of $0.125°(\sim 14 \text{km})$ over the western United States, with a quasi-uniform $1°$ mesh over the remainder of the globe. Three simulations were conducted on this grid: The historical run covered the period from October 1st, 1979 to December 31st, 2000, with the last three months of 1979 discarded as the spin-up period, for a total of 21-years of three-hourly output. This historical time period was chosen to provide an adequate sampling of the inter-annual variability. For projections of future wind energy change, our mid-century and end-of-century simulations ran with the "business as usual" Representative Concentration Pathway 8.5 (RCP8.5) [17] from October 1st, 2029 to December 31st, 2050, and from October 1st, 2079 to December 31st, 2100, respectively. In each case the first three months of the simulation were discarded, yielding two additional 21-year-long simulations. Analogous simulations with VR-CESM have also been conducted by Rhoades, Ullrich, and Zarzycki [18] and Huang and Ullrich [19] for assessing snowpack and future precipitation, respectively. Greenhouse gas (GHG) and aerosol forcings are prescribed based on historical or RCP8.5
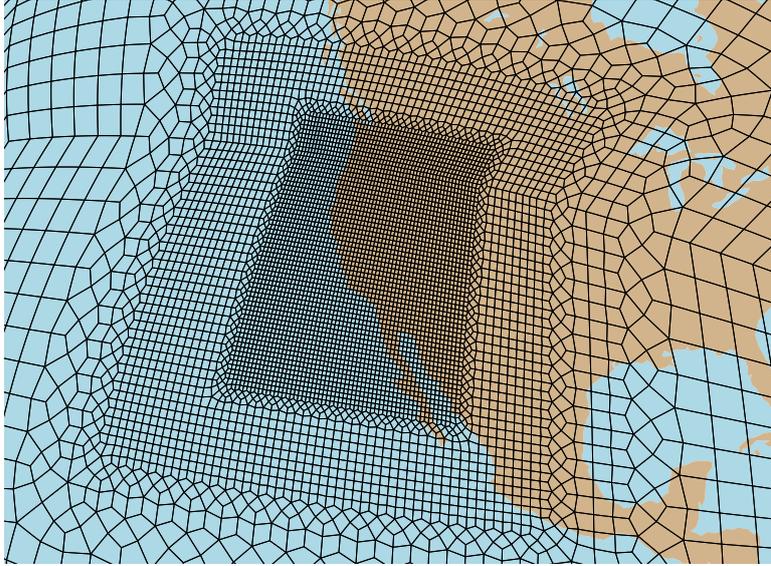
Figure 4.2: The VR-CESM grid used in this study, constructed by first successively refining a cubed-sphere grid with a 1°(111km) quasi-uniform resolution to a resolution of 0.125°(∼14km) over the western USA. This figure is a reproduction of Figure 2 from Wang, Ullrich, and Millstein [1].

concentrations for each simulation. The historical and mid-century VR-CESM simulations were previously validated and analyzed in Wang, Ullrich, and Millstein [1]. Here we expand the time horizon through the end of the 21st century, and analyze the potential changes on localized wind regimes. Note that in Wang, Ullrich, and Millstein [1], we found that although the large-scale patterns are captured, there is nonetheless a low wind speed bias from VR-CESM which leads to an under estimation of capacity factors. In order to calibrate the wind speed from VR-CESM, we estimated a bias correction factors of 1.3 in Wang, Ullrich, and Millstein [1]. This bias-correction factor was calculated based on a comparison between VR-CESM and a high-resolution regional simulation (referred to as DNV GL in Wang, Ullrich, and Millstein [1]). Linear bias correction factors have been applied in past efforts in order to match global modeling or reanalysis outputs with operational data, for example, see Staffell and Pfenninger [20] and Olauson, Edström, and Rydén [21]. Capacity factors, which are analyzed in section 4.3.3, were therefor calculated from the bias-corrected wind speed.

4.2.2 Agglomerative clustering

In the nomenclature of machine learning, the output data from the CESM model simulations is referred to as "unlabeled" – namely, there is no prior knowledge of the different wind patterns and their associated frequencies. In order to develop such a labeling, we apply an unsupervised machine learning algorithm to group and distinguish different wind patterns. Specifically, we use the agglomerative clustering algorithm with Ward's method [22] to minimize the total within-cluster variance. Under this algorithm, each data point is initialized as a single-item cluster. At each iteration of the method, smaller nearby clusters are chosen to merge and form larger clusters; the particular choice of merged clusters minimizes a global inter-cluster distances metric (i.e., Ward's method minimizes the variance of clusters being merged). This "bottom-up" algorithm then iterates to create a dendrogram, which is tree-like structure, illustrating the arrangement of clusters. The number of clusters used in the subsequent analysis can then be varied by halting the iteration procedure at a particular level. Typically this choice is made through inspection of the resulting clusters at each iteration, so as to identify the earliest point at which there is sufficient distinction between all clusters in the set. This algorithm's primary advantage over k-means clustering [23] is that it does not require the parameter k (how many clusters to generate) to be specified beforehand. Since we did not have prior knowledge of the number of distinct wind patterns before execution of the clustering algorithm, agglomerative clustering provided a natural mechanism to tune this value.

In this study, clustering is solely applied to 80m wind vector fields (composed of horizontal and meridional wind magnitudes). This particular height was chosen as it is typical of the hubs of large wind turbines. Principal component analysis (PCA) was first applied to 3-hourly (eight times daily) 80m wind vector fields to reduce dimensionality, similar to the approach taken in Millstein et al. [7], Jin, Harley, and Brown [24], Berg et al. [25], Conil and Hall [26] and Ludwig, Horel, and Whiteman

[27]. We retained the first ten principal components for clustering, as they accounted for over 80% of the total variance. Then, each day was categorized into a particular cluster based on a set of (8×10) 80 PCA components. For each region (NC and SC), regridded data from all three time periods (historical, mid-century, and end-of-century) was simultaneously provided as input to the clustering algorithm. This was to ensure the consistency of clusters across all three time periods. We ran the clustering algorithm separately on NC and SC domains since the synoptic-scale wind patterns produce distinct localized effects in these regions. To determine how many wind patterns would be needed to distinguish wind regimes, we leveraged the dendrogram produced by the agglomerative clustering algorithm and determined the point when distinctly different wind patterns were merged [28]. After examination of the clustering output, we concluded that for each of NC and SC domains, ten clusters provided a good representation of different wind regimes – namely, five clusters did not sufficiently distinguish various qualitatively different wind patterns, and fifteen clusters produced several instances of cluster pairs with only subtle differences. A quantitative assessment using the CH index [29], which measures the overall within-cluster variance and the overall between-cluster variance, confirmed the optimality of ten clusters in each region. Namely, ten clusters produced a higher CH index than the index from either five and fifteen clusters – indicating that the clusters have larger between-cluster variance, and smaller within-cluster variance. Therefore, for both NC and SC we analyzed the different wind patterns from ten clusters. Note that in the remainder of the text the numbers associated with each cluster do not bear meaning, and are only for labeling purposes. Each cluster is labeled by its domain and cluster number (e.g. NC 6 is cluster 6 from NC domain).

4.2.3 Decomposition of changes in wind clusters

Climate change can impact wind clusters through two principal avenues: First, through the modification of the frequency of the wind cluster, and second, through

the modification of the wind patterns within each cluster. The change in either the total wind field or the wind field of each cluster can be decomposed into these two contributions as follows. We denote the historical frequency of a given cluster i as f_i^h, the end-of-century frequency as f_i^e, the historical average wind field within the cluster by U_i^h, and the end-of-century wind field within the cluster by U_i^e. Thus the average historical U^h and end-of-century U^e wind fields can be written as:

$$U^h = \sum_i U_i^h f_i^h, \qquad\qquad U^e = \sum_i U_i^e f_i^e. \qquad (4.1)$$

The average frequency of the cluster f_i and average wind field within the cluster U_i (combining both historical and end-of-century) are given by:

$$f_i = \frac{1}{2}(f_i^h + f_i^e), \qquad\qquad U_i = \frac{U_i^h f_i^h + U_i^e f_i^e}{f_i^h + f_i^e} \qquad (4.2)$$

Then defining change in cluster frequency by $\Delta f_i = f_i^e - f_i^h$ and change in the average wind field by $\Delta U = U^e - U^h$, we can write

$$\Delta U = \sum_i U_i^e f_i^e - U_i^h f_i^h \qquad (4.3)$$

$$= \sum_i \underbrace{U_i \Delta f_i}_{(a)} + \underbrace{(U_i^e - U_i^h) f_i}_{(b)} - \underbrace{\frac{\Delta f_i^2 (U_i^e - U_i^h)}{4 f_i}}_{(c)}, \qquad (4.4)$$

where (4.4a) denotes the change in average wind speed due to the change in frequency of cluster i, (4.4b) denotes the change in average wind speed due to the change in the wind field within each cluster i, and (4.4c) denotes nonlinear changes associated with simultaneous changes in frequency and wind field. In this wind speed decomposition, U represents the wind speed magnitude from VR-CESM, not the wind vector field. Note that such a decomposition is independent of our choice of clustering technique, and can be performed for any grouping of fields from two periods.

4.3 Results

Section 4.3.1 describes the wind patterns associated with each cluster. Section 4.3.2 then examines the climatological synoptic-scale fields from clusters with significant trends. In section 4.3.3, we analyze the future projections of wind clusters from the end-of-century VR-CESM simulation, and their impact on wind energy output.

4.3.1 Trends in cluster frequency

As described in section 4.2.2, days from historical and end-of-century time periods were grouped into ten clusters per region (NC and SC) based solely on wind vector fields (twenty clusters total). A qualitative summary of these clusters, their dominant seasonality, and end-of-century minus historical frequency change (annual and broken down by season) is given in Table 4.1. By using a combined dataset of historical and end-of-century daily wind fields as input for the cluster analysis, we would generally expect that changes in cluster frequency will dominate the total change in the wind field. Namely, since the cluster analysis is, in effect, grouping days with similar wind fields, we expect that the wind field for days in a particular cluster to be more similar to one another than to the wind field of days in another cluster. For each of these twenty clusters, Figures S3-S5 and S6-S8 show the magnitudes of each of the three terms in Equation (4.4) for the northern and southern California clusters, respectively. In general, we observe that change in cluster frequency is the dominant contributor to change in wind patterns, followed by changes in wind fields within each cluster (except in those cases where the change in cluster frequency is small). In each case the nonlinear term is not a significant contributor to the overall change. The remainder of this section focuses on analysis of select clusters, with additional discussion on the large-scale drivers that could influence the wind climatology in each case.

Table 4.1: Top: Dominant seasons, historical frequency, end-of-century frequency changes, and qualitative summary for NC and SC clusters. Bottom: Historical frequency and end-of-century frequency change broken down by season. Frequency changes indicated in bold are significant under the two-proportion z-test at the 95% significance level. The seasonal frequency of these clusters is also depicted in Figures S1 and S2. Seasons are March-April-May (MAM), June-July-August (JJA), September-October-November (SON), and December-January-February (DJF).

Cluster	Dominant Seasons	Annual f_i^h	Δf_i	Qualitative summary
NC 1	DJF MAM	13.6%	**-1.5%**	Westerly wind
NC 2	DJF	10.2%	**-1.3%**	Stronger westerly wind w/ offshore trough
NC 3	DJF SON	11.2%	**- 3.2%**	Offshore blocking
NC 4	SON MAM	13.4%	- 0.5%	Low wind
NC 5	JJA	5.3%	+ 0.3%	Strong northerly wind
NC 6	JJA MAM	12.7%	**+ 2.4%**	Marine air penetration
NC 7	JJA MAM	12.3%	+ 0.2%	Marine air penetration
NC 8	JJA SON	8.0%	**+ 2.1%**	Marine air penetration
NC 9	DJF MAM	9.2%	+ 0.6%	Low southerly wind
NC 10	JJA	4.0%	**+ 0.8%**	Marine air penetration
SC 1	MAM DJF	14.1%	**- 1.1%**	Strong alongshore wind
SC 2	JJA SON	23.1%	- 0.3%	Weak onshore flow
SC 3	DJF MAM	12.5%	+ 0.4%	Low wind
SC 4	JJA MAM	15.5%	**+ 2.8%**	Onshore flow
SC 5	DJF	3.8%	- 0.5%	Southwesterly wind
SC 6	DJF SON	8.8%	**- 2.3%**	Santa Ana winds
SC 7	JJA SON	7.3%	**+ 2.0%**	Weakened onshore flow
SC 8	DJF MAM	7.2%	**- 1.7%**	Westerly wind
SC 9	SON MAM	4.9%	**+ 1.0%**	Low wind
SC 10	DJF MAM	2.8%	- 0.4%	Onshore flow

Cluster	MAM f_i^h	Δf_i	JJA f_i^h	Δf_i	SON f_i^h	Δf_i	DJF f_i^h	Δf_i
NC 1	17.5%	- 0.9%	1.1%	**- 0.8%**	15.7%	**- 5.1%**	20.5%	+ 0.8%
NC 2	9.3%	**- 2.6%**	0.1%	0.0%	7.0%	-1.1%	24.5%	- 1.3%
NC 3	7.1%	**- 1.7%**	1.0%	**- 0.9%**	15.2%	**- 6.4%**	21.7%	**- 3.9%**
NC 4	17.8%	+ 0.5%	5.8%	**- 4.0%**	20.8%	- 0.1%	9.4%	+ 1.6%
NC 5	2.3%	**+ 1.3%**	15.7%	- 0.9%	3.0%	+ 0.7%	0.0%	+ 0.1%
NC 6	17.5%	-0.5%	19.1%	**+ 6.2%**	11.7%	**+ 3.8%**	2.3%	+ 0.2%
NC 7	11.7%	**+ 2.4%**	27.3%	**- 3.1%**	8.1%	+ 0.7%	1.8%	+ 0.8%
NC 8	4.3%	**+ 1.9%**	16.8%	**+ 3.3%**	10.5%	**+ 3.2%**	0.3%	+ 0.1%
NC 9	10.6%	- 1.1%	0.2%	- 0.1%	6.7%	**+ 2.1%**	19.5%	+ 1.7%
NC 10	1.9%	+ 0.7%	12.9%	+ 0.3%	1.2%	**+ 2.3%**	0.0%	0.0%
SC 1	22.7%	- 1.4%	2.2%	- 0.6%	13.3%	**- 3.5%**	18.4%	+ 1.0%
SC 2	19.5%	**+ 3.1%**	45.9%	**- 6.4%**	21.4%	+ 2.0%	5.2%	0.0%
SC 3	12.2%	**- 2.6%**	0.2%	0.0%	16.5%	- 1.3%	21.4%	**+ 5.5%**
SC 4	17.2%	**+ 4.0%**	30.8%	**+ 5.3%**	12.9%	+ 1.5%	0.8%	+ 0.7%
SC 5	2.5%	+ 0.2%	0.0%	0.0%	1.9%	- 0.5%	10.7%	- 1.7%
SC 6	4.0%	**- 1.7%**	0.0%	+ 0.1%	10.4%	**- 4.4%**	21.1%	**- 3.0%**
SC 7	2.8%	**+ 2.1%**	18.0%	+ 1.9%	7.7%	**+ 4.2%**	0.5%	- 0.2%
SC 8	10.8%	**- 3.9%**	0.2%	+ 0.2%	4.9%	- 0.3%	12.9%	**- 2.8%**
SC 9	5.8%	+ 0.8%	2.7%	- 0.5%	8.8%	**+ 2.8%**	2.3%	**+ 1.1%**
SC 10	2.4%	- 0.5%	0.0%	+ 0.1%	2.0%	- 0.5%	6.8%	- 0.5%

4.3.2 Synoptic-scale character of prominent clusters

This section describes the synoptic-scale character of the select clusters from Table 4.1. We focus on analyzing the mean meteorological fields, including the 700hPa geopotential height, and the wind field at 80m above the ground. The 700hPa geopotential height field was chosen as it is reflective of the general circulation, with wind flow at this level being largely geostrophic but still strongly connected with near-surface winds. Because of the terrain-following coordinate, the lowest model level in CESM is everywhere below the 80m level, and so all wind speeds are interpolated.

NC 1 and NC 2: Reduced ventilation from westerly winds

Clusters NC 1 (westerly wind) and NC 2 (stronger westerly wind) in the NC domain are frequent (13.6% and 10.2%) wind patterns that peak in frequency during the winter season (20.5% and 24.5% frequency in DJF). They are accompanied by relatively large annual frequency changes (-1.5% and -1.3%), with the largest decreases occurring in the spring and fall. Further analysis of these patterns is beneficial to explain decreases in wind energy output during DJF, described later in the paper (Table 4.6). Among the two, cluster 2 shows higher wind speed in NC domain than cluster 1, and thus will be analyzed here.

The synoptic-scale fields for NC 2 are depicted in Figure 4.3. The 700hPa geopotential height field shows a trough over the Gulf of Alaska that promotes flow directed perpendicular to the coast and hence on-shore ventilation through the NC domain. As discussed later, NC 2 tends to produce the highest wind speeds at the Shiloh and Altamont Pass wind plants among all clusters, and so a reduction in the frequency of this pattern will be associated with decreasing NC capacity factors in DJF. Comparing end-of-century to historical within this cluster, two effects appear to be prominent: First there is an increase in the geopotential gradient in the mid-Pacific which drives up wind speeds over the open ocean. However, simultaneously increased overland temperatures (not shown) appear to be promoting an increase
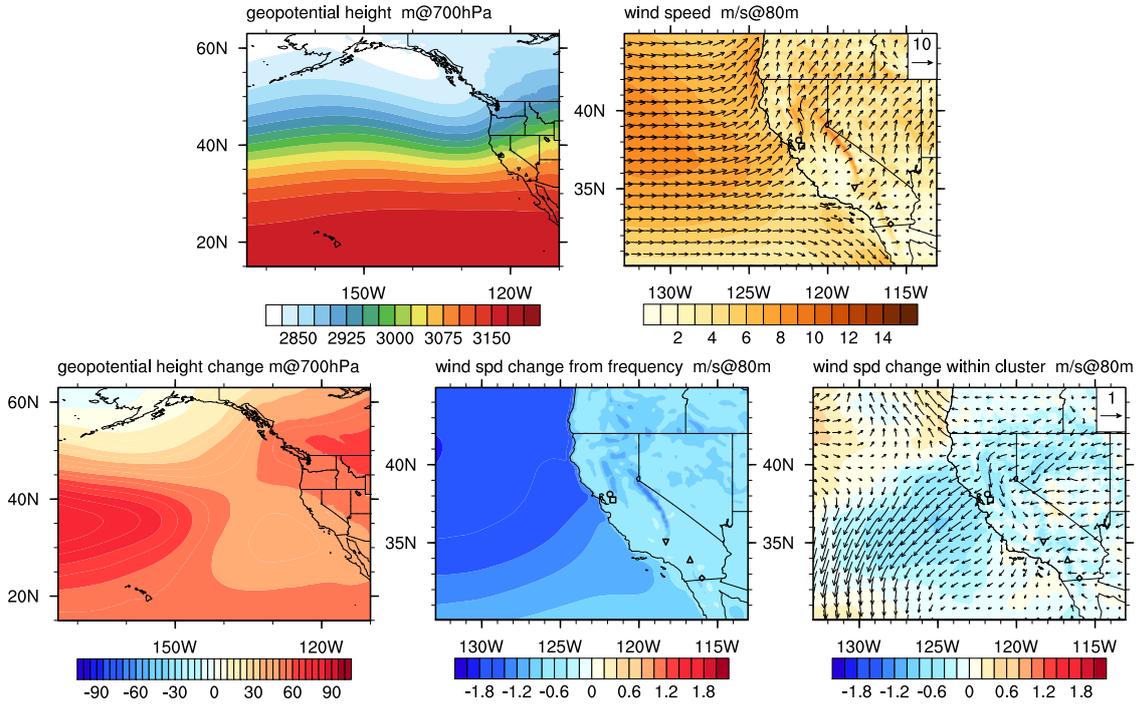
Figure 4.3: Meteorological fields from cluster NC 2. (top left) Historical mean 700hPa geopotential height; (top right) 80m historical wind field; (bottom left) 700hPa geopotential height change; (bottom middle) end-of-century minus historical wind speed change due to change in cluster frequency ($U_i \Delta f_i / f_i$); and (bottom right) end-of-century minus historical wind speed change within cluster ($U_i^e - U_i^h$).

in the overland geopotential height (thicker air masses from warmer temperature). This second factor drives a reduction in onshore flow, and consequently we observe decreasing wind speeds within this cluster across the NC domain.

NC 3: Reduced offshore blocking

Figure 4.4 depicts the synoptic-scale fields from NC 3, which again peaks in the winter season and exhibits a frequency decrease of 3.2% through end-of-century. This cluster corresponds to offshore blocking along the California coast. In opposition to NC 6 (associated with summertime marine air penetration), this cluster exhibits a pronounced ridge over the Eastern Pacific, leading to a strong northerly

wind flow parallel to the California coastline that is associated with the second largest wind speeds at the NC wind plants. Within this cluster, the 700hPa geopotential height field exhibits a broad increase in end-of-century; however, the change in geopotential height is larger at lower latitudes and smaller over the Northern Pacific. This leads to a weakening of the northerly flow, in turn causing an overall decrease in offshore and onshore wind speeds. Overall, the decrease in frequency and character of this pattern drives weaker wind speeds at both Shiloh and Altamont Pass.

Note that other studies (i.e., Wang and Schubert [30]) noted an increased trend in blocking over the 20th century, particularly in the Gulf of Alaska, which seems contrary to our observations in this section (particularly given that NC 3 is representative of this offshore blocking pattern). To assess if this trend is present in the VR-CESM data, we counted blocking days at each grid point over each DJF season, defined as days where the geopotential at a given point exceeded the climatological geopotential for that period plus one standard deviation (separately calculated for historical and end-of-century). Note that the blocking days were selected outside the clustering framework, using only the aforementioned criterion. The results of this analysis are plotted in Figure 4.5, and are inconsistent with an increased blocking frequency.

NC 6-8 and NC 10: Increased summertime marine air penetration (MAP)

Figure 4.6 depicts the synoptic-scale fields of cluster 6 in the NC domain, which is expected to increase in frequency by 2.4% through end-of-century. The change in frequency of this cluster appears to occur in conjunction with a decreasing frequency of the NC 4 cluster, associated with low wind events. NC 6 is indicative of a typical summertime marine air penetration (MAP) condition [31, 32, 33]. Notably, the increasing frequency of summertime MAP events from these clusters agrees with the findings of Wang and Ullrich [31]. MAP events feature an off-shore trough
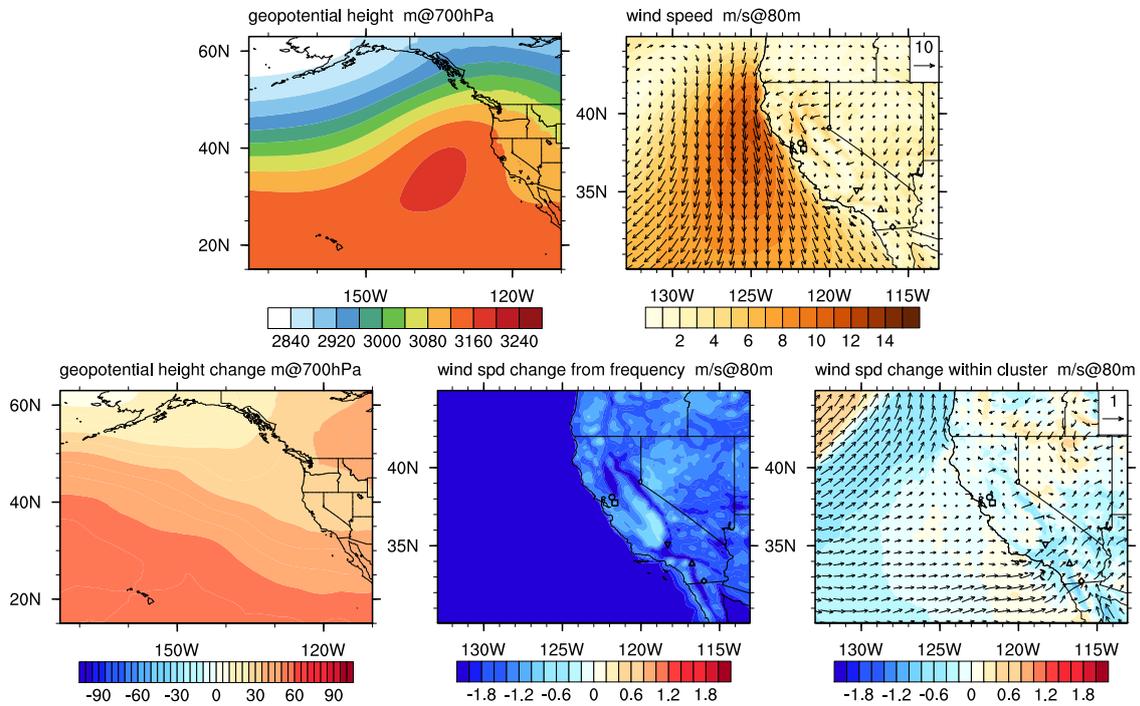
Figure 4.4: As Figure 4.3 but for NC domain cluster 3.
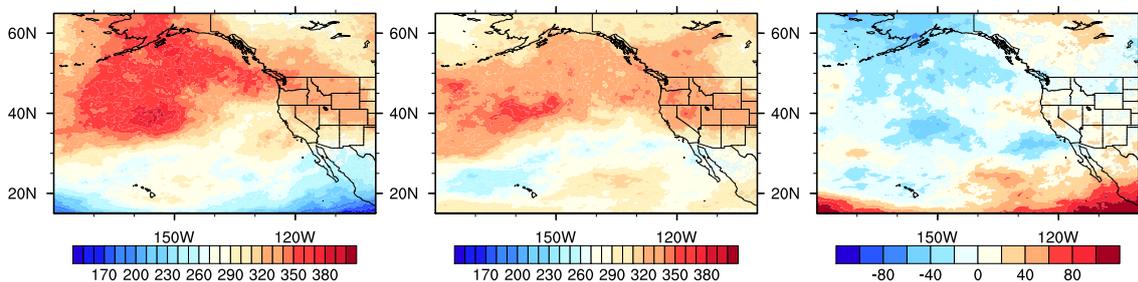
Figure 4.5: Total number of days each grid point exceeds the mean plus one standard deviation of 500hPa geopotential height field for (Left) historical and (Center) end-of-century. (Right) Difference between end-of-century and historical.

and geopotential height contour lines perpendicular to coastline, allowing cool and moist marine air to penetrate inland. It is the location of the off-shore trough that is directly responsible for driving marine air through the San Francisco Bay Delta.

Within this cluster and relative to the historical period, the magnitude of the 700hPa geopotential height field under the end-of-century increases, as a direct consequence of low-level warming (not shown). This low-level warming drives a thickening of air layers and thus an increase in the 700hPa geopotential height field. However, this increase is less pronounced over the Northern Pacific, which drives a weakening of the typically northerly wind pattern that traces the coastline in Northern California, and an increase in the on-shore flow pattern driven by the general circulation. This in turn leads to an increase in wind speeds through the San Francisco Delta region during MAP days (and at Shiloh and Altamont Pass in NC domain). A shift in this particular synoptic-scale pattern also drives increased ventilation in the SC domain.

These changes to frequency and wind pattern suggest the tendency towards more MAP days and more intense MAP winds are primary drivers for increased summertime wind speeds in the San Francisco Bay region.

SC 1: More seasonally concentrated strong alongshore wind

Moving to the SC domain, cluster SC 1 captures days of strong alongshore wind off the U.S. west coast (Figure 4.7) that appear most prominently between the fall and spring seasons. The alongshore flow weakens south of the SC domain, and so is associated with a strong inland flow of the marine air through the Los Angeles region. This pattern is associated with some of the highest historical capacity factors for the Alta wind plant (see table 4.8). Due to the location of Alta wind plant, which sits in the pass between in the Tehachapi mountains, the ventilation from the San Joaquin valley to the Mojave also contributes to the high capacity factors. It is also a frequent pattern, and one that has been projected to decrease in frequency by 1.1%
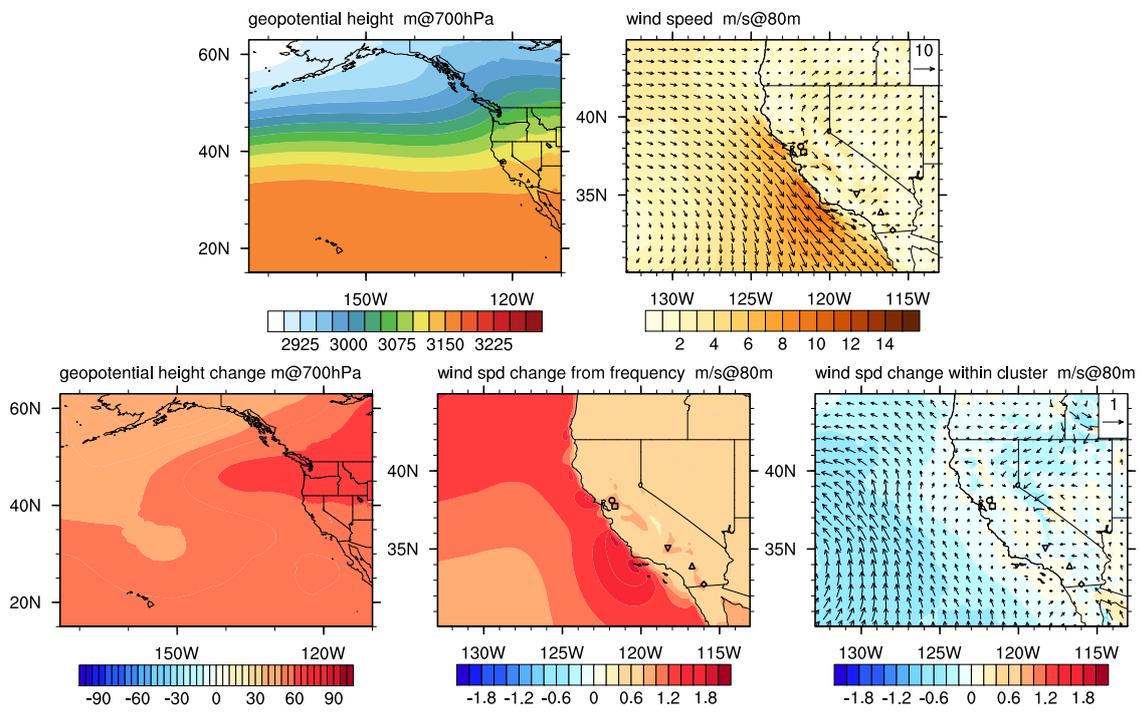
Figure 4.6: As Figure 4.3 but for NC domain cluster 6.
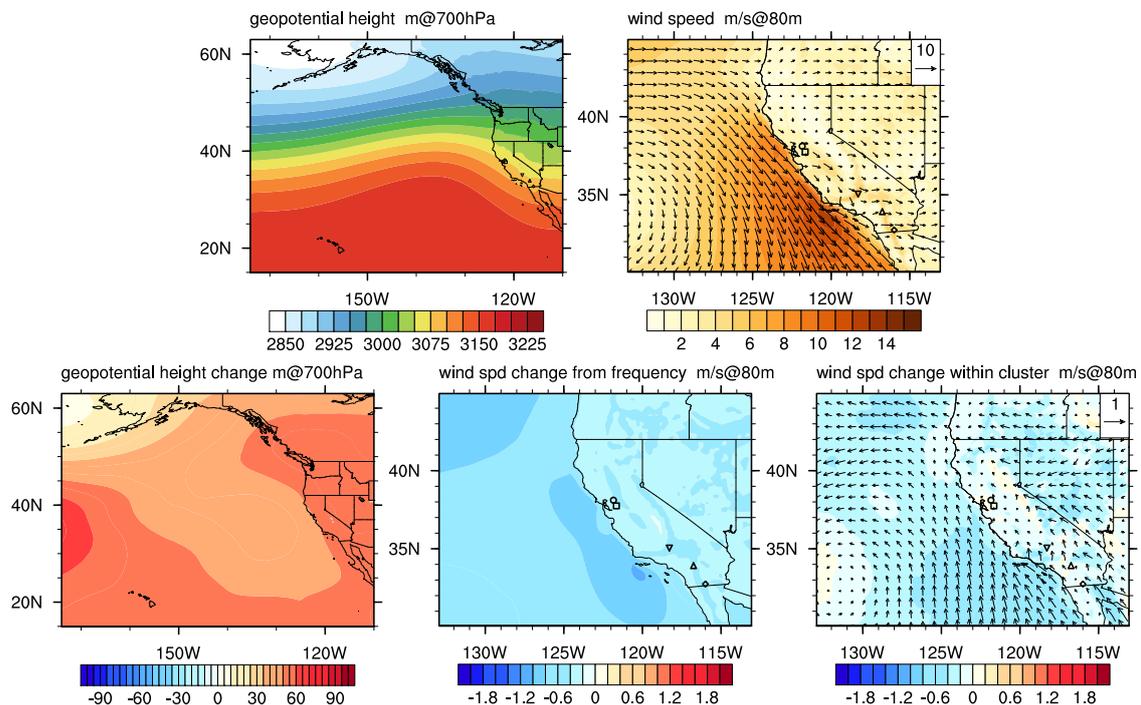
Figure 4.7: As Figure 4.3 but for SC domain cluster 1.

annually; however, this change in frequency is primarily because of an increase in seasonality – the pattern sees an increase in frequency in DJF but decrease in MAM and SON. Within this cluster, the 700hPa geopotential height field change shows an inhomogenous pattern that favors overland warming, and reduces the alongshore gradient, thus leading to a weakening of the flow. The net result of these changes is a reduction in spring and winter wind speeds in the SC region.

SC 4: Increased summertime marine air penetration

Spring and summertime marine air penetration is also reflected in the SC domain via cluster SC 4, and its increased frequency through end-of-century supports our prior observations with cluster NC 6 (marine air penetration). As shown in Figure 4.8, a local trough sits off-shore with a 700hPa geopotential contour perpendicular to the shoreline in SC domain, leading to onshore marine air. The end-of-century
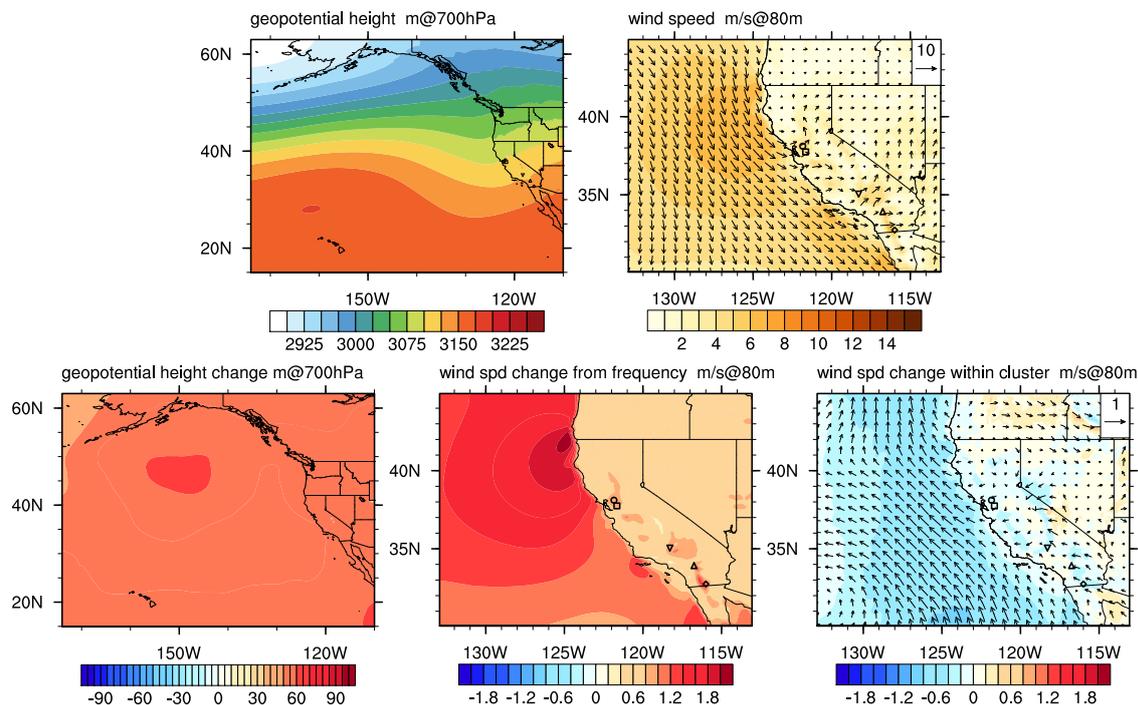
Figure 4.8: As Figure 4.3 but for SC domain cluster 4.

change to the 700hPa geopotential height surface also produces a small enhancement in wind speeds parallel to the shore. Consequently both the increased frequency of SC 4 and slightly increased onshore winds within SC 4 leads to increased ventilation of the SC domain.

SC 5: Less frequent wintertime southwesterly wind

SC 5 represents wintertime southwesterly wind from an offshore trough sitting near the U.S. west coast. This cluster brings relatively high wind speeds, but is becoming less frequent during the winter season. By the end-of-century, the offshore trough intensifies, leading to higher wind speeds over the Pacific. Simultaneously, the 700hPa geopotential height anomaly center over the SC domain acts to block the onshore wind, leading to wind speeds decreasing over almost all areas within California.
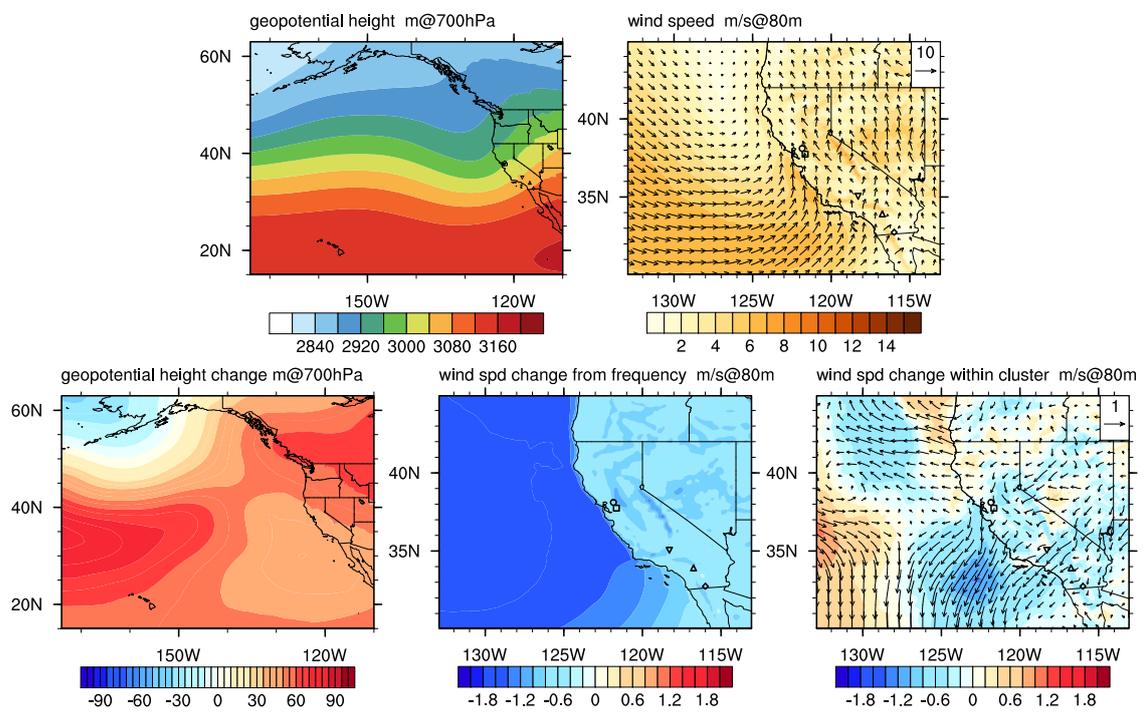
Figure 4.9: As Figure 4.3 but for SC domain cluster 5.
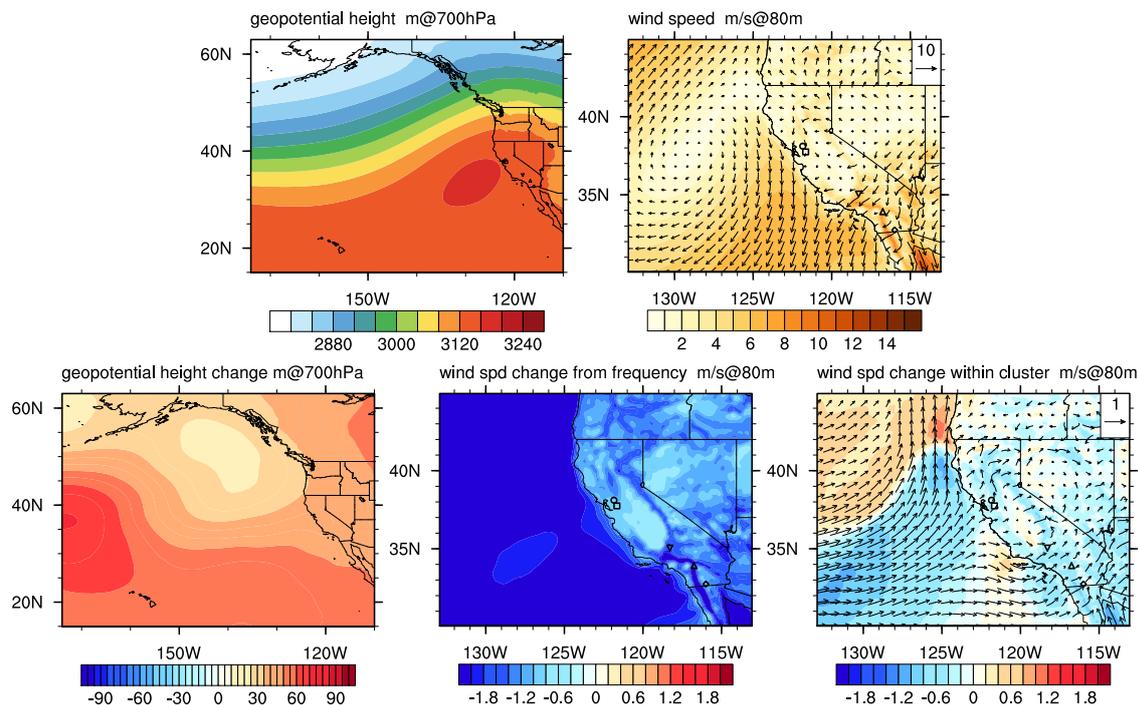
Figure 4.10: As Figure 4.3 but for SC domain cluster 6.

SC 6: Less frequent and weaker Santa Ana winds in fall/winter

The second largest change in cluster frequency for the SC domain occurs in cluster 6, which is 2.3% less frequent by end-of-century. The synoptic fields for these days is depicted in Figure 4.10, and corresponds to a typical wind pattern from Santa Ana events [34, 35]. The relatively high 700hPa geopotential height field over the western US, along with the high center sitting off-shore, leads to the northeasterly wind field throughout the SC region. The end-of-century change in 700hPa geopotential height field indicates a weakening of the onshore ridge, in turn producing slightly weaker winds during Santa Ana events. The decrease in cluster frequency around Fall season is also consistant with findings from Miller and Schlegel [3], where decreasing frequency of Santa Ana occurrence was also projected in early Fall through the end-of-century.
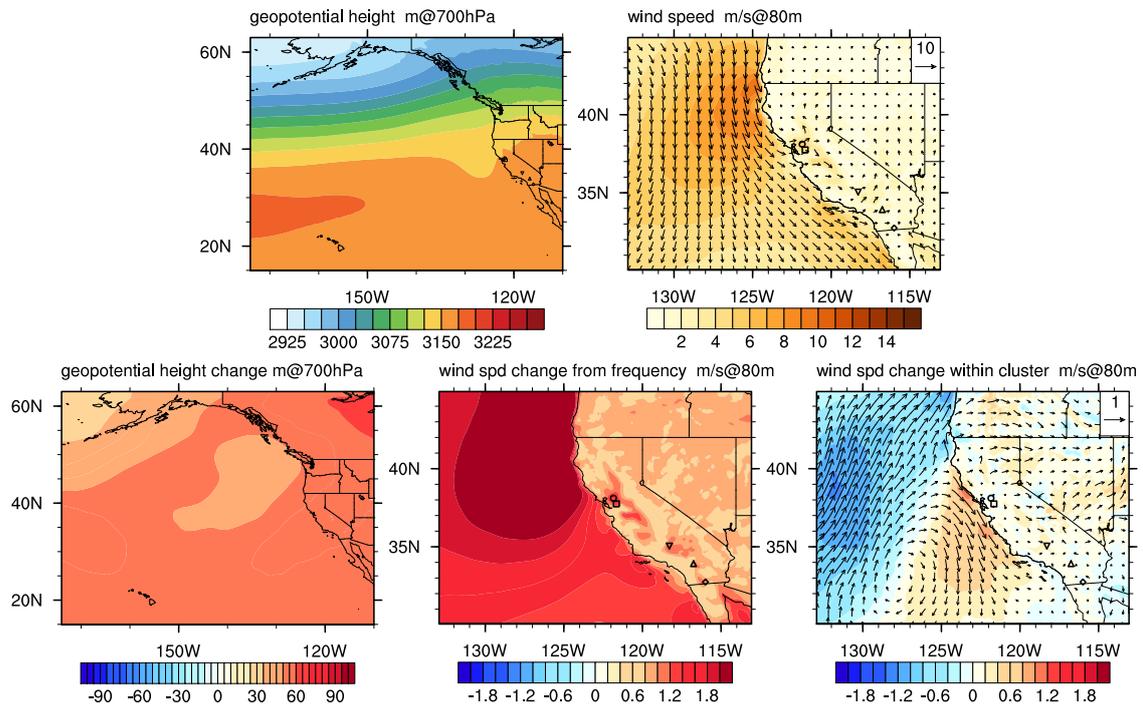
Figure 4.11: As Figure 4.3 but for SC domain cluster 7.

SC 7: More frequent and less seasonal weakened onshore flow

SC cluster 7, which corresponds to weakened onshore flow in the summer and fall seasons, also shows a significant increase in frequency by 2.0%. The synoptic-scale fields of this cluster are depicted in Figure 4.11. By the end-of-century, the high 700hPa geopotential height anomaly center sitting offshore to the California coast acts to increase the northerly flow parallel to the coastline in Northern California, and blocks northerly flow in SC domain. This leads to a weakening of the offshore flow throughout the SC domain.

SC 8: Less frequent westerly wind in winter/spring

SC cluster 8 represents a steady westerly marine flow directed onshore (Figure 4.12), and appears most prominently in the winter season. This cluster is less
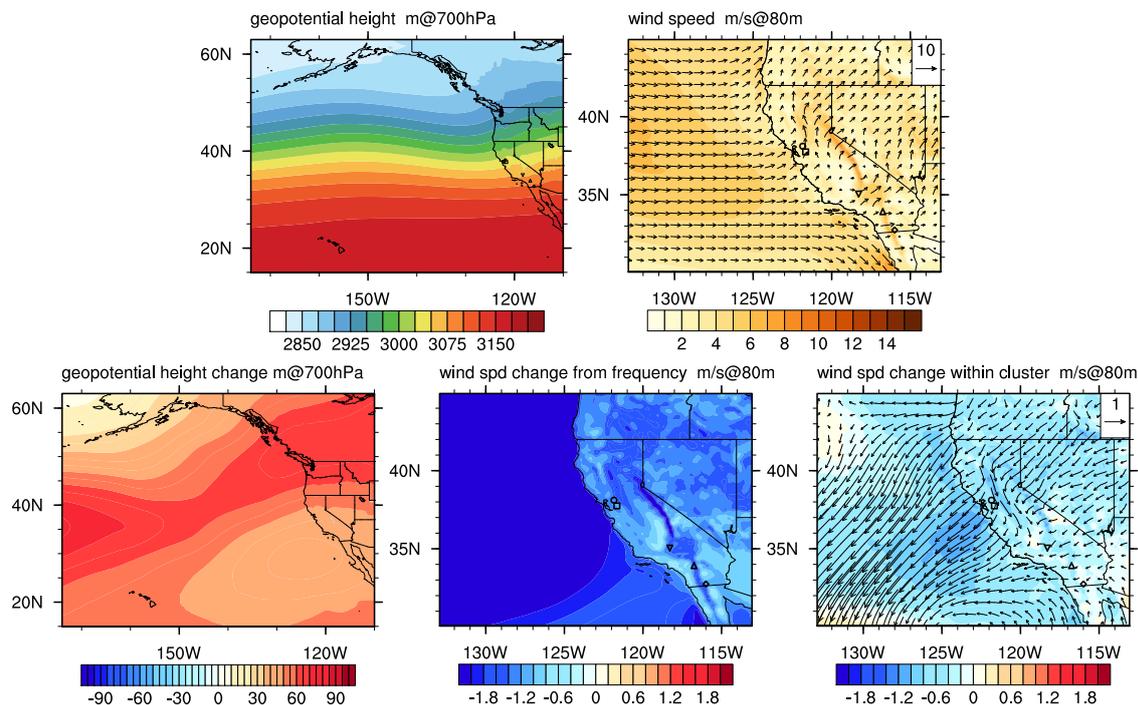
Figure 4.12: As Figure 4.3 but for SC domain cluster 8.

frequent (7.2%) but has been projected to decrease by 1.7% in its frequency under end-of-century, with most of the decrease occurring in winter and spring. Similar to the previously described clusters, the 700hPa geopotential height field in cluster 8 is also increasing, although with a magnitude that is reduced over the area centered around the offshore region near Baja California. The net result of this change in the geopotential height field is a reduced wind field throughout the whole California, and also a reduction in onshore marine flow. Consequently the changes in this cluster produce a reduction in wind speeds throughout the SC domain.

4.3.3 Trends in wind energy production

In this section, projected changes in wind energy production are considered in light of the cluster analysis. Before proceeding, we first assess projected changes

in wind energy production from model output. Wind fields from VR-CESM runs were interpolated to each wind plant location so as to directly compute wind energy capacity factor (CF in %) changes between historical and end-of-century (details of this calculation and mid-century differences can be found in Wang, Ullrich, and Millstein [1]). Before calculating CF based on the wind fields from VR-CESM, a constant bias correction factors of 1.3 (Section 4.2.1) was applied to the wind fields to reduce the low wind speed bias from VR-CESM. Then CF were calculated from the bias-corrected wind fields. Table 4.2 through 4.9 are all based on the bias-corrected CF values. CFs are commonly defined as actual power output divided by the maximum wind power output that can be generated through the wind turbine system. The relationship between wind speed and CF is nonlinear, and is calculated via different characteristic power curves at each wind plant location (see Table 3.2), and do not include electrical losses during the power generation process. Table 4.2 lists overall seasonal and annual CF differences at each location without using the clustering methodology. Percentage changes in the lowermost table are calculated with end-of-century CF minus historical CF, divided by historical CF, and written as a percentage change by multiplying 100. Overall, CFs are observed to increase in summer season (JJA), whereas winter (DJF) seasons exhibit a CF decrease. Here the overall seasonal trends from end-of-century during JJA and DJF are consistent with mid-century trends reported in Wang, Ullrich, and Millstein [1], but with an increased magnitude. CF changes based on the original wind fields (without bias correction) are given in section 3 in Table 4.3.

Our goal is to now explain the statistically significant CF changes observed in Table 4.2. In each of the following subsections we decompose the CF from each wind plant into the contribution from each cluster, and further decompose the change in CF into frequency changes and within-cluster changes following section 4.2.3.

Table 4.2: Historical seasonal and annual capacity factor (%) (upper table), absolute change in capacity factors (middle table), and percentage capacity factors changes under end-of-century comparing to historical (lower table) at each wind plant sites across California. Absolute changes are calculated with end-of-century CF minus historical CF. Percentage changes are calculated with end-of-century CF minus historical CF, divided by historical CF, and multiplied by 100 to write as percentages. Shiloh and Altamont Pass are located in NC domain, and the other three wind plants are in SC domain. All CF values are based on bias-corrected wind fields from VR-CESM.

Boldface indicates a percent change above the 95% significance level.

Wind plant	MAM	JJA	SON	DJF	Annual
Shiloh	33.45	50.41	30.60	27.47	35.53
Altamont Pass	23.84	40.67	19.22	14.11	24.52
Alta	44.43	40.02	34.25	38.75	39.38
San Gorgonio	19.87	23.59	12.70	11.77	17.02
Ocotillo	37.06	39.82	20.67	12.09	27.50

Wind plant	MAM	JJA	SON	DJF	Annual
Shiloh	+ 0.98	**+ 2.44**	**- 1.65**	**- 3.68**	- 0.46
Altamont Pass	**+ 1.63**	**+ 3.81**	+ 0.39	**- 1.36**	**+ 1.13**
Alta	- 1.54	+ 1.02	**- 5.29**	**- 3.67**	**- 2.35**
San Gorgonio	+ 0.10	**+ 1.91**	**- 1.32**	**- 2.14**	- 0.35
Ocotillo	+ 1.21	**+ 3.57**	- 1.33	- 0.47	+ 0.76

Wind plant	MAM	JJA	SON	DJF	Annual
Shiloh	+ 2.92%	**+ 4.84%**	**- 5.39%**	**- 13.39%**	- 1.29%
Altamont Pass	**+ 6.82%**	**+ 9.37%**	+ 2.04%	**- 9.65%**	**+ 4.62%**
Alta	- 3.46%	+ 2.54%	**- 15.44%**	**- 9.47%**	**- 5.98%**
San Gorgonio	+ 0.52%	**+ 8.09%**	**- 10.37%**	**- 18.14%**	- 2.04%
Ocotillo	+ 3.27%	**+ 8.97%**	- 6.42%	- 3.89%	+ 2.77%

Table 4.3: Historical seasonal and annual capacity factor (%) (upper table), absolute change in capacity factors (middle table), and percentage capacity factors changes under end-of-century comparing to historical (lower table) at each wind plant sites across California. Absolute changes are calculated with end-of-century CF minus historical CF. Percentage changes are calculated with end-of-century CF minus historical CF, divided by historical CF, and multiplied by 100 to write as percentages. Shiloh and Altamont Pass are located in NC domain, and the other three wind plants are in SC domain. Boldface indicates a percent change above the 95% significance level under t-statistics.

Wind plant	MAM	JJA	SON	DJF	Annual
Shiloh	15.51	23.62	14.61	15.23	17.26
Altamont Pass	10.40	20.17	8.36	6.17	11.31
Alta	24.50	19.45	17.80	20.51	20.57
San Gorgonio	8.21	9.89	5.01	4.77	6.98
Ocotillo	20.28	24.72	11.11	5.31	15.42

Wind plant	MAM	JJA	SON	DJF	Annual
Shiloh	+ 0.30	**+ 1.56**	**- 1.62**	**- 2.67**	- 0.60
Altamont Pass	+ 0.91	**+ 2.51**	+ 0.24	- 0.84	**+ 0.72**
Alta	**- 1.47**	+ 0.77	**- 3.81**	**- 2.23**	**- 1.68**
San Gorgonio	0.00	**+ 0.94**	- 0.57	**- 1.05**	- 0.16
Ocotillo	**+ 1.84**	**+ 3.00**	- 0.45	- 0.22	**+ 1.06**

Wind plant	MAM	JJA	SON	DJF	Annual
Shiloh	+ 1.92%	**+ 6.59%**	**- 11.12%**	**- 17.54%**	- 3.46%
Altamont Pass	+ 8.75%	**+ 12.46%**	+ 2.88%	- 13.55%	**+ 6.34%**
Alta	- 6.00%	+ 3.97%	**- 21.39%**	**- 10.89%**	**- 8.15%**
San Gorgonio	- 0.01%	**+ 9.51%**	- 11.42%	**- 22.06%**	- 2.36%
Ocotillo	**+ 9.06%**	**+ 12.14%**	- 4.06%	- 4.15%	**+ 6.85%**

Namely, we apply

$$\Delta CF = \sum_i \underbrace{CF_i \Delta f_i}_{(a)} + \underbrace{(CF_i^e - CF_i^h) f_i}_{(b)} + h.o.t., \qquad (4.5)$$

where CF_i^h and CF_i^e are the historical and end-of-century average CF for cluster i and $CF_i = (CF_i^h + CF_i^e)/2$. Here $h.o.t.$ denotes higher-order terms that are negligible in the decomposition.

NC JJA (Shiloh and Altamont Pass)

Both NC wind plant locations experience a significant increase in JJA CF, driven by essentially two factors. First, from Table 4.1 we see that there is a significant reduction in the frequency of low wind days (NC 4), and an accompanying increase in summertime MAP days (NC 6 and NC 8). Second, there is a significant increase in the wind speeds on MAP days (NC 6, 7, and 8), as explained in section 4.3.2 – in fact, the increase in wind speeds actually compensates for a reduced frequency of the NC 7 cluster of MAP days. Table 4.4 identifies the 6 clusters responsible for 98.1% and 98.6% of the historical wind energy production for Shiloh and Altamont Pass.

NC SON (Shiloh)

In accordance with Table 4.1, there is a decrease in the frequency of NC 1 and 3, associated with westerly wind and blocked offshore wind, and a compensating increase in the frequency of NC 6, 8, and 9, corresponding to MAP days and low southerly wind. As discussed in sections 4.3.2 and 4.3.2 inhomogeneity in the changing geopotential field has the further effect of reducing the wind speeds within the NC 1 and NC 3 clusters, further driving down CFs. Curiously, Altamont Pass does not experience a corresponding decrease in total CF, as historical CF at this wind plant during NC 1 and NC 3 days are much lower than NC 6 and NC 8 and so the

Table 4.4: Historical mean CF in select clusters (CF_i^h)(%), historical contribution to total seasonal CF ($CF_i^h f_i^h$), end-of-century CF change due to changes in cluster frequency (ΔCF (a)), and within-cluster change in wind speeds (ΔCF (b)) for the NC JJA season. Boldface in the (ΔCF (a)) column indicates clusters with significant change in frequency (see Table 4.1). Boldface in the (ΔCF (b)) column indicates a significant within-cluster change in CF at the 95% significance level obtained from t-statistics. The values in the "Total" row indicate how much total CF and CF change is attributed to this subset of clusters (compared to Table 4.2).

NC JJA (top 6 clusters)					
Cluster	Wind plant	CF_i^h	$CF_i^h f_i^h$	ΔCF (a)	ΔCF (b)
4	Shiloh	36.79	2.13	**- 1.55**	+ 0.12
	Altamont Pass	26.80	1.55	**- 1.16**	+ 0.14
5	Shiloh	53.71	8.45	- 0.46	**- 0.34**
	Altamont Pass	25.39	3.99	- 0.22	**- 0.12**
6	Shiloh	52.11	9.95	**+ 3.25**	+ 0.31
	Altamont Pass	49.27	9.41	**+ 3.10**	+ 0.44
7	Shiloh	47.51	12.96	**- 1.52**	**+ 0.80**
	Altamont Pass	52.10	14.21	**- 1.70**	**+ 1.32**
8	Shiloh	60.09	10.08	**+ 2.00**	**+ 0.05**
	Altamont Pass	38.12	6.39	**+ 1.34**	**+ 0.86**
10	Shiloh	45.58	5.87	+ 0.15	**+ 0.32**
	Altamont Pass	35.14	4.53	+ 0.11	**+ 0.09**
Total	Shiloh		49.45	+ 1.85	+ 1.27
	Altamont Pass		40.09	+ 1.47	+ 2.74

Table 4.5: As Table 4.4, except for NC SON.

NC SON (top 6 clusters)					
Cluster	Wind plant	CF_i^h	$CF_i^h f_i^h$	ΔCF (a)	ΔCF (b)
1	Shiloh	24.66	3.87	**- 1.12**	**- 0.73**
	Altamont Pass	16.39	2.57	**- 0.74**	**- 0.51**
2	Shiloh	38.15	2.67	- 0.39	- 0.35
	Altamont Pass	22.07	1.55	- 0.22	- 0.24
3	Shiloh	38.49	5.84	**- 2.33**	- 0.49
	Altamont Pass	13.76	2.09	**- 0.78**	- 0.36
6	Shiloh	37.67	4.42	**+ 1.42**	- 0.15
	Altamont Pass	33.97	3.98	**+ 1.30**	- 0.01
8	Shiloh	43.05	4.53	**+ 1.33**	- 0.32
	Altamont Pass	25.53	2.68	**+ 0.82**	+ 0.05
9	Shiloh	13.95	0.93	**+ 0.29**	- 0.03
	Altamont Pass	7.77	0.52	**+ 0.16**	+ 0.04
Total	Shiloh		22.27	- 0.80	- 2.06
	Altamont Pass		13.40	+ 0.53	-1.12

shifting cluster frequencies actually drive up average CF. Unlike the summer and winter seasons, the transitional fall and spring seasons do not feature a prominent subset of wind clusters. However, low wind days (NC 4) are much more likely to occur in the future during these seasons – we thus see that Shiloh is projected to see a decrease in CF in the fall. The breakdown of the contributions from the six most prominent clusters to Shiloh's CF is given in Table 4.5, which accounts for 72.8% of the wind energy production for this season. However, changes in these six clusters effectively explain the observed change in wind speed in this season.

NC DJF (Shiloh and Altamont Pass)

Both wind plants experience a significant decline in total CF over this season. The observed change can be largely attributed to a decrease in the frequency of NC 2 and NC 3 (strong westerly wind and blocked offshore wind), which have the highest average CF at Shiloh, and an increase in the frequency of NC 1, 4, and 9 clusters, which are each associated with lower wind speeds and CF. There is further a significant decrease in the wind speeds of cluster NC 2, the most frequent wintertime pattern, as described in section 4.3.2 to be attributed to higher overland pressures.

Table 4.6: As Table 4.4, except for NC DJF.

Cluster	Wind plant	CF_i^h	$CF_i^h f_i^h$	ΔCF (a)	ΔCF (b)
NC DJF (top 5 clusters)					
1	Shiloh	19.96	4.08	+ 0.16	- 0.29
	Altamont Pass	12.24	2.50	+ 0.10	- 0.07
2	Shiloh	48.93	11.98	- 0.62	**- 1.47**
	Altamont Pass	27.62	6.76	- 0.34	**- 1.14**
3	Shiloh	27.14	5.90	**- 1.05**	- 0.05
	Altamont Pass	8.54	1.85	**- 0.34**	+ 0.05
4	Shiloh	11.32	1.06	+ 0.16	**- 0.25**
	Altamont Pass	4.97	0.47	+ 0.08	**- 0.02**
9	Shiloh	19.07	3.72	+ 0.29	**- 0.74**
	Altamont Pass	10.12	1.98	+ 0.16	**- 0.07**
Total	Shiloh		26.74	- 1.06	- 2.80
	Altamont Pass		13.56	- 0.34	- 1.24

NC wintertime is associated with 5 clusters that describe 97.4% and 96.1% of total seasonal wind energy productions at Shiloh and Altamont Pass, respectively.

SC JJA (San Gorgonio and Ocotillo)

These two wind plants experience a pronounced increase in CF over this season attributed to two factors. First, a strengthening of the onshore flow (when it occurs) that leads to a reclassification of SC 2 days (weak onshore flow) to SC 4 and SC 7(onshore flow) days (Table 4.1). Second, an increase in the overall strength of SC 2 days when they do occur and SC 7 days, generally associated with an increase in onshore flow speeds associated with a stronger land/sea temperature gradient. The three clusters in Table 4.7 describe 97.1% and 96.9% of total JJA wind energy productions for San Gorgonio and Ocotillo, respectively.

SC SON (Alta and San Gorgonio)

Wind speeds are projected to decrease throughout the SC domain in the fall season leading to a significant decrease in CF at Alta and San Gorgonio. As observed in Table 4.8 this can be attributed to a widespread drop in wind speeds within essentially all clusters. This is accompanied by a significant drop in frequency of

Table 4.7: As Table 4.4, except for SC JJA.

SC JJA (top 3 clusters)					
Cluster	Wind plant	CF_i^h	$CF_i^h f_i^h$	ΔCF (a)	ΔCF (b)
2	San Gorgonio	19.15	8.78	- 1.33	+ 1.34
	Ocotillo	33.00	15.13	- 2.26	+ 1.89
4	San Gorgonio	32.99	10.16	**+ 1.73**	- 0.19
	Ocotillo	56.36	17.36	**+ 2.99**	+ 0.16
7	San Gorgonio	19.39	3.48	**+ 0.37**	**+ 0.01**
	Ocotillo	29.36	5.27	**+ 0.58**	**+ 0.40**
Total	San Gorgonio		22.42	+ 0.77	+ 1.15
	Ocotillo		37.76	+ 1.31	+ 2.45

Table 4.8: As Table 4.4, except for SC SON.

SC SON (top 7 clusters)					
Cluster	Wind plant	CF_i^h	$CF_i^h f_i^h$	ΔCF (a)	ΔCF (b)
1	Alta	61.71	8.20	**- 2.10**	- 0.45
	San Gorgonio	15.77	2.10	**- 0.56**	+ 0.03
2	Alta	38.25	8.19	+ 0.71	**- 1.08**
	San Gorgonio	11.75	2.51	+ 0.23	- 0.11
3	Alta	19.32	3.19	- 0.22	**- 0.71**
	San Gorgonio	4.89	0.81	- 0.06	**- 0.15**
6	Alta	43.08	4.49	**- 1.90**	- 0.05
	San Gorgonio	18.03	1.88	**- 0.74**	- 0.22
7	Alta	16.16	1.24	**+ 0.72**	+ 0.22
	San Gorgonio	7.03	0.54	**+ 0.32**	+ 0.12
8	Alta	40.18	1.98	- 0.09	**- 0.37**
	San Gorgonio	16.89	0.83	- 0.04	- 0.14
9	Alta	22.25	1.97	**+ 0.58**	**- 0.38**
	San Gorgonio	7.93	0.70	**+ 0.19**	**- 0.26**
Total	Alta		29.26	- 2.30	- 2.81
	San Gorgonio		9.37	- 0.66	- 0.72

SC 1 (strong alongshore winds) and SC 6 (Santa Ana winds) and accompanying increase in SC 7 (weak onshore wind) and SC 9 (low wind) – whereas SC 1 and SC 6 days correspond to the highest and third-highest CFs, SC 7 and SC 9 are the lowest and third lowest producers.

SC DJF (Alta and San Gorgonio)

As in the NC region, overland warming across SC leads to a widespread weakening of the within-cluster winds and a reduction in CF across the board. This process

Table 4.9: As Table 4.4, except for SC DJF.

SC DJF (top 6 clusters)

Cluster	Wind plant	CF_i^h	$CF_i^h f_i^h$	ΔCF (a)	ΔCF (b)
1	Alta	55.26	10.14	+ 0.54	+ 0.06
	San Gorgonio	13.97	2.56	+ 0.13	- 0.24
3	Alta	19.31	4.12	**+ 1.00**	- 0.48
	San Gorgonio	4.20	0.90	**+ 0.21**	- 0.19
5	Alta	43.82	4.67	- 0.65	**- 1.02**
	San Gorgonio	9.31	0.99	- 0.14	**- 0.21**
6	Alta	41.27	8.73	**- 1.23**	- 0.37
	San Gorgonio	18.27	3.86	**- 0.52**	- 0.44
8	Alta	39.31	5.06	**- 1.05**	- 0.39
	San Gorgonio	13.12	1.69	**- 0.32**	- 0.38
9	Alta	19.46	0.44	**+ 0.20**	-0.09
	San Gorgonio	3.48	0.08	**+ 0.04**	+ 0.03
Total	Alta		33.16	- 1.19	- 2.29
	San Gorgonio		11.22	- 0.60	- 1.43

further drives an increase in the frequency of SC 3 (low wind), which is associated with one of the lowest CF values, at the expense of SC 6 (Santa Ana winds) and SC 8 (westerly winds), which have among the highest CF values. There is further a substantial drop in the within-cluster wind speeds of SC 5 (southwesterly winds), as explained in section 4.3.2. Table 4.9 identifies the six clusters responsible for 85.6% and 85.7% of wind energy productions at Alta and San Gorgonio, respectively.

4.4   Discussion and Summary

This study utilized the state-of-the-art climate model CESM in its variable-resolution configuration to analyze California wind patterns change under the future climate. The agglomerative clustering algorithm was applied to the climate model output to group different weather patterns into separate clusters within the NC and SC domains. We defined ten wind clusters from each domain, and analyzed changes to within cluster wind speeds and also changes to the frequency of occurrence of each cluster by the end-of-century. Additionally, we analyzed the synoptic-scale patterns that accompany each cluster. The changes to these patterns can then be used to

identify some of the causes of changes to within cluster wind speeds. Moreover, some of these synoptic scale changes (e.g., changes to the land – sea temperature contrast) are directly tied to global warming, which allows us to tie a specific portion of the forecasted future change in wind resources directly to identified climate change phenomena. That being said, much of the forecasted change to wind resources is linked to changing frequency of weather patterns or clusters. The changes to frequency of each cluster type is tied to global circulation patterns, and possibly to climate modes and other teleconnections. Determining the specific mechanisms that cause the shifts to the cluster frequency is therefore out of scope within this study, but remains an intriguing target for future work.

Below we list the most important changes we observe to clusters by the end-of-century.

4.4.1 Northern California

Westerly winds (NC 1 and NC 2): These two clusters are among the most frequent winter season cluster, and have been projected to become less frequent with lower within cluster wind speed. The reduction in within-cluster wind speed is associated with the change in geopotential height field over the Pacific, and overland warming under the future climate. Both factors contribute to the decrease in within cluster wind speed.

Offshore blocking (NC 3): This is another wintertime cluster with a projected decreasing frequency and weaker within cluster wind speeds. The latter is related to the change in geopotential height pattern, driving a weaker northerly flow offshore, thus leading to weaker within cluster wind speeds.

Marine air penetration (NC 6-8 and NC 10): These clusters peak in frequencies during summertime. All have been projected to become more frequent with stronger within cluster wind speeds. The increase in within-cluster wind speeds is

123

associated with changes in the geopotential height pattern, which leads to a weakening of the offshore northerly wind, and promoting the onshore flow pattern. This increase in wind speeds contributes to the projected greater wind power during the summer season.

4.4.2 Southern California

Strong alongshore wind (SC 1): This cluster produced some of the highest capacity factors due to its frequent occurrences in all seasons only except summer, and its high within cluster wind speed. It has been projected to become more seasonally concentrated with its occurrences shifting from spring and fall to winter. For within-cluster wind speeds change, the change in the geopotential height field pattern reduces the alongshore gradient, leading to a weaker alongshore flow, and a decrease in wind speeds statewide.

Marine air penetration (SC 4): This cluster peaks in frequency during summertime. It has been projected to become more frequent with slightly increased onshore winds. The latter is caused by the increase in the geopotential height pattern which drives up wind speeds offshore, creating a better ventilation condition.

Santa Ana winds (SC 6): This is the second most frequent wintertime cluster, and has been projected to decrease in frequency with weaker within-cluster wind speeds. This reduction of the within-cluster wind speeds during Santa Ana events is associated with the weakening of the onshore ridge during end-of-century.

Weakened onshore flow (SC 7): This cluster is the third most frequent summertime cluster, with a projected increase in frequency. Under end-of-century, the geopotential height anomaly acts to strengthen the northerly wind offshore in Northern California, while blocks the offshore flow in Southern California.

Westerly wind (SC 8): This is a prominent cluster during winter and spring seasons, and its frequencies during these two season both decrease under end-of-century, along with weaker within cluster wind speeds. The latter is driven by large-scale dynamical changes that cause a weakening of wind speeds across California, including suppressed onshore flow in Southern California.

4.4.3   Changes in capacity factor

Along with changes to cluster frequency and within cluster wind speeds, we found statistically significant changes to energy generation (specifically to estimated capacity factor, or CF) at all wind plants.

There is an increase in the within cluster wind speeds during JJA driven by an increase land/sea temperature contrast and a subsequent tendency towards more frequent marine air penetration events for both NC and SC. This increasing frequency in marine air penetration events is accompanied by a frequency decrease from NC 4 (low wind) and SC 2 (weak onshore flow). Therefore, beside the within cluster wind speed increase, this frequency shift from low wind cluster to high wind clusters further contributes to the capacity factors increase during summertime.

This pattern is reversed in the winter season, with a smaller land/sea contrast that contributes to a decrease in within cluster wind speeds in both NC and SC. During the winter season, we observe an overland warming, that leads to an increase in the geopotential height field, and decrease in wind speeds statewide. The 700hPa geopotential height over Northern Pacific decreases in winter. This change in the general circulation also contributes to the wind speed decrease in winter. There is also a clusters frequency shift from high wind speed clusters to low wind speed clusters during winter season for both two domains (a frequency shift from NC 2 and NC 3 to NC 1, NC 4 and NC9 in the NC domain, and from SC 6 and SC 8 to SC 3 in the SC domain). So both the cluster frequency changes, and the within cluster wind speed changes contribute to the decrease in capacity factors during the

winter season.

The overall seasonal CF trends in JJA and DJF from the end-of-century were consistent with the trends from the mid-century [1], though the magnitudes of the changes are larger. Findings from this study are also consistent with the increasing frequency of marine air penetration events from Wang and Ullrich [31], decreasing wind speed during fall and winter seasons from Duffy et al. [5], and decreasing frequency of Santa Ana winds during early fall from Miller and Schlegel [3].

Bibliography

[1] Meina Wang, Paul Aaron Ullrich, and Dev Millstein. "The future of wind energy in California: Future projections with the Variable-Resolution CESM". In: *Renewable Energy* 127 (2018), pp. 242–257.

[2] SC Pryor and RJ Barthelmie. "Climate change impacts on wind energy: A review". In: *Renewable and sustainable energy reviews* 14.1 (2010), pp. 430–437.

[3] Norman L Miller and Nicole J Schlegel. "Climate change projected fire weather sensitivity: California Santa Ana wind occurrence". In: *Geophysical Research Letters* 33.15 (2006).

[4] Paul B Breslow and David J Sailor. "Vulnerability of wind power resources to climate change in the continental United States". In: *Renewable Energy* 27.4 (2002), pp. 585–598.

[5] Phillip B Duffy et al. "Climate Change Impacts on Generation of Wind, Solar, and Hydropower in California". In: *California Energy Commission (2014) CEC-500-2014-111.* (2014).

[6] Lejiang Yu et al. "Temporal and spatial variability of wind resources in the United States as derived from the Climate Forecast System Reanalysis". In: *Journal of Climate* 28.3 (2015), pp. 1166–1183.

[7] Dev Millstein et al. "Wind energy variability and links to regional and synoptic scale weather". In: *Climate Dynamics* (2018), pp. 1–16.

[8] Richard B Neale et al. "Description of the NCAR community atmosphere model (CAM 5.0)". In: *NCAR Tech. Note NCAR/TN-486+ STR* (2010).

[9] James W Hurrell et al. "The community earth system model: a framework for collaborative research". In: *Bulletin of the American Meteorological Society* 94.9 (2013), pp. 1339–1360.

[10] John M Dennis et al. "CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model". In: *The International Journal of High Performance Computing Applications* 26.1 (2012), pp. 74–89.

[11] Colin M Zarzycki et al. "Aquaplanet experiments using CAM's variable-resolution dynamical core". In: *Journal of Climate* 27.14 (2014), pp. 5481–5503.

[12] Alan M Rhoades et al. "Sensitivity of Mountain Hydroclimate Simulations in Variable-Resolution CESM to Microphysics and Horizontal Resolution". In: *Journal of Advances in Modeling Earth Systems* 10.6 (2018), pp. 1357–1380.

[13] Alan M Rhoades et al. "Characterizing Sierra Nevada snowpack using variable-resolution CESM". In: *Journal of Applied Meteorology and Climatology* 55.1 (2016), pp. 173–196.

[14] Xingying Huang et al. "An evaluation of the variable-resolution CESM for modeling California's climate". In: *Journal of Advances in Modeling Earth Systems* 8.1 (2016), pp. 345–369.

[15] P. A. Ullrich. "SQuadGen: Spherical quadrilateral grid generator". In: *University of California, Davis, Climate and Global Change Group software.* (2014). URL: http://climate.ucdavis.edu/squadgen.php.

[16] O Guba et al. "The spectral element method on variable resolution grids: Evaluating grid sensitivity and resolution-aware numerical viscosity". In: *Geosci. Model Dev. Discuss* 7 (2014), pp. 4081–4117.

[17] Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. "An overview of CMIP5 and the experiment design". In: *Bulletin of the American Meteorological Society* 93.4 (2012), p. 485.

[18] Alan M Rhoades, Paul A Ullrich, and Colin M Zarzycki. "Projecting 21st century snowpack trends in western USA mountains using variable-resolution CESM". In: *Climate Dynamics* 50.1-2 (2018), pp. 261–288.

[19] Xingying Huang and Paul A Ullrich. "The Changing Character of Twenty-First-Century Precipitation over the Western United States in the Variable-Resolution CESM". In: *Journal of Climate* 30.18 (2017), pp. 7555–7575.

[20] Iain Staffell and Stefan Pfenninger. "Using bias-corrected reanalysis to simulate current and future wind power output". In: *Energy* 114 (2016), pp. 1224–1239.

[21] Jon Olauson, Per Edström, and Jesper Rydén. "Wind turbine performance decline in Sweden". In: *Wind Energy* 20.12 (2017), pp. 2049–2053.

[22] Joe H Ward Jr. "Hierarchical grouping to optimize an objective function". In: *Journal of the American statistical association* 58.301 (1963), pp. 236–244.

[23] John A Hartigan and Manchek A Wong. "Algorithm AS 136: A k-means clustering algorithm". In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1 (1979), pp. 100–108.

[24] Ling Jin, Robert A Harley, and Nancy J Brown. "Ozone pollution regimes modeled for a summer season in California's San Joaquin Valley: A cluster analysis". In: *Atmospheric environment* 45.27 (2011), pp. 4707–4718.

[25] Neil Berg et al. "El Niño-Southern oscillation impacts on winter winds over Southern California". In: *Climate dynamics* 40.1-2 (2013), pp. 109–121.

[26] Sebastien Conil and Alex Hall. "Local regimes of atmospheric variability: A case study of Southern California". In: *Journal of Climate* 19.17 (2006), pp. 4308–4325.

[27] Francis L Ludwig, John Horel, and C David Whiteman. "Using EOF analysis to identify important surface wind patterns in mountain valleys". In: *Journal of Applied Meteorology* 43.7 (2004), pp. 969–983.

[28] Daniel S Wilks. "Cluster analysis". In: *International geophysics*. Vol. 100. Elsevier, 2011, pp. 603–616.

[29] Tadeusz Caliński and Jerzy Harabasz. "A dendrite method for cluster analysis". In: *Communications in Statistics-theory and Methods* 3.1 (1974), pp. 1–27.

[30] Hailan Wang and Siegfried D Schubert. "Causes of the extreme dry conditions over California during early 2013". In: (2014).

[31] Meina Wang and Paul A. Ullrich. "Marine air penetration in California's Central Valley: Meteorological drivers and the impact of climate change". In: *Journal of Applied Meteorology and Climatology* 2017 (2017).

[32] Scott Beaver and Ahmet Palazoglu. "Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality". In: *Journal of Applied Meteorology and Climatology* 45.12 (2006), pp. 1710–1726.

[33] Michael A Fosberg and Mark J Schroeder. "Marine air penetration in central California". In: *Journal of Applied Meteorology* 5.5 (1966), pp. 573–589.

[34] MN Raphael. "The santa ana winds of california". In: *Earth Interactions* 7.8 (2003), pp. 1–13.

[35] Anthony L Westerling et al. "Climate, Santa Ana winds and autumn wildfires in southern California". In: *Eos, Transactions American Geophysical Union* 85.31 (2004), pp. 289–296.

Chapter 5

Conclusions and Future Work

This thesis bridges the gap between the relatively well-understood future change of the synoptic-scale weather systems and the less well-studied climate change impact on the regional-scale. By studying a typical summertime wind regime, the MAP (Chapter 2), in the San Francisco Delta region, we project its future change by analyzing the localized meteorological features of MAP, and associate the change to the synoptic backgrounds. Then, we analyze the regional climate change impact from the perspective of California wind energy resources (Chapter 3). This study improves our understanding of the physical mechanisms that are driving the changes in wind resources under the future climate. Last but not the least, we expand our horizon from Chapter 3 by extending our analysis to the end-of-century (Chapter 4). A statistical method is developed to help distinguish different wind patterns, and have a detailed classification of future change in each wind cluster at each wind farm site. Below sections describe conclusions from each Chapter. Future work and potential research topics are also discussed.

5.1 Marine Air Penetration in California's Central Valley

5.1.1 Summary

In Chapter 2, we developed an objective criteria to identify marine air penetration (MAP) events in the San Francisco Delta from both observational and reanalysis data. This criteria includes a temperature difference along the Central Valley, which captures the cooling effect from MAP; and a onshore wind component, which measures the wind speed coming inland. We then characterized the synoptic-scale fields that are associated with MAP, and developed five large-scale indicators of MAP: a positive geopotential anomaly in the Gulf of Alaska; a negative geopotential tendency off the Oregon coast; an inland surface temperature anomaly in the northwest part of the continental United States; an offshore geopotential height difference; an offshore wind speed anomaly. With this set of large-scale feature, we built a logistic regression model to predict the occurence of MAP under future climate, and isolated its frequency trend. Based on the results from this probabilistic model, we detected a weak positive trend of 0.2 days/century under RCP8.5 of sufficiently strong summertime MAP days.

5.1.2 Future work

This study provides insight into the meteorological features associated with MAP events around the San Francisco delta. The suite of synoptic-scale features, accompanied by local meteorological measurements, can possibly be used to provide short-term forecast of MAP events. This short-term forecasting will be valuable for local electric power planning. Though the synoptic-scale features developed in this study may be tuned to the localized MAP events in the San Francisco delta, we argue that similar methodologies can be applied to study such events in other places around

the globe.

Possible future studies can focus on detecting factors that are important for the emergence of MAP but not investigated here, including but not limited to local meteorological conditions, large-scale teleconnections, coastal upwelling, and interactions with climate modes such as North Pacific Oscillation, Pacific Decadal Oscillation, and El Niño–Southern Oscillation. Furthermore, since fog along coastlines is often associated with sea breezes, interactions between the two meteorological features remains a topic for future studies.

5.2 The future of wind energy in California

5.2.1 Summary

In Chapter 3, our goal is to assessed the future change of wind field and its impact on wind energy. To approach this, we utilized the state-of-the-art climate model variable-resolution CESM. We first thoroughly evaluated VR-CESM model performance from the historical simulation against a high-resolution simulation from WRF, several widely used reanalysis datasets, and publicly available observations. Our comparison suggested that even though VR-CESM exhibits a bias towards slower wind speeds inland, it is able to capture the monthly climatology and the spatial patterns of wind fields. This led us to conclude that VR-CESM is correctly representing the key regional and synoptic-scale processes that are relevant for wind speed forecasts. Based on the model simulations, we then interpolated the wind speed to hub-height level at 80m, and analyzed the meteorological patterns associated with large-scale shifts in wind character. Significant seasonal changes in capacity factors were observed at almost all five wind farm locations during mid-century under RCP8.5, with an increase in summertime (JJA) resources and a decrease in fall (SON) and winter (DJF). Synoptic-scale drivers were also identified, and suggested climate change may favor synoptic patterns that lead to higher wind speed

during JJA, and lower wind speed during SON and DJF.

5.2.2   Future work

This study utilized the state-of-the-art climate model to improve the characterization of the future change in California wind energy resources till the mid-century. By analyzing the synoptic-scale meteorological patterns that are associated with large shifts in wind fields, the result improves our understanding of the physical mechanisms related to the trends in wind resource variability.

Besides the synoptic-scale patterns that were analyzed in this study, climate modes such as Pacific Decadal Oscillation, and El Niño–Southern Oscillation, North Atlantic Oscillation may have impact on changes in California wind resources as well, and will require further investigation. Intense and extreme winds can cause damage to wind turbines, and future study can focus on analyzing the capacity of models to capture future change of extreme winds. Forecasting trends and variations in wind energy using other statistical time series model, such as ARIMA or Prophet, remain a possible direction for future studies as well.

5.3   Clustering analysis of wind patterns in California

5.3.1   Summary

In Chapter 4, we expands the horizon from Chapter 2 and Chapter 3 by extending the time frame of our analysis to the end-of-century, and including a complete set of California wind patterns besides the marine air penetration. We utilized the same model as in Chapter 3 and extended the simulation time frame to the end-of-century (2080-2100). To distinguish different wind patterns from unlabeled modeling output, we applied the agglomerative clustering algorithm to 80m wind fields in

Northern and Southern California domains separately, which produced ten wind clusters for each domain. We then analyzed the analyzed the frequency changes of each cluster by the EoC and the synoptic-scale patterns that accompany each cluster. CF changes at all five wind farm locations under each cluster were also assessed. Our result suggests the overall seasonal CF at all five wind farms are increasing during summertime (JJA), and decreasing in winter (DJF), which is consistent with the trends in Chapter 3. Statistically significant CF changes are identified at all wind farms from different clusters.

5.3.2 Future work

This study provides a statistical method to identify different wind patterns from unlabeled data without requiring any prior knowledge of various wind types. The synoptic-scale analysis on wind clusters improves our understanding of the variability in space and time of California wind energy resources by end-of-century.

The statistical method in this study can be applied to future studies involving pattern detection from unlabeled data. Future work can focus on detecting the impact from various climate modes on wind energy resources in California. Other statistical models and machine learning algorithms can be applied for short-term wind energy forecasting from meteorological fields.