UNIVERSITY OF CALIFORNIA
Los Angeles

The Santa Ana Winds of Southern California in the Context of Fire Weather

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Atmospheric and Oceanic Sciences

by

Yang Cao

2015

Abstract of the Dissertation

The Santa Ana Winds of Southern California in the Context of Fire Weather

by

Yang Cao
Doctor of Philosophy in Atmospheric and Oceanic Sciences
University of California, Los Angeles, 2015
Professor Robert G. Fovell, Chair

The Santa Ana winds represent a high-impact weather event owing to the intimate relationship between the extremely dry, fast winds and the wildfire threat. The winds can be locally gusty, particularly in the complex terrain of San Diego county, where the airflow has characteristics of downslope windstorms. These winds can cause and/or rapidly spread wildfires, the threat of which is particularly acute during the autumn season before the onset of winter rains. It remains a day-to-day challenge to accurately predict wind gust speed, especially in the mountainous regions.

Our study employs large physics ensembles composed of high-resolution simulations of severe downslope windstorms that involve an exhaustive examination of available model physical parameterizations. Model results are calibrated and validated against the San Diego Gas and Electric (SDG&E) mesonet observations, a dense, homogenous, and well-positioned network with uniform high quality. Results demonstrate model horizontal resolution, model physics, random perturbations and landuse database can have a material effect on the strength, location and timing of Santa Ana winds in real-data simulations. A large model physics ensemble reveals the land surface model to be most crucial in skillful wind predictions, which are particularly sensitive to the surface roughness length. A surprisingly simple gust parameterization is proposed for the San Diego network, based on the discovery that this homogeneous mesonet has a nearly invariant network-averaged gust factor. The gust forecast technique is of special interest in the context of routine weather combined with atmospheric humidity and fuel moisture information.

A real-time wildfire threat warning system, the Santa Ana Wildfire Threat Index (SAWTI), has been developed to effectively communicate the upcoming Santa Ana wind strength with respect to the anticipated fire danger to first responders and the public. In addition to the wind and gust forecast techniques, attempts have been made to model two essential elements that SAWTI is heavily dependent on, i.e., the live fuel moisture and the greenness, using meteorological information. The models developed can skillfully determine these essential elements from both forecast and reanalysis data.

The dissertation of Yang Cao is approved.

Brian J. D'Agostino

Frederic P. Schoenberg

Qinbin Li

Jonathan L. Mitchell

Robert G. Fovell, Committee Chair

University of California, Los Angeles
2015

To my family and teachers

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgments

My deepest gratitude goes to my advisor, Professor Robert Fovell, for his guidance, inspiration and generous support in my research life. I could never thank him enough for his tireless effort in teaching me how to strengthen myself in both ability and personality and how to be an independent researcher. I have benefited a lot from his profound knowledge, physical ideas, and perpetual enthusiasm on scientific research and college education, and learned how to approach scientific problems from a variety of perspectives and be strong and patient when confronted with obstacles and failures in my research. He has set a perfect example for me as a scholar, teacher, and mentor.

I wish to express my sincere thanks to the rest members of my thesis committee, Brian D'Agostino and Professors Jonathan Mitchell, Qinbin Li, and Rick Schoenberg. I thank them for sharing with me their expertise and insights from different perspectives, for helping me clarify my thinking with their inspiring questions, and for their insightful comments and constructive criticisms at every stage of the dissertation. I would also like to extend my appreciation to all my other teachers at UCLA, especially Professors Kristen Corbosiero, Kuo-Nan Liou, Alex Hall, Yongkang Xue, Jim McWilliams, Yi Chao, Roberto Mechoso, David Neelin and Richard Thorne. They all have inspired and helped me in many ways. I also learned a lot from the discussions with my fellow colleagues Travis Wilson, Yizhe Bu, Hsien-Liang Tseng, Omar Nava, Seth Saslo, Chris Collimore, Huihui Yuan and Seung Hee Kim in the past years. Special thanks to Travis Wilson for sharing his MET codes.

The dissertation greatly benefited from collaborations with people outside the department. Specially, I greatly enjoyed collaborations with Tom Rolinski of the United States Forest Service's Predictive Services unit, San Diego Gas and Electric (SDG&E) meteorologists Brian D'Agostino and Steven Vanderburg, and Dr. Scott Capps of Vertum Partners. I would like to thank them for many helpful discussions and constructive suggestions.

Many thanks to all my dearest friends Luan Xu, Jianjing Kuang, Shasha Zhang, Frank Li, Peng Wang & Lunjin Chen, Junhong Liang, Ling Qi, Mei Gao, Lin Song & Xiang Zhai, Chao Yue, Ross Cheung, Daniel Walton, Fengpeng Sun & Aihua Zhu, Wen Li, Yu Gu, Qing Yue, Stephanie Steele and many other friends for their company in the past years and sharing all my pains and joys. We emotionally supported each other and celebrated each other's achievements, all of which are truly precious memories.

I cannot leave out my family, to whom I owe so much. I would like to sincerely thank my husband Wenbo Xin, my parents Xinwen Cao & Aiqiu Chu and brother Hui Cao for their boundless love and eternal support through all my good and hard times. I would also like to thank Professor Jeng-Gong Duh and Ms. Su-Yueh Tsai from Taiwan, for their encouragement and continuous support over many years. I am very grateful to Ms. Mei-Ying Fovell, who shares a lot of cherished moments with me, kindly invites me and my husband to the Thanksgiving dinner every year and always shows her great support.

This research was sponsored by San Diego Gas and Electric company. I would like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

Vita

2004-2008	Bachelor of Science in Agricultural Resources and Environment Nanjing University of Information Science and Technology, Nanjing, China.
2008-2010	Master of Science in Atmospheric and Oceanic Sciences University of California, Los Angeles, USA.
2009-2014	Graduate Research Assistant Department of Atmospheric and Oceanic Sciences University of California, Los Angeles, USA.
2011-2012	Teaching Assistant Department of Atmospheric and Oceanic Sciences University of California, Los Angeles, USA.
2014	Dissertation Year Fellowship University of California, Los Angeles, USA.

Publications and Presentations

Cao, Y., R. G. Fovell, and K. L. Corbosiero, 2011: Tropical cyclone track and structure sensitivity to initialization in idealized simulations: A preliminary study. *Terr. Atmos. Ocean. Sci.* 22, 559-578, doi:10.3319/TAO.2011.05.12.01(TM)

Cao, Y., and R. G. Fovell (August 2011), Tropical Cyclone sensitivity to initialization in semi-idealized simulations, **talk at 14th Conference on Mesoscale Meteorology 2011**, Los Angeles, CA

Cao, Y., and R. G. Fovell (September 2012), The Santa Ana winds of Southern California, **talk at Meeting with Chinese Civil Aviation Administration Delegation 2012**, Los Angeles, CA

Cao, Y., and R. G. Fovell (October 2012), The effect of initialization on Tropical Cyclone track, intensity and structure, **talk at Chinese-American Oceanic and Atmospheric Association Southern California Chapter Annual Workshop 2012**, Los Angeles, CA

Cao, Y., and R. G. Fovell (August 2013), Predictability and sensitivity of downslope windstorms in San Diego County, **talk at 15th Conference on Mesoscale Meteorology 2013**, San Diego, CA

Cao, Y., and R. G. Fovell (January 2014), Predictability and sensitivity of downslope windstorms in San Diego County, **talk at Big Weather Workshop at NCAR 2014**, Boulder, CO

Cao, Y., and R. G. Fovell (August 2014), Wind gust forecasting in complex terrain, **talk at 16th Conference on Mountain Meteorology 2014**, San Diego, CA

Cao, Y. (Oct 22nd, 2014), Predictability of the Santa Ana winds in the San Diego County, **talk at AOS department seminar, UCLA**, Los Angeles, CA

Cao, Y., R. G. Fovell, Downslope windstorms of San Diego County. Part I: A Case Study. *Mon. Wea. Rev.*, in preparation

Cao, Y., R. G. Fovell, Downslope windstorms of San Diego County. Part II: Physics ensembles and sensitivity tests. *Mon. Wea. Rev.*, in preparation

Cao, Y., R. G. Fovell, Downslope windstorms of San Diego County. Part III: Gust Forecasting. *Mon. Wea. Rev.*, in preparation

Fovell R. G., and **Y. Cao** (January 2013), Santa Ana Severity Index: Formulation, validation, and limitations, **poster at AMS 93rd Annual Meeting 2013**, Austin, TX

Fovell, R. G., Y. P. Bu, K. L. Corbosiero, W.-W. Tung, **Y. Cao**, H.-C. Kuo, L.-H. Hsu, and H. Su, 2015: Influence of cloud microphysics and radiation on tropical cyclone structure and motion: A review. *Michio Yanai Symposium Monogr., Amer. Meteor. Soc, in press*

Rolinski, T., R. G. Fovell, **Y. Cao**, S. B. Capps, B. J. D'Agostino and S. C. Vanderburg (August 2014), Santa Ana wildfire threat assessment, **poster at 16th Conference on Mountain Meteorology 2014**, San Diego, CA

Rolinski, T., S. B. Capps, R. G. Fovell, **Y. Cao**, B. J. D'Agostino and S. C. Vanderburg, The Santa Ana Wildfire Threat Index methodology explained. *Bull. Amer. Meteor. Soc.*, in preparation

CHAPTER 1

Introduction

In this chapter, we outline the problems to be addressed, provide background on their nature, and motivate our methodology.

1.1 The Santa Ana winds

Southern California (**Fig. 1.1**) is known for its "Santa Ana" winds, which were named after a city and canyon in Orange County. Santa Anas are very dry, sometimes hot, offshore winds directed from the Great Basin and Mojave Desert over the mountains and through the passes of Southern California [cf. *Small*, 1995, *Sommers*, 1978] that can produce gusts exceeding 45 m s^{-1} (100 mph) in favored areas[1]. The winds evince terrain amplification of the mountain gap and downslope varieties [*Huang et al.*, 2009, *Hughes and Hall*, 2010]. Santa Ana events occur most frequently between October and February, with December being the peak month [**Fig. 1.2**; *Raphael*, 2003]. Its season is typically thought of as extending from September through April [*Raphael*, 2003], although recent years (2013 and 2014) have seen events of significant strength during the month of May.

Although Santa Anas tend to form most frequently in midwinter, the most dangerous events often occur in autumn, before the winter rains have begun [*Sommers*, 1978, *Westerling et al.*, 2004]. At that time, the vegetation tends to be extremely dry, and fire

---

[1] Examples: On 21 October 2007, the weather station on Laguna Peak, overlooking Pt. Mugu, recorded a 50 m s^{-1} (111.5 mph) wind gust. More recently, on 30 April 2014, a station in San Diego County (Sill Hill, SILSD) reported a 45 m s^{-1} (101 mph) gust, and remained above 40 m s^{-1} (90 mph) for a total of five nonconsecutive hours.

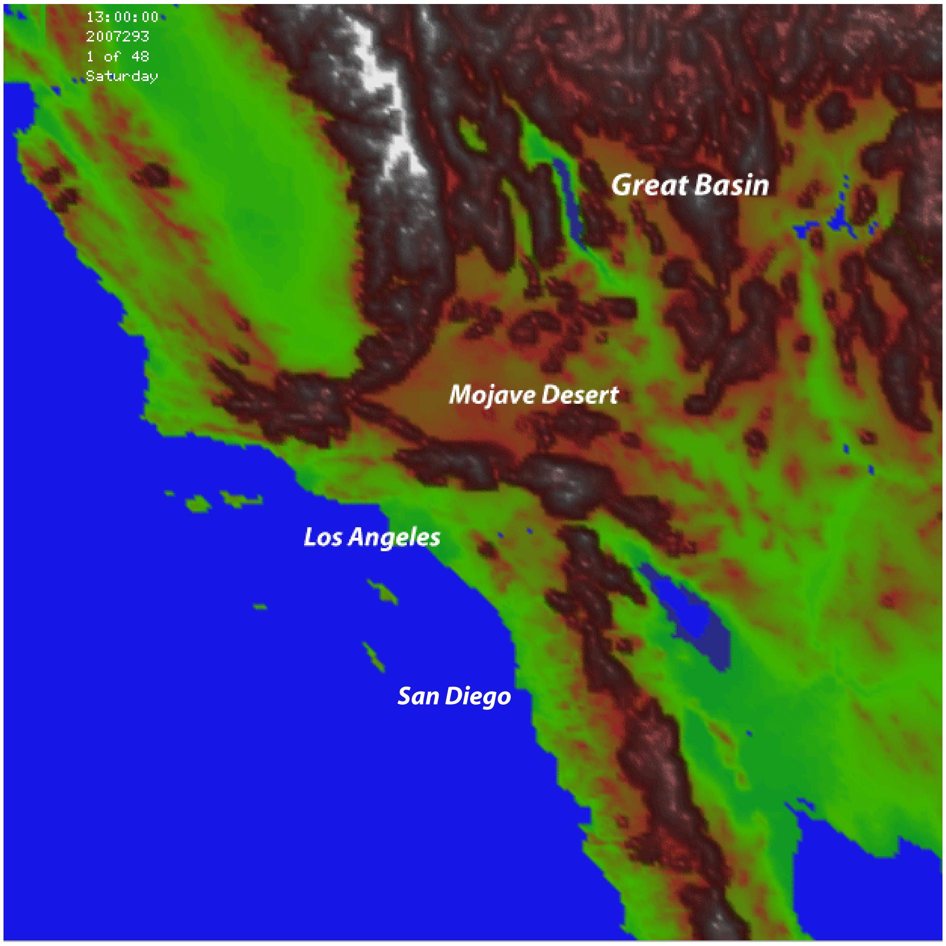


Figure 1.1: Southern California topography.

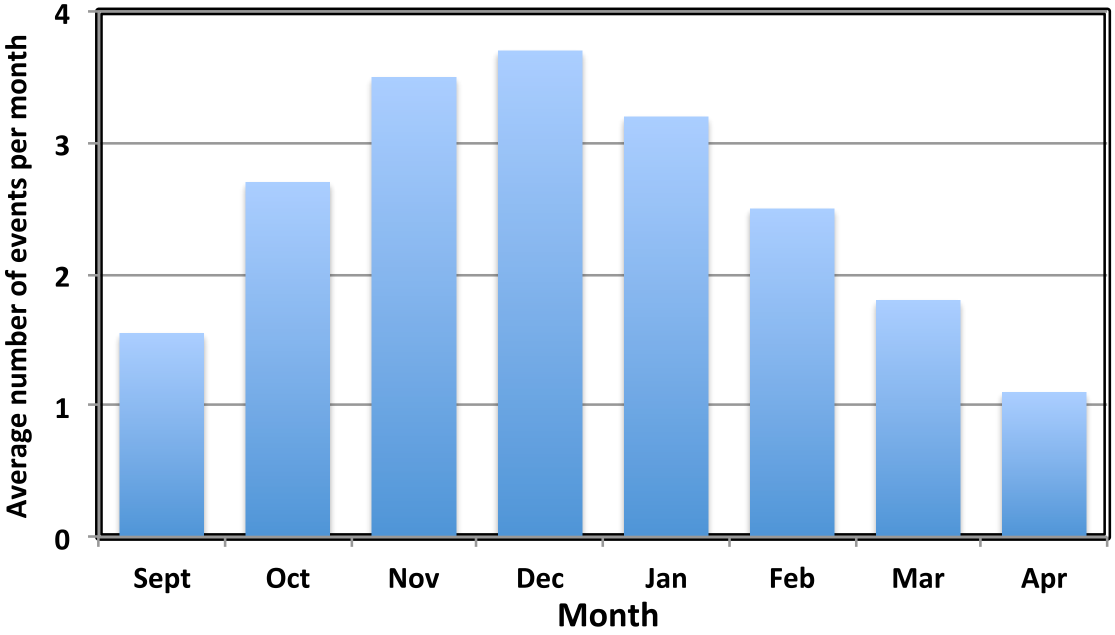


Figure 1.2: Average number of occurrences of Santa Ana events per month [reproduced from *Raphael*, 2003, Fig. 3].

danger is elevated owing to the combination of low-to-very low humidity and strong winds that can spark and spread flames. Autumn fires historically have the potential to be very large in area, being fanned by the Santa Ana winds, which shown rather strikingly in **Fig. 1.3** from *Chang and Schoenberg* [2011]. Each dot on the figure represents a Los Angeles county fire that took place during a multi-year period, plotted with respect to month and (a logarithmic function of) fire size. Note that while fires are most common during the summer months, very large fires are more likely occur in the autumn.

Santa Ana events result when cooler air spills across the Great Basin, becoming partially dammed by the mountains that separate Southern California from the inland deserts. This increases the horizontal gradient in sea-level pressure (SLP) and helps enhance flow speeds through prominent terrain gaps such as the Cajon Pass (leading to Santa Ana) and through the Soledad Gap (northwest of Los Angeles; see **Figure**

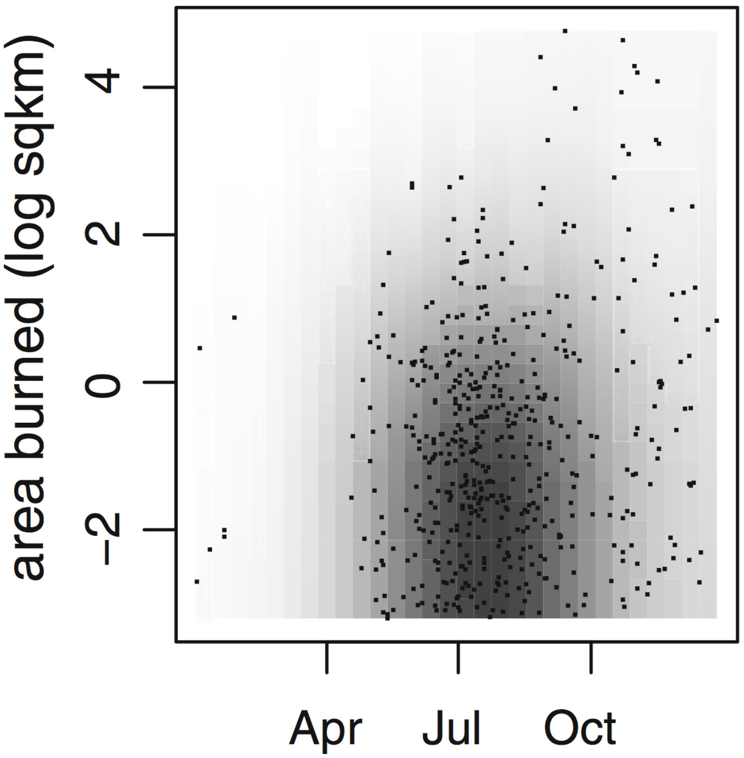


Figure 1.3: Area burned vs. month of year. Points on the plots indicate the 513 wildfires of at least 0.0405 km^2 (10 acres) in the Los Angeles County Fire Department dataset between January 1976 and December 2000 [from *Chang and Schoenberg*, 2011].

1.1), creating prominent wind corridors in the northern part of the Los Angeles basin [*Jackson et al.*, 2013]. Wind speeds can also be very large in San Diego County, where the terrain gaps appear less prominent but mountain heights are also generally lower. We will see that in this part of Southern California, the flow across the topography shares many characteristics of classic downslope windstorms.

Downslope windstorms are a type of large amplitude mountain wave that can produce strong, often gusty winds on the lee side of a mountain barrier [*Durran*, 2003]. Subsidence can cause very low relative humidities near the surface, particularly if the air mass starts with low absolute humidity. The necessary ingredients for downslope windstorms are a sufficiently large mountain barrier, as well as strong cross-barrier winds and a stable atmosphere, both near the mountaintop level [*Jackson et al.*, 2013]. Downslope windstorms are observed in many areas of the world, and carry such names as the Bora, Chinook, Foehn, Zonda and Taku winds [*Durran*, 2003, *Schamp*, 1964].

In complex terrain, the wind can vary greatly over small distances and gustiness is common in downslope windstorms, which may be caused by rotors and subrotors embedded in the flow [*Doyle and Durran*, 2004, *Jackson et al.*, 2013], as illustrated in **Fig. 1.4** from [*Doyle and Durran*, 2004]. Terrain-amplified winds and gusts can knock down trees and power lines, starting and spreading fires, making accurate forecasts in this region extremely important. Previous studies have demonstrated that a wide range of behaviors can result from relatively subtle changes in environmental conditions, as illustrated in **Fig. 1.5**, which combines information from *Durran* [1986] and *Vosper* [2004].

Proper model validation, however, can be hampered by the sparseness of the surface network, the absence of stations in wind-prone areas, as well as deficiencies in anemometer placement. As an example, on 21 October 2007, the Witch Creek fire was sparked by wind-whipped power lines located about 20 m AGL, and was driven by an especially strong Santa Ana winds to become one of the largest fires in California his-

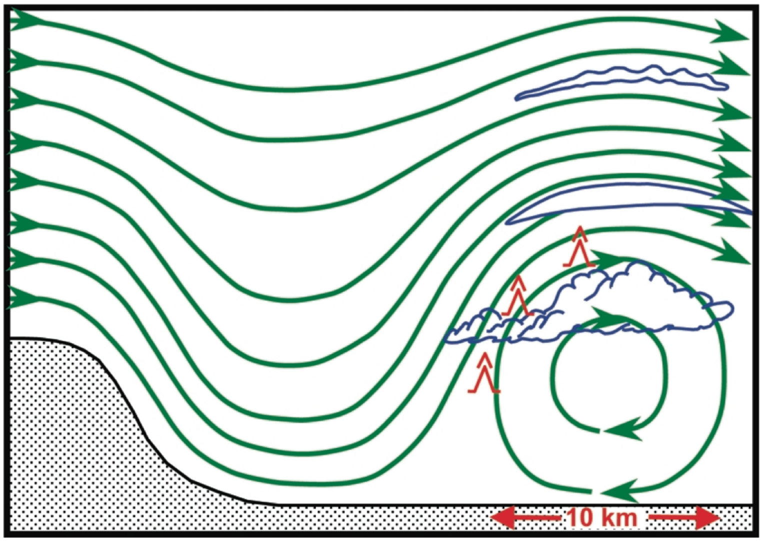


Figure 1.4: Schematic streamlines illustrating a rotor circulation and attendant cloud features (adapted from Ludlam and Scorer 1957). Regions of clear-air turbulence associated with the rotor circulation are denoted by the red symbols [Figure and caption from *Doyle and Durran*, 2004].

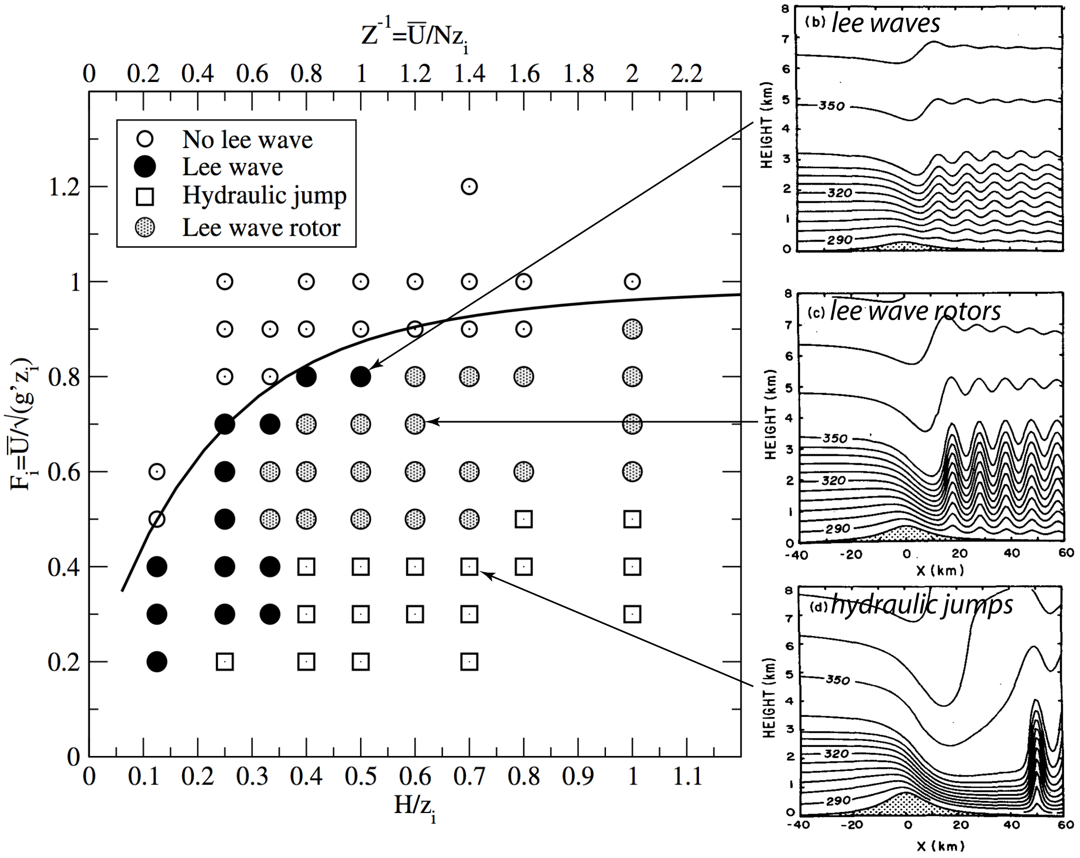


Figure 1.5: The regime diagram showing the F_i and H $/z_i$ dependence of the flow when the no-slip lower boundary condition is applied. The non-dimensional parameters F_i = U/NL, NH/U and S = H/L are fixed at 0.08, 0.5 and 0.04, respectively [left panel, figure and caption from *Vosper*, 2004]. Also shown (solid curve) is the linear prediction of the maximum Froude number for which a steady trapped lee-wave field is present on the inversion. Isentropes for the air in a two-layer atmosphere flowing over a 600 m high mountain are showing (a) lee waves, (b) lee wave rotors, and (c) hydraulic jumps. The airflow is from left to right [right panel, adapted from *Durran*, 1986].

tory[2]. It is certain that the meteorological stations that existed at the time did not fully capture the ferocity of the winds experienced at the initiation site of that or other fires that started during this windstorm.

1.2 The Santa Ana Wildfire Threat Index

For many years, fire warnings in Southern California have been communicated to first responders and the public in the form of "red flag warnings" issued by the National Weather Service. Although the triggering criteria varies from place to place, red flag warnings are issued in Southern California when relative humidities are expected to be low ($< 15\%$) and wind gusts strong (> 35 mph or 15.6 m s^{-1}) for more than three hours[3]. The state of the available "fuel" for fires may also be taken into account. However, the red flag warning is a "one size fits all" categorization, with all such warnings ostensibly being created equal.

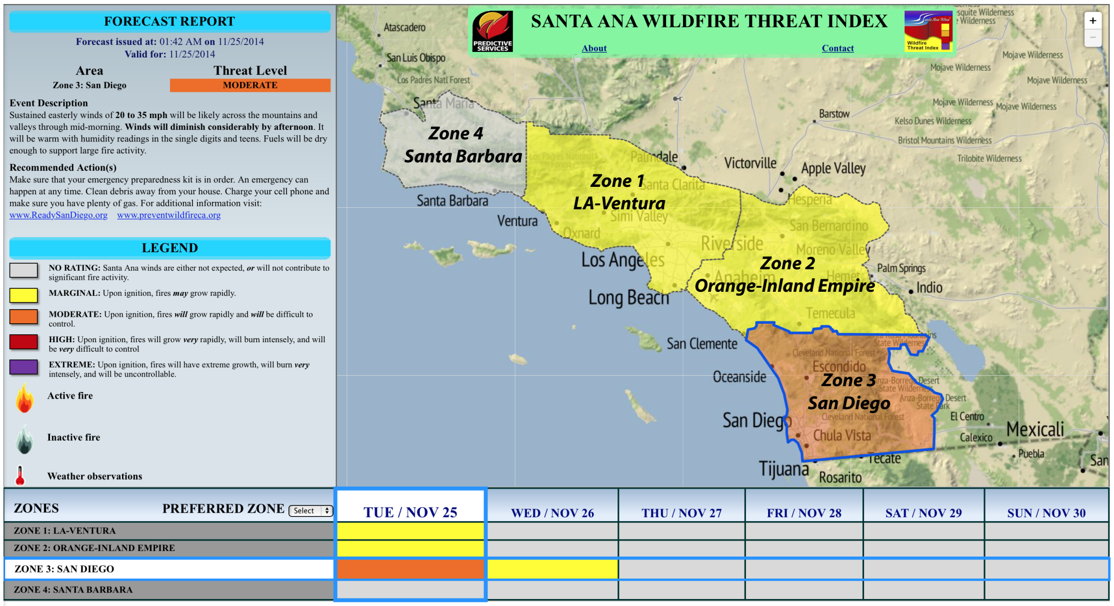
To quantify the fire danger posed by Santa Ana winds, a threat index was recently proposed by Tom Rolinski of the United States Forest Service's Predictive Services unit, in collaboration with San Diego Gas and Electric (SDG&E) meteorologists Brian D'Agostino and Steven Vanderburg. Termed the Santa Ana Wildfire Threat Index (SAWTI), the product provided to first responders and the public consists of a scale consisting of four categories, denoted "marginal", "moderate", "high" and "extreme" depending on conditions[4]. SAWTI also has a "no rating" classification that is issued when either Santa Ana winds are not expected or winds are not anticipated to contribute significantly to fire behavior. The specific threat level is customized for each of four areas in Southern California, extending from Santa Barbara to San Diego (**Fig. 1.6**). For

---

[2] According to information obtained from the California Department of Forestry and Fire Protection (Cal Fire), the Witch Creek fire was the 3*rd* largest California wildfire since 1932 upon its containment, and is ranked 6*th* largest as of this writing.

[3] Information from the National Interagency Fire Center website, http://gacc.nifc.gov/oscc/predictive/weather/myfiles/Watches_and_Warnings_for_California.htm

[4] http://santaanawildfirethreat.com



FORECAST REPORT

Forecast issued at: 01:42 AM **on** 11/25/2014
Valid for: 11/25/2014

Area	**Threat Level**
Zone 3: San Diego	MODERATE

Event Description
Sustained easterly winds of **20 to 35 mph** will be likely across the mountains and valleys through mid-morning. **Winds will diminish considerably by afternoon**. It will be warm with humidity readings in the single digits and teens. Fuels will be dry enough to support large fire activity.

Recommended Action(s)
Make sure that your emergency preparedness kit is in order. An emergency can happen at any time. Clean debris away from your house. Charge your cell phone and make sure you have plenty of gas. For additional information visit:
www.ReadySanDiego.org www.preventwildfireca.org

LEGEND

⬜	**NO RATING:** Santa Ana winds are either not expected, ***or*** will not contribute to significant fire activity.
🟨	**MARGINAL:** Upon ignition, fires ***may*** grow rapidly.
🟧	**MODERATE:** Upon ignition, fires ***will*** grow rapidly and ***will*** be difficult to control.
🟥	**HIGH:** Upon ignition, fires will grow ***very*** rapidly, will burn intensely, and will be ***very*** difficult to control
🟪	**EXTREME:** Upon ignition, fires will have extreme growth, will burn ***very*** intensely, and will be uncontrollable.
🔥	**Active fire**
🔥	**Inactive fire**
🌡	**Weather observations**

ZONES PREFERRED ZONE Select	TUE / NOV 25	WED / NOV 26	THU / NOV 27	FRI / NOV 28	SAT / NOV 29	SUN / NOV 30
ZONE 1: LA-VENTURA	🟨	⬜	⬜	⬜	⬜	⬜
ZONE 2: ORANGE-INLAND EMPIRE	🟨	⬜	⬜	⬜	⬜	⬜
ZONE 3: SAN DIEGO	🟧	🟨	⬜	⬜	⬜	⬜
ZONE 4: SANTA BARBARA	⬜	⬜	⬜	⬜	⬜	⬜

Figure 1.6: Santa Ana Wildfire Threat Index real time forecast issued at 01:42AM on 11/25/2014. A moderate Santa Ana wildfire threat level was shown in Zone 3 and a marginal threat level was shown in Zone 1 and 2 on that day.

historical reasons, Zone 1 consists of Ventura and Los Angeles counties; Orange county and the portions of Riverside and San Bernardino counties west of Palm Springs fall into Zone 2; Zone 3 mainly consists of San Diego county; and Santa Barbara is Zone 4.

While the disseminated product is categorized, SAWTI actually employs a numerical function called large fire potential (LFP), which is computed using this equation:

$$LFP = LFP_w \times FC = \left(\frac{1}{1000} windspeed^2 \times Tdd\right) \times \left\{\frac{1}{10}\left[\left(\frac{DL}{LFM} - 1\right) + G\right]\right\}^{1.7}. \quad (1.1)$$

LFP consists of two components, representing the state of the weather (LFP_w) and fuel conditions (FC). LFP_w is proportional to the dew point depression (Tdd, the spread between the temperature and dew point) multiplied by the wind speed squared. This captures the weather threat represented by low humidities and fast winds, principal components of the red flag warning. The wind speed is squared to increase its impact on the resulting equation, and to mimic kinetic energy. The data used for LFP_w comes from weather stations and/or numerical model output simulating conditions near the surface.

The fuel conditions component is a rather complex function of three factors, the *dryness level* (DL), *live fuel moisture* (LFM) and *greenness* (G). The DL term incorporates the *energy release component* (ERC) [*Bradshaw et al.*, 1984], which represents the amount of heat per unit area generated by a fire, combined with dead fuel moisture (DFM) to produce an integer scale with three values (1-3, inclusive), with higher levels representing drier conditions. DFM represents the moisture content of dead, woody vegetation, which responds to changing environmental conditions, such as temperature, humidity, available solar radiation, and precipitation, depending on its thickness [*Nelson Jr*, 2000]. DFM can be computed using sticks that are weighed to determine their moisture content, or (as in the SAWTI) using the "Nelson model" [*Carlson et al.*, 2007, *Nelson Jr*, 2000], which consists of equations describing the heat and moisture balance for woody materials of various diameters.

With respect to LFM, the moisture content of live vegetation, the SAWTI is concerned specifically with the moisture of chamise (Adenostoma fasciculatum), a member of the chaparral family that is sensitive to drought conditions [*Dennison and Moritz*, 2009, *Dennison and Roberts*, 2003]. Chamise is a common sight in the wooded areas of Southern California and its LFM is determined by weighing plant samples before and after drying. Greenness represents the state of grasses and other ground cover. Other factors being equal, the greener the vegetation, the slower the fire spread might be, because the moisture acts as a heat sink [*Kożuchowski and Żmudzka*, 2002, *Rothermel*, 1972].

In the LFP equation, G is represented by integer values between 0 and 5, with the maximum value actually representing the least green conditions. DFM and LFM are both expressed as percentages and the FC component as a whole is scaled so it varies between 0 and 1. Thus, fuel conditions can be seen as acting as a check on the weather component of the threat. When grasses are brown, LFM is low, and dead fuels are also dry, the FC term is close to one, which leaves the weather portion of the index

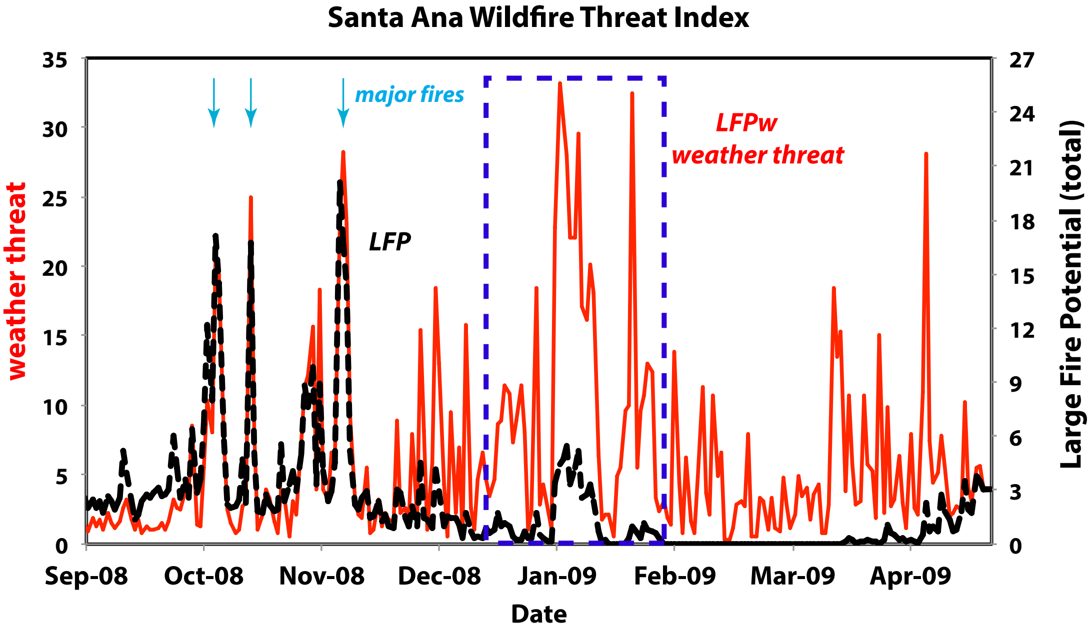


Figure 1.7: Validation of Large Fire Potential (LFP) against observed fire activities between September 2008-April 2009. LFP_w (red line) is the weather component of the total LFP (dotted black line). LFP_w values are scaled by a factor of 0.001.

undiminished. However, when fuels are moist, even moderately severe Santa Ana wind conditions are not expected to spread fires rapidly.

Figure 1.7 presents a time series of LFP (dashed black line) and its weather component (red line) for the period from September 2008 through April 2009. These data were computed from meteorological, satellite, on site sampling information and the Nelson model, and represents conditions existing in the aforementioned Zone 1. LFP_w exhibits several spikes during the period as Santa Ana wind events developed and dissipated. The threat peaked during a Santa Ana episode in January, with LFP_w values reaching just short of 35. This is a relatively high value, and by itself represents a significant weather threat.

When fuel conditions are factored in, LFP remained high during the October and November events. However, LFP values were very small after that time, a consequence of significant precipitation that raised fuel moisture levels. Naturally, if there is no ignition, there will be no fire, no matter how large the fire threat is. As it happened, ignitions that developed into major fires did occur during the first three LFP spikes, including the Sesnon, Foxborough, Sayre, Tea, and Freeway fires. In contrast there were no large fires during the extended period of elevated LFP_w in January, despite five recorded ignitions.

1.3 Ensemble forecasting

Ensemble Forecasting (EF), opposite to the commonly known deterministic forecasting, is fumdamentally stochastic in nature. Instead of making a single simulation, EF performs multiple parallel forecasts or simulations that differ with respect to their initial conditions, or/and numerical model configurations, consisting of different model parameter values, parameterization types and/or multi-model systems. The underlying assumption of EF is that the subsequent sets of forecasts may be taken as a representative random sample from the evolved Probability Distribution Function (PDF) [*Hamill and Snyder*, 2000]. EF systems provide an objective way to estimate uncertainty in weather and climate forecasts. Since the early 1960s, it has been known that the ensemble average of a number of forecasts is generally more accurate than the forecast from a skilled individual forecaster [*Bosart*, 1975, *Gyakum*, 1986, *Sanders*, 1963, 1973]. EF aims to improve the forecast through ensemble averaging, to provide an evaluation of the reliability of the forecasts, and to provide a quantitative basis for probabilistic forecasting.

Typically, the ensemble forecasting system includes the following four elements [*Kalnay*, 2003]:

1. the *true evolution* of the atmosphere (unknown in real time);

2. the *control forecast* starting from the analysis (from the best estimate of the initial

state of the atmosphere);

3. *perturbed ensemble forecasts* with initial perturbations added and/or subtracted from the control;

4. the *ensemble average*.

It is reasonable to assume the control forecast is the most accurate because it has been optimized in terms of the initial condition state, model physics parameterization tuning and numerics [*Warner*, 2011]. The spread (variance) among the ensemble members can be an indication of the flow-dependent quantitative uncertainty in the ensemble forecast. The PDF of the frequency distribution of a spread can provide information about extreme events, which is very useful from a practical point of view. Ensemble forecasting was implemented for operational use in 1992, at both the National Centers for Environmental Prediction (NCEP) and the European Centre for Medium-Range Weather Forecasts (ECMWF) [*Molteni et al.*, 1996, *Palmer et al.*, 1993, *Steven Tracton and Kalnay*, 1993, *Toth and Kalnay*, 1993, 1997].

The forecast error can be divided into two groups: the *initial condition error* and *model error*. Even if initial condition errors are accurately accounted for, ensemble forecasts tend to be underdispersive and underestimate the true uncertainty of the atmospheric evolution [*Buizza et al.*, 2005]. A second major contributor to forecast uncertainty is attributed to model errors, such as parameter uncertainties, parameterization deficiencies, and improperly represented subgrid-scale processes. Inclusion of model perturbations is not intended to add artificial spread to an underdispersive ensemble, but to introduce actual uncertainty to the ensemble that was previously omitted.

The model errors have been proposed to fall into two distinct classes, i.e., systematic and stochastic [*Hamill et al.*, 2000]. Systematic error refers to model bias originating from poorly tuned parameterizations, while stochastic error is the remainder of the forecast error that is random in nature, and is referred to as “model uncertainty”.

Previous study suggested several possible approaches to the representation of model errors, including stochastic-dynamic models [*Epstein*, 1969], physics tendency perturbations [*Buizza et al.*, 1999], multiple physics schemes [*Murphy et al.*, 2004], parameter variations in the physics packages [*Stainforth et al.*, 2005], the multi-model ensemble or the super-ensemble [*Hagedorn et al.*, 2005, *Krishnamurti et al.*, 2000]. These methods have been shown to contribute to more reliable probabilistic forecasts.

The chief advantage unique to the stochastic perturbation approach is that each ensemble member has the same climatological and model bias, in contrast to multiparameter, multiparameterization, and multimodel ensembles in which each member is essentially a different model with its own dynamical attractor [*Berner et al.*, 2011]. In this study we implemented a stochastic kinetic-energy backscatter (SKEB) scheme [*Shutts*, 2005] to capture the "internal" model error component of the forecasting system. The SKEB scheme aims at representing model uncertainty from interactions with unresolved scales.

As we know, the atmosphere is a multi-scale system, and the stochastic backscattering of subgrid-scale fluctuations [*Mason and Thomson*, 1992] was originally developed in the context of large-scale simulations. The turbulent dissipation rate is the difference between upscale and downscale spectral transfer, with the upscale component being the kinetic energy source to the resolved flow. These concepts were adapted to the ECMWF ensemble forecast system by *Shutts* [2005]. The input of small-scale kinetic energy, by the backscatter algorithm also helps to correct a known problem with the energy spectrum in the ECMWF model - the absence of the observed $-5/3$ spectral slope in the mesoscale portion of the spectrum *Shutts* [2005].

Berner et al. [2009] found improved probabilistic skill for medium-range forecasts up to 10 days following implementation of a SKEB scheme in the ECMWF global ensemble system. *Berner et al.* [2011] simplified the SKEB of *Berner et al.* [2009] and implemented it into version 3.3 of the Weather Research and Forecasting (WRF) model, assuming a spatially and temporally constant dissipation rate. Another difference between *Berner et al.* [2011] and *Berner et al.* [2009] is that the former followed the argument of *Shutts* [2005] who claimed that the energy in the subgrid-scale should be backscattered onto not only the horizontal velocity components (u and v), but also the temperature field. *Berner et al.* [2011] tested this simplified SKEB scheme and determined it outperformed the multi-physics scheme (or what we call the "physics ensemble" in this dissertation), except near the surface. They concluded that model error can be best captured when both schemes are used simultaneously.

1.4 Organization of this thesis

One of the goals of our work was to contribute to the SAWTI by making it possible to compute the index from forecast and reanalysis model outputs. In contrast to the DL component of FMC, there is no available equation or model for LFM or G, at least one that is readily and easily applicable to Southern California conditions. Yet, such models are needed not only to make forecasts for the future, but also to reconstruct past events, so that a historical record of LFP can be constructed. Thus, we have developed equations for LFM and G that employ observed or simulated meteorological information to make skillful estimates of these important factors for the past, present or future. The statistical models we developed for these FMC components are presented in **Chapter 6**.

Our major effort, however, is to assess how predictable the Santa Ana winds are in the San Diego area region and how skillfully a regional-scale weather prediction model can forecast the winds and especially the gusts that they cannot even resolve. For this part, we examine the skill of the WRF model's Advanced Research WRF (WRF-ARW) [*Skamarock et al.*, 2008] core in forecasting Santa Ana winds in San Diego County. Model forecasts are validated and calibrated against sustained wind observations reported by the SDG&E mesonet, a recently-installed surface observing network of more than 140 stations sited primarily in well-exposed, wind-prone areas on the west-facing slopes of the county's mountains. This is one of the highest-density surface mesonets in the world.

After discussing the model setup and initialization procedure in **Chapter 2**, we examine a specific case study in **Chapter 3** that occurred in February 2013. During that event, wind gusts exceeding 40 m s^{-1} (90 mph) were recorded in the SDG&E network. We will show that the WRF model can do a good job of capturing the spatial and temporal variation of the winds observed during that multi-day episode, although it does better at some stations and in some areas than others. The model provides the important third dimension (height) that is completely missing in a surface-based observing network.

The WRF model has an enormous large number of potential model configurations, as a consequence of the numerous physical parameterization options ("model physics") that are available. Physical processes requiring parameterization include land surface and subsurface (soil) processes, mixing in the planetary boundary layer, radiative transfer, cloud microphysical processes, and the influence of subgrid turbulence and cloud activity. There are several viable options for each of these processes, and some are undoubtedly better than others for this application. The model acquires information regarding surface and vegetation conditions from external databases, and this choice also influences model skill, as shown in **Chapter 4**. Model resolution and numerical choices (such as horizontal diffusion and incorporating the effect of subgrid scale topography) are also considered.

The ultimate goal of this work is to forecast wind gusts, the short-period bursts that cause most of the wind damage. As models like WRF cannot resolve gusts when operated at reasonable spatial and temporal resolutions, a strategy for parameterizing these dangerous bursts is needed. In **Chapter 5**, we present a surprisingly simple but effective wind gust parameterization, motivated by the unique aspects of the SDG&E mesonet. Finally, in **Chapter 7**, we provide a brief summary of this dissertation.

CHAPTER 2

Data sources and modeling strategy

In this chapter, we describe the numerical model we will use for our study in **Chapters 3, 4 and 5**, how it will be initialized and configured, our source of observations for validation, and issues that need to be addressed to insure a fair and successful validation.

2.1 Available observations

Observations are crucial for vetting a numerical model, but there are several significant challenges involved. First of all, most of the information available for validation is located very close to the surface. Even this information has historically been relatively sparse and, worse still, each network tends to measure the wind differently with respect to sensor hardware, mounting height, intervals employed for sampling, averaging and reporting, and station sitting philosophies. All of these can dramatically impact the magnitudes of winds and gusts that are reported, complicating the validation process. As an example, most Automated Surface Observing System (ASOS) stations report winds measured at the World Meteorological Organization (WMO) standard height of 10 m [*WMO*, 2010] above ground level (AGL), and averaged over a two-minute period, with data available at one-minute intervals. The mesoscale model we will use (described presently) provides a wind diagnostic for this height, which typically resides between the lowest model level and the surface. However, most measurements in complex terrain come from Remote Automated Weather Stations (RAWS) that have anemometers mounted closer to the surface (at 20 ft. or about 6.1 m AGL) and transmit ten-minute average winds once per hour (leaving over 80% of the hour unsampled for the sustained wind). Thus, regardless of other factors, contemporaneous and co-located RAWS and ASOS reports can be expected to disagree. In validation exercises, adjustments dependent on vertical stability and surface roughness have to be made to the model's standard 10-m wind diagnostic to avoid a false conclusion of overprediction.

Furthermore, the unfortunate reality is that some anemometers are improperly shielded by obstacles such as buildings and trees, or simply were not installed in the areas of greatest wind and/or hazard. Even a cursory examination of RAWS site photos hosted by the Desert Research Institute (DRI)[1] reveals numerous examples of problematic anemometer placement. During the Witch Creek Fire of October 2007 (**Chapter 1**), the RAWS station at Goose Valley (GOSC1) occupied an important location immediately downwind and downslope from the ignition location, but at the time was closely surrounded by significant obstacles (verified by inspection). It is not known how much larger its event maximum sustained wind (15 m s^{-1}) and gust (25 m s^{-1}) might have been had the station not been sited close to large trees.

Since 2009, SDG&E has deployed over 140 stations in wind-prone areas across San Diego County (**Fig. 2.1**). These stations conform to the RAWS standard[2] with respect to anemometer height (20 ft. or about 6.1 m AGL) and wind averaging interval (10 min averages from 3 sec samples), but report every 10 min instead of hourly. Stations identifiers consist of five characters, terminating with "SD". (This suffix will be ignored when convenient or necessary.)

As a test, SDG&E station GOSSD was purposely placed at a better-exposed location 0.7 km along Black Canyon Road from GOSC1's original location[3]. For the month of December 2011, which included several moderate Santa Ana wind events,

---

[1] http://www.raws.dri.edu

[2] National Wildfire Coordinating Group,"Interagency Wildland Fire Weather Station Standards and Guidelines", publication PMS 426-3, June 2012.

[3] Before November 2011, GOSSD was sited even closer to GOSC1, in a less well-exposed area intended to mimic the RAWS installation (Steven Vanderburg, personal communication). GOSC1 was subsequently moved.

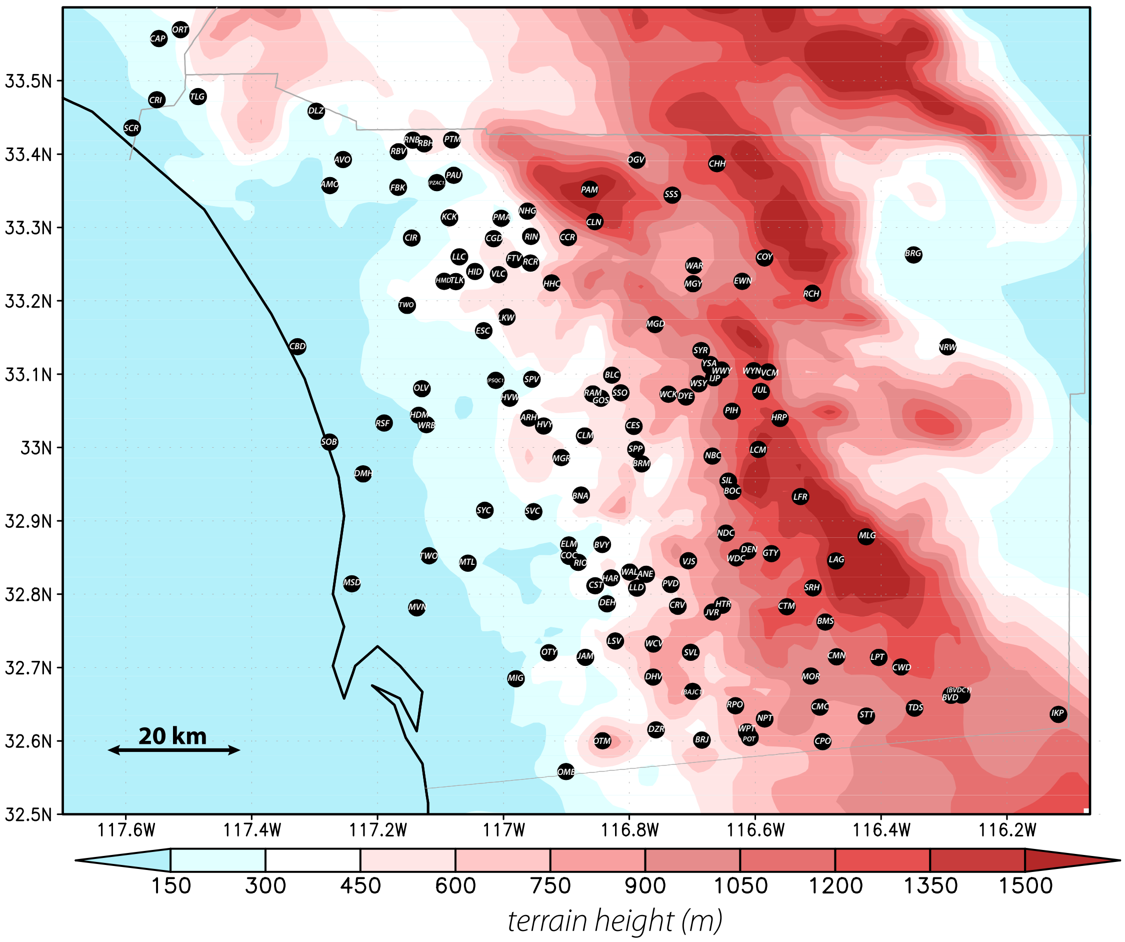


Figure 2.1: SDG&E surface station locations (black dots), with underlying topography shaded. Stations in place as of March 2013.

the 10-min averaged sustained winds at GOSSD were about 50% stronger than at the more sheltered RAWS station (**Fig. 2.2**), even though they were measured at the same height. Indeed, among the 744 contemporaneous observations of sustained wind during that month, 639 SDG&E observations were larger than their corresponding RAWS wind speed, 48 observations were the same, and only 56 of the RAWS observations (< 8% of the total) exceeded the SDG&E reports. As demonstrated clearly in subsequent chapters, even closely spaced and well-exposed stations can exhibit wind variability of this magnitude, so part of the GOSSD-GOSC1 discrepancy might have been due to an unappreciated terrain effect. However, this result motivated us to use the SDG&E network exclusively to validate our model results, owing to its high density, instrumental uniformity and optimal siting philosophy. The purpose of our work, after all, is to forecast winds impinging upon electrical lines at risk in well-exposed terrain.

2.2 Model experimental design

The simulations examined herein were made using the WRF model's Advanced Research WRF (ARW) core, version 3.5. We will closely examine three Santa Ana wind events that took place in the year 2003. To represent an operational environment, the model was initialized with the North American Mesoscale (NAM) model gridded analysis and forecasts from its 1200 UTC 14 February 2013 cycle for the 14-16 February 2013 Santa Ana event; the 0600 UTC 4 October 2013 model run for the October 2013 episode; and the 0600 UTC 13 May 2014 cycle for the May 2014 Santa Ana. All model integrations were for 54 hours.

A five-domain telescoping grid arrangement (denoted D1-D5) was used with horizontal grid spacings of 54, 18, 6, 2, and 0.667 km, respectively (**Fig. 2.3**). The innermost 667 m nest extends about 80 km west-east by 70 km north-south and covers roughly 70% of the SDG&E mesonet, while its parent 2 km grid encompasses the entire network. The highest resolution USGS (United States Geological Survey) terrain

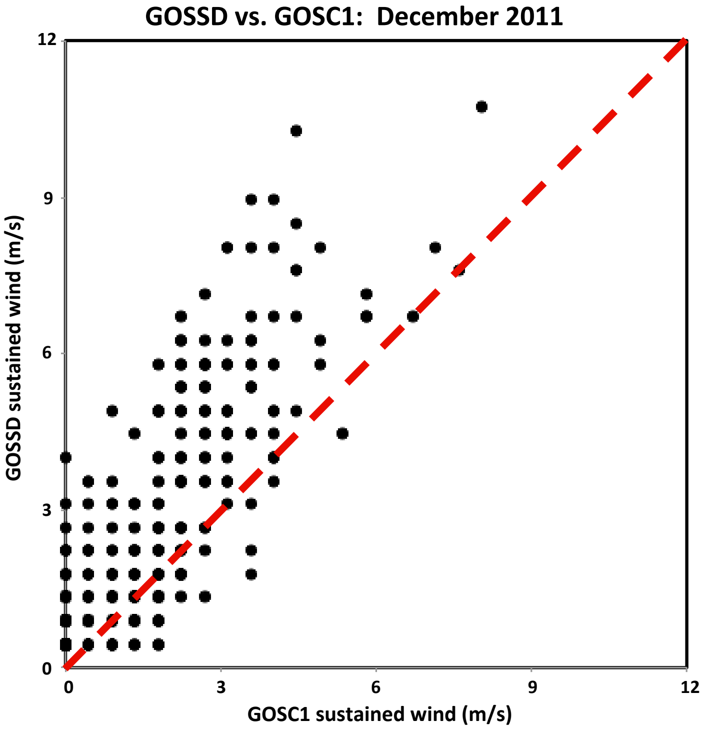


Figure 2.2: Scatterplot of hourly sustained winds measured at the Goose Valley RAWS (GOSC1) and SDG&E (GOSSD) sites for December 2011, with a 1:1 correspondence line (red).

database available is used in the innermost two domains, permitting the model to capture fairly fine-scale topographic features (see **Fig. 2.3 inset**). The model top is 10 hPa, with 50 layers (51 full-sigma vertical levels) employed, focusing highest resolution in the lower troposphere in the usual fashion.

2.3 Validation strategy

The SDG&E wind data were employed to validate model output available at hourly intervals. Mesonet data were obtained at full temporal (10 min) resolution from the MADIS (Meteorological Assimilation Data Ingest System) archive. We elected to replace observed winds and gusts on the hour with the largest values of each reported in the previous 50 min, motivated by the relatively larger high-frequency variability present in the observations. However, this was found to have relatively little impact on the results and no influence on the conclusions.

Of much greater significance is the fact that, by default, the WRF model computes a wind diagnostic relating the lowest model level wind (V_a) at height z = Z_a to the WMO standard height of 10 m (V_{10m}) via the logarithmic wind profile assumption [e.g., *Oke*, 1987]:

$$V_{10m} = V_a \frac{\ln \frac{10}{z_0} - \psi_{10m}}{\ln \frac{Z_a}{z_0} - \psi_a}, \tag{2.1}$$

where z_0 is the surface roughness length, and ψ_a and ψ_{10m} represent stability correction functions at Z_a and 10 m that vanish when the surface layer is neutrally stratified. Neutrality is often presumed when wind speeds exceed about 5 m s^{-1} or so [e.g., *Verkaik*, 2000, *Wieringa*, 1976], which does appear valid among our model simulations. Proper comparison with the SDG&E network winds, however, requires further adjustment to its anemometer mounting height at 6.1 m AGL level ($V_{6.1m}$), i.e.,

$$V_{6.1m} = V_{10m} \frac{\ln \frac{6.1}{z_0} - \psi_{6.1m}}{\ln \frac{10}{z_0} - \psi_{10m}}, \tag{2.2}$$

where $\psi_{6.1m}$ is the stability correction computed at anemometer height. Although some-

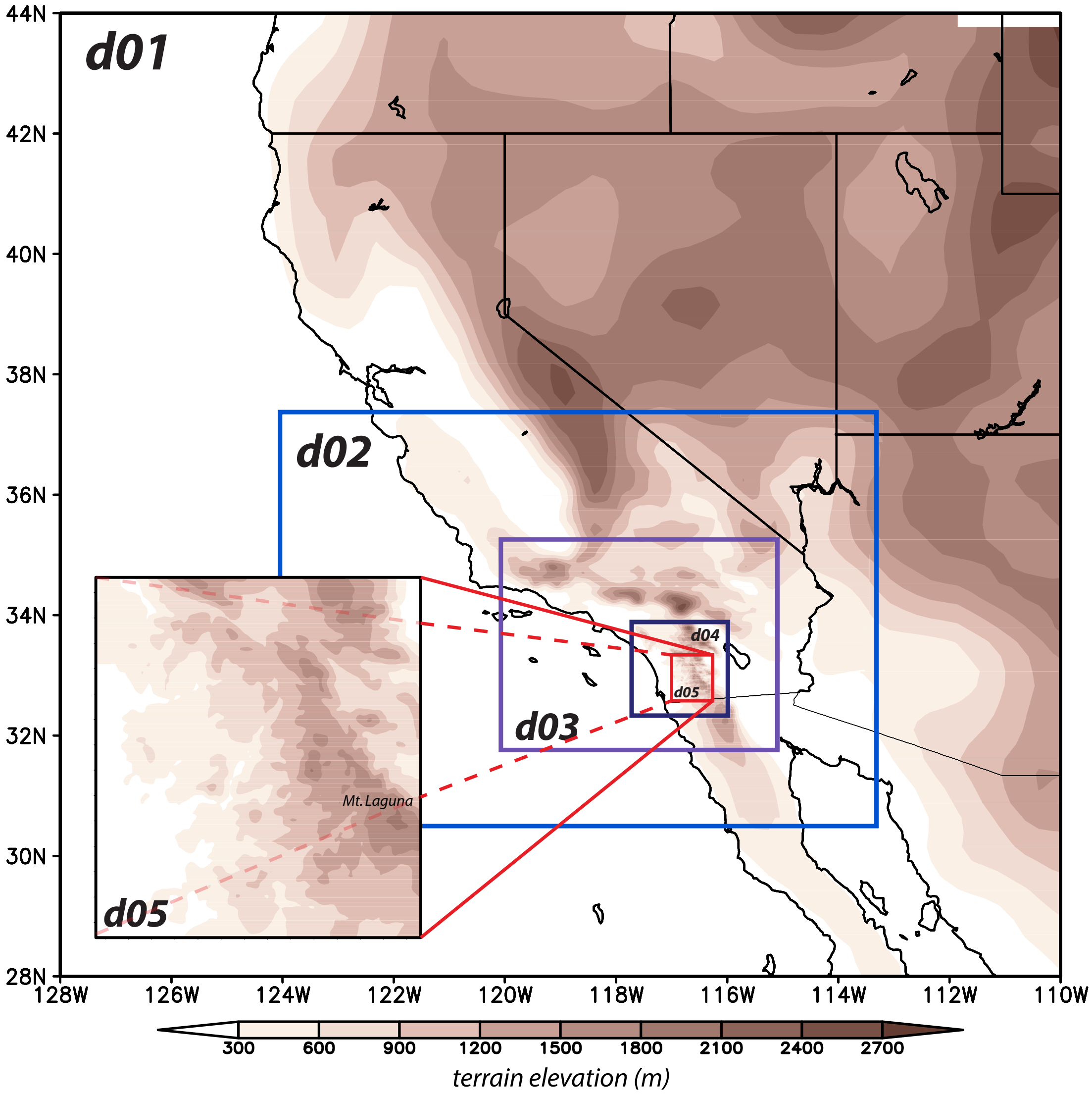


Figure 2.3: Domain configuration for the WRF-ARW simulations, with topography shaded. Domains 1-5 employ horizontal grid spacings 54, 18, 6, 2, and 0.667 km over Southern California, respectively. Inset shows an enlarged version of domain 5.

what dependent on the land surface model and surface layer scheme, landuse database [e.g., USGS vs. MODIS (Moderate Resolution Imaging Spectroradiometer)] employed and season, z_0 values range between 0.05 and 0.9 at SDG&E stations, resulting in wind speed reductions of 10% to 20%. [The stability corrections in Eq. (**2-2**) were retained for simulations examined in detail in this report, but these was found to have relatively little impact on the results and no influence on the conclusions.] Finally, it is noted that Eqs. (**2-1)** and (**2-2**) could have been written with the zero-plane displacement modification of the anemometer heights that is sometimes used in areas with significant obstacles; we neglect this adjustment because of the siting characteristics of the SDG&E mesonet.

We will show that most model physics configurations generate a high wind bias relative to the observed sustained winds, even after the anemometer height adjustment is applied. The worst offenders were ostensibly those employing the Mellor-Yamada-Janjić (MYJ) PBL scheme [*Janjić*, 1994]. However, we discovered the MYJ code was recomputing the 10-m wind values, specifying smaller roughness lengths than actually employed in the model calculations. This purely cosmetic adjustment (shared by the QNSE PBL scheme) exacerbated the high wind bias, and removing the code made physics ensemble members employing the MYJ scheme much more competitive.

CHAPTER 3

The 14-16 February 2013 Santa Ana wind event

In this chapter, we examine a moderately strong Santa Ana wind event that occurred during February 2013, utilizing observations from the SDG&E mesonet. These observations are used to validate the WRF model simulations we will use to understand how the event evolved in San Diego county (see **Fig. 1.1**). The domain setup and important background on the validation strategy were introduced in **Chapter 2**.

3.1 14-16 February 2013 event observations

Certain synoptic-scale conditions interacting with local topography contribute to Santa Ana occurrence [*Hughes and Hall*, 2010, *Sommers*, 1978, *Yoshino*, 1975]. This mid-February Santa Ana wind event commenced around 0000 UTC 15 February 2013, as maximum sea-level pressures exceeded 1028 hPa in the Great Basin (**Fig. 3.1 a**), and a mid-level ridge approached the Western U.S, bringing northeast winds over the mountains encircling Southern California (**Fig. 3.1 d**). Some stations reported their fastest offshore winds around 1800 UTC 15 February 2013, when the Great Basin High and the 700 hPa ridge reached peak magnitudes (**Figs. 3.1 b and e**). During the next 24 h, the high pressure migrated eastward (**Fig. 3.1 c**), away from Southern California, the surface offshore winds weakened and the 700 hPa ridge flattened (**Fig. 3.1 f**).

Although only moderate in overall strength as a Santa Ana episode, some very impressive winds and gusts were recorded in the SDG&E network during the event. **Figure 3.2** presents the maximum wind gusts observed in the SDG&E network for the

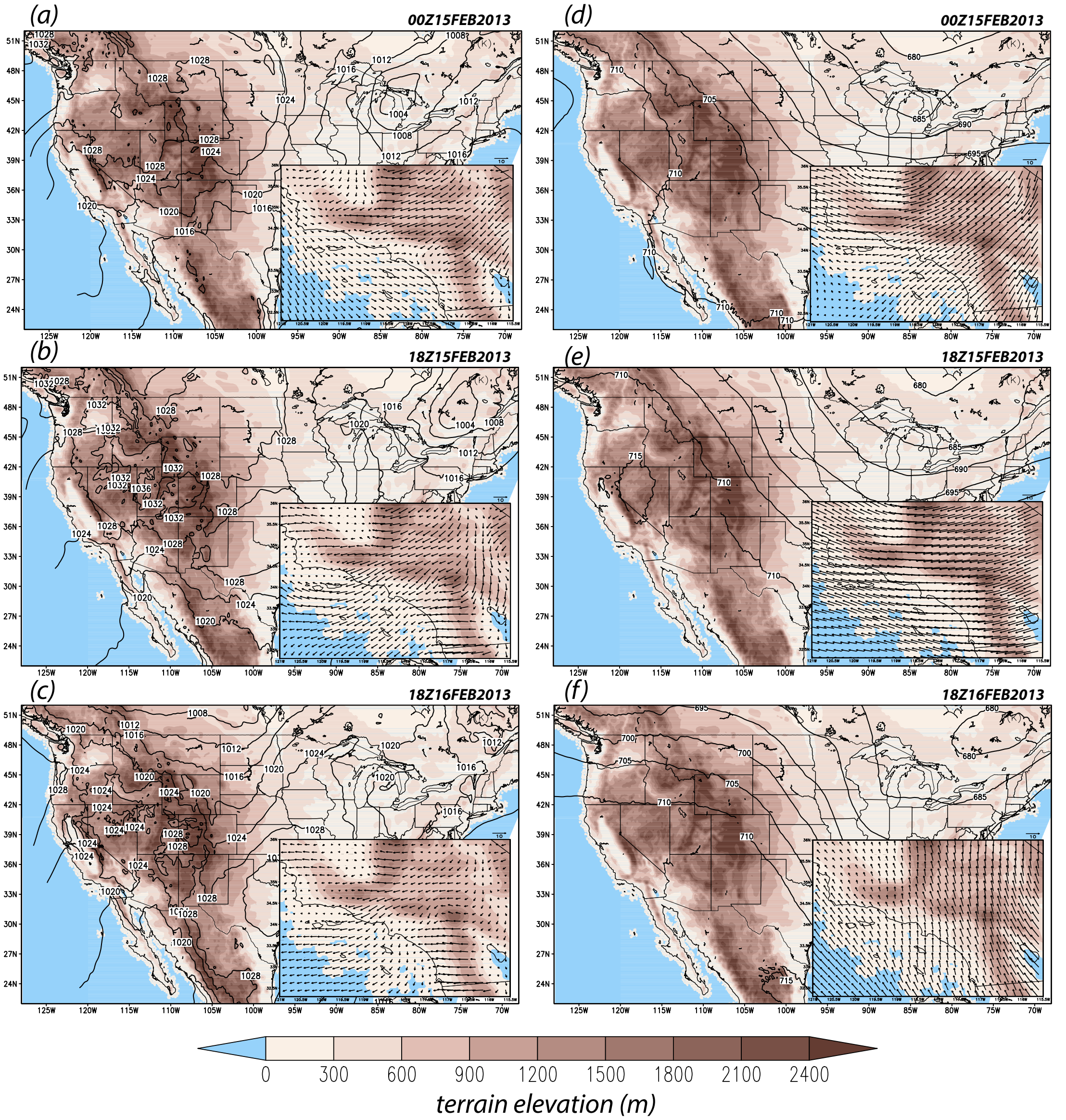


Figure 3.1: NAM model sea level pressure analyses for (a) 0000 UTC 15 February 2013; (b) 1800 UTC 15 February 2013; and (c) 1800 UTC 16 February 2013, and 700 mb analyses for (d) 0000 UTC 15 February 2013; (e) 1800 UTC 15 February 2013; and (f) 1800 UTC 16 February 2013. Insets show the total 10 m winds of Southern California. Only every fourth grid point is plotted for clarity. Topography is shaded.

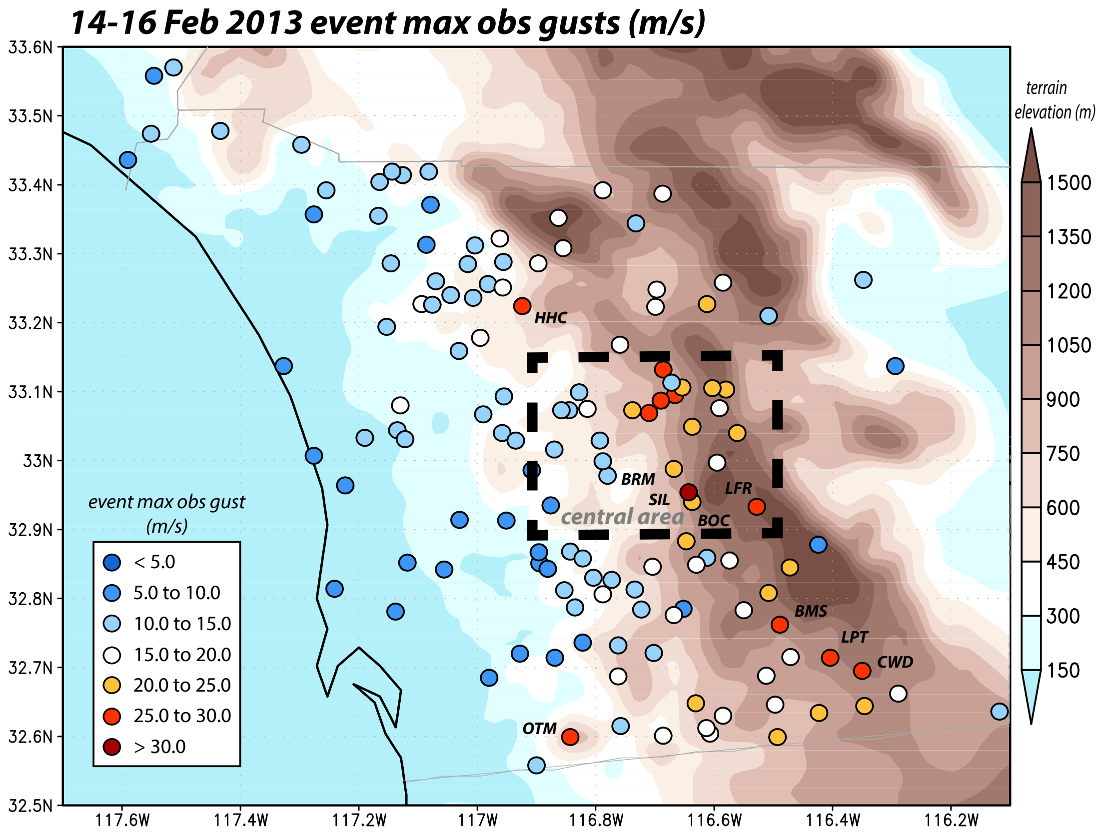


Figure 3.2: Spatial distribution of 14 February 2013 event maximum observed wind gusts (m s^{-1}) for SDG&E stations. Brown shading indicates model topography. Thick black contour shows coastline at 2 km resolution. Black dotted square denotes the "central area" region.

14-16 February 2013 event. The strongest gusts are found to be located along the western slopes, close to but not right at the ridgelines. The great spatial variability of the winds can be detected in **Fig. 3.3**, which focuses on a "central area" that comprises the stations of greatest present interest. In that relatively small region, peak gusts varied between 10 and 30 m s^{-1} within a 5 km distance, suggesting each station is representative only of a small local area, at least with respect to the winds. The event-maximum sustained winds (not shown) are similar in pattern although naturally weaker in magnitude.

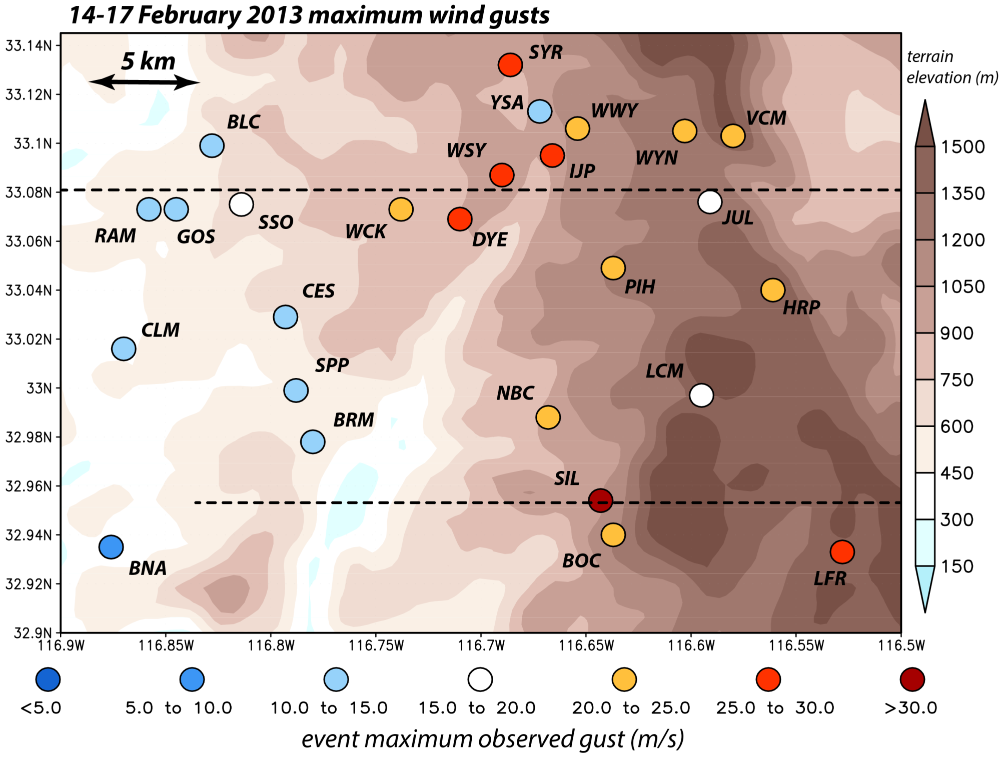


Figure 3.3: As in **Fig. 3.2**, but zoomed into the "central area" region. Labels indicate names of SDG&E stations, with "SD" suffix omitted. Black dotted lines denote locations of cross-sections across WSY and SIL.

At 1830 UTC (1030 AM PST) on 15 February 2013, SDG&E station Sill Hill (SIL) recorded a 41 m s^{-1} wind gust (**Fig. 3.4**), at a time when no other stations in this region had a gust exceeding 26 m s^{-1}. Indeed, the winds were 50% weaker at Boulder Creek (BOC), the SDG&E station just 1.6 km to the south (**Fig. 3.2**). (Keeping in mind that the sustained wind represents 10-min averages and the gusts are single 3-sec samples, note how similar the sustained wind at SIL is to the wind gusts from BOC.) It would be easy to dismiss such a high wind observation, but the station record shows that gusts exceeding 36 m s^{-1} were a frequently recorded occurrence (**Fig. 3.4**), and eye-level winds of 33 m s^{-1} had been measured with hand-held anemometers at the site about an hour before the winds peaked (Brian D'Agostino and Steven Vanderburg, personal communication). A close inspection at the topographic map in the vicinity of SIL and BOC (not shown) indicates that SIL is sited on a small local ridge while BOC is in a local terrain crease, very subtle factors that may be relevant to the wind speeds and exposures and illustrate the challenge that is faced in simulating and validating the winds across this area.

We now shift focus to the Witch Creek (WCK) area, where the local SDG&E station density is high (**Fig. 3.2**). Wind gusts observed over a two-day period at West Santa Ysabel (WSY), located on the west-facing slope about 9-10 km down from the ridge, reveal a Santa Ana episode consisting of two pulses or phases separated by a protracted lull (**Fig. 3.5 a**). Gusts during the first phase peaked at 26 m s^{-1} at 1800 UTC (10 AM PST) on the 15th. After a marked weakening during the afternoon, the gusts regained comparable strength by midnight local time before finally slowing as the event wound down.

Although about 40% weaker than the gusts, the sustained winds at WSY followed a similar trend. At SDG&E station Julian (JUL), close to the ridge, the gusts were significantly weaker than WSY's during the first phase, stronger (although still fairly slow) during the afternoon lull, and markedly weaker during the second phase. This provides a hint that there is something structurally and/or dynamically different about

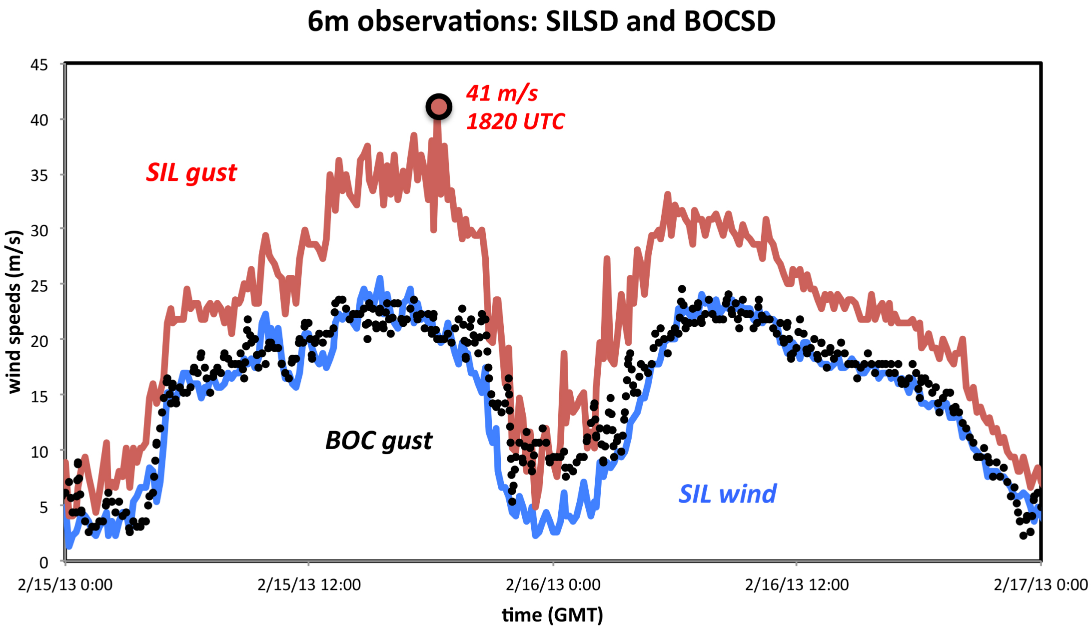


Figure 3.4: Time series of observed winds (m s^{-1}) at SIL and BOC over 2 days. Red and blue lines depict SIL gust and sustained wind, respectively; black dots denotes BOC gust. The red dot marks the SIL gust peak.

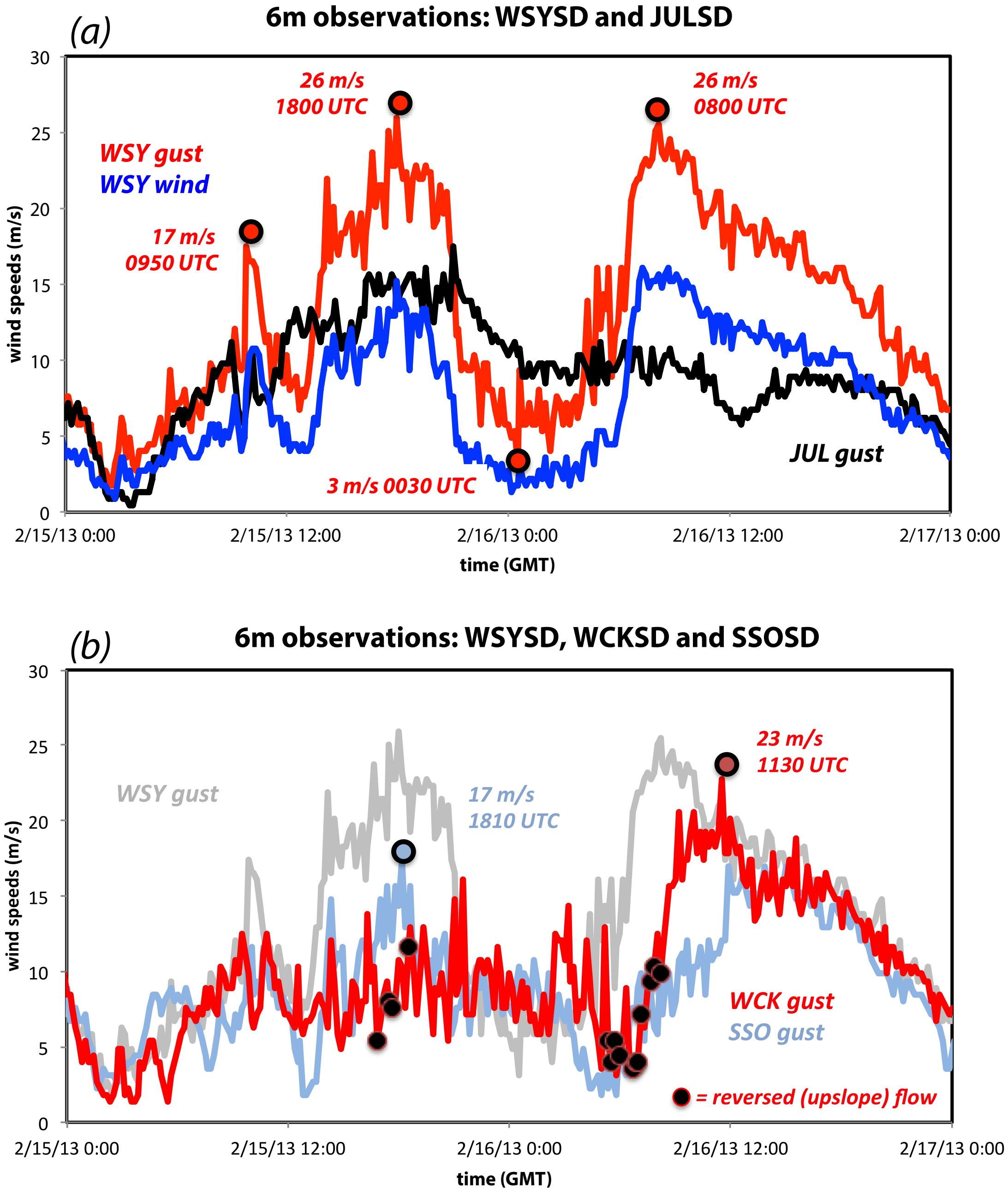


Figure 3.5: (a) Time series of observations of WSY gusts (red), sustained winds (blue), and JUL gusts (black) over 2 days. Red dots mark largest and smallest gusts at WSY; and (b) As in (a), but for WSY gusts (grey), WCK gusts (red), and SSO gusts (blue), with the red and blue dots marking WCK and SSO gust peaks, respectively. Black dots indicate times when winds were directed upslope at WCK.

the second half of the event.

The winds also behaved very differently at the WCK station (**Fig. 3.5 b**), which is less than 5 km downslope from WSY. Through the first phase, WCK's gusts remained much weaker than WSY's. Note the wind direction at WCK occasionally reversed to upslope (at times indicated by the black dots) during this period. During the lull between the two phases, the WSY and WCK winds were comparably weak. More wind reversals occurred at WCK during the onset of the second phase before downsloping became firmly reestablished there. WCK recorded its event maximum gust of 23 m s^{-1} at 1130 UTC 16 February, during the second pulse and about 3 hours after the winds at WSY started to decline.

Station Sunset Oaks (SSO) is located 7 km farther downslope from WCK. Note that, during the first pulse, its gusts were weaker than, but in phase with, WSY. The wind reversals at WCK during this time occurred when downslope flow was observed both uphill (at WSY) and downhill (at SSO), suggesting a rotor or hydraulic jump may have formed there. Station SSO emerged from the lull last, and its second peak was reached after the gusts at both WSY and WCK had started to decline. Taken together, these stations suggest a two-part Santa Ana event in which winds were largely in phase early in the event, apart from the suspected jump at WCK, and had a second part consisting of a marked downslope progression as the overall winds abated.

3.2 Model simulations and validations

3.2.1 Control run setup

As mentioned in the previous chapter, our WRF-ARW simulations utilize five telescoping domains down to a horizontal resolution of 667 m and are initialized with NAM gridded analyses and forecasts. Based on a systematic validation of model vs. observed winds over a set of events, which will be explored in **Chapter 4**, the physics ensemble member that appears to best represent the sustained wind observations with respect to magnitude and temporal and spatial variation employed the Pleim-Xiu (PX) [*Pleim and Xiu*, 1995, *Xiu and Pleim*, 2001] land surface model (LSM) and surface layer schemes, the Asymmetric Convection Model version 2 (ACM2) [*Pleim*, 2007a,b] planetary boundary layer (PBL), and the Rapid Radiative Transfer Model for General Circulation Models [RRTMG; *Iacono et al.*, 2008] radiation parameterization. The selected configuration, labeled PX/ACM2, also utilized the landuse database derived from MODIS satellite observations, and explicit horizontal diffusion was not applied. Unsurprisingly, given the nature of the weather regime under study, we found that the simulations were nearly insensitive to other physics choices, such as the microphysics and cumulus schemes.

3.2.2 Spatial and temporal variation of the winds in the control run

To a large extent, the control PX/ACM2 simulation captured the magnitude and temporal evolution of the network-averaged sustained winds, at least after the first 12 hours (**Fig. 3.6**). The spatially- and temporally-averaged mean absolute error (MAE) was 2.23 m s^{-1} for the SDG&E network, and stations with relatively large and small errors appear to be randomly dispersed in space (**Fig. 3.7**). The largest MAEs occurred at a handful of particularly windy stations, as part of a moderate ($R^2 \sim 0.5$) positive (and somewhat curvilinear) relationship between MAE and average observed wind that exists only due to stations with event-averaged winds exceeding 8 m s^{-1} (**Fig. 3.8 a**). In contrast, although the mean network bias (**Fig. 3.9**) is nearly zero (0.07 m s^{-1}), it is negatively correlated ($R^2 \sim 0.5$) with the event mean wind (**Fig. 3.8 b**). The mean bias in the aforementioned "central area" is also nearly zero, but very variable in space (**Fig. 3.10**). This hints at the value of high network density and the danger of drawing conclusions from a small number of stations. (A physically-based explanation for the fact that the bias itself is biased will be in **Chapter 5**.)

Figure 3.11 compares hourly time series of observed and simulated sustained wind

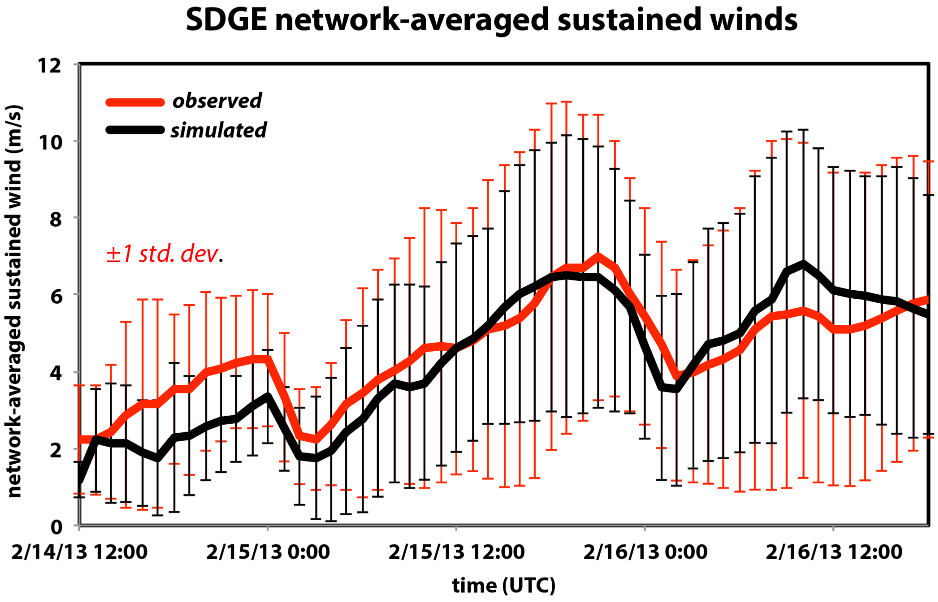


Figure 3.6: Time series of SDG&E network-averaged sustained wind (m s^{-1}) observations (red line) at 6.1 m AGL over 2 days, for comparison with predictions from the control (PX/ACM2) run (black line). Red and black color bars are plus and minus one standard deviation for observations and the control run.

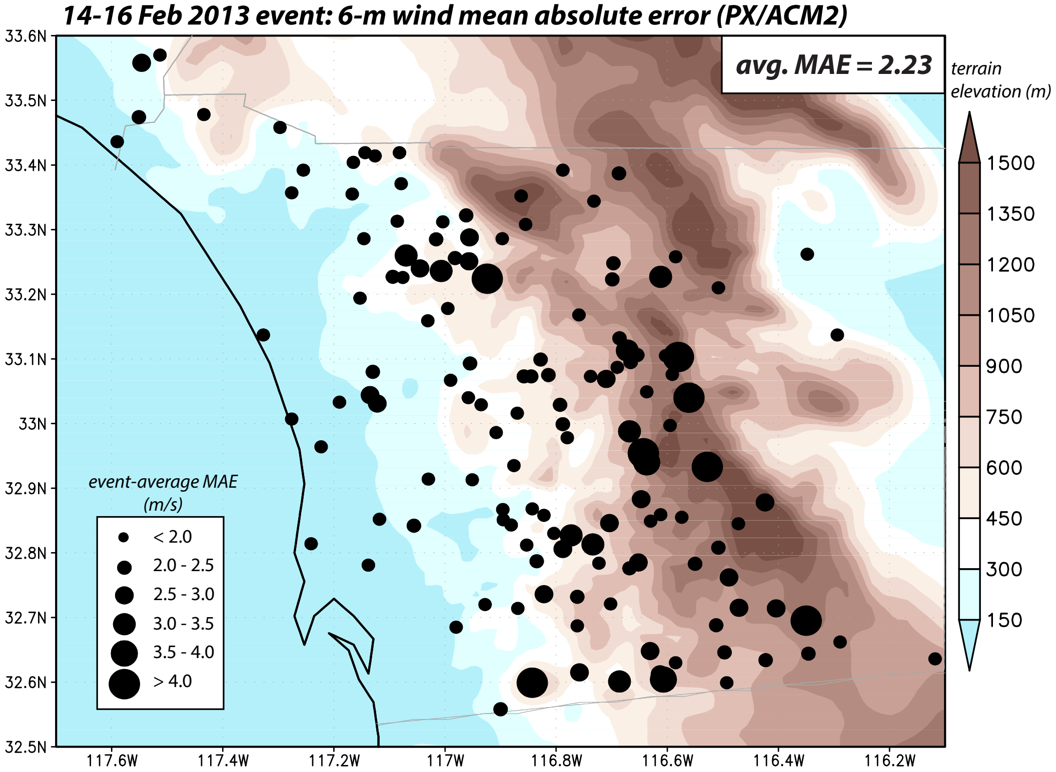


Figure 3.7: As in **Fig. 3.2**, but for event mean sustained wind MAE for PX/ACM2 simulation, using the MODIS landuse database. The average value over the entire SDG&E network is 2.23 m s^{-1}.

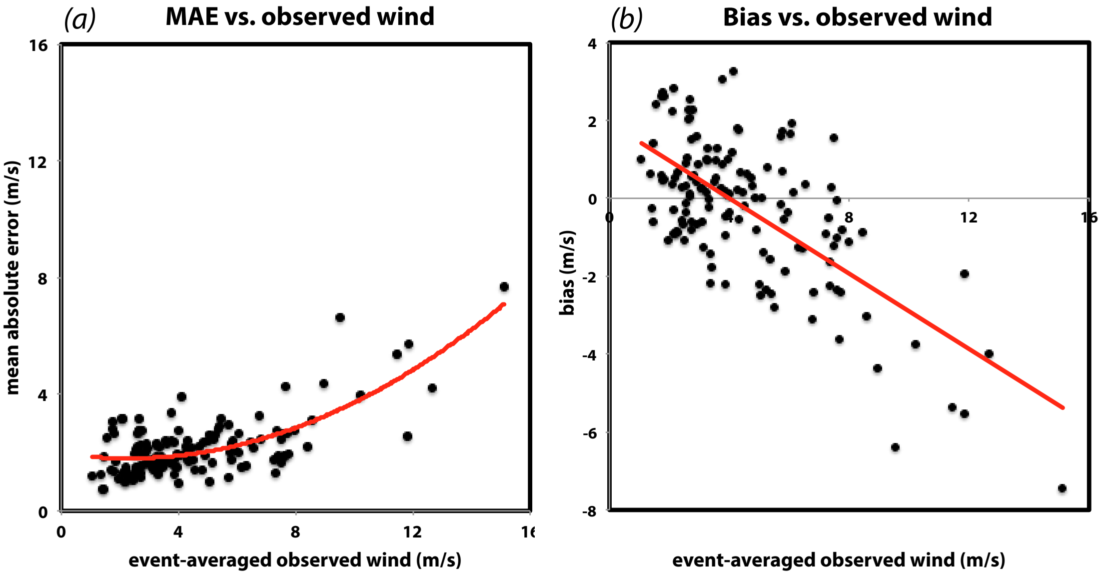


Figure 3.8: (a) Scatterplot of event-mean observed wind vs. MAE in the PX/ACM2 simulation for SDG&E stations. A curvilinear least-squares fit (red line) is shown for reference; and (b) As in (a), but for event-mean observed wind vs. mean bias.

at stations WSY, WCK and SSO. Overall, the simulation captures the evolution and magnitude of the winds at each station to a good degree, although there are some clear timing issues. At WSY, the magnitude of the second pulse was underpredicted, although the phasing was skillful (**Fig. 3.11 a**). The lull ended too early at both WCK and SSO (**Figs. 3.11 b and c**). However, when viewed in combination (**Fig. 3.11 d**), we see the model captured the overall behavior: during the first pulse, the winds remained markedly weaker at WCK than those both upslope and downslope, and the second pulse was characterized by a clear downslope progression of the winds with time.

Having demonstrated reasonable fidelity with the available observations, the simulation will be used to complete the horizontal wind field (**Figs. 3.12 and 3.14**), and provide insight into the vertical dimension that is missing from the observations (**Figs. 3.13 and 3.15**). At 0800 UTC 15 February 2013 (**Fig. 3.12 a**), the downslope windstorm had started, but the winds near the ground at WSY and stations farther downslope

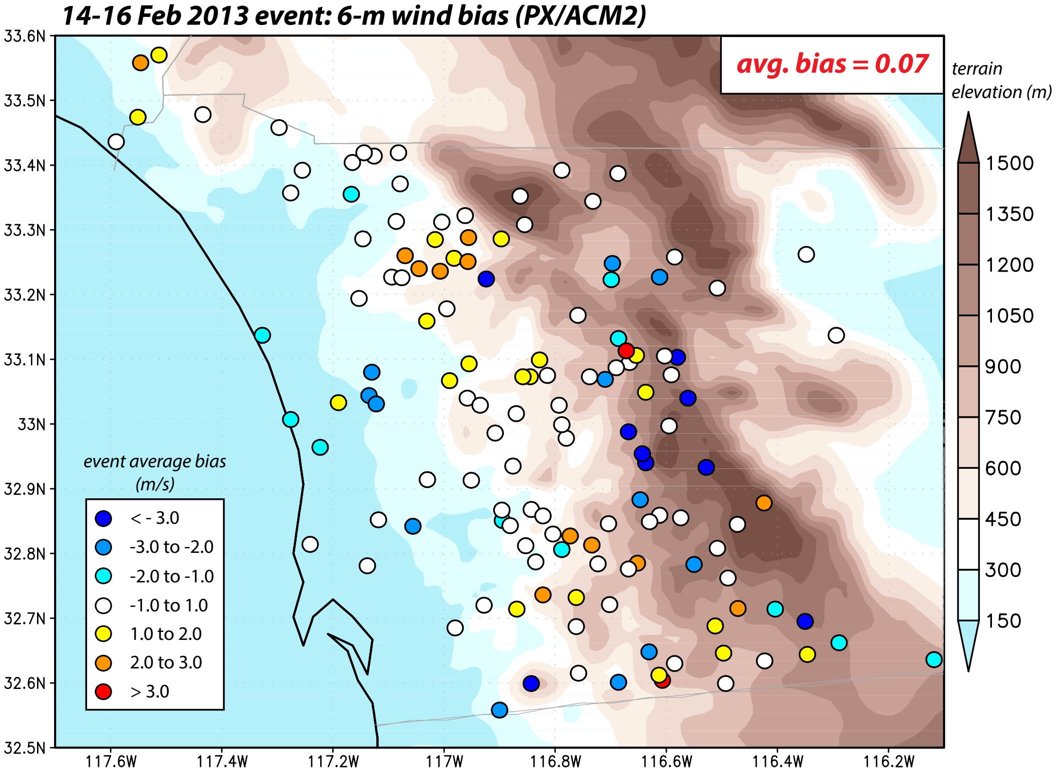


Figure 3.9: As in **Fig. 3.7**, but for event mean sustained wind bias. The average value over the entire SDG&E network is 0.07 m s^{-1}.

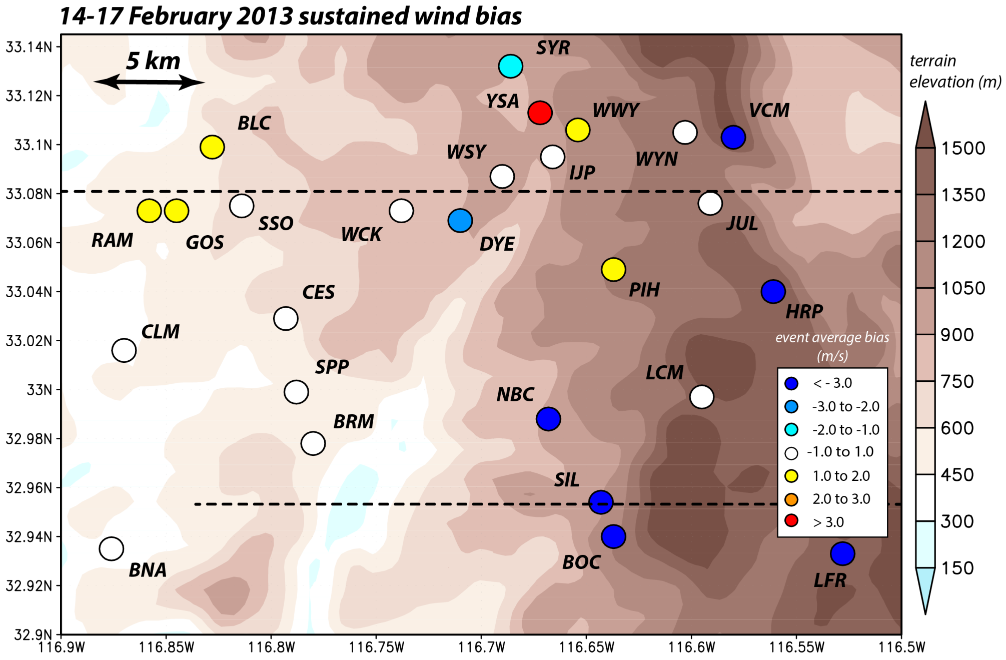
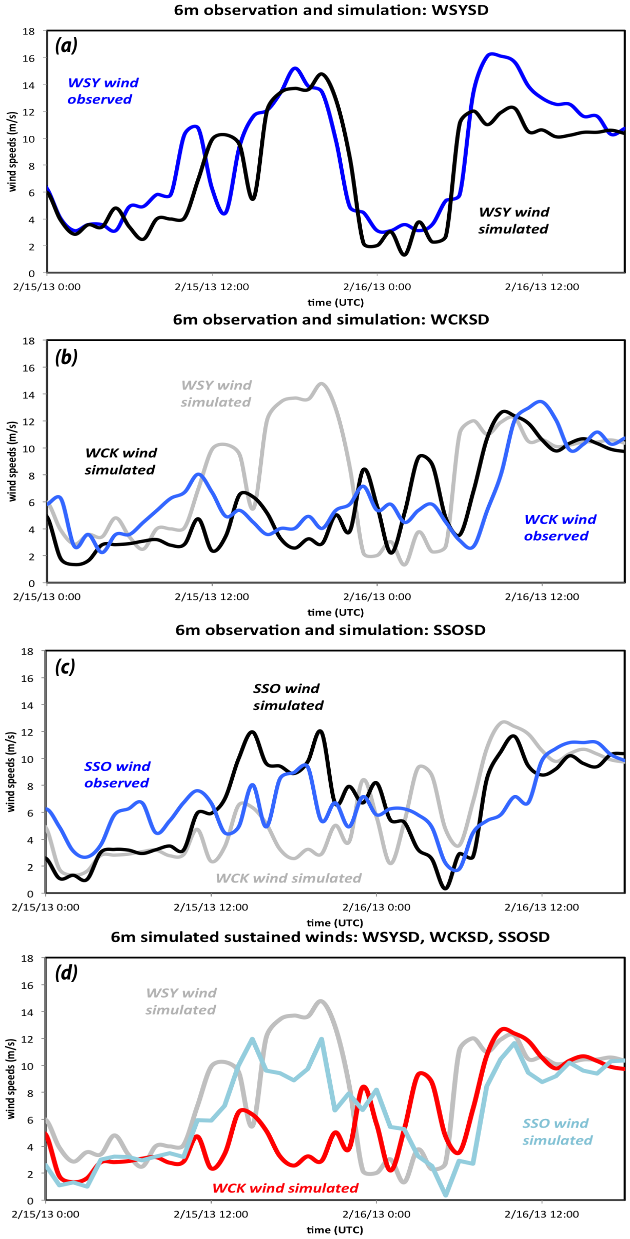


Figure 3.10: As in **Fig. 3.9**, but zoomed into the "central area" region. Labels indicate names of SDG&E stations, with "SD" suffix omitted. Black dotted lines denote cross-sections across WSY and SIL.


6m observation and simulation: WSYSD

6m observation and simulation: WCKSD

6m observation and simulation: SSOSD

6m simulated sustained winds: WSYSD, WCKSD, SSOSD

Figure 3.11: Time series of observed (blue curve) and predicted (black curve) 6 m sustained winds (m s^{-1}) at (a) WSYSD; (b) WCKSD; and (c) SSOSD; and (d) comparison of predicted 6 m sustained winds (m s^{-1}) at WSYSD (grey curve), WCKSD (red curve), and SSOSD (cyan curve). Note these are sustained winds, and not gusts, with a time resolution of 1 hour. Note also the time interval plotted here ends earlier than in **Fig. 3.5**.

had not yet begun to rise. Actually, significant easterly flow was already present above WSY and WCK, but had not yet reached the surface (**Fig. 3.13 a**). Recall that by 1740 UTC, winds recorded at WSY and SSO had reached their first-phase peaks, but WCK's winds remained quite weak (**Figs. 3.5 b and 3.11b**). The model simulation has indeed developed a jump-like feature almost directly above WCK at this time (**Fig. 3.13 b**), rendering relatively weak winds there and upslope reversals just to the south (see square and red arrows in **Fig. 3.12 b**). Note also that, as expected, the wind speeds had not strengthened very much at JUL, which is located at the top of the ridge and at the very edge of the terrain amplification.

Five hours later, there was a brief period (around 2130 UTC) during which the observed winds at WCK were actually stronger than at the other stations (**Fig. 3.5 b**), as the winds at WSY and SSO were entering the lull period while the gusts at WCK reached their first-phase peak of 16 m s^{-1}. While the timing is not perfect, the model suggests this occurred as the jump-like feature retreated upslope, passing over WCK and relocating the upslope flow to WSY (**Fig. 3.13 c**; see square in **Fig. 3.12 c**). As the windstorm subsequently retreated farther upslope, it also weakened and became more elevated (**Fig. 3.13 d**). The model reveals that strong near-surface winds still existed during the lull, but became concentrated close to the ridge and in an area where there were no stations (**Fig. 3.12 d**). The retreat occurred during the afternoon hours, likely in response to diurnal variation in the environment. This is a subject of continuing research.

The second phase of the Santa Ana event ensued as the reintensifying flow began progressing downslope again after 0500 UTC (**Figs. 3.13 e and 3.14 a**). Note another, smaller amplitude jump formed in the vicinity of WCK, again consistent with the observations (**Fig. 3.5 b**). By midnight, however, that feature had disappeared and the downsloping flow became "flatter" and, eventually, shallower as the Santa Ana event eventually wound down (**Figs. 3.13 f-h**). The observations indicated a westward progression in the peak near-surface wind speeds (**Fig. 3.5 b**) occurred, and the model has

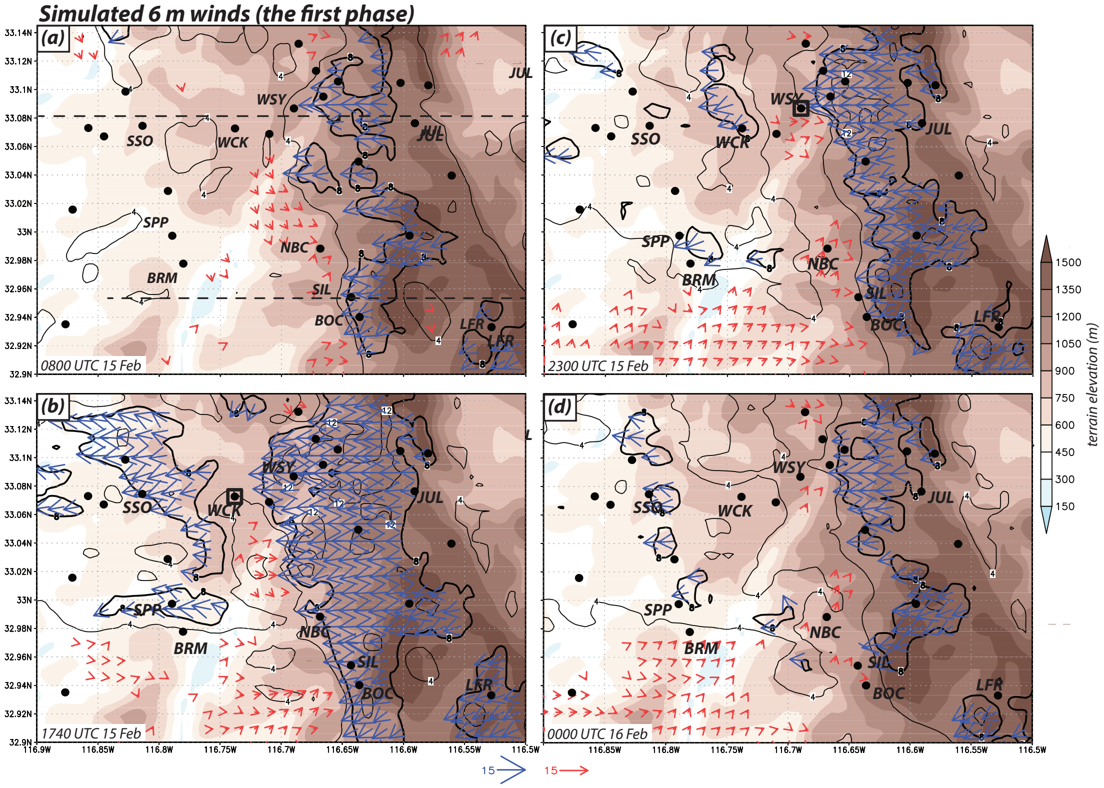


Figure 3.12: Spatial distribution of simulated 6 m horizontal wind speed (4 m s^{-1} interval, 8 m s^{-1} contour bolded) for (a) 0800 UTC 15 February 2013; (b) 1740 UTC 15 February 2013; (c) 2300 UTC 15 February 2013; and (d) 0000 UTC 16 February 2013, with topography shaded. Dashed lines in (a) denote locations of vertical cross-sections shown in **Figs. 3.13** and **3.15**. Red arrows denote winds with a westerly component exceeding 0.5 m s^{-1} in magnitude. Blue arrows denote winds with a easterly component, and is greater than 8 m s^{-1} .

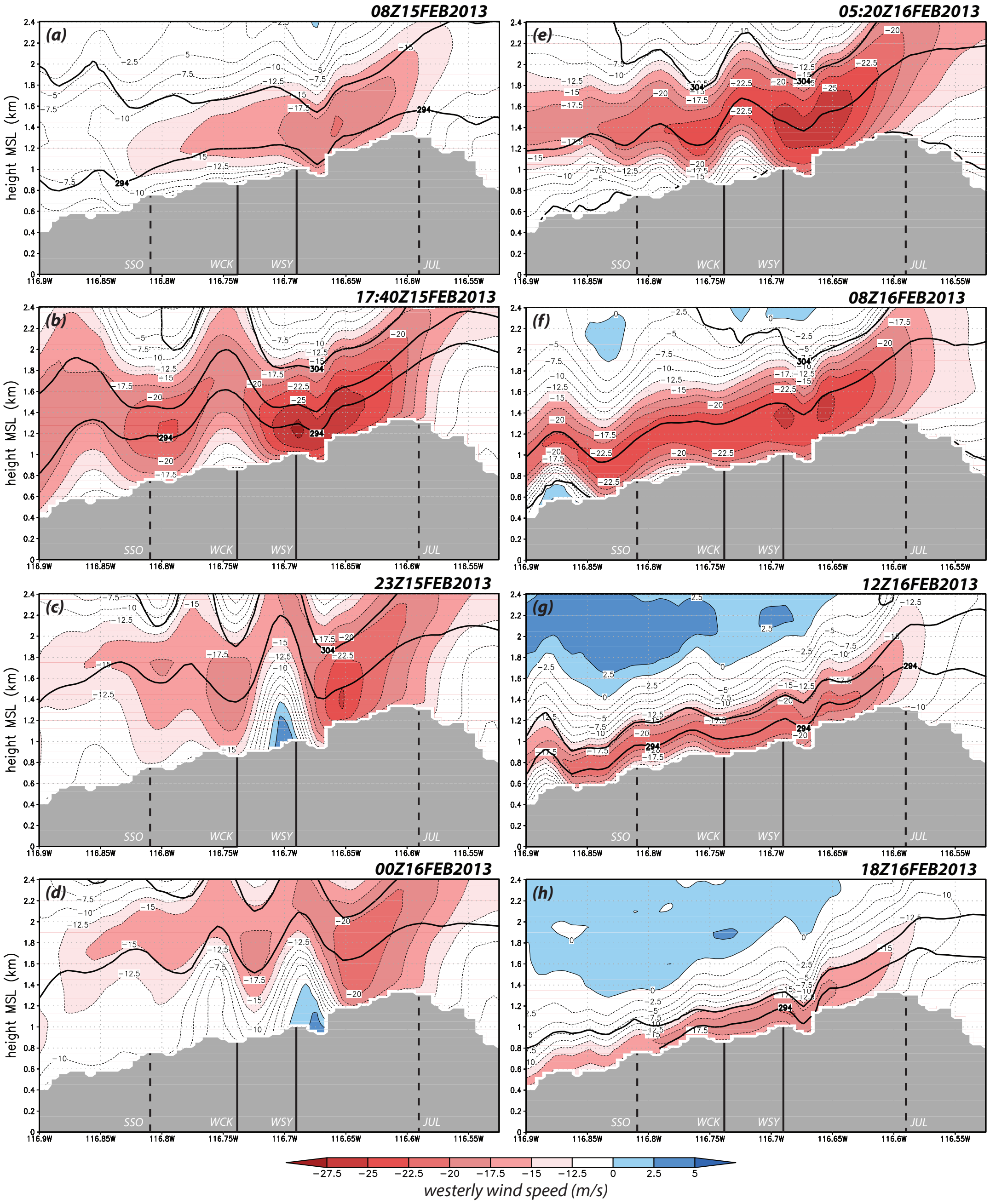


Figure 3.13: Vertical cross-section of horizontal wind speed (shaded, with 2.5 m s^{-1} interval thin contours), taken west-east across WSY with underlying topography in grey (see **Fig. 3.12 a**). Thick contours denote isentropes (5K interval). Approximate locations of stations JUL, WSY, WCK, and SSO are marked. WCK, SSO, and JUL are displaced somewhat from the vertical plane depicted.

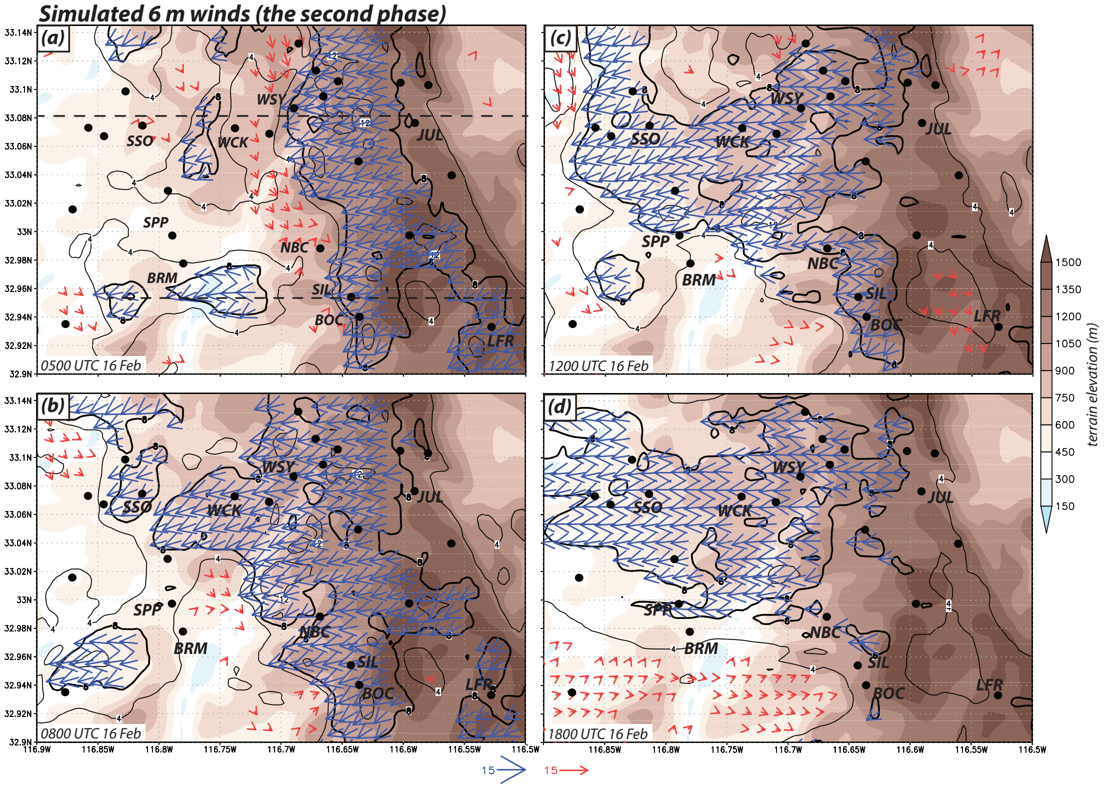

Simulated 6 m winds (the second phase)

Figure 3.14: As in **Fig. 3.12**, but for (a) 0500 UTC 16 February 2013; (b) 0800 UTC 16 February 2013; (c) 1200 UTC 16 February 2013; and (d) 1800 UTC 16 February 2013.

largely captured this behavior (**Figs. 3.11 d and 3.14 b-d**).

Figure 3.15 switches focus to the west-east vertical cross-sections across another station SIL (see **Figs. 3.12 a and 3.14 a**), the focus of **Fig. 3.4**. During the first phase of the event (**Figs. 3.15 a-d**), the downslope winds were not able to progress beyond this station, at least at this latitude, prior to the afternoon retreat upslope. It is recalled that SIL's observed peak gust (41 m s^{-1}) occurred at 1820 UTC (**Fig. 3.4**), the time of **Fig. 3.15 c**. The winds extended farther downslope during the second pulse, fitfully forming jump-like features (**Figs. 3.15 e-g**) in areas lacking stations (**Figs. 3.14 a-c**). The event winds waned more quickly in this portion of the central area than the subzone around WCK (**Figs. 3.14 d and 3.15 h**).

SIL and BOC were among the most severely underpredicted sites (**Figs. 3.10 and**

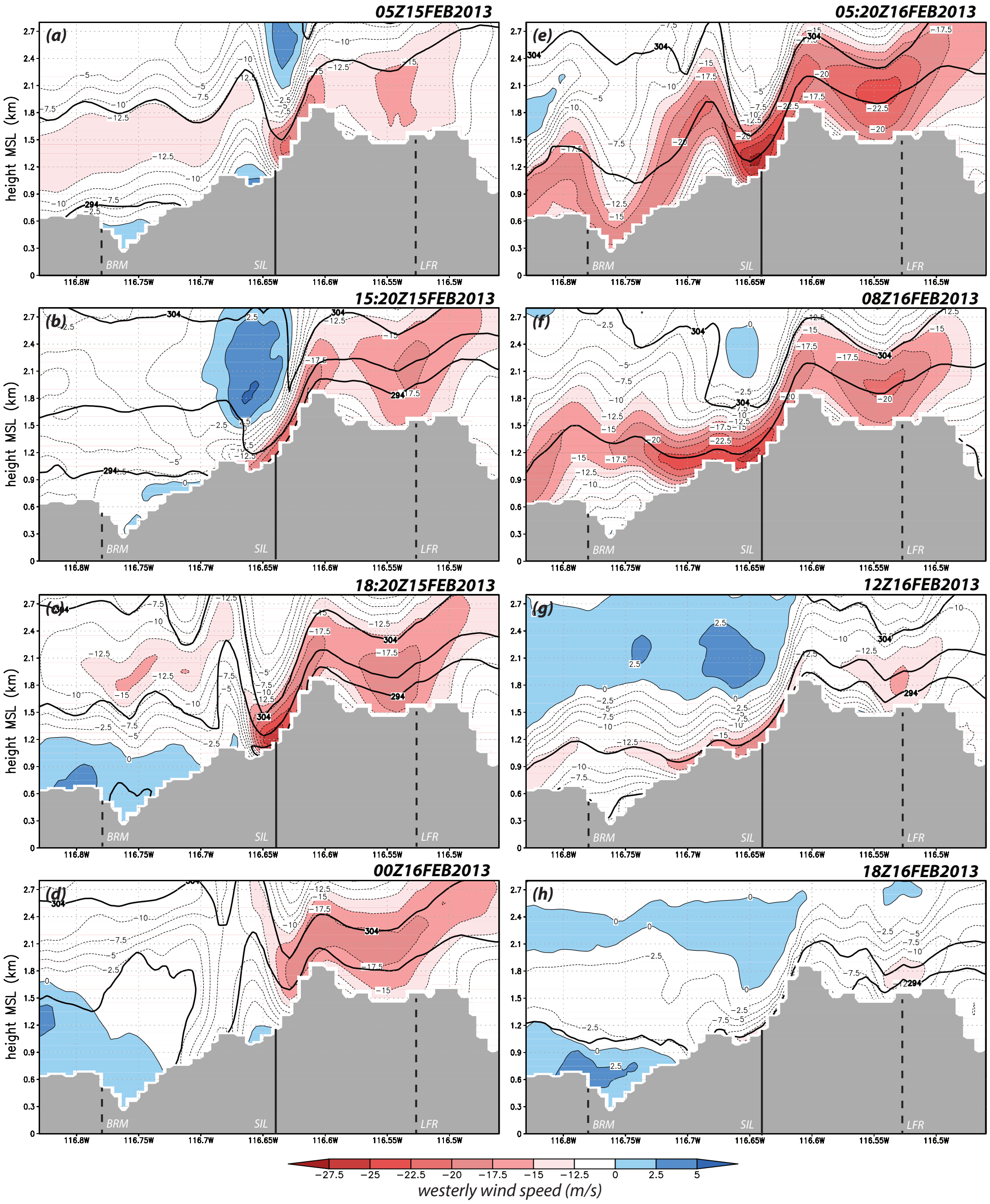


Figure 3.15: As in **Fig. 3.13**, but for station SILSD. Note some times do not match those in **Figs. 3.13 a-h**.

3.16 a). Reconstructions for these two stations are very similar, which is unsurprising due to their small separation (1.6 km) relative to the 667 m resolution of D5. North Boulder Creek (NBC) was also underpredicted (**Fig. 3.16 b**), with a delayed onset of the windstorm, although the model accurately captured the fact the NBC site was less windy than both SIL and BOC. At Barona Mesa (BRM), located farther downslope (**Fig. 3.10**), both the simulated and observed winds remained generally weak during the episode (**Fig. 3.16 c**).

3.2.3 Sensitivity to model physics

As mentioned earlier, PX/ACM2 was selected for the control run owing to its small MAE and nearly-zero network-integrated bias for the event-averaged sustained wind, in this as well as other cases. Nearly all other physics combinations resulted in a positive wind bias as well as larger MAE for this event (**Fig. 3.17**). The members clustered with respect to LSM, with the choice of the PBL scheme having only a secondary effect. Although other variables such as temperature and humidity are undoubtedly important, especially for windstorms in fire-prone areas, we have limited this evaluation to sustained winds.

The most commonly used LSM/PBL combination is probably Noah with the Yonsei University (YSU) parameterization, along with the surface layer scheme derived from MM5 (Noah/YSU). This combination resided in the middle of our 48-member physics ensemble (**Fig. 3.17**), with obviously larger MAE at most stations (compare **Fig. 3.18** to **Fig. 3.7**). While SIL was still substantially underpredicted, over three-quarters of the sites had a positive wind bias (**Fig. 3.19**). Keep in mind that these winds have been adjusted to the 6.1 m level; a straight comparison with the model's 10-m wind diagnostic would have suggested an even *larger* overprediction. The reasons why some LSMs outperform others will be explained the next chapter.

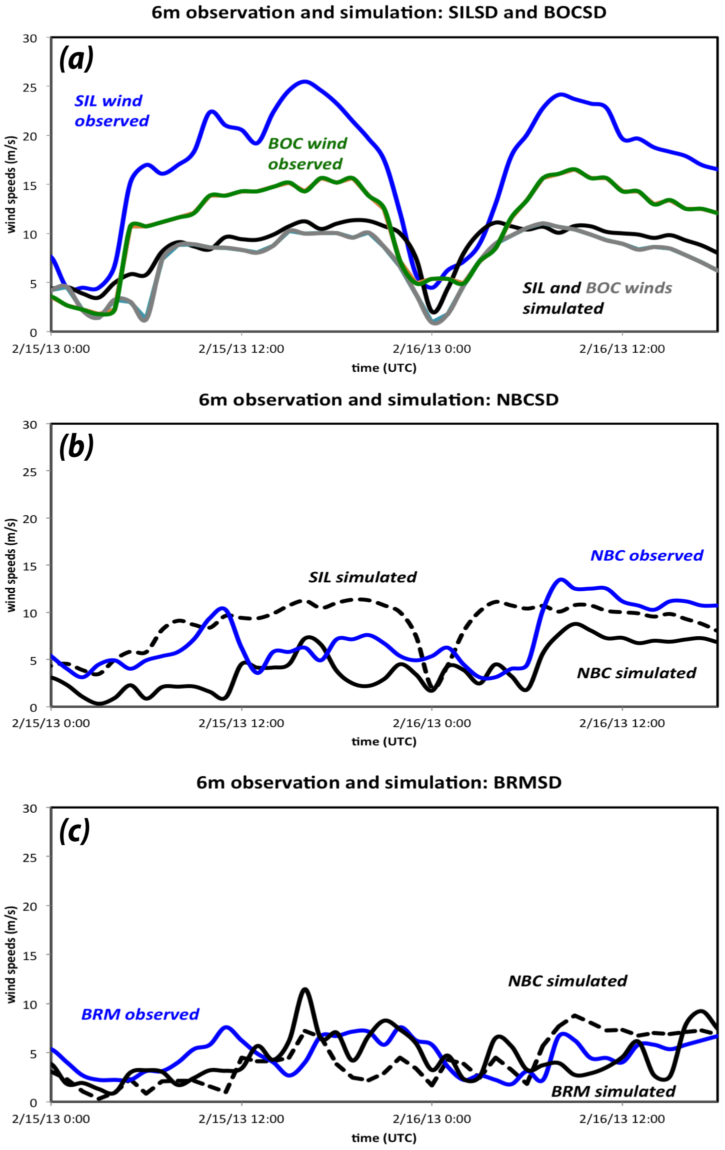


Figure 3.16: (a) Time series of observed (blue curve) and predicted (black curve) sustained winds at SILSD, and observed (green curve) and predicted (grey curve) sustained winds at BOCSD; (b) Time series of observed (blue curve) and predicted (black curve) sustained winds at NBCSD, with simulated SIL winds overlaid; and (c) Time series of observed (blue curve) and predicted (black curve) sustained winds at BRMSD, with simulated NBC winds overlaid. All winds measured or valid at 6 m AGL, in m s^{-1}.

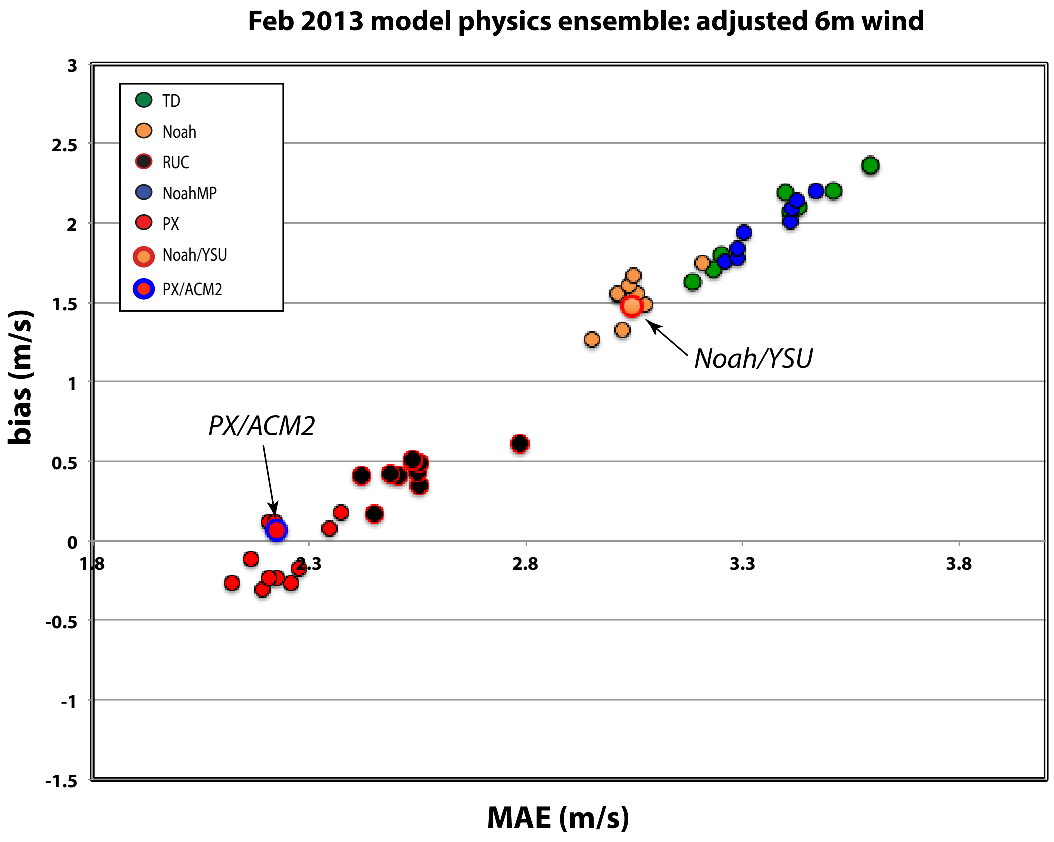


Figure 3.17: Scatterplot of event-averaged bias vs. MAE (both m s^{-1}) from the 48 physics ensemble members for the 14-16 February 2013 episode, color-coded by LSM. For each PBL scheme, the recommended and/or most frequently adopted surface layer parameterization was employed. For members using the MYJ PBL scheme, a standard but cosmetic recalculation of the near-surface winds was removed, as noted in text. Diffusion option is turned off. Landuse database is derived from MODIS. See **Chapter 4** for more information.

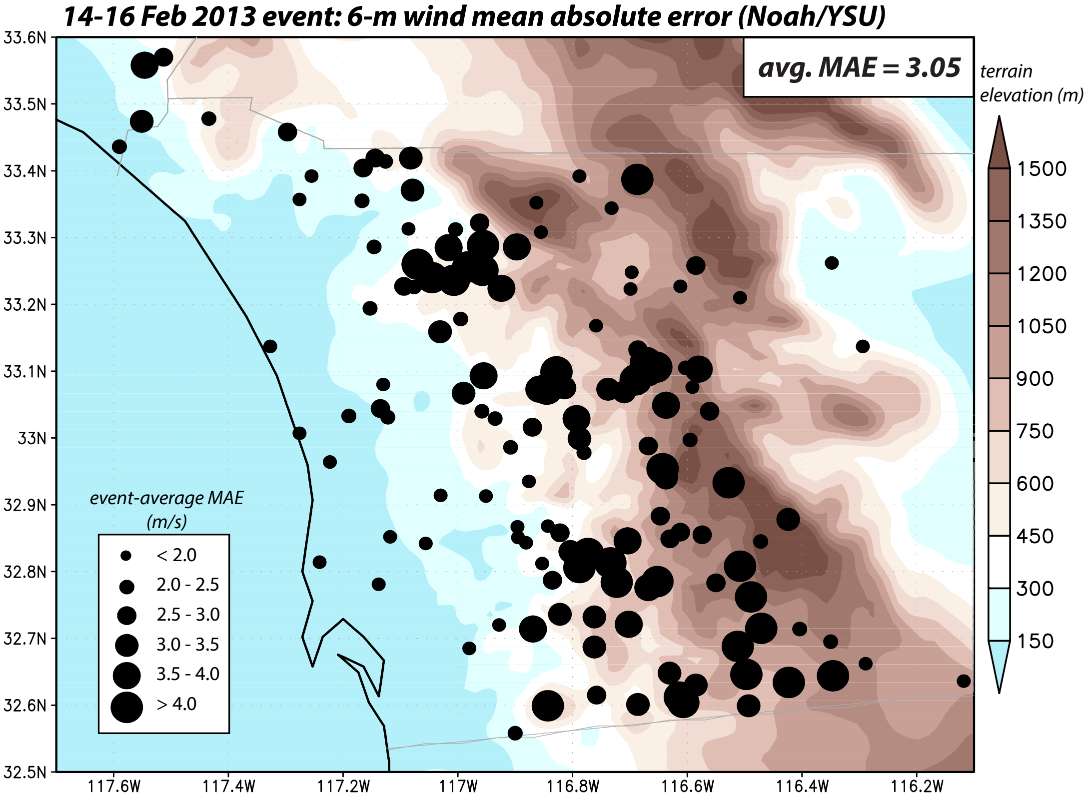


Figure 3.18: As in **Fig. 3.7**, but for the Noah/YSU simulation. The average value over all SDG&E network is 3.05 m s^{-1}.

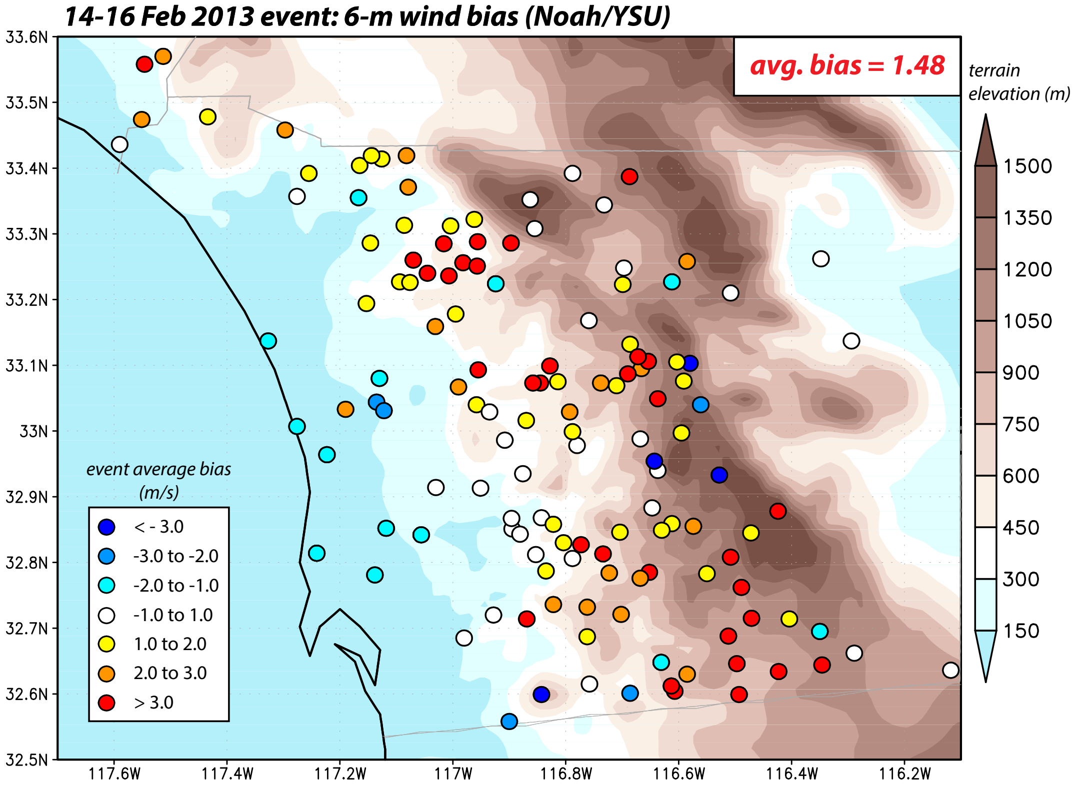


Figure 3.19: As in **Fig. 3.9**, but for the Noah/YSU simulation. The average value over all SDG&E network is 1.48 m s^{-1}.

3.3 Summary and concluding remarks

We have examined the 14-16 February 2013 Santa Ana event, which possessed many characteristics of a moderately intense downslope windstorm on the west-facing slopes of San Diego County. As expected, the event was associated with a surface Great Basin High and a passage of 700 hPa ridge aloft. This study was made possible by observations from the SDG&E mesonet, a dense, homogeneous and reliable observation network of ~ 140 stations sited in wind-prone areas, especially in the mountainous backcountry of San Diego County. These observations revealed that the 14-16 February 2013 Santa Ana episode consisted of two pulses separated by a protracted lull, and suggested the first phase possessed a hydraulic jump-like flow in part of the network, while the second was characterized by a clear downslope progression of the winds with time as the event itself wound down.

Mesonet data greatly facilitated validation and calibration of the model. Most configurations were found to consistently overpredict the winds at most stations in the SDG&E network, even after adjustments for the non-standard anemometer height (6.1 m AGL) were made. Even the best-performing model physics combination had some issues, including having larger MAE for winder stations, and the tendency to simultaneously overpredict less windy sites and underpredict flow speeds in windier areas of the network. Overall, however, the chosen physics combination, including the PX LSM and ACM2 PBL schemes, did a reasonable job of capturing the evolution and characteristics of this event.

The model was then used to fill in gaps in the observations, especially the vertical structure of the wind field. Vertical cross-sections revealed that a jump-like feature did form on the west-facing slope, which progressed upslope during the conclusion of the first phase of the event. The model also showed that while the winds were observed to be weak across the network during the afternoon lull, they actually stayed relatively strong near the ridgeline, in an area lacking stations.

In the next parts of this multi-part study, we will examine how and why model physics influences forecast skill with respect to the sustained winds, the sensitivity of the downslope flow to random perturbations, and address the important issue of gust parameterization, as the greatest concern is the impact of these high-frequency, small-scale wind bursts that models of the present time cannot resolve. As in the present study, a key role will be played by the exceptionally dense and homogeneous SDG&E network.

CHAPTER 4

Ensemble analysis

Hundreds of simulations were made for this study, which has included variations in the initialization data source [including the NCEP North American Regional Reanalysis (NARR), the NCEP/NCAR Reanalysis Project (NNRP) dataset, the NCEP Final (FNL) operational global analysis dataset, the NAM analyses and forecasts, the Rapid Update Cycle (RUC) analyses and forecasts, the NCEP Climate Forecast System Reanalysis (CFSR), and the European Reanalysis (ERA)-INTERIM datasets], WRF-ARW model version (3.2 through 3.5), model start time, grid and topographic resolution, and the size, configuration, placement of the model nests. In this chapter, we examine the sensitivity of the Santa Ana winds to model resolution, model physics, random perturbations, diffusion options, and land use database, as well as evaluate the "topo_wind" option, which was developed to parameterize the effects of the unresolved topography exerts over the surface circulations.

4.1 The 4-6 October 2013 and 13-15 May 2014 events

In addition to the mid-February 2013 Santa Ana from **Chapter 3**, we examined several other episodes, including the 4-6 October 2013 and the 13-15 May 2014 events. The 13-15 May 2014 event was a major Santa Ana wind event, which sparked several fires in the Rancho Bernardo, Oceanside and Camp Pendleton areas. The first fire to ignite was the Bernardo fire, which occurred as strong winds and gusts pushed to the coastline. Recall that the February 2013 event was characterized by two pulses separated by a protracted lull, with the first phase possessing a hydraulic jump-like feature in part of the network, while the second showing a more spatially uniform downslope flow. We will see that the October event was a single-peak episode that resembled the mid-February episode's second phase, while the May event was a much stronger two-peak episode.

The time evolution of the SDG&E network-averaged 6-meter sustained winds of the three events (**Fig. 4.1 a**) reveals that the February and the October events were comparable in strength while the May episode was stronger, with two peaks of similar strength[1]. **Figure 4.1 b** shows the SDG&E event-averaged (over 54 h) 6-meter sustained winds of all the stations, ranked with respect to their event-averaged wind from the February 2013 episode. Site Rancho Santa Fe (RSF) received the weakest winds in the February and October events, while station SIL generally experienced the strongest flow. While the May episode obviously had stronger wind speeds overall, the ordering of the stations is roughly similar for the three events, as evidenced by the relatively small degree of scatter with respect to the February ranking.

4.2 Sensitivity to resolution

High-resolution simulations require significantly more resources than lower-resolution ones, which may make them impractical in an operational environment. To put it simply, we do not wish to use finer grid spacing than is actually necessary. However, it is well appreciated that terrain gap and downslope flows are significantly modulated by the shape of the topography, which is in turn dependent on the resolution of the model grid and the topographic database. In this subsection, we explain why we elected to place a sub-1 km nest over the heart of the SDG&E network.

Resolution sensitivity is demonstrated using vertical cross-sections taken west-east across station WCK for simulations employing grid spacings of 667 m, and 2, 6 and 10

---

[1] The time interval shown for each case is to some degree arbitrary.

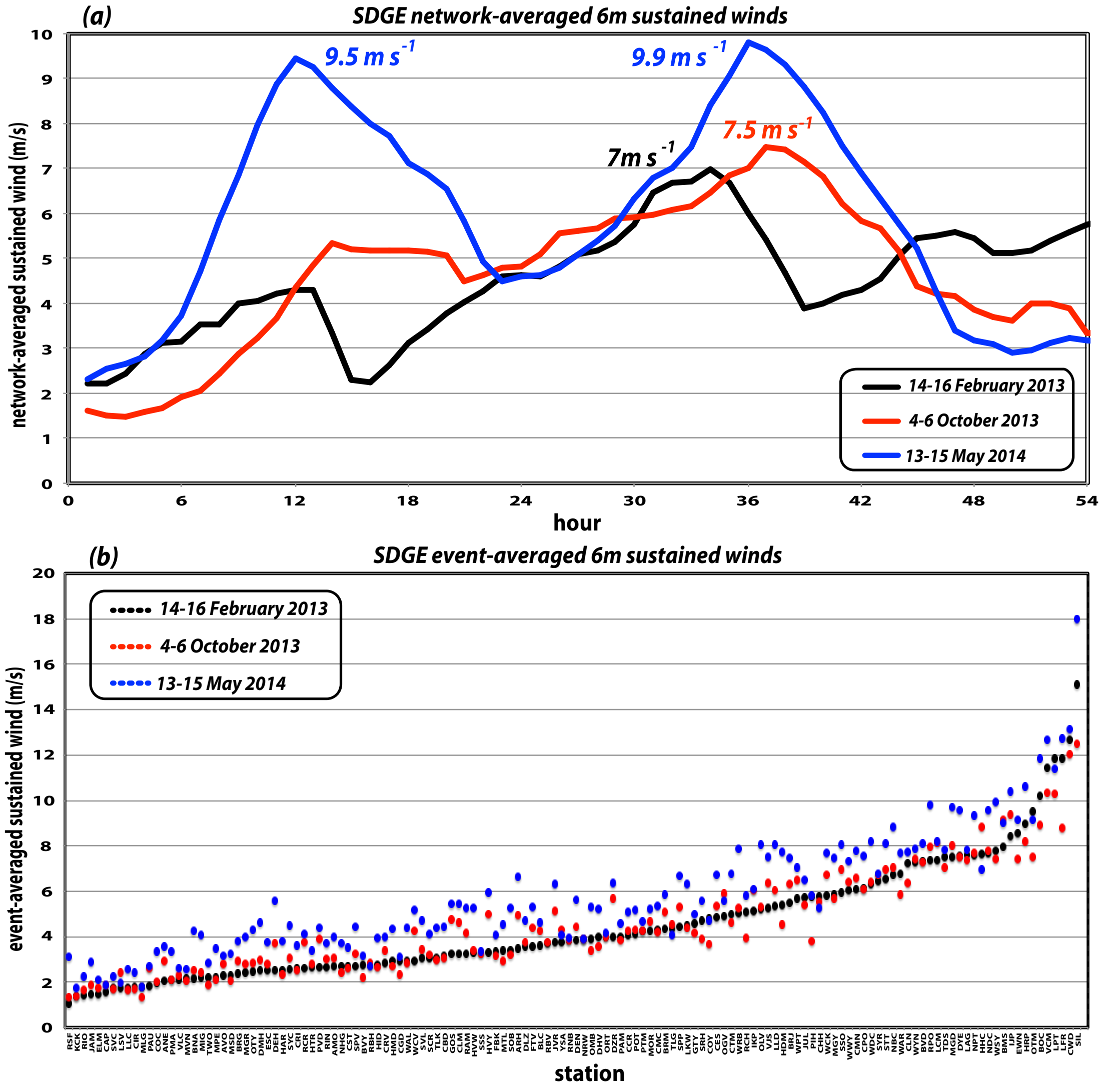


Figure 4.1: (a) Time series of SDG&E network-averaged observed 6-meter sustained wind (m s^{-1}) over 54 hours for the 14 February 2013 event (black line), the 4 October 2013 event (red line), and the 13 May 2014 event (blue line); and (b) SDG&E event-averaged observed 6-meter sustained wind (m s^{-1}) over 54 hours for the 14 February 2013 event (black dots), the 4 October 2013 event (red dots), and the 13 May 2014 event (blue dots).

km (**Fig. 4.2**). The finest grid, shown in **Fig. 4.2 a**, employs a very high resolution (~10 m) terrain database that is available from USGS[2]. (By comparison, the USGS terrain database distributed with the WRF model has a maximum resolution of about 30 sec, or ~1 km.) The new dataset sharpens up some of the topographic features, and overall it compares very well with the terrain as rendered within Google Maps (not shown). In particular, note the presence of a narrow northeast-southwest oriented canyon immediately east of WSY, and the steeper slope at the windward side. It is conceivable that these terrain features are important to the winds, especially at the ignition sites, in which case they should be resolved. These terrain features are notably absent in the 2 km topography (**Fig. 4.2 b**), but otherwise the rendition appears acceptable[3].

The fields shown in **Fig. 4.2** represent four-hour averages spanning the peak of the February 2013's first phase at WSY (around 1700 UTC on February 15^{th}; see **Fig. 3.5 a**), from simulations using the same model physics (PX/ACM2) as Chapter 3's control run. In that chapter, it was demonstrated that the persistent, jump-like feature (**Fig. 4.2 a**) in the easterly flow that developed in the control simulation is realistic (see also **Fig. 3.13 b**). This feature only appeared when sub-1 km spacing was used. The flow in the 2 km simulation (**Fig. 4.2 b**) is not unreasonable, resembling a spatially smoothed version of the 667 m run's flow. The strongest winds are still placed in approximately the right location, being close to station WSY.

Further resolution degradation, however, profoundly alters the slope of the terrain west of the ridge and improperly changes the location of maximum winds. At 6 km, the mountaintop has been flattened, and WSY has been pushed onto it (**Fig. 4.2 c**). The temporally-averaged winds are relatively weaker, and now located between SSO and WCK. The coarsest grid widens the ridge into a mesa, and the strongest winds are shifted even farther downslope (**Fig. 4.2 d**).

As the grid spacing becomes coarser, the terrain-accelerated flow becomes substan-

---

[2] http://nationalmap.gov/viewer.html

[3] The local rise in the topography that occurs at around 116.5°W longitude in the 2 km rendition (**Fig. 4.2 b**) results from a nearby topographic feature encroaching into the plane depicted owing to smoothing.

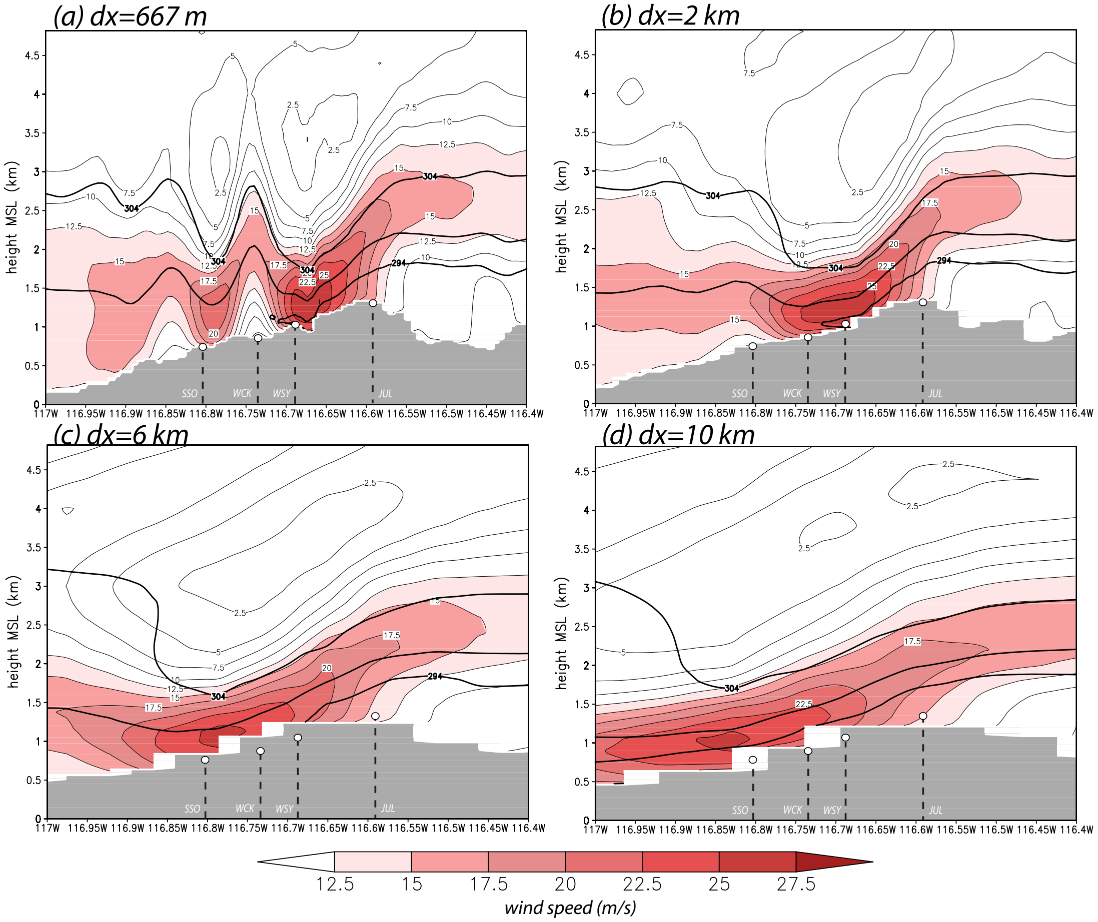


Figure 4.2: Vertical cross-sections of 4-hour averaged horizontal wind speed (2.5 m s^{-1} contours and red shaded fields) for 1500-1900 UTC 15 February 2013, taken west-east across Witch Creek for the (a) 667 m; (b) 2 km; (c) 6 km and (d) 10 km horizontal grid spacing simulations. Thick black contours are the 294, 299, and 304 K isentropes. Grey shaded area depicts topography.

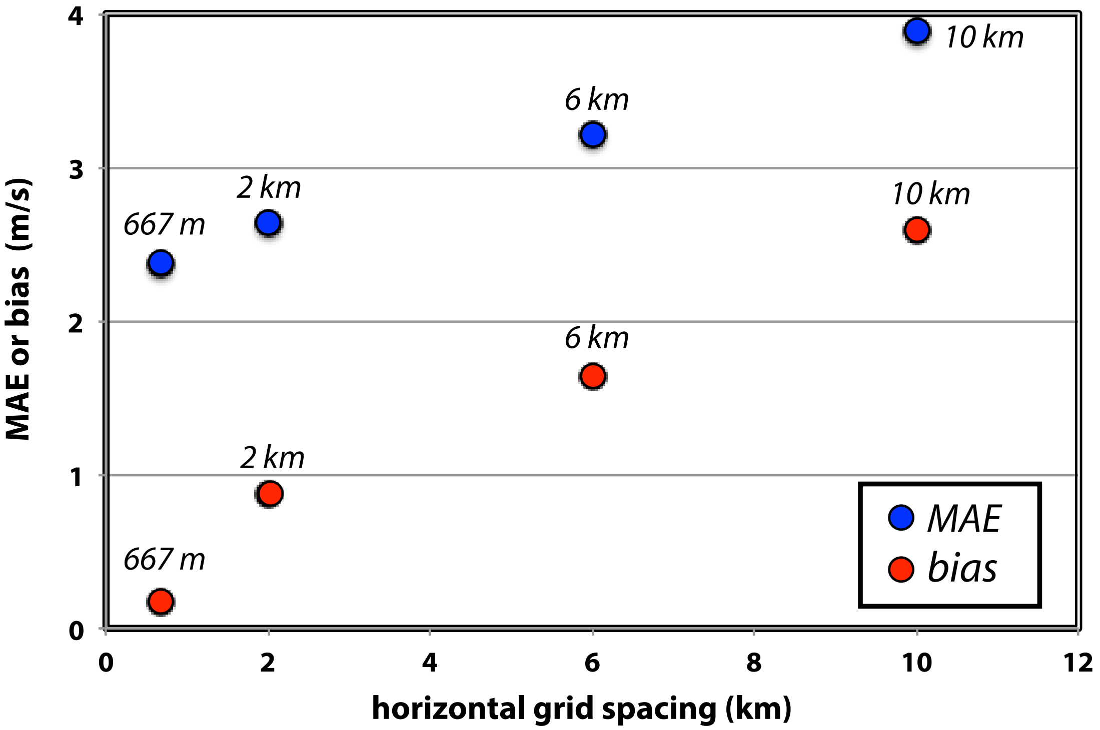


Figure 4.3: Horizontal grid spacing (km) vs. MAE or bias (both m s^{-1}) of the 6 m sustained winds for the 667 m, 2, 6, and 10 km horizontal grid spacing runs respectively. The MAEs and the biases are calculated based on the stations within the 667 m-domain (d05) boundary as shown in **Fig. 2.3** for all the runs. Blue points represent MAE, and red points represent bias.

tially larger in horizontal extent (although not stronger), and "fits" the observations more poorly. As a consequence, both MAE and bias increase with the grid spacing (**Fig. 4.3**). The October 2013 case yields similar conclusions (**Fig. 4.4**). If one's goal is to assess wind speeds at a particular time and place, accurately resolving the terrain shape is of paramount importance. Based on these results, it is concluded that grid spacing wider than 2 km cannot reliably place the fastest winds at the most likely correct locations. As a consequence, we chose to deploy a sub-1 km domain centered over the west-facing slopes that form the heart of the SDG&E mesonet.

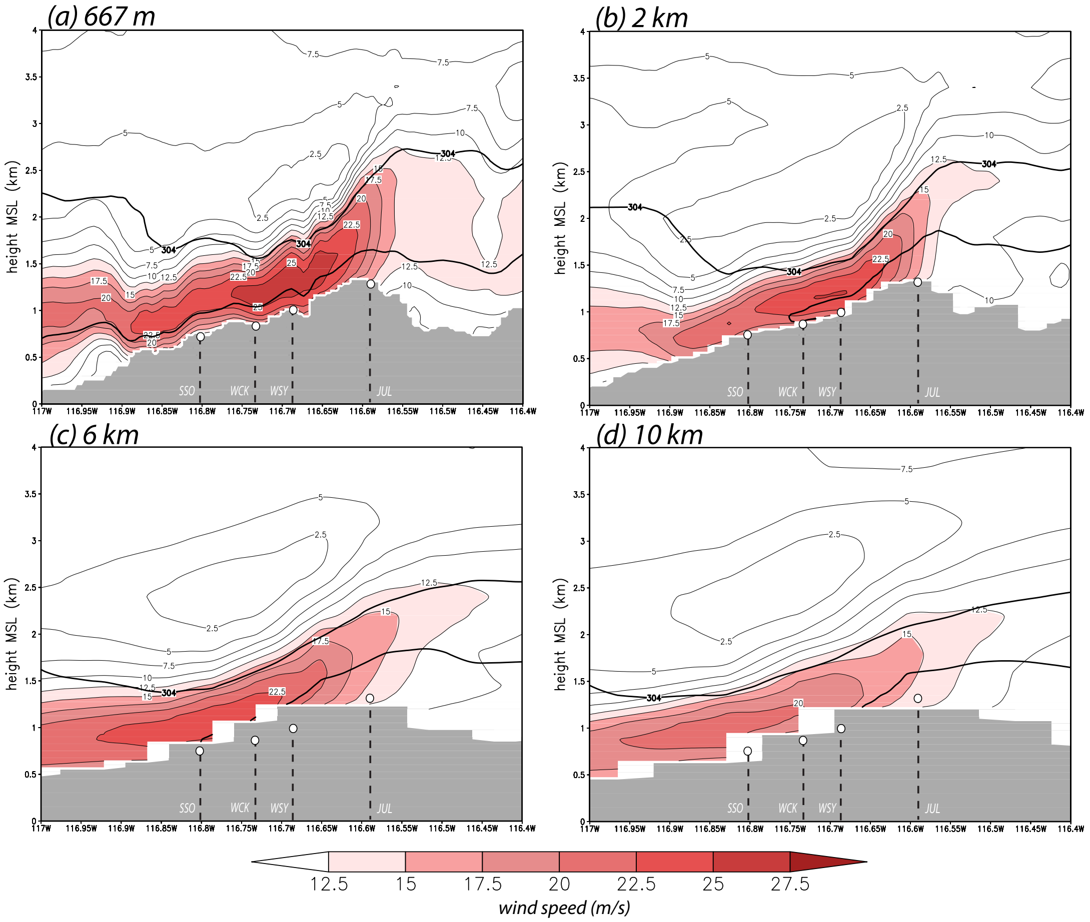


Figure 4.4: As in **Fig. 4.2**, but for 1400-1800 UTC 5 October 2013.

4.3 Sensitivity to model physics

In this section, we demonstrate the sensitivity of the terrain-amplified flow to the WRF model's physical parameterizations. The three Santa Ana episodes noted in **Sec. 4.1**, above, are examined. As in **Chapter 3**, the simulations in this section were made with WRF version 3.5 using five domains telescoping to 667 m grid spacing in the Santa Ysabel vicinity (**Fig. 3.3**). All were initialized with NAM analyses, either at 1200 UTC 14 February 2013, 0600 UTC 4 October 2013, or 0600 UTC 13 May 2014. All model integrations were for 54 hours.

Our physics ensemble involves an exhaustive examination of available model physical parameterizations, conducted to create many parallel realizations of the three events. The ensemble consists of variations of the LSM, surface layer, and PBL schemes. The LSMs included the Noah, Noah-MP ("Multi-parameterization"), RUC, PX, and Thermal Diffusion (TD) schemes. Ten PBL parameterizations were included, being the YSU, MYJ, Quasi-Normal Scale Elimination (QNSE), Mellor-Yamada Nakanishi and Niino level 2.5 (MYNN2), ACM2, Bougeault-Lacarrere (BouLac), Bretherton-Park (UW), Total Energy - Mass Flux (TEMF), Grenier-Bretherton-McCaa (GRM) and Medium Range Forecast (MRF) options. For each PBL scheme, the surface layer treatment cost commonly employed with it was selected. Some of the combinations were not workable, which left us with a total of 48 viable combinations.

Unsurprisingly, given the weather associated with Santa Ana events, we found that the downslope windstorms are not very sensitive to the treatment of microphysics or cumulus convection, and the sensitivity of the winds to the radiation parameterization was also very small. Therefore, all simulations examined herein employed the WRF Single-Moment 3-class scheme, a simple ice-bearing suitable for mesoscale grid sizes, the Kain-Fritsch cumulus schemes (in the 54 and 18 km domains only) and the RRTMG package for longwave and shortwave radiation. For the present, only the MODIS-based landuse database is employed. Sensitivity to this database is examined later in this chapter.

4.3.1 The 14-16 February 2013 case

4.3.1.1 The physics ensemble

As seen in **Chapter 3**, not all model configurations are created equal. The PX/ACM2 member was selected for the control configuration owing to its small MAE and nearly zero network-integrated bias (**Figs. 4.5 and 4.6 a**) for the event-averaged sustained wind, and because it did a good job of capturing the evolution of the February 2013 event. The popular Noah/YSU combination was not as skillful (**Fig. 4.6 b**) and the worst simulations were those employing the Noah-MP and TD LSMs (**Fig. 4.5**). Note that **Fig. 4.5** also reveals that both bias and MAE were most influenced by the LSM, with the PBL scheme being of secondary importance. Note further that nearly all of the ensemble members overpredicted the winds averaged over all stations and times, despite careful adjustments to the mesonet's 6.1 m anemometer mounting height (see **Chapter 2**). Had those adjustments not been made, the positive biases would have appeared even larger.

Variation among ensemble members with respect to the wind was concentrated on the lee (western) slope of the mountains, especially in the Santa Ysabel area and near the Mexican border (**Fig. 4.7**). Vertical cross-sections taken west-east across WCK (similar to **Fig. 4.2**) for three members, PX/ACM2, Noah/YSU, and TD/YSU (**Figs. 4.8 a, b, and c**), demonstrate the sensitivity of the vertical wind structure to model physics variations. These are again 4-hour averages, centered on the time of WSY's first peak and WCK's first wind reversals (**Figs. 3.5 a and b**), and thereby represent temporally persistent features. The PX/ACM2 and Noah/YSU members develop hydraulic jumps but place them differently; the former's siting is more consistent with the observations (**Chapter 3**).

However, many of the simulations, like the TD/YSU member (**Fig. 4.8 c**), fail to

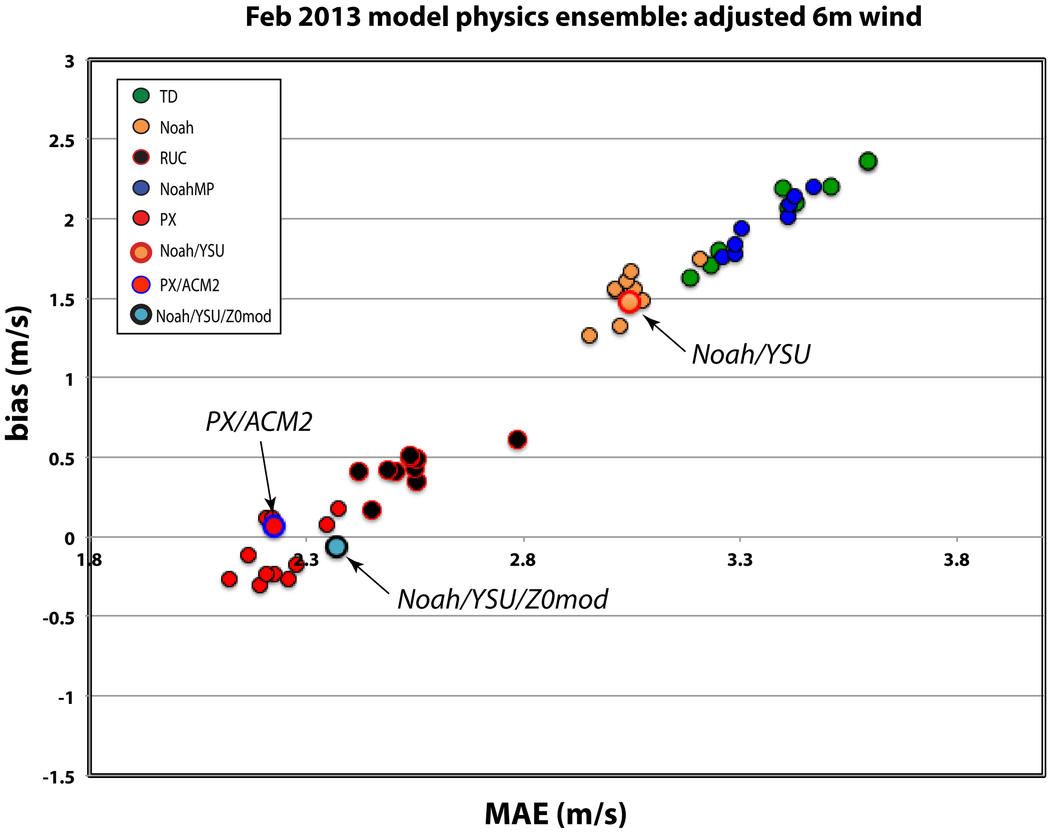


Figure 4.5: Mean absolute error (MAE) vs. bias (both m s^{-1}) of the 6 m sustained winds from a February 2013 event's physics ensemble incorporating 5 LSMs and 10 PBL schemes. Points represent event-averaged values, and are color-coded by LSM, with the PX/ACM2, Noah/YSU, Noah/YSU/$z_0$mod members highlighted.

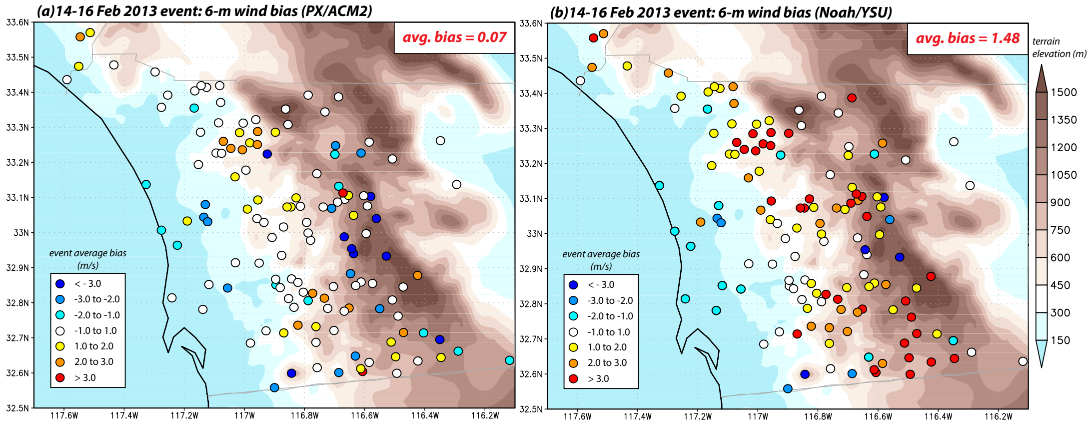


Figure 4.6: (a) Same as **Fig. 3.9**; and (b) same as **Fig. 3.19**.

develop much of a jump, which is why the feature does not appear in the 48-member ensemble average for this same time period (**Fig. 4.9 a**). As anticipated from **Fig. 4.7**, the ensemble spread is largest at and above WCK (**Fig. 4.9 b**). A secondary region of variation resides just downslope of SSO, representing differences with respect to the downslope extent of the strong winds during this time window. At least all of the ensemble members place the strongest winds above and just upslope of WSY, which reflects the model resolution more than the particular physics employed (cf. **Fig. 4.2**).

Figure 4.10 a focuses in on the near-surface (10-m) winds during this same 4-h window[4]. Note the variation among the 48 ensemble members was quite small upwind of, and beyond, the ridge, until the easterly flow passed the narrow canyon just upslope from WSY. From that point downslope, the variation has become quite substantial (2-18 m s^{-1}), in the very region where is a crucial need for skillful forecasts. As suggested by the figure, only a subset of the physics ensemble members captured the weak winds observed at or very near WCK. For contrast, we can also examine a 4-h window around the second phase's peak (**Fig. 4.10 b**), during which time a jump did not appear in the observations. Note that the region of largest variation on the lee side has shifted farther down the slope, and there is more spread in the flow on the east side of the ridge,

---

[4] As this is a comparison among ensemble members, these winds are not adjusted to the SDG&E anemometer height.

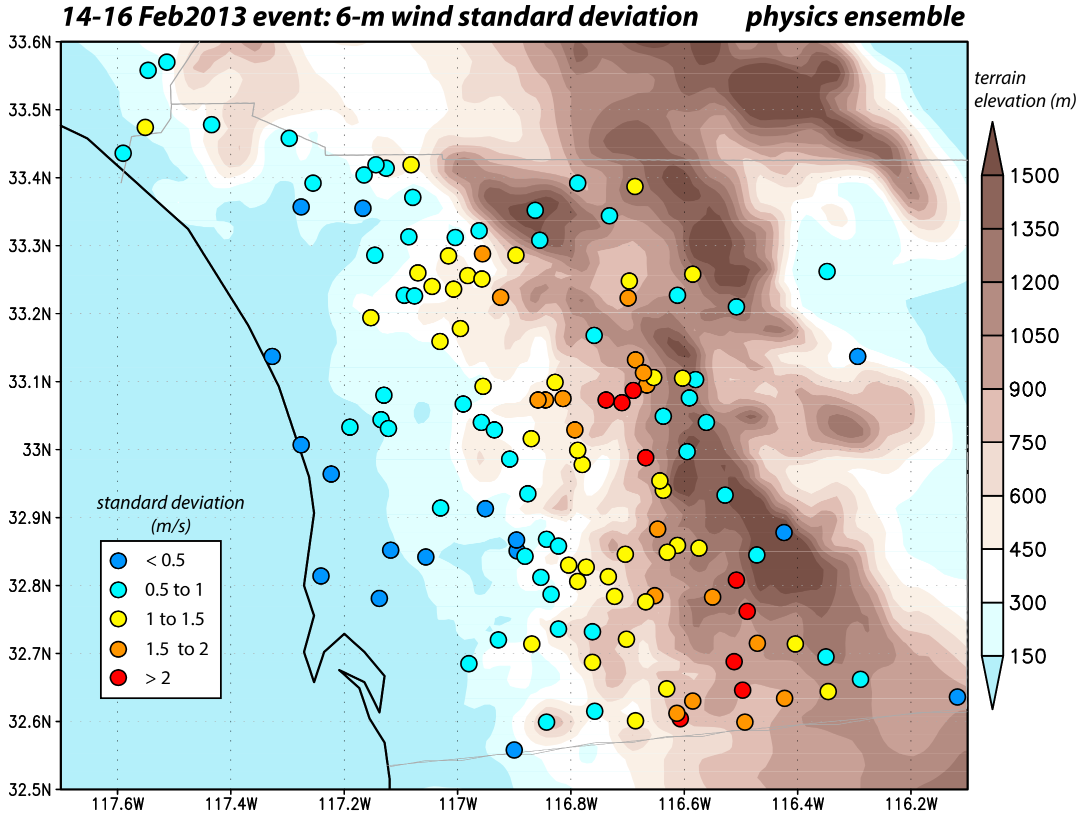


Figure 4.7: Spatial distribution of the physics ensemble standard deviation of 6 m sustained winds for SDG&E stations, using the MODIS landuse database. Deep red and blue colors indicate the physics ensemble standard deviation is relatively large and small. Brown shading indicates model topography.

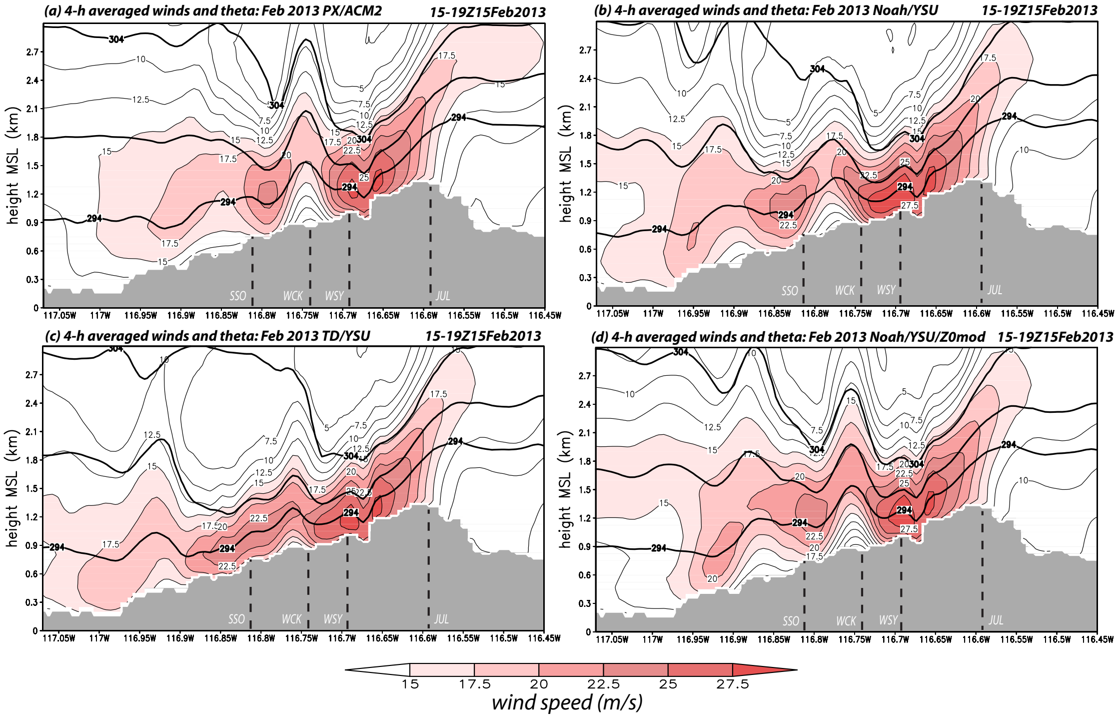


Figure 4.8: Similar to **Fig. 4.2**, but for four members of the 667 m physics ensemble. Members are: (a) PX/ACM2; (b) Noah/YSU; (c) TD/YSU and (d) Noah/YSU/$z_0$mod.

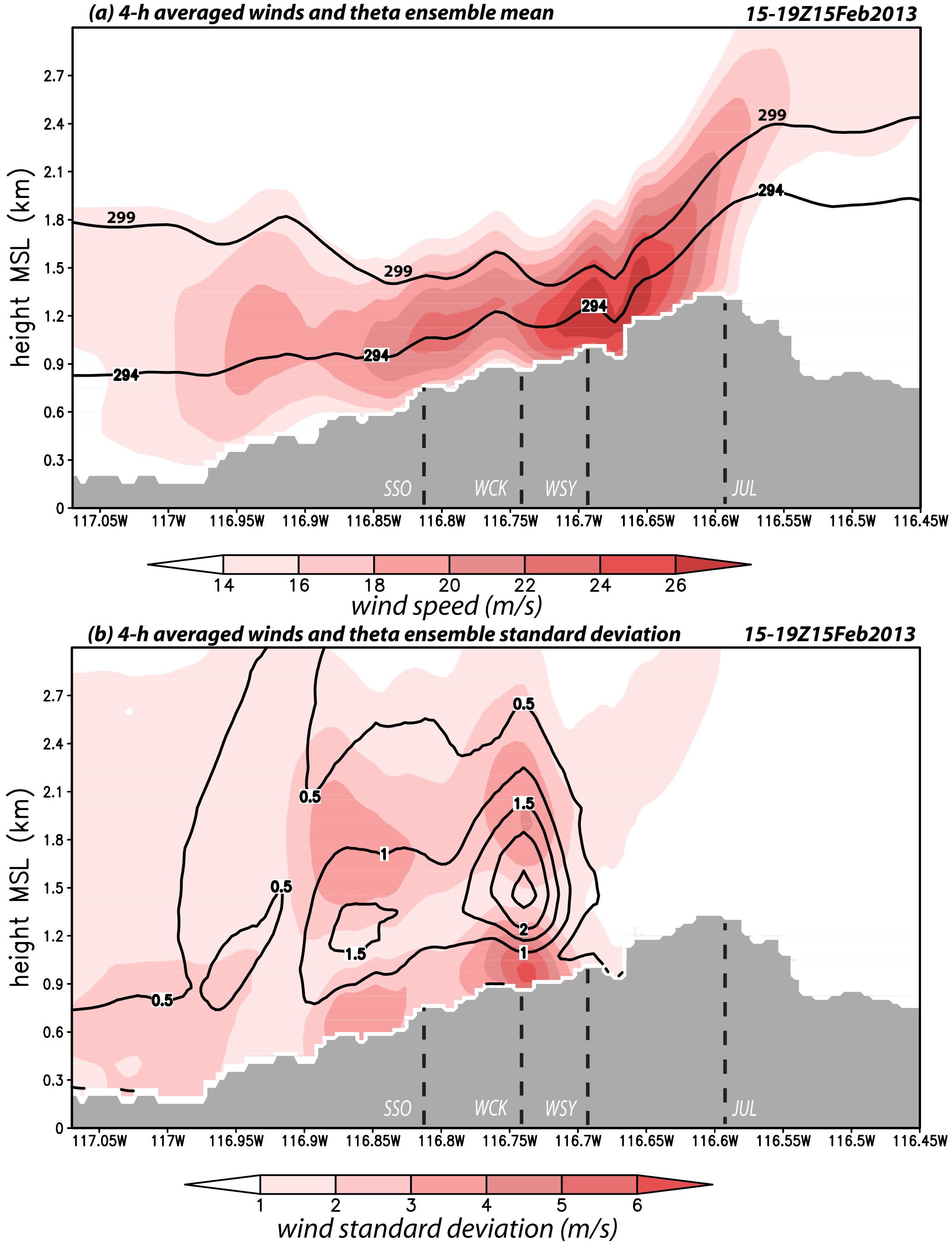


Figure 4.9: Similar to **Fig. 4.2**, but for the physics ensemble (a) mean horizontal wind speed (m s^{-1}) and isentropes (K); and (b) their standard deviation.

perhaps reflecting the natural tendency for nonlinear simulations to diverge with time.

4.3.1.2 The influence of surface roughness

The February 2013 physics ensemble has revealed that the single most important physics option controlling the quality of the mean wind reconstruction is the LSM. With respect to the near-surface wind, our analysis indicates that LSMs differ most with respect to how they handle the surface roughness, z_0. In theory, z_0 is related to the height of surface obstacles, and in practice represents the height at which the logarithmic wind profile (cf. **Chapter 2**) reaches zero wind speed. In the WRF model, the roughness for a particular location depends on its landuse categorization, which depends on the database (e.g., USGS or MODIS) being used and, for most schemes, the time of year (winter vs. summer). (Although the logarithmic wind profile is calculated in the surface layer scheme, z_0 values provided by the LSM are used. Hence, the sensitivity is attached to the LSM.)

In the MODIS representation, 67% of the SDG&E network's landscape (**Fig. 4.11**) is represented by either its "closed shrublands" (category 6) or "open shrublands" (category 7) classifications (see **Table 4.1**). The default z_0 value for these categories is 0.05-0.06 m in summertime, and 0.01 m in winter. The TD and Noah-MP schemes use the default z_0 assignments, while others make modifications, which is what creates the physics ensemble diversity with respect to the wind reconstructions. The Noah LSM permits z_0 to vary through the year for some landuse categories, including the shrublands, so that each event might be utilizing somewhat different surface characteristics; values for category 7 are 0.017 m in early October and 0.06 m in mid-May. The RUC and PX schemes increase the roughness of the shrubland categories considerably, with the latter using $z_0 = 0.15$ throughout the year (**Table 4.1**), fifteen times larger than the default wintertime value. **Figure 4.12** compares the roughness lengths of PX and Noah for the February event. Note the PX LSM presumes a rougher surface virtually everywhere within the SDG&E network.

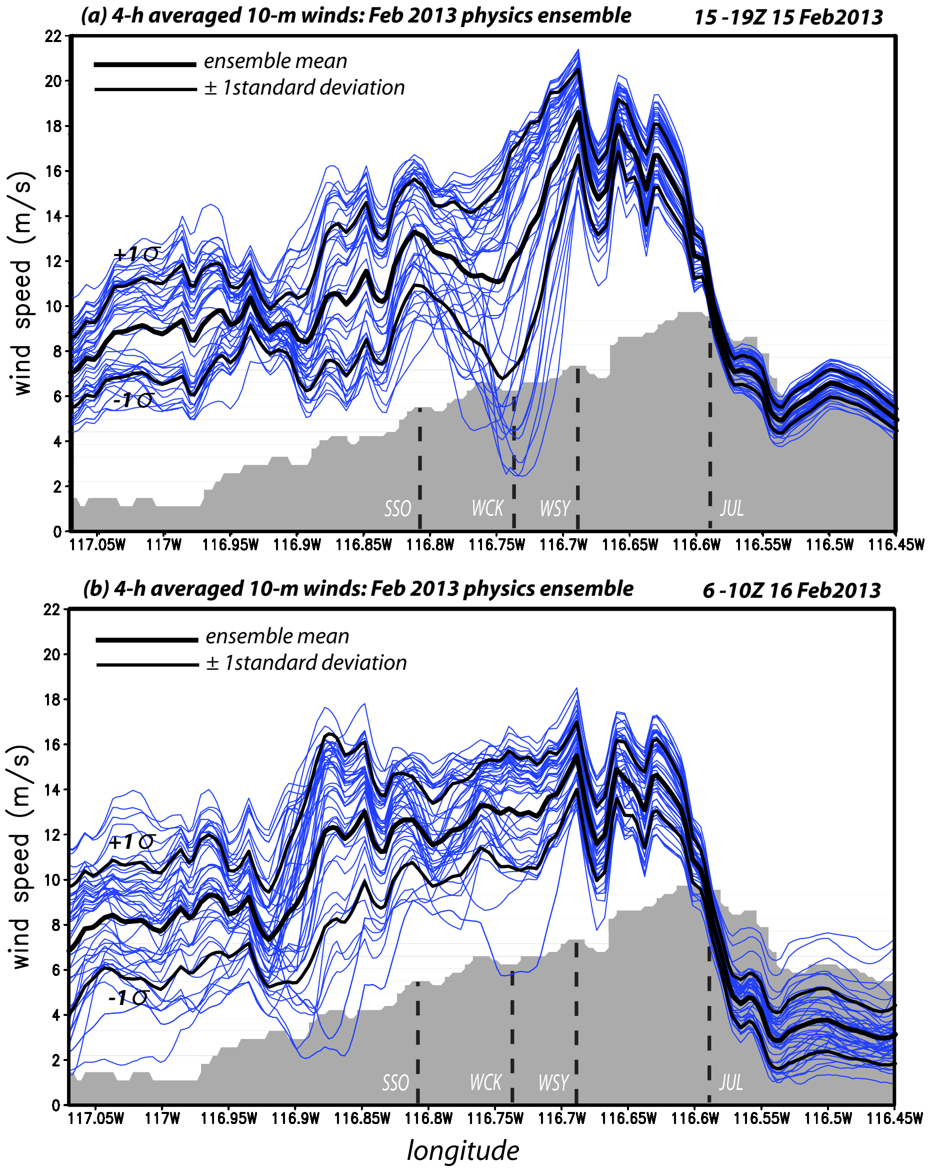


Figure 4.10: Vertical cross-sections of the physics ensemble mean (thick black line) and ± 1 standard deviation (thin black lines) of the 4-hour averaged horizontal 10 m wind speed (blue lines) for (a) the first phase 1500-1900 UTC 15 February 2013; and (b) the second phase 0600-1000 UTC 16 February 2013 of this event, taken west-east across Witch. Grey shaded area depicts topography. Approximate locations of Juline, West Santa Ysabel, Witch Creek, and Sunset Oaks are shown.

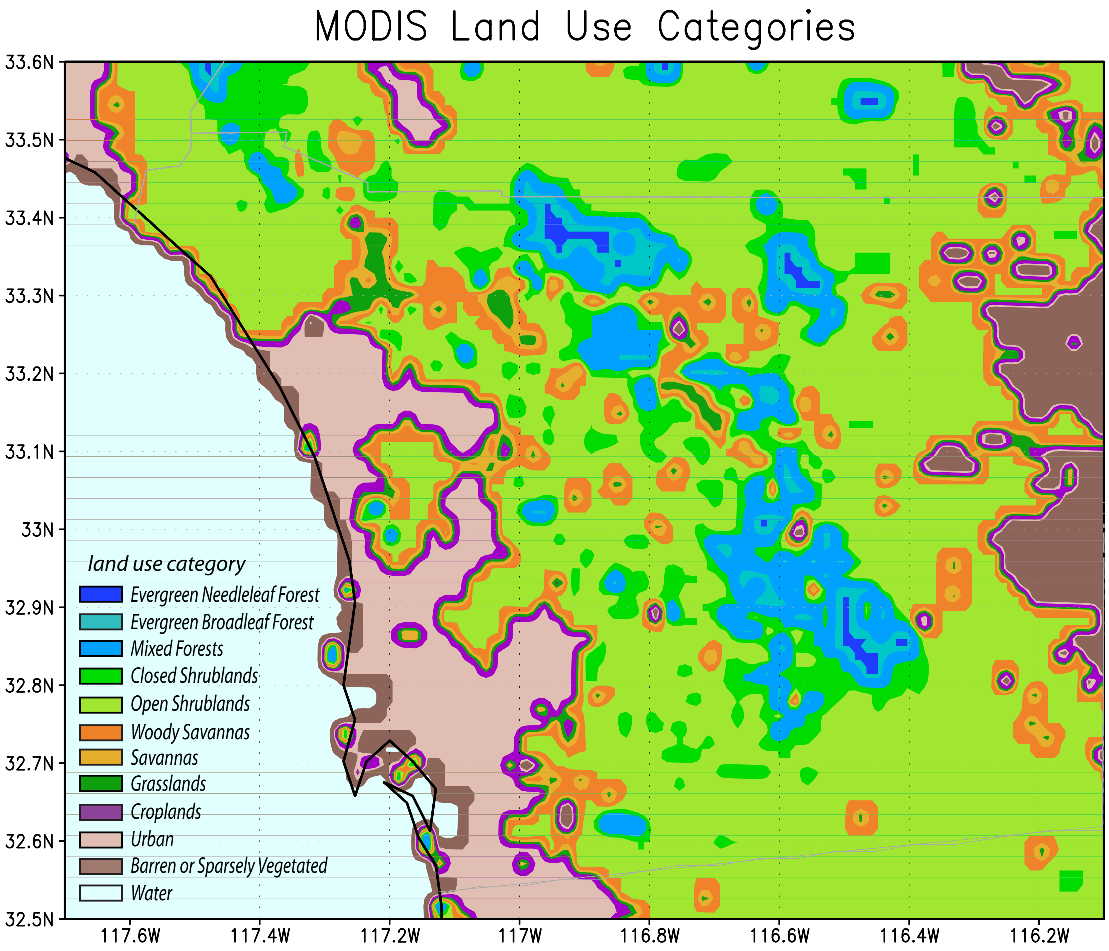


Figure 4.11: Land use category map for MODIS, color coded by different land use categories.

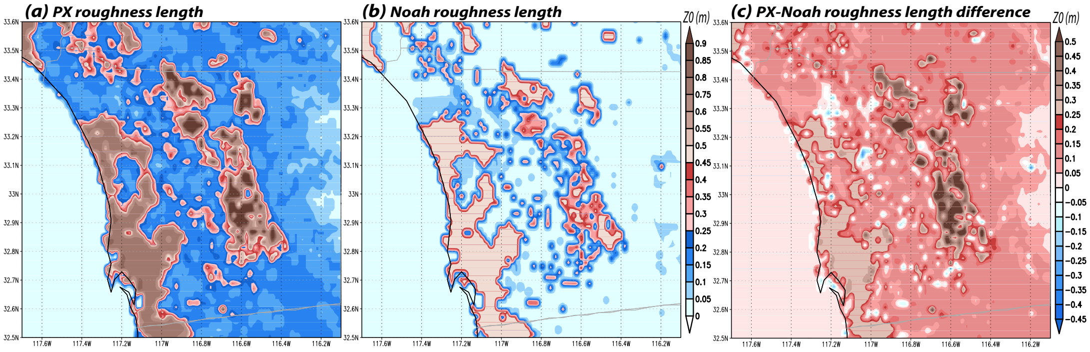


Figure 4.12: Wintertime roughness length (z_0) map for LSMs (a) PX; (b) Noah; and (c) the difference between PX and Noah.

The importance of z_0 in an LSM is demonstrated by modifying the Noah scheme to mimic PX by using its surface roughness assignments. This simulation, dubbed “YSU/Noah/$z_0$mod”, yields a much more faithful reconstruction of the network-averaged wind with respect to event mean bias and MAE (**Figs. 4.5, 4.6 b, and 4.13 b**). Indeed, the correspondence between PX/ACM2 and YSU/Noah/$z_0$mod (compare **Figs. 4.13 a and b**) is very high, demonstrating that z_0 is the principal controlling factor. The match is not likely to be perfect, in part because the PBL scheme does influence the results and also because the PX and Noah LSMs handle fractional landuse differently. Still, using PX-inspired roughness values in Noah clearly resulted in superior wind performance and a very small network-averaged bias.

The surface roughness does not just impact the anemometer-level wind estimates, but can also influence the structure and behavior of the downsloping flow. We note that only the simulations that presumed relatively larger surface roughness on the west-facing slopes developed the jump-like feature above WCK. Given the PX LSM’s z_0 values, the Noah/YSU/$z_0$mod member produced a downslope wind structure that strongly resembles the PX/ACM2 (control) run (compare **Fig. 4.8 d** with **4.8 a**). One would not like to make firm conclusions based on a single event, so the next two subsections explore physics diversity and sensitivity to z_0 in two other Santa Ana episodes.

Table 4.1: Default roughness lengths employed by surface schemes for MODIS landuse categories occurring in the SDG&E network for winter and summer season simulations. Water areas of the 2 km nest excluded.

MODIS LU INDEX	LAND FRAC (%)	PX	Noah (Feb)	MODIS (winter)	Noah (Oct)	Noah (May)	MODIS (summer)	TYPE
1	5.9	1	0.5	0.5	0.5	0.5	0.5	Evergreen Needleleaf Forest
2	0.2	0.9	0.5	0.5	0.5	0.5	0.5	Evergreen Broadleaf Forest
5	6.0	1	0.3	0.2	0.23	0.5	0.5	Mixed Forests
6	11.8	0.15	0.03	0.01	0.02	0.05	0.05	Closed Shrublands
7	54.4	0.15	0.04	0.01	0.02	0.06	0.06	Open Shrublands
8	0.2	0.25	0.05	0.01	0.06	0.05	0.05	Woody Savannas
9	0.3	0.15	0.10	0.15	0.1	0.1	0.15	Savannas
10	2.3	0.07	0.1	0.1	0.2	0.1	0.12	Grasslands
11	0.1	0.2	0.18	0.3	0.3	0.3	0.3	Permanent Wetlands
12	0.4	0.1	0.3	0.05	0.4	0.4	0.15	Croplands
13	11.2	0.8	0.5	0.8	0.5	0.5	0.8	Urban
16	7.3	0.05	0.01	0.01	0.01	0.01	0.01	Barren/Sparse Vegetated

4.3.2 The 4-6 October 2013 case

Maximum wind gusts observed in the SDG&E network for the 4-6 October 2013 event are shown in **Fig. 4.14**. We have seen this was a single-peak event, in contrast to the February 2013 episode's two-phase structure (**Fig. 4.1 a**). Similar to the February 2013 event, the strongest gusts observed were found to be located along the western slopes, close to but not right at the ridgelines, although the maximum gusts were relatively weaker during this episode. The largest winds and spatial variability still resided in the Santa Ysabel and Sill Hill areas, while gusts recorded near the coastline were rather weak.

As in the February 2013 episode, the plot of event-averaged MAE vs. bias revealed a roughly linear relationship, with ensemble members differentiated primarily with respect to LSM (**Fig. 4.15**). The PX and RUC LSMs have again performed better with respect to event-averaged MAE and bias, with the PX/ACM2 physics combination still being one of the best. For this event, the Noah runs were the *worst* performers over-

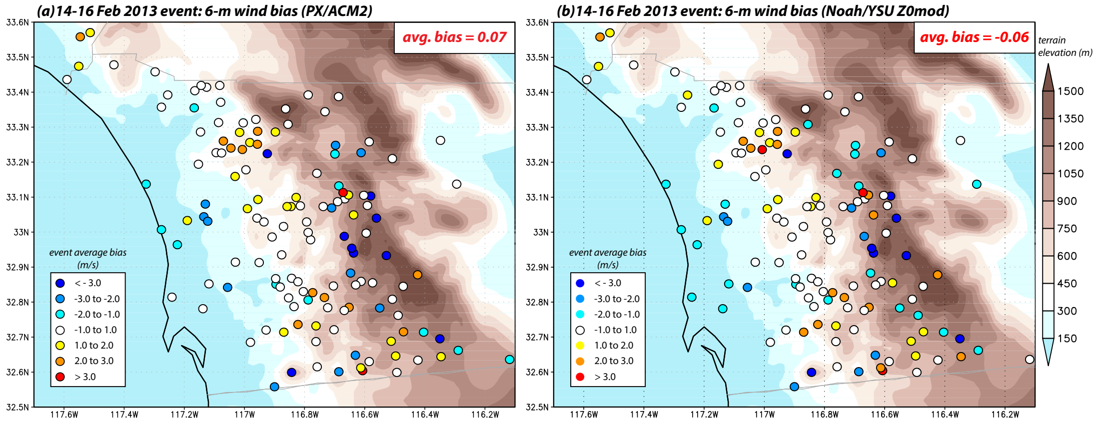


Figure 4.13: (a) Same as **Fig. 3.9**; and (b) As in (a), but for the Noah/YSU/$z_0$mod member.

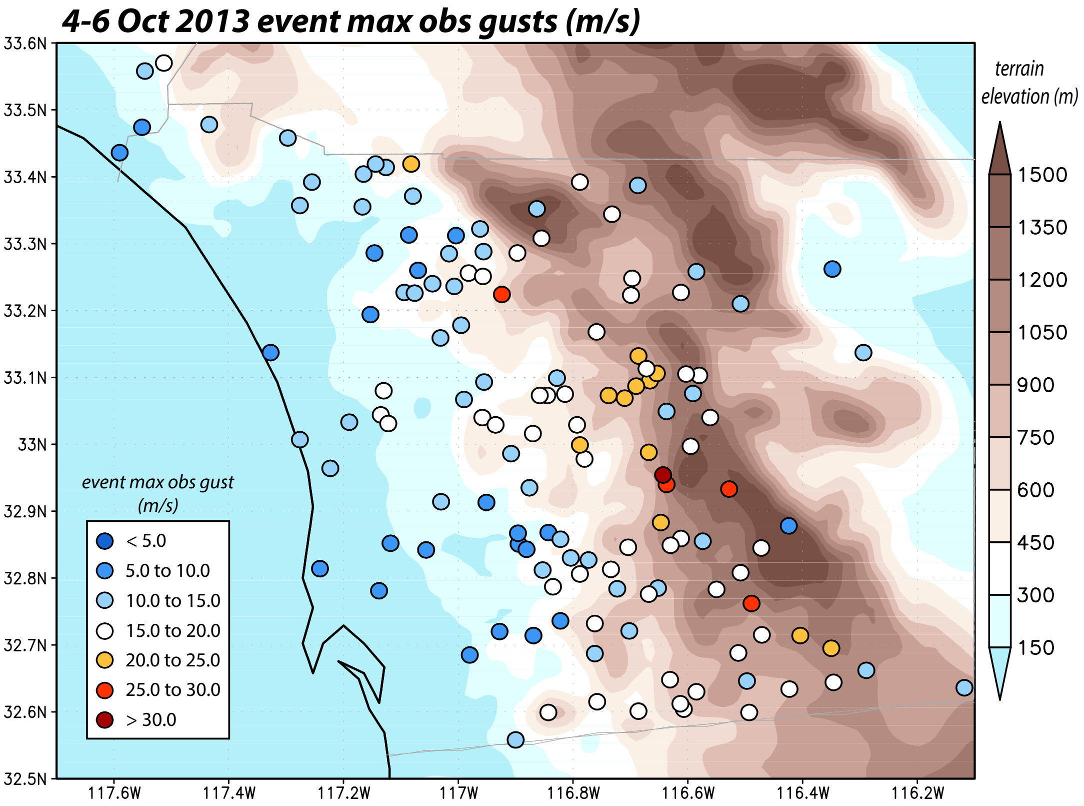


Figure 4.14: As in **Fig. 3.2**, but for the 4 October 2013 event.

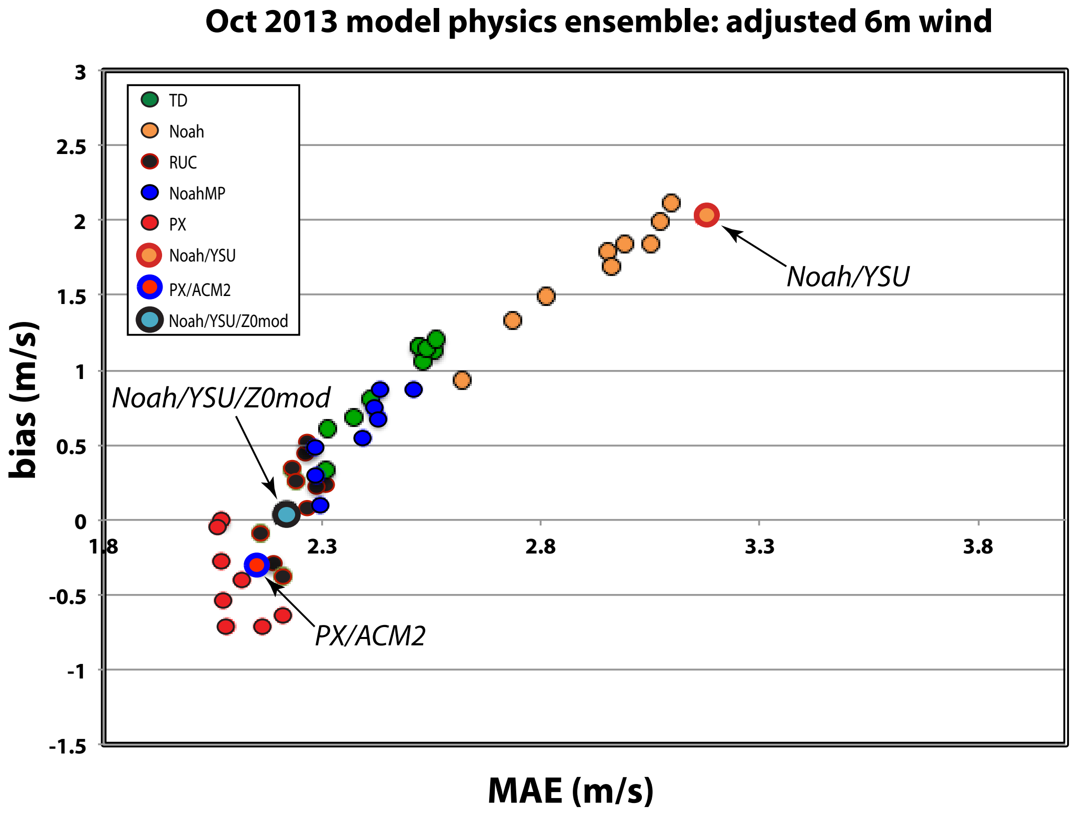


Figure 4.15: As in **Fig. 4.5**, but for the 4 October 2013 event.

all with the Noah/YSU member having the largest MAE and 2nd largest bias. Note that many more stations had large MAEs in the Noah/YSU simulation relative to its PX/ACM2 counterpart (**Figs. 4.16 a and b**).

In this case, the Noah LSM's seasonal variation appears to be to blame. The early October event occurred during the summer half-year, which in WRF runs from April 14 to October 14, so the TD and Noah-MP used that season's default roughness values of 0.05 and 0.06 m for categories 6 and 7, respectively, which are larger than the winter values (0.01 m) applied to the shrublands in the February episode (**Table 4.1**). The Noah scheme's z_0 values were 0.015 and 0.017 for these same locations, about 3.5 times smaller than the MODIS default for summer and 9-10 times smaller than PX's temporally-invariant assignments (**Fig. 4.17**). As a consequence, the Noah scheme was

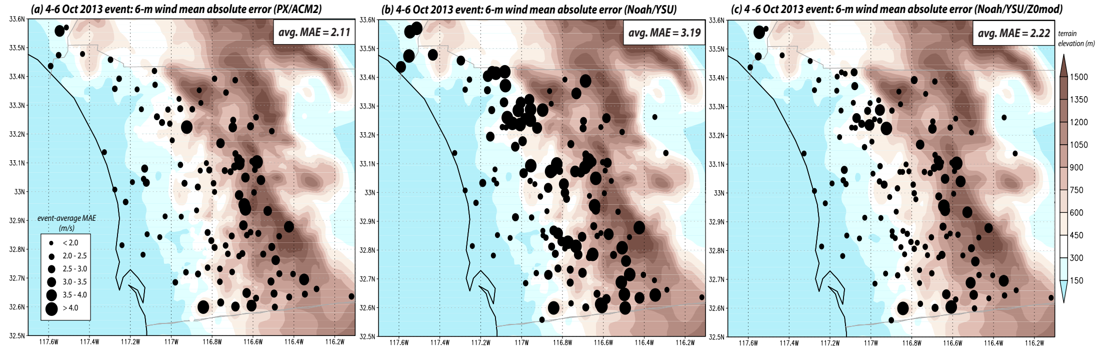


Figure 4.16: As in **Fig. 3.18**, but for (a) PX/ACM2; (b) Noah/YSU; and (c) Noah/YSU/$z_0$mod of the 4 October 2013 event.

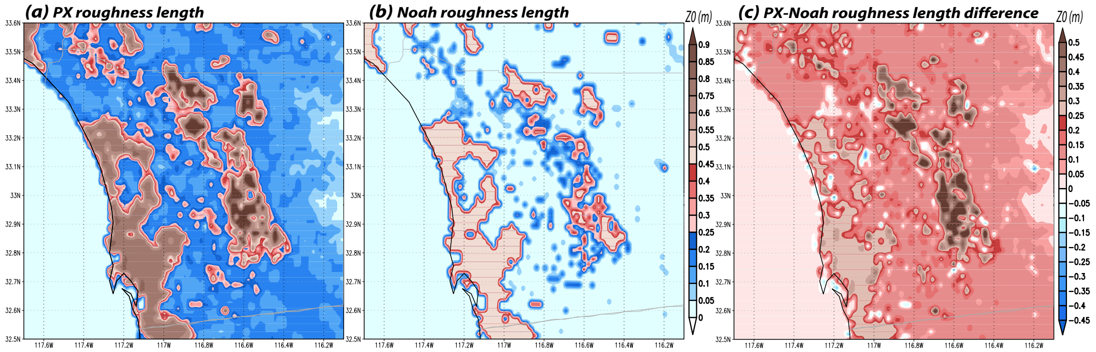


Figure 4.17: As in **Fig. 4.12**, but for early October.

employing the ensemble's smallest surface roughnesses in the area where the downslope winds were strong, and this led to the positive biases and sizable MAEs of its members.

As with the February 2013 episode, also examined the modified version of the Noah/YSU member (Noah/YSU/$z_0$mod) that adopted the PX scheme's roughness values. While the Noah/YSU/$z_0$mod and PX/ACM2 statistics differed a bit more with respect to event-integrated forecast skill (**Fig. 4.15**) and the spatial distribution of sustained wind bias (**Fig. 4.18**), they were still very similar and represented a considerable improvement relative to the original Noah/YSU simulation.

Regarding the rest of the physics ensemble, **Fig. 4.19** shows that the ensemble

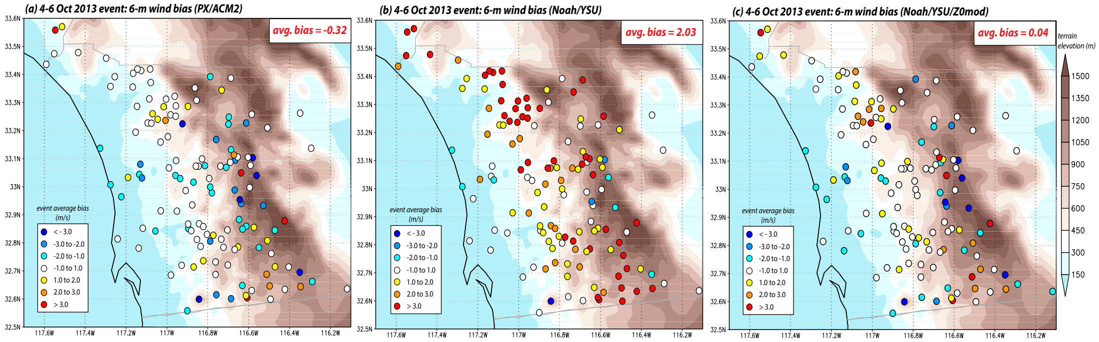


Figure 4.18: As in **Fig. 3.9**, for (a) PX/ACM2; (b) Noah/YSU; (c) Noah/YSU/$z_0$mod of the 4 October 2013 event.

spread in wind speed is smaller relative to the February 2013 episode, but still centered mainly on the west-facing slopes. Consistent with the observations, the downsloping flow during this event did not develop a hydraulic jump. Instead, the flow pattern resembled that seen during the February episode's second phase (**Fig. 4.20**), consisting of a relatively deep and spatially extensive tongue of easterly flow. These fields are again 4-h averages, centered around the time of the event's peak intensity.

As confirmed by the ensemble average and standard deviation fields for this same time window (**Fig. 4.21**), variations among the ensemble members were relatively subtle. The obvious differences among the ensemble members shown in this figure and in **Fig. 4.20** were with respect to how far westward the offshore flow was able to progress. This extent was largest in the Noah/YSU case (**Fig. 4.20 b**), accounting for its sizable positive biases (**Fig. 4.18 b**). Overall, the spatial variation of the ensemble spread was relatively small across the Santa Ysabel area (**Fig. 4.21 b**), and the spread of the near-surface wind (**Fig. 4.22**) largely reflects differences in z_0 rather than essential structural variations.

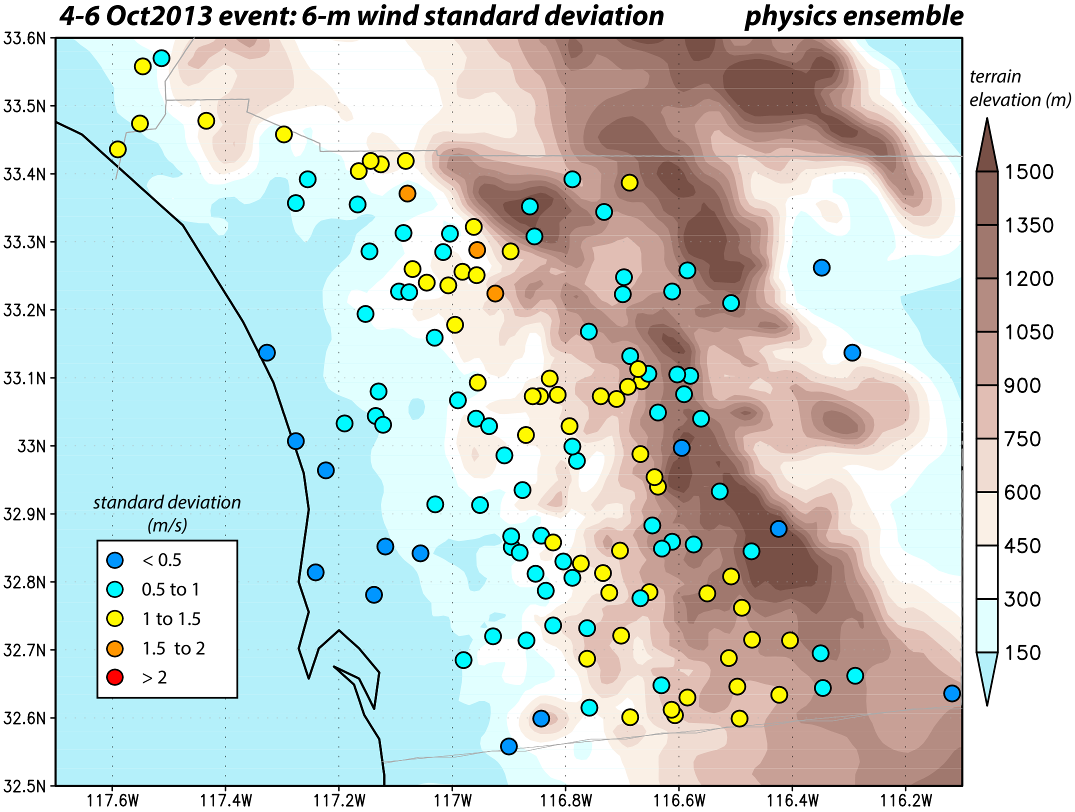


Figure 4.19: As in **Fig. 4.7**, but for the 4 October 2013 event.

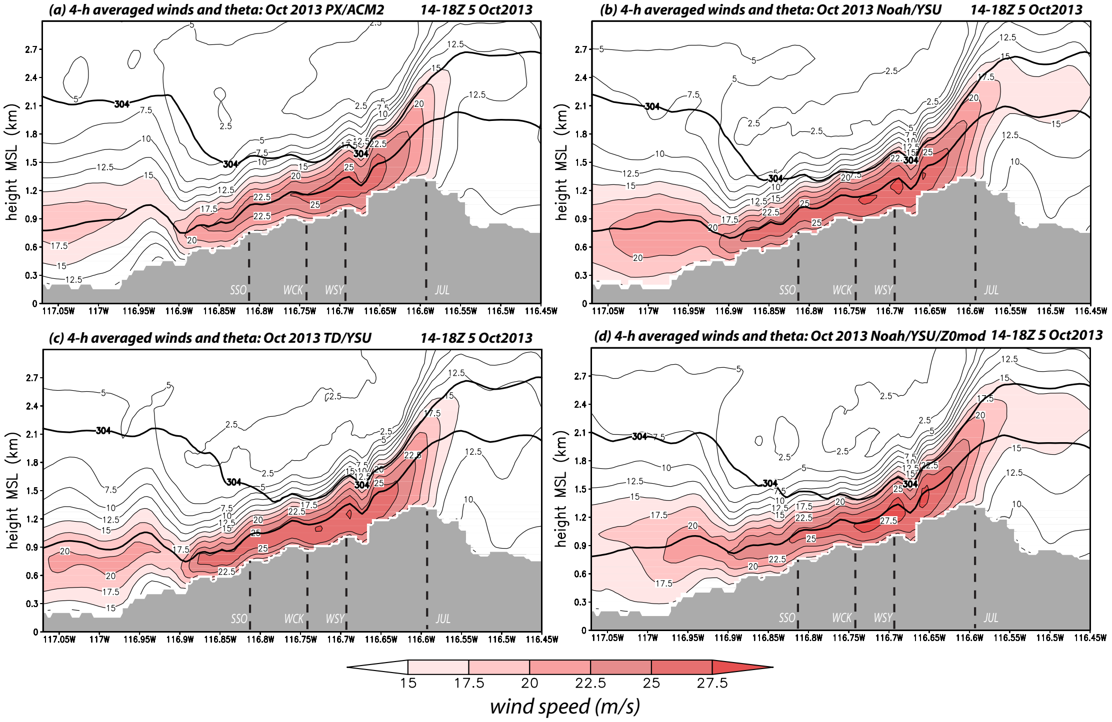
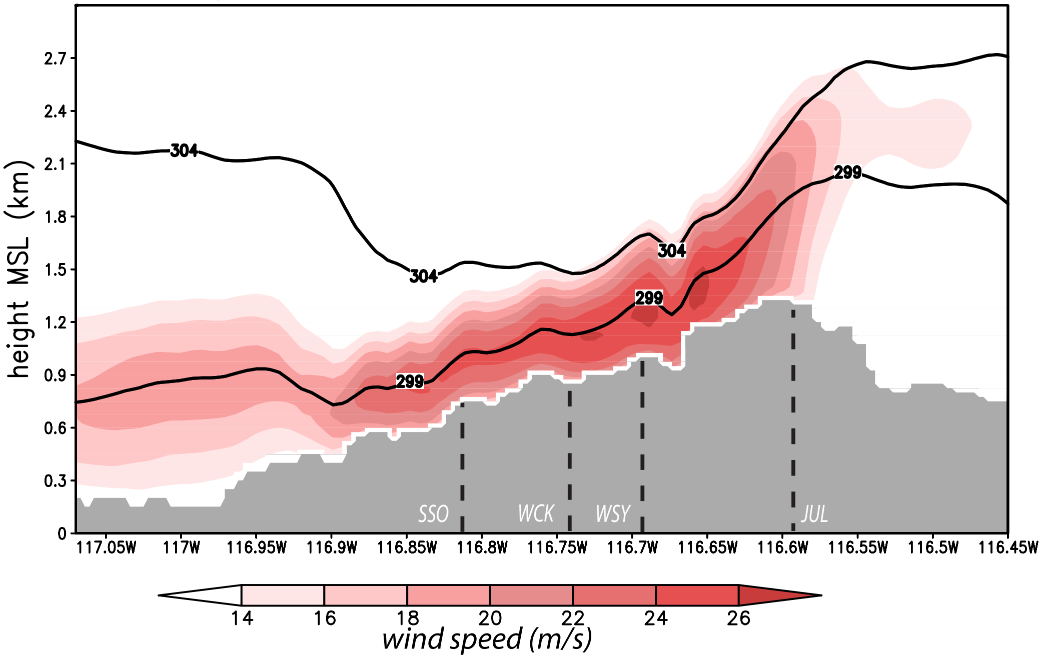


Figure 4.20: As in **Fig. 4.8**, but for 1400-1800 UTC 5 October 2013.

(a) 4-h averaged winds and theta ensemble mean 14-18Z 5Oct2013



(b) 4-h averaged winds and theta ensemble standard deviation 14-18Z 5Oct2013

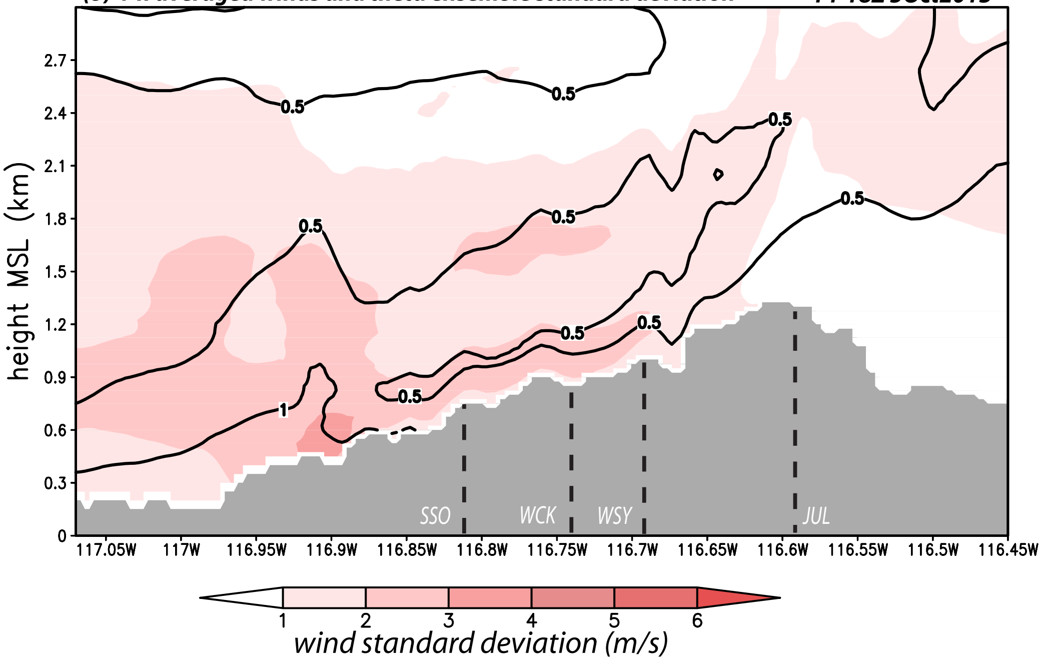


Figure 4.21: As in **Fig. 4.9**, but for 1400-1800 UTC 5 October 2013.

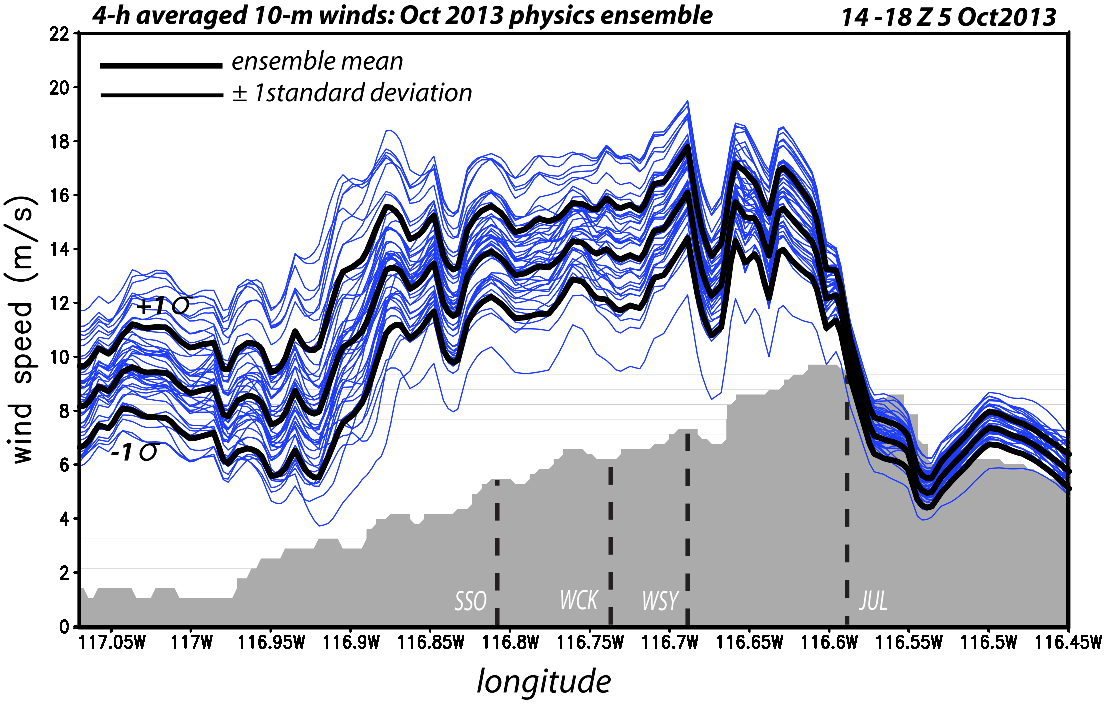


Figure 4.22: As in **Fig. 4.10**, but for 1400-1800 UTC 5 October 2013.

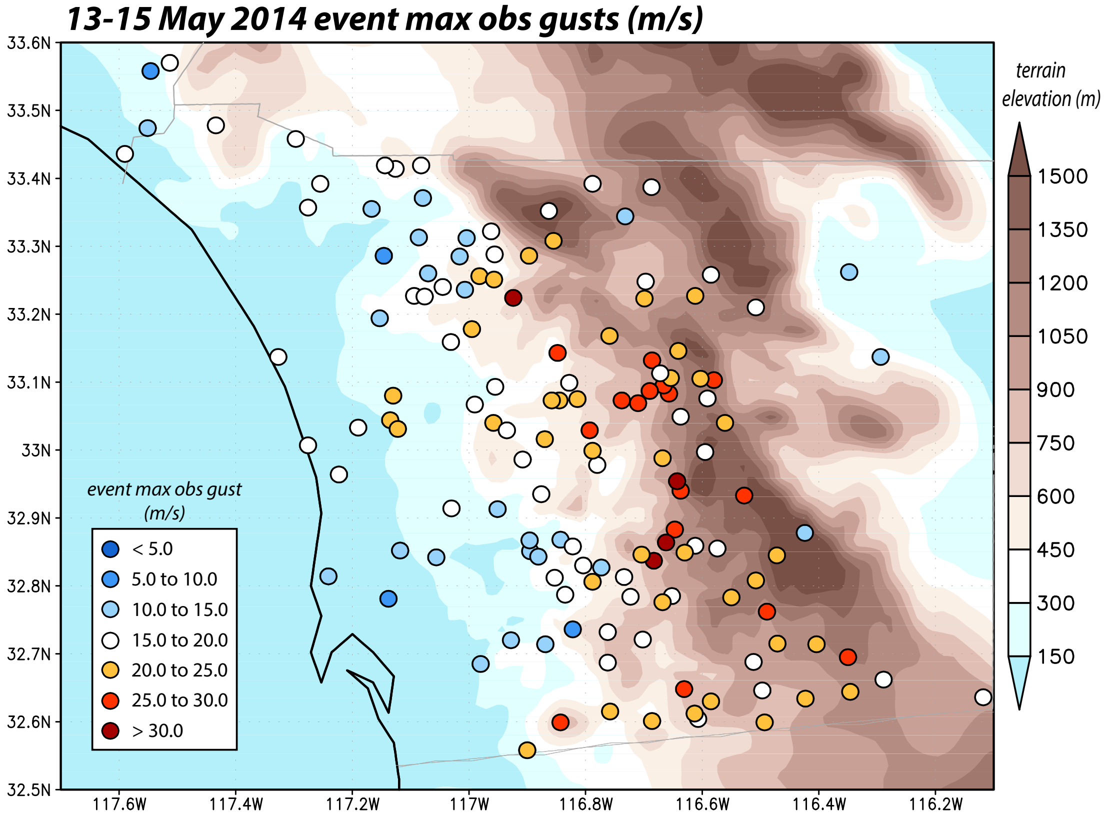


Figure 4.23: As in **Fig. 3.2**, but for the 13 May 2014 event.

4.3.3 The 13-15 May 2014 case

The 13-15 May 2014 event was clearly the strongest of these three events with respect to network-averaged winds (**Fig. 4.1**) as well as in maximum gusts observed through the network (**Fig. 4.23**). The strongest gusts were observed in the Santa Ysabel and Sill Hill areas.

Again, the plot of event-averaged MAE vs. bias revealed a roughly linear relationship, with LSM being the primary controlling factor (**Fig. 4.24**). The PX and RUC LSMs have again performed better with respect to event-averaged MAE and bias, with the PX/ACM2 physics combination still being one of the most skillful ones (**Figs. 4.25 a and d**). The worst simulations were those employing the TD, Noah and Noah-MP LSMs, with the Noah/YSU member having a relatively large MAE and a positive bias

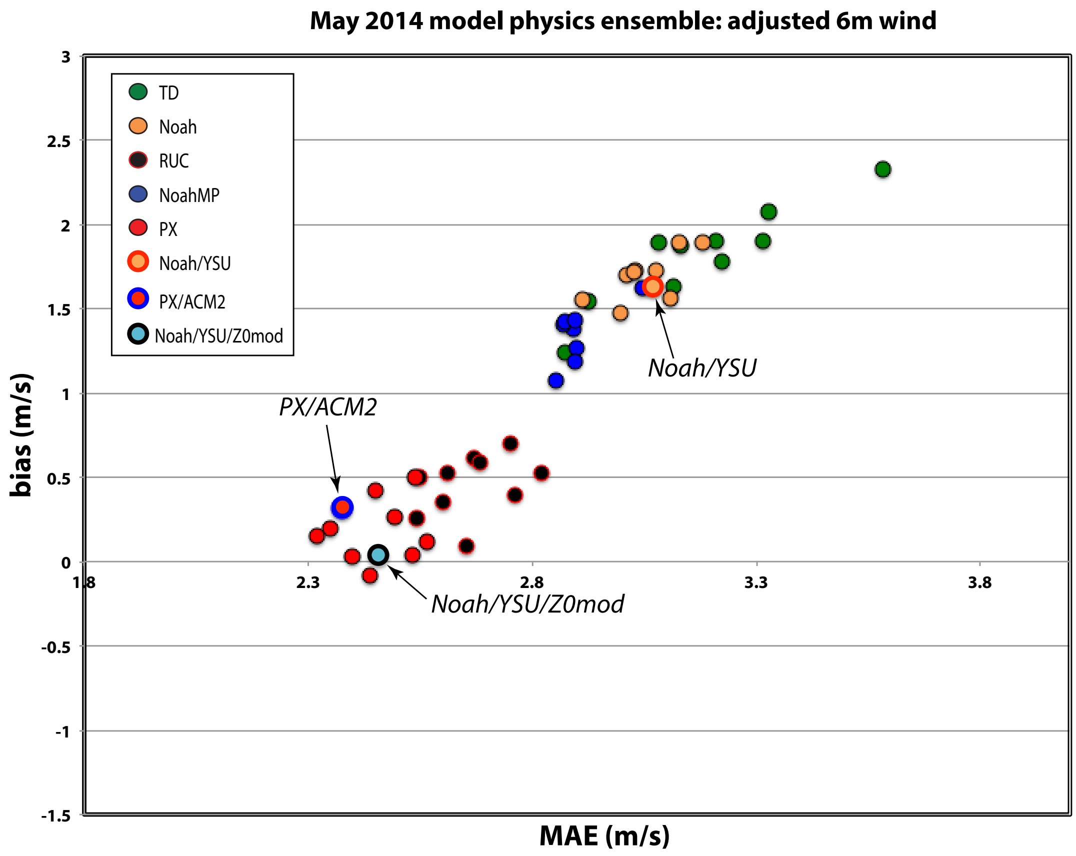


Figure 4.24: As in **Fig. 4.5**, but for the 13 May 2014 event.

compared to PX/ACM2 (**Figs. 4.25 b and e**). The mid-May event also occurred during the summer half-year, so the TD and Noah-MP used the summer's default roughness values of 0.05 and 0.06 m for categories 6 and 7, respectively, while the Noah scheme's z0 values were roughly the same for these same locations (**Table 4.1**). The modified version of the Noah/YSU member (Noah/YSU/$z_0$mod) that adopted the PX scheme's roughness values is similar to PX/ACM2 and improved the forecast skill to a large extent (**Figs. 4.24, 4.25 c, f, 4.27 a, and d**).

The ensemble spread in wind speed is small and similar to the October 2013 episode, and the largest disagreement among the ensemble members was also largely centered mainly on the lee slopes (**Fig. 4.26**). The mid-May event was a two-phase episode (**Fig. 4.1 a**), however, as with the October episode, the downsloping flow during this event

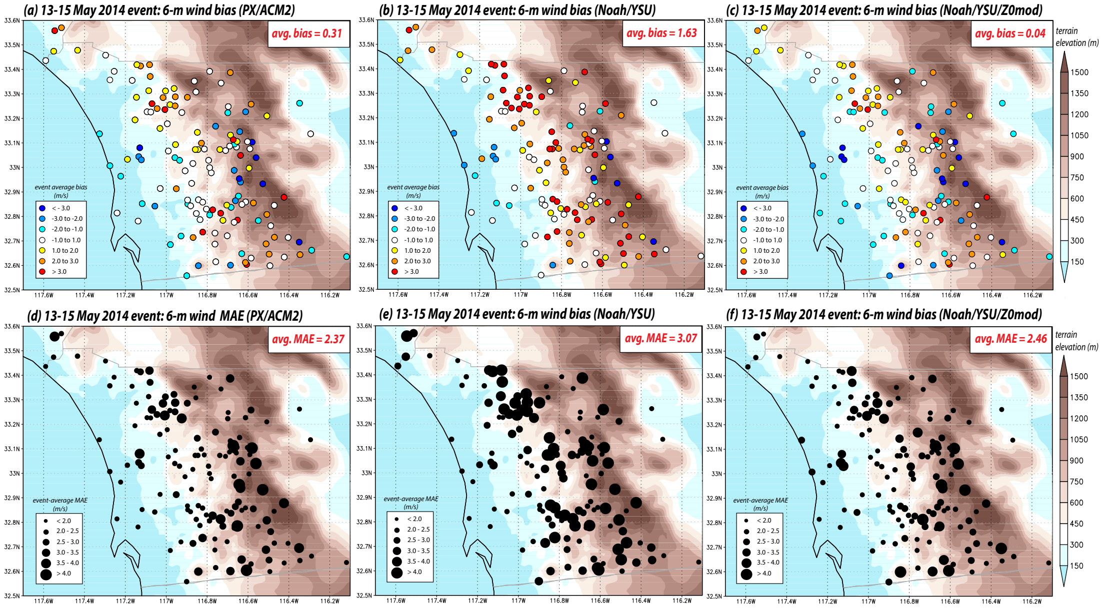


Figure 4.25: (a-c) As in **Fig. 4.18**, but for the 13-15 May2014 event; and (d-f), as in (a-c), but for the MAE.

did not develop a hydraulic jump, and the easterly flow pattern was deeper, stronger and spatially more extensive than the October event (**Fig. 4.27**). These fields are again 4-h averages, centered around the time of the event's first peak.

Variations among these runs were relatively subtle during both the first and second phases, and mainly exist in the downslope extent and vertical depth of the easterly flow. This is confirmed by the ensemble means and standard deviations for this same time window for the two phases of this event (**Fig. 4.28**), and the corresponding variation of the near-surface winds again largely reflects differences in z_0 rather than essential structural variations (**Fig. 4.29**).

4.4 Sensitivity to random perturbations

In the preceding sections, we have demonstrated that substantial sensitivity to model physics exists with respect to near-surface sustained wind forecasts, the spatial extent

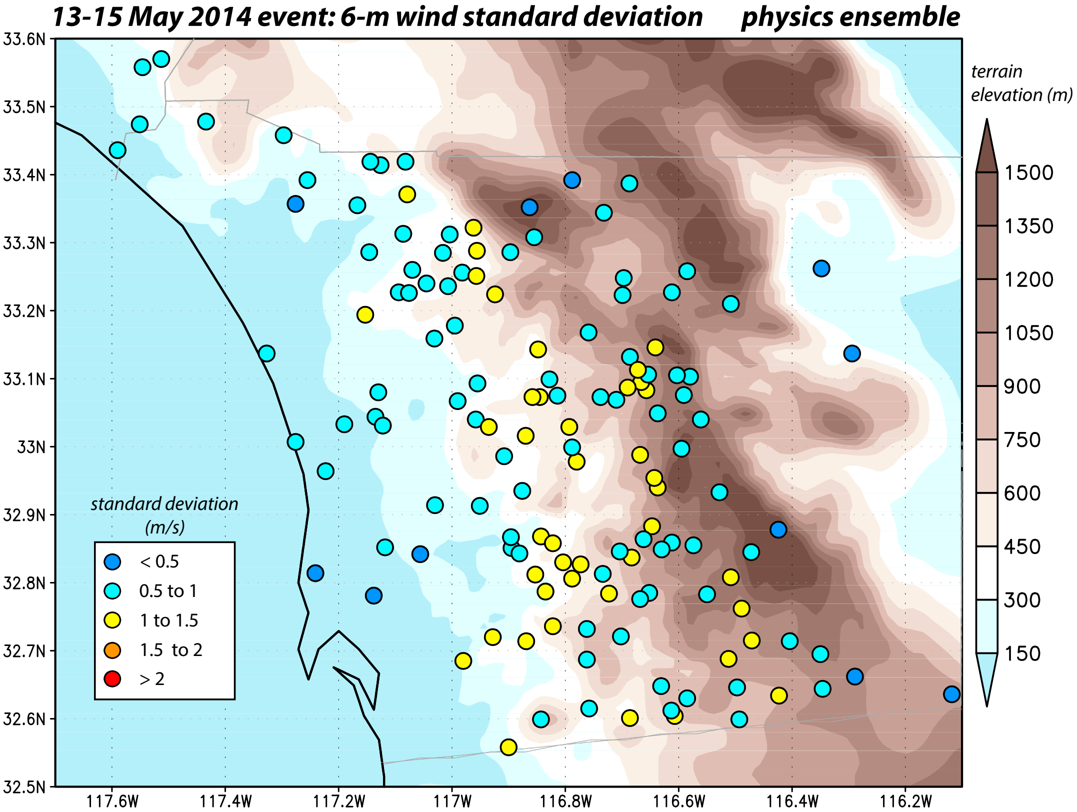


Figure 4.26: As in **Fig. 4.7**, but for the13 May2014 event.

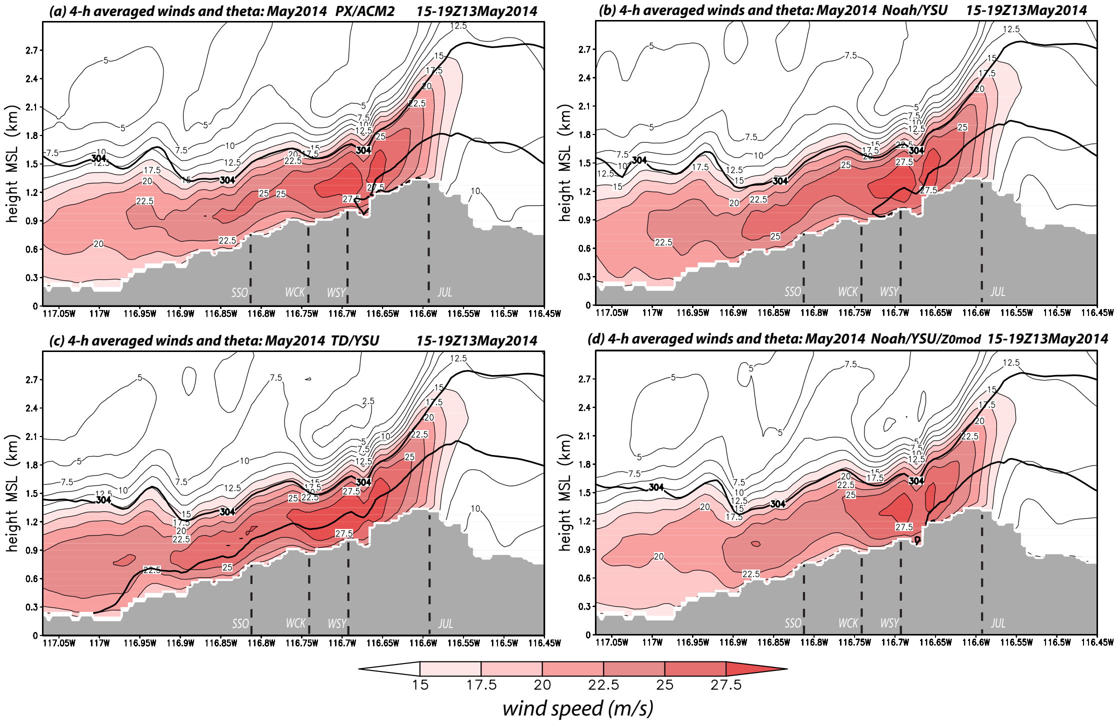


Figure 4.27: As in **Fig. 4.8**, but for the first phase of the 13 May 2014 event.

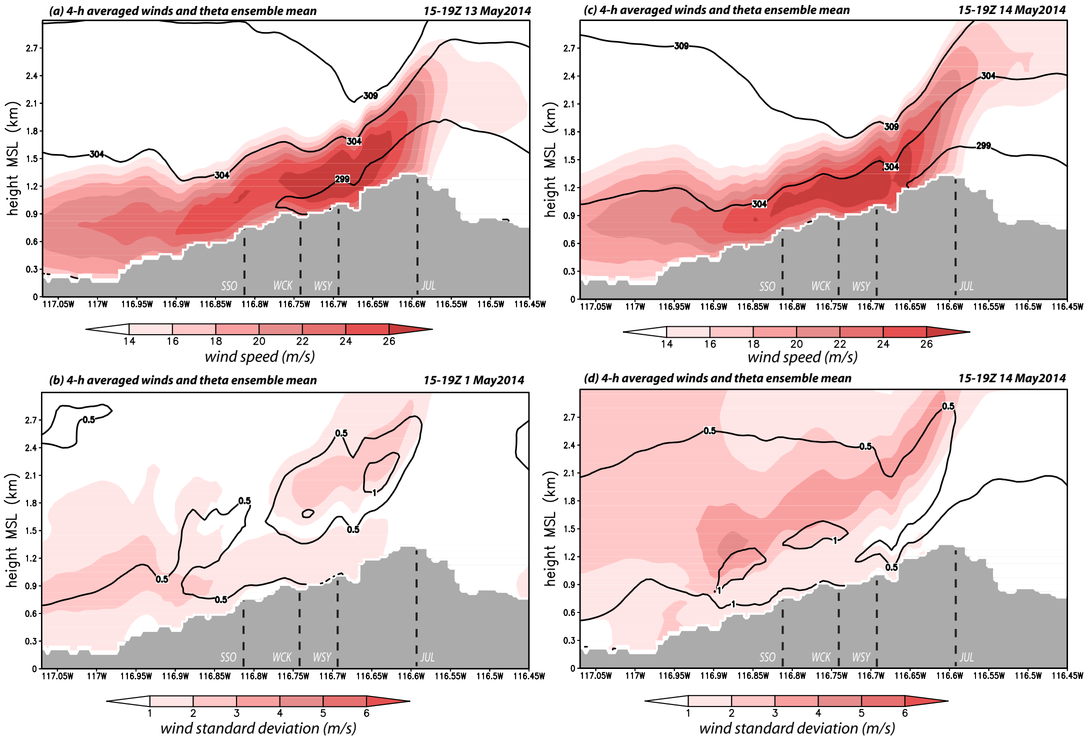


Figure 4.28: As in **Fig. 4.9**, but for the13 May2014 event. (a-b) are for the first phase of the event; and (c-d) are for the second phase.

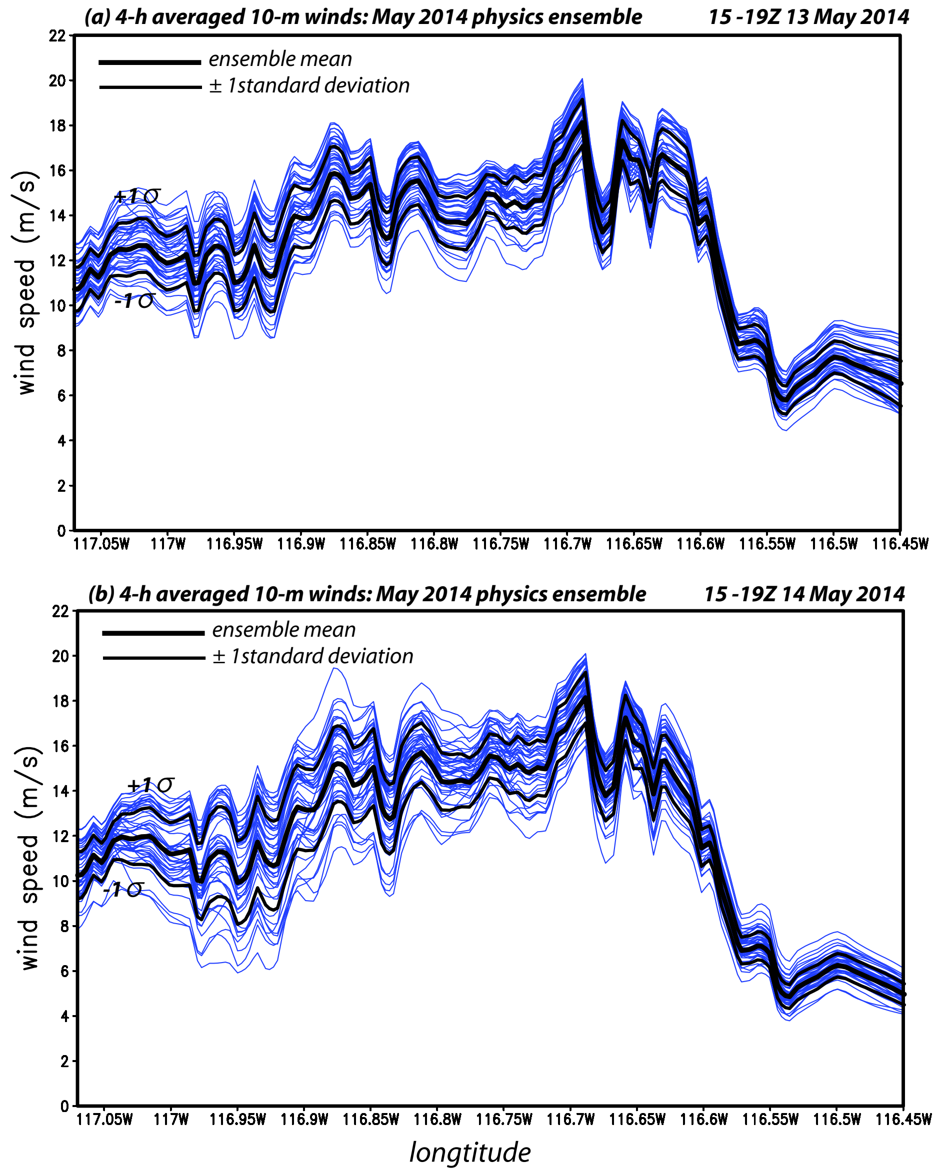


Figure 4.29: As in **Fig. 4.10**, but for (a) 1500-1900 UTC 13 May 2014; and (b) 1500-1900 UTC 14 May 2014.

of the offshore winds, and structural features of the downslope flow, particularly for the strong February 2013 episode. In this section, we explore the sensitivity to random perturbations, via ensembles created with WRF's SKEB scheme [*Shutts*, 2005]. This technique inserts random noise perturbations into the rotational horizontal wind components and the potential temperature field where and when turbulence is diagnosed. SKEB ensembles have been made for each of the three Santa Ana events examined in this chapter, based on the Noah/YSU ensemble member. Each SKEB ensemble consists of one control run and 20 perturbed members, created by only varying the random number seed used as input to the SKEB procedure.

Figure 4.30 presents ensemble plots of near-surface wind speed for the February, October 2013 and May 2014 episodes, for comparison with **Figs. 4.10, 4.22**, and **4.29**. These again represent 4-h averages spanning the peaks of each event, and includes both phases of the February case. It is immediately obvious that the first pulse of the February 2013 event was very sensitive to perturbations. The ensemble spread in the vicinity of WCK, where the hydraulic jump was observed to form, is substantial and, indeed, strongly reminiscent of that from the physics ensemble (**Fig. 4.10 a**). In both cases, the wind speed spread was small until the easterly flow passed the narrow canyon just upslope from WSY, and was widest around WCK (with a secondary maximum near station SSO; see **Chapter 3**). The spatial distribution of ensemble variance also resembles the physics ensemble's (compare **Figs. 4.31** with **4.7**).

Recall that the Noah/YSU member was only partially successful in reproducing the jump near WCK, as it was less well formed and located a little too far downslope (**Fig. 4.8 b**). Close inspection of **Fig. 4.30 a** reveals that a subset of this physics combination's perturbed members did a better job with respect to jump magnitude and placement. In many others, the jump was either sited even farther downslope, or there was no jump to be found anywhere along the slope. To some degree, it was serendipitous that the unperturbed Noah/YSU simulation performed as well as it did, at least with respect to the structure of the offshore flow. It is noted in passing that while each member of

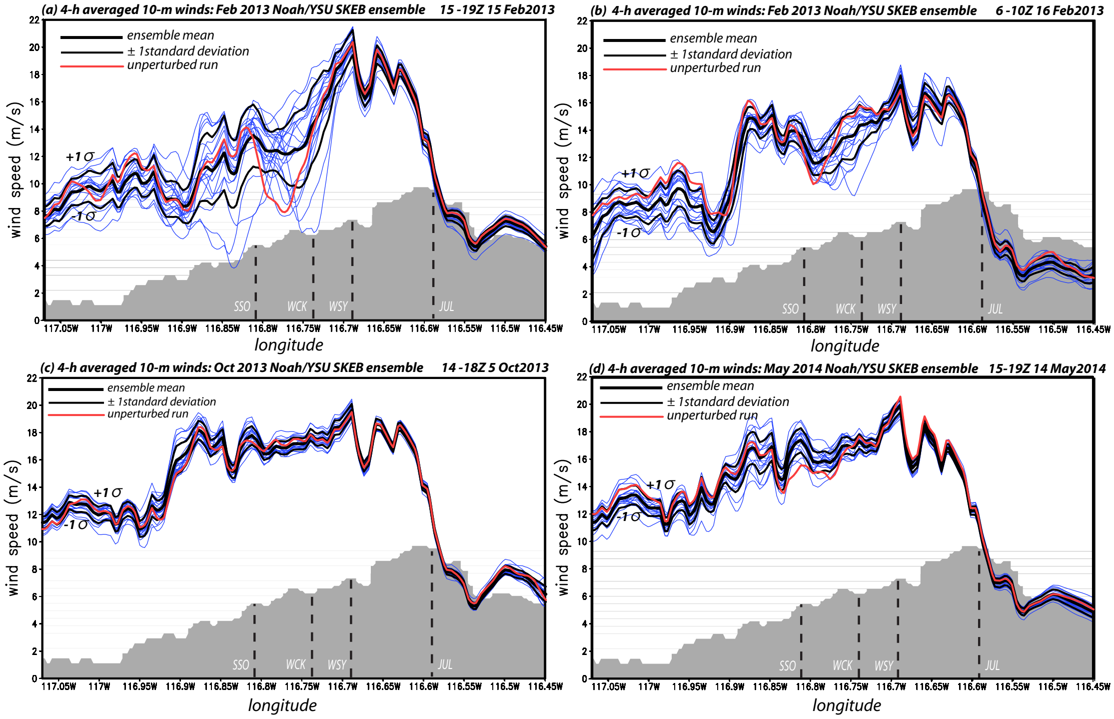


Figure 4.30: Similar to **Fig. 4.10**, but for the SKEB perturbation ensembles of (a) 1500-1900 UTC 15 February 2013; (b) 0600-1000 UTC 16 February 2013; (c) 1400-1800 UTC 5 October 2013; and (d) 1500-1900 UTC 13 May 2014. The red line overlaid represents the unperturbed run.

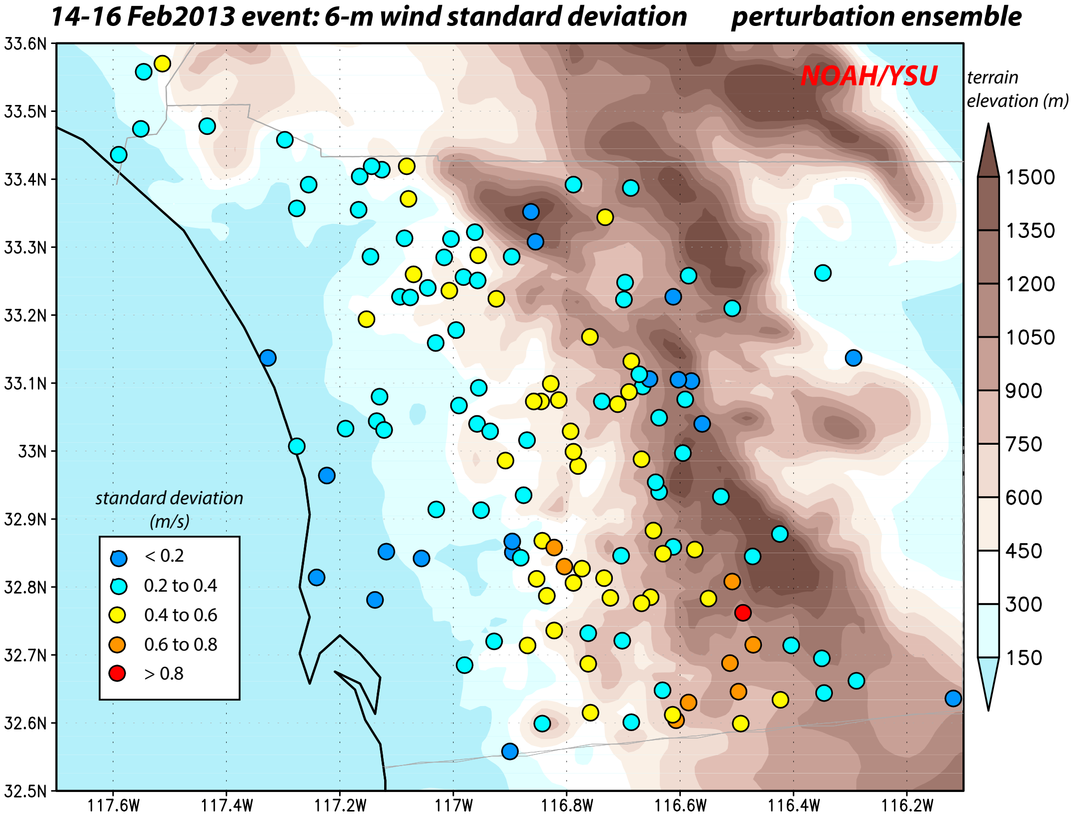


Figure 4.31: As in **Fig. 4.7**, but for the 14 February 2013 event SKEB perturbation ensemble, with random perturbations added to Noah/YSU.

the February Noah/YSU perturbation ensemble had a different structure, their event-integrated forecast skill scores were little different from their unperturbed counterpart (not shown). This is not only because the jump feature occurred only during part of the episode, but also because shifting the jump placement merely moved the error around without changing it very much.

For the other time periods, including the February episode's second phase, the sensitivity to perturbations was much smaller (**Figs. 4.30 b-d**). This is not just a function of the flow strength, as the October and May events actually pushed faster winds farther down the slope than occurred during either phase of the February Santa Ana. The structure of the February episode's first pulse was more amenable to perturbation, as the better-performing members of the physics ensemble tended to produce steeply sloped isentropes during this period of time (see, for example **Fig. 4.8**). The shape and spacing of isentropes reveal the vertical stability, which is smaller when the contours are more vertical and widely separated. Decreased stability increases susceptibility to turbulent motions. Since the SKEB scheme infuses random noise where turbulence is expected, it is logical that the perturbation ensemble spread is larger at and immediately downwind of where the isentropes have a quasi-vertical structure, and its influence is more limited at times when the flow is more strongly stable.

Recall that for February's first phase, the physics combinations with the larger z_0 values were most likely to produce jumps. Yet, the SKEB ensemble based on the PX/ACM2 member revealed less sensitivity to random noise (not shown). It seems somewhat contradictory that enhanced roughness makes the downslope flow more likely to exhibit high-amplitude jump-like features but also less susceptible to turbulence, since the jumps are associated with more vertically-oriented isentropes. It is possible that the enhanced surface roughness assisted in "locking" the downslope flow pattern into place, such that larger magnitude perturbations would have been needed to exert an influence.

4.5 Evaluation of other WRF options

4.5.1 Evaluation of the “topo_wind” subgrid terrain parameterization

Topography obviously plays a strong role in creating small-scale wind variations, affecting not only wind magnitude but also its direction. As a consequence, a wind observation at a single point may not be truly representative of a very large area beyond its immediate surroundings. Forecast skill can suffer when topographic features that cannot be resolved on the grid exert important influences on wind behavior at observation points. One way to mitigate this issue is to conduct higher-resolution simulations, in the hopes of capturing more of the relevant terrain features. This may not be a practical strategy, and can still leave important topographic structures unresolved.

Another approach to the problem is to use information relating to the unresolved topography to modulate the wind, accounting for the enhancement that might be expected at hilltops and speed decreases that could occur elsewhere owing to terrain drag. Two versions of the “topo_wind” strategy, pioneered by *Jiménez and Dudhia* [2012], are implemented in the WRF model. These parameterizations employ knowledge of subgrid topography to modify the grid-resolved wind. As of WRF version 3.5, the original “topo_wind” implementation can only be used with the Noah LSM.

We tested the impact of both available versions of “topo_wind” on the wind reconstructions, for comparison with the PX/ACM2 and Noah/YSU/$z_0$mod simulations from the three Santa Ana wind episodes. Recall that in the three windy events, February and October of 2013, and May 2014, the Noah members had high wind biases (that were especially large relative to the other members in the autumn episode). For simplicity, only the newer version (option 2) of the strategy, dubbed “TW2”, is examined herein. This option is applied to the original Noah/YSU simulation as well as the Noah/YSU/$z_0$mod member, to see how it impacts the simulations before and after explicit revision of the surface roughness lengths.

Time series of the network-averaged sustained wind confirm that TW2 causes a reduction in the Noah/YSU wind bias, which is obvious in all three cases (**Fig. 4.32**). In each event, the wind speed reduction was about 30%, and a linear function of the Noah/YSU member's original wind speeds (averaged R^2 = 0.97; not shown). As a consequence, the "topo_wind" option is acting to reduce positive biases, to the point of making the bias negative in two cases (**Figs. 4.32, 4.33 b, 4.34 b, and 4.35 b**). As might be anticipated from prior discussion, a change of wind speed might influence the structure of the downsloping flow, especially for the February episode's first phase, but in this case the impacts were judged relatively small (not shown). All in all, the "topo_wind" option appears to be accomplishing much the same thing as the roughness length enhancement motivated by the success of the PX LSM members. The figures show that doing both, however, reduces the wind speeds too far (**Figs. 4.32, 4.33 c, d, 4.34 c, d, 4.35 c, and d**).

4.5.2 Evaluation of the horizontal diffusion option

The WRF model can typically be integrated without explicit horizontal diffusion, at least when the standard, odd-order advection scheme is employed (which has some implicit smoothing function). When used for "real data" applications, at the horizontal grid resolutions being examined herein, the recommended configuration for WRF is to apply horizontal diffusion on model surfaces, if it is used at all. Yet, this may represent a problem when topography is included. Horizontal diffusion should operate in physical space. When diffusion is computed along model surfaces, however, the smoothing is being applied up and down the mountain slope. This can cause unrealistic vertical transports of momentum, heat and moisture in areas with complex terrain.

Based on experiences in modeling other areas with significant topography, our standard WRF real-data configuration does not employ horizontal diffusion at all. However, we tested to see whether diffusion has an impact on our forecast wind skill in the SDG&E mesonet. **Figure 4.36** presents event-mean bias vs. MAE plots for the

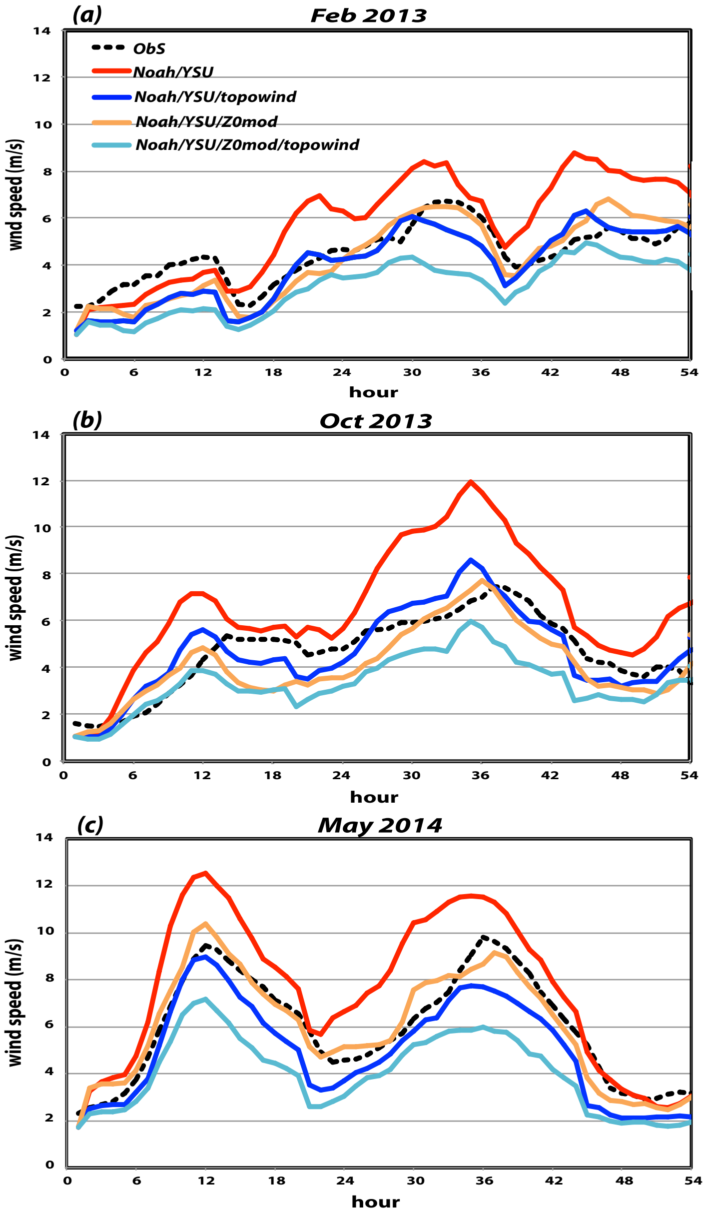


Figure 4.32: Time series of SDG&E network-averaged 6 m sustained wind (m s^{-1}) observations (black line), for comparison with network-averaged 6 m sustained wind predictions from the Noah/YSU (red line), Noah/YSU/$z_0$mod (orange line), along with the same two runs but made with a “topo_wind” (topo_wind option=2) option (blue and cyan lines) for (a) the Feb 2013 event; (b) the Oct 2013 event; and (c) the May 2014 event. All of the simulations used the MODIS landuse database.

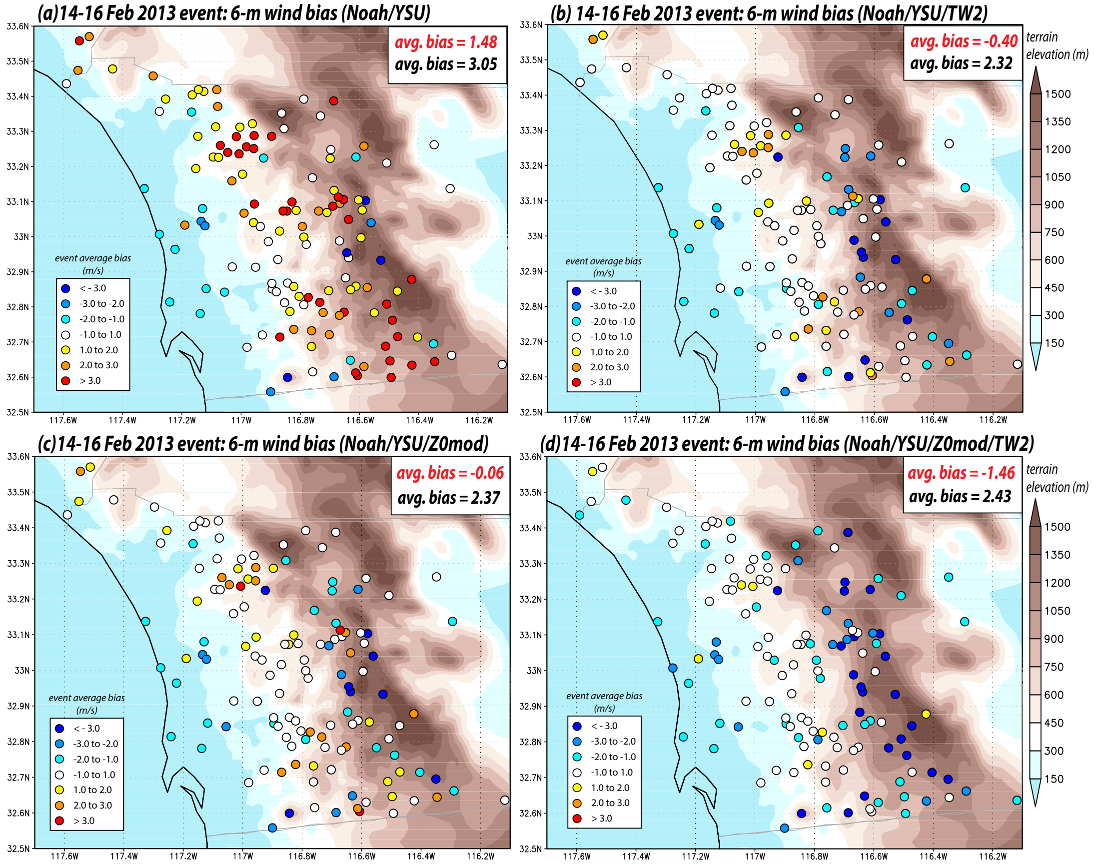


Figure 4.33: (a) Same as **Fig. 3.9**; (b) as in (a), but for the Noah/YSU/TW2; (c) same as **Fig. 4.13 b**; and (d) as in (c), but for Noah/YSU/$z_0$mod/TW2.

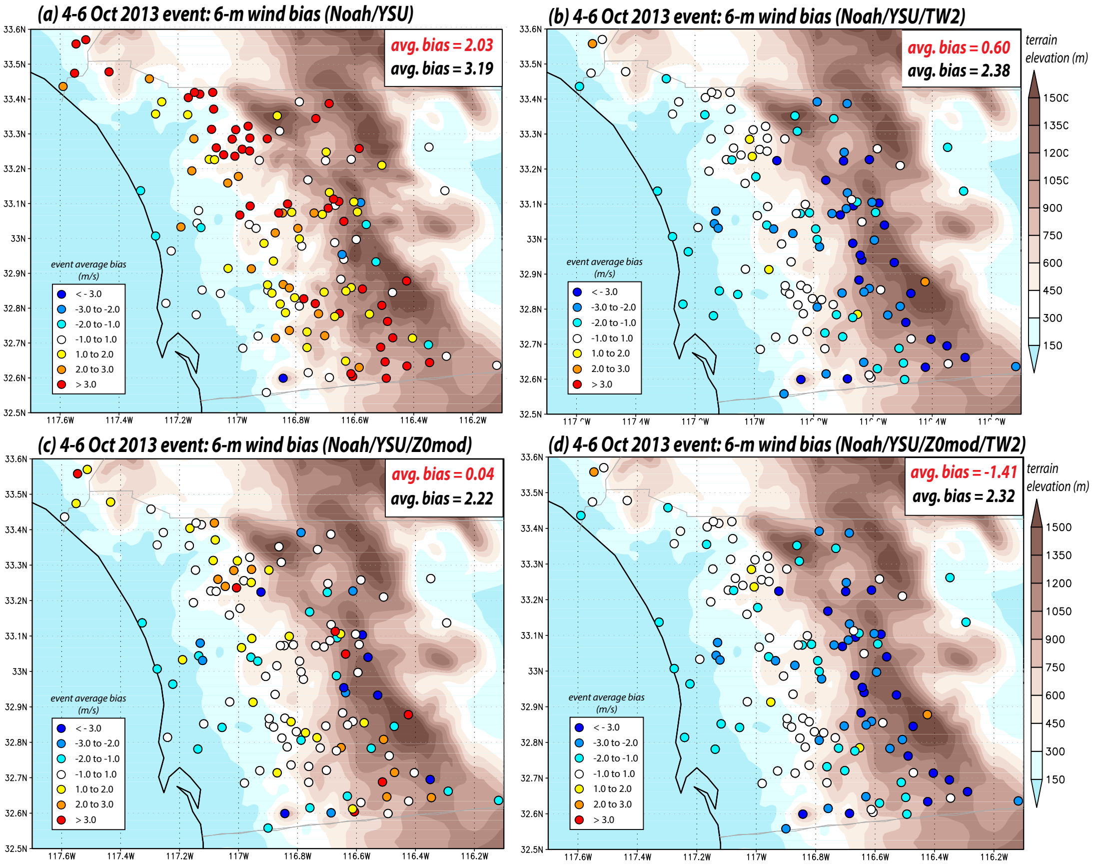


Figure 4.34: As in **Fig. 4.33**, but for the 4 October 2013 event.

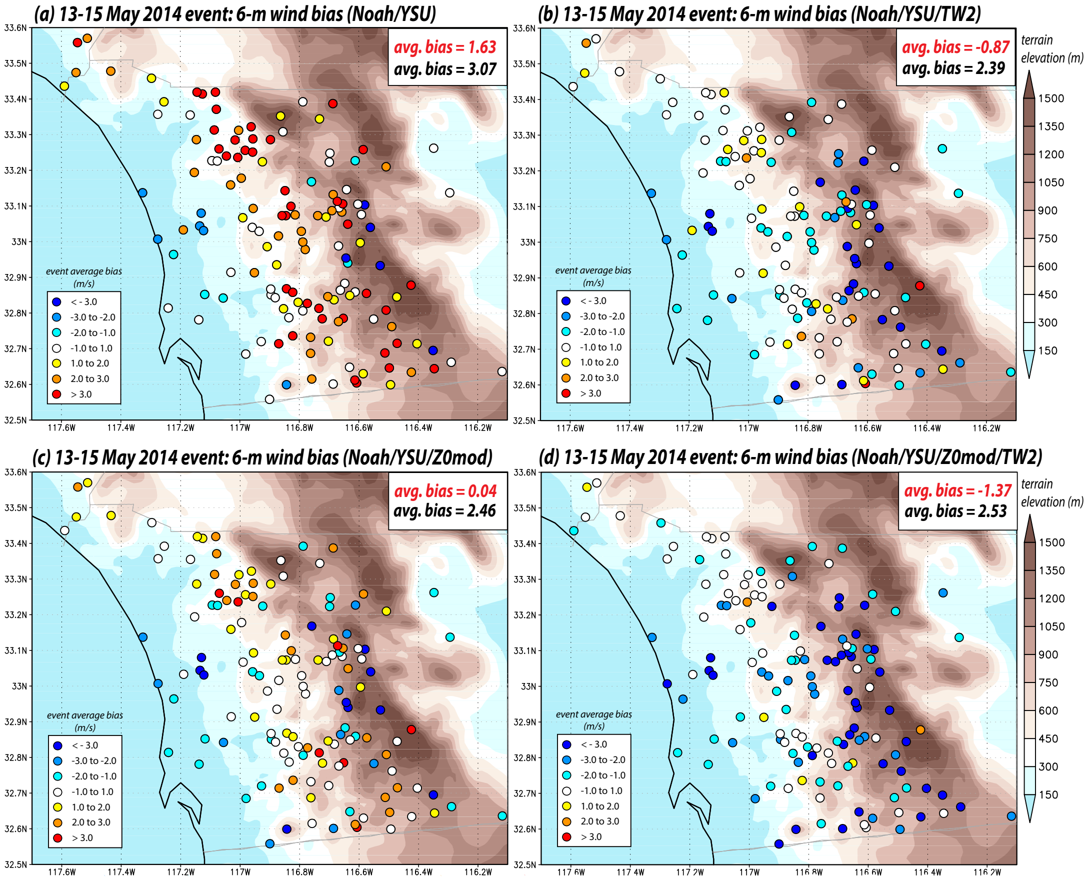


Figure 4.35: As in **Fig. 4.33**, but for the 13 May 2014 event.

Figure 4.36: Similar to **Fig. 3.17**, but the points are color-coded by the diffusion options (on vs. off) for (a) the February 2013; (b) the October 2013; and (c) the May 2014 events.

three events, including representative members from the physics ensembles. Simulations without and with horizontal diffusion are compared, so there are two points for each physics combination. We can observe that the influence of horizontal diffusion is small.

4.5.3 Sensitivity to the land use database

We have conclusively shown that the surface roughness plays a critical role in determining the forecast skill and represents the single largest difference among the physics ensemble members. Even in events where downslope wind structure did not vary much among the many physics combinations, it remained that the bias and MAE for the wind reconstructions were mainly determined by z_0. Although the roughnesses may be modified by the LSM, the values applied at a particular point in space depend on the horizontal resolution of the model grid and the source of the landuse classifications. Up to this point, we have been using a database that originated with MODIS observations. In this subsection, we compare a subset of our results with those made using the older, more commonly used USGS database.

The USGS and MODIS databases categorize the landscape of the SDG&E network somewhat differently (**Fig. 4.37**). The most remarkable difference is in the spatial extent of the forested region in the mountains, which is smaller on the MODIS map, while the urbanized area near the coastline is much more extensive. These have implications for surface roughness, since this varies with landuse assignment. MODIS treats the forested area as rough as USGS (**Fig. 4.38**), but since it occupies a smaller area, its default z_0 values are actually smaller than USGS over a relatively wide region (**Figs. 4.38 c and f**). Meanwhile, the MODIS z_0 values are larger on the west-facing slopes occupied by shrublands in both seasons, especially in the summer (which affects October and April/May Santa Anas). In effect, the MODIS map changes the gradient of the surface roughness, which can be expected to influence the magnitude and structure of the downslope winds. Perhaps importantly, the high-resolution MODIS categorization also recognizes more features in the vicinity of WCK, including some locally rough patches.

We investigated the sensitivity of wind speed MAEs and biases to the landuse database for the February (“winter”), October (“summer”) and May (“summer”) events. We first focus on the TD LSM because it does not alter the z_0 values provided by the databases, thereby maximizing the potential distinctiveness between the two sources. With the MODIS categorization, the TD scheme generally has relatively large positive wind speed bias, particularly during the winter half-year when shrubland roughness is presumed to be very small (0.01 m for MODIS categories 6 and 7). For all three events, however, the positive bias is smaller in the USGS runs (**Fig. 4.41**).

For the Noah scheme, the picture is more mixed, as the MODIS version is more skillful in the February and May events relative to USGS, but less so in October (**Fig. 4.42**). This again reflects Noah’s seasonal cycle of roughness length. **Figure 4.39** suggests that the MODIS-USGS roughness length difference in the mountainous regions are greatest in October. The PX LSM substantially modifies the roughness lengths irrespective of the landuse database (**Fig. 4.40**), so the sensitivity with respect to the

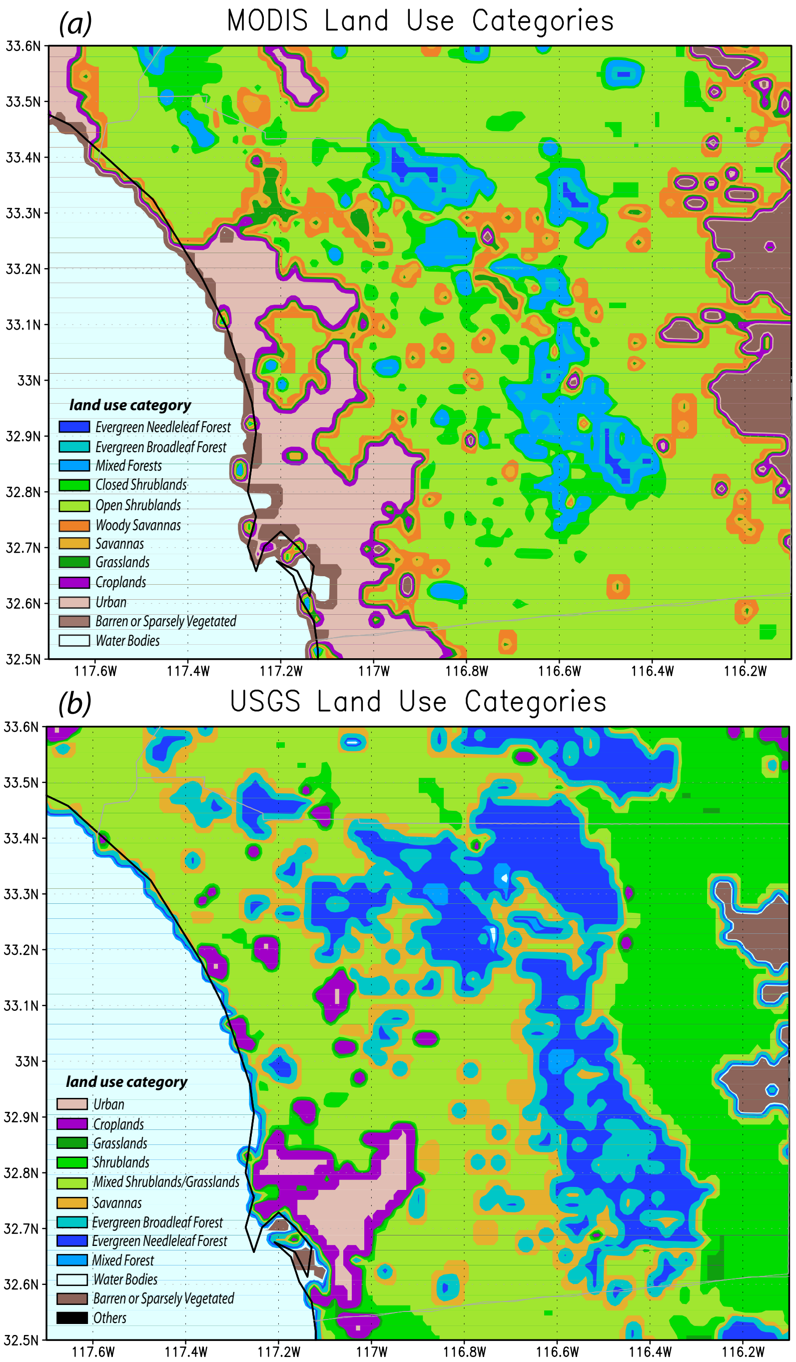


Figure 4.37: (a) Same as **Fig. 4.11**; and (b), as in (a), but for the USGS database.

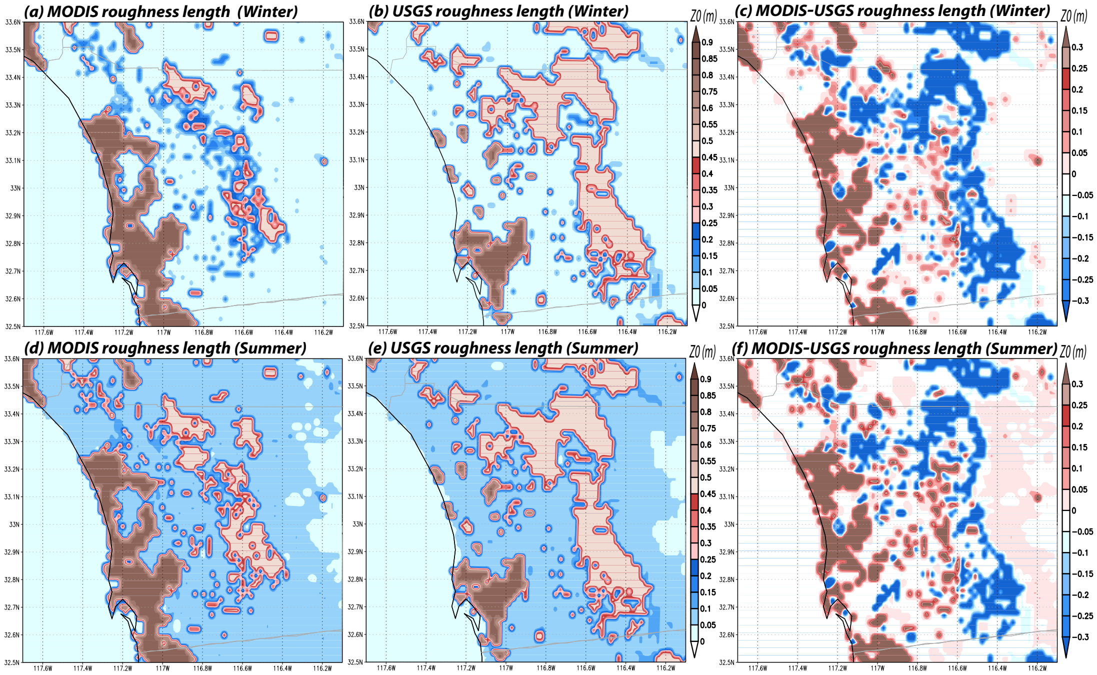


Figure 4.38: As in **Fig. 4.12**, but for wintertime (a) MODIS; (b) USGS; and (c) the difference between MODIS and USGS; and summertime (d) MODIS; (e) USGS; and (f) the difference between MODIS and USGS.

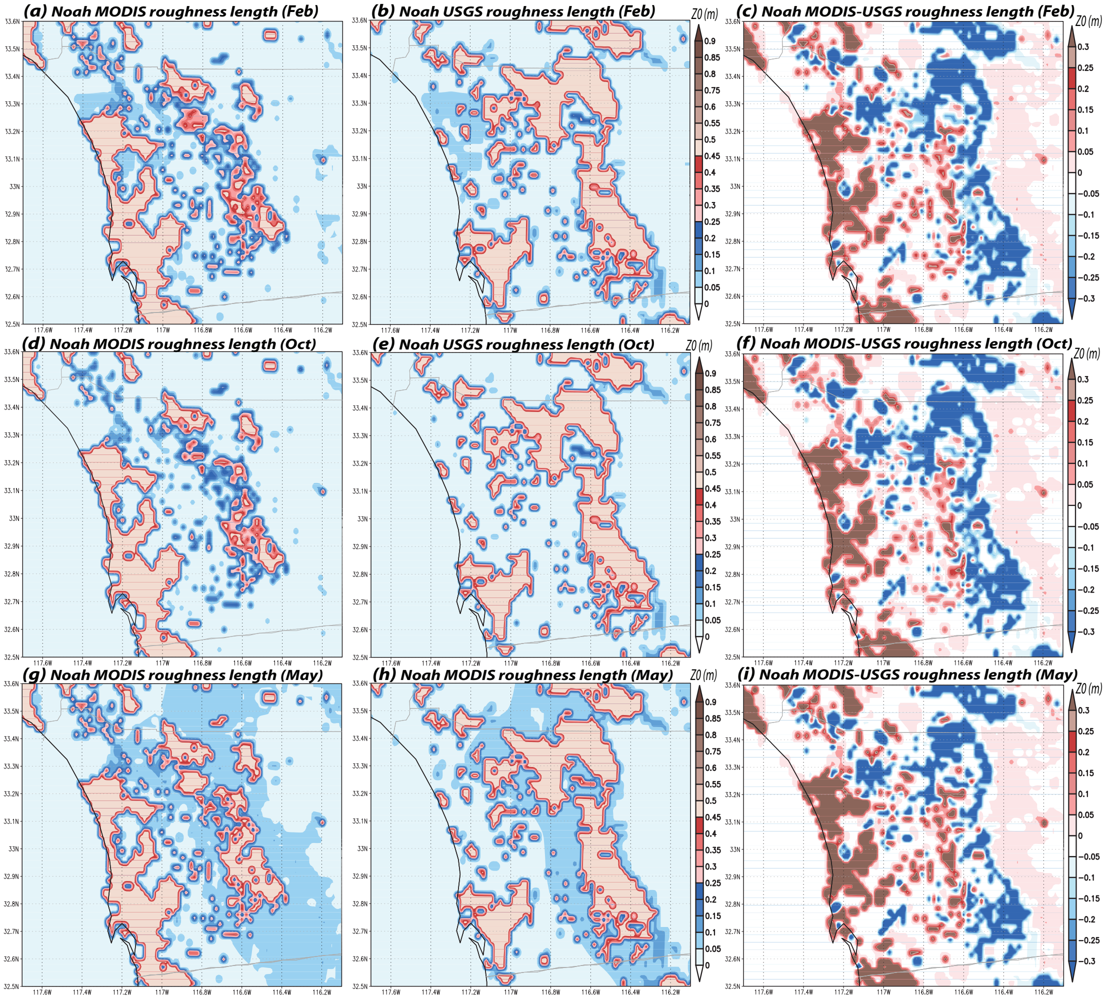


Figure 4.39: As in **Fig. 4.38**, but for the Noah LSM for the three events.

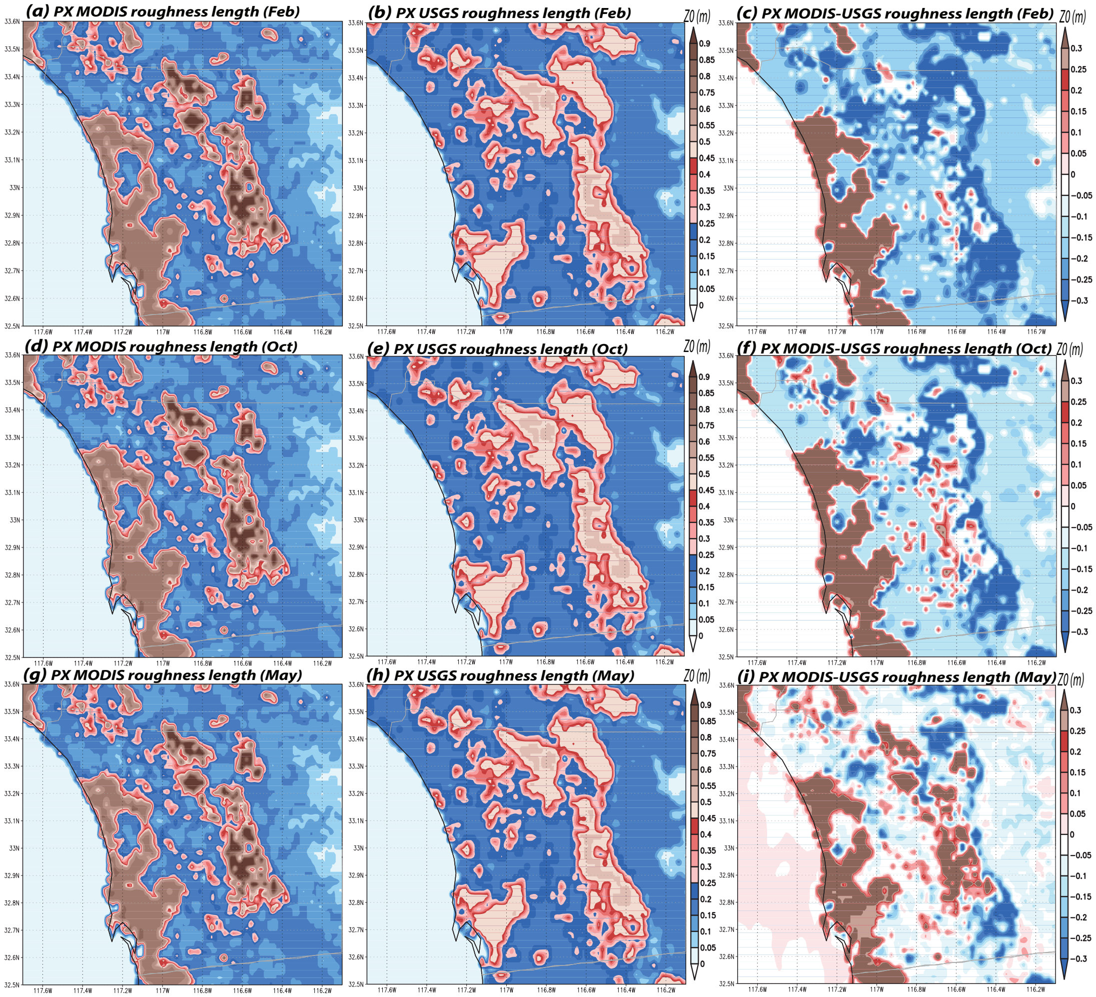


Figure 4.40: As in **Fig. 4.38**, but for the PX LSM for the three events.

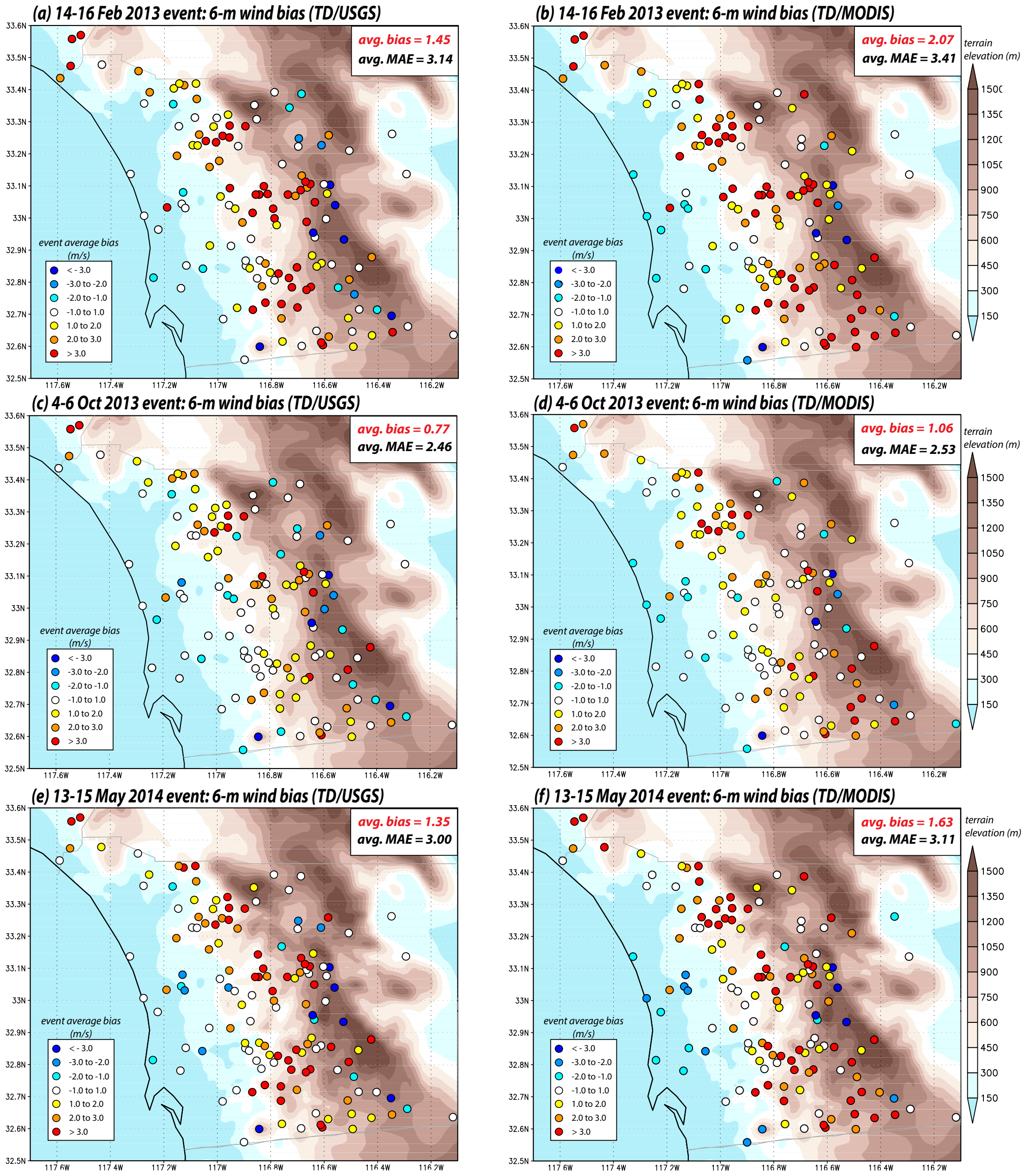


Figure 4.41: As in **Fig. 3.9**, but for (a) TD/USGS; and (b) TD/MODIS for the February 2013 event; (c) TD/USGS; and (d) TD/MODIS for the October 2013 event; (e) TD/USGS; and (f) TD/MODIS for the May 2014 event.

specific database employed is relatively small (**Fig. 4.43**). This is also true for the Noah/YSU/$z_0$mod member (not shown).

4.6 Summary

From close inspection of hundreds of WRF model simulations, we have demonstrated that the wind speeds and flow patterns during moderately strong and strong Santa Ana episodes are sensitive to horizontal grid spacing, random perturbations and, especially, to the land surface schemes that determine the surface roughnesses employed in the simulations. Model horizontal grid spacings wider than 2 km were determined insufficient to properly capture the terrain shape, thereby locating the regions of faster and slower winds improperly. The simulations were sensitive to model physics, but mainly due to variations in the surface roughness that determined how fast the near-surface winds were as well as how far downslope the offshore flow could remain close to the ground. PBL treatments exerted only secondary influences on the model forecast skill, as determined by bias and MAE with respect to observations made in the SDG&E mesonet.

The PX LSM scheme emerged as the best overall. Our analysis demonstrated its success is largely due to its unique treatment of surface roughness in the shrubland landuse categories that dominate the west-facing slopes in the SDG&E network. In particular, PX increases z_0 for those categories, and holds them fixed with time. Most WRF simulations resulted in a positive wind bias because the surface is treated as being too smooth. We have found that altering other LSMs to increase their roughnesses improved their MAE and bias scores. This is important, because other land surface treatments tend to provide superior temperature and humidity reconstructions (which were beyond the scope of this work).

It is intuitive that increasing the surface roughness should slow down the winds. However, it may also change the nature of the downsloping flow, at least in the February

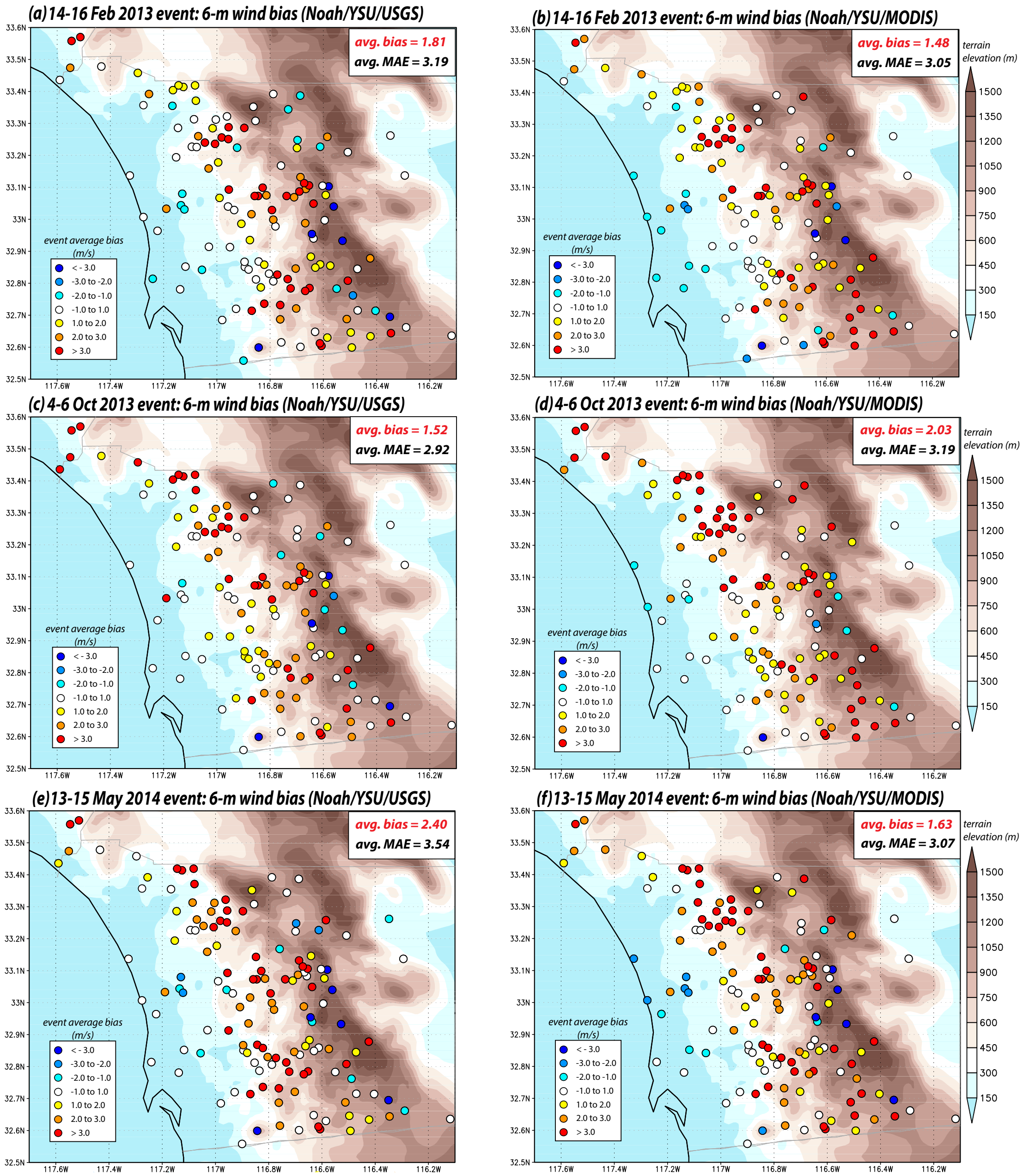
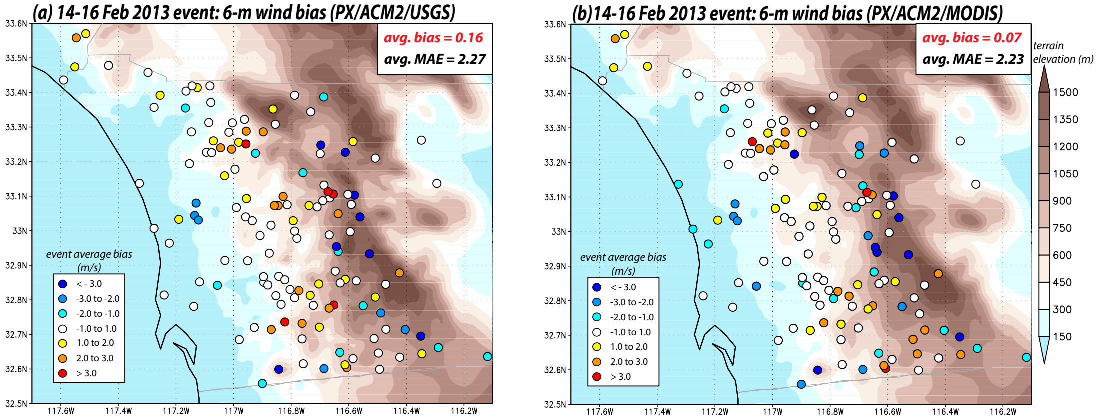
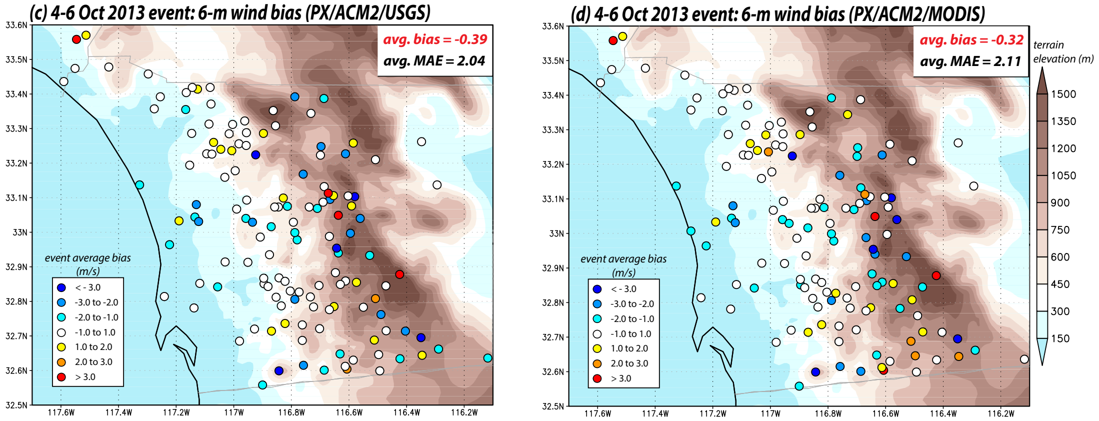


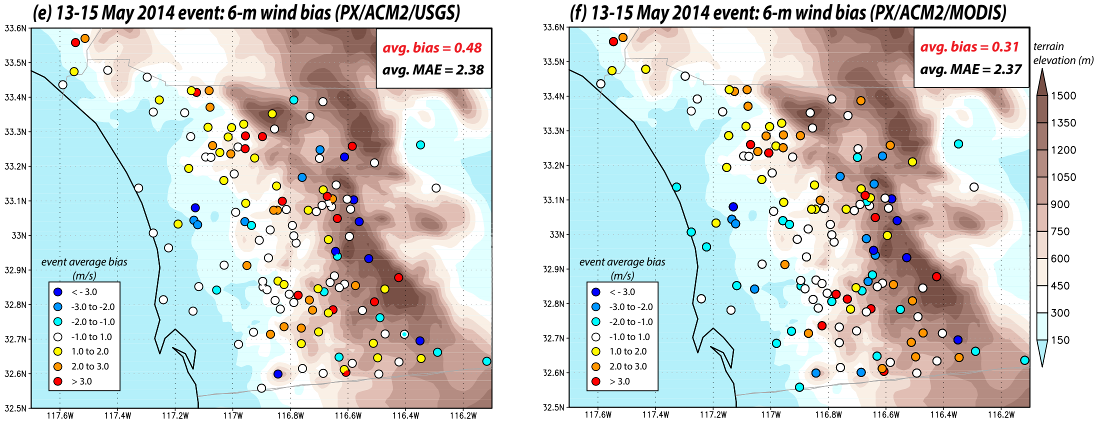
Figure 4.42: As in **Fig. 4.41**, but for the Noah LSM.



(a) 14-16 Feb 2013 event: 6-m wind bias (PX/ACM2/USGS) (b)14-16 Feb 2013 event: 6-m wind bias (PX/ACM2/MODIS)



(c) 4-6 Oct 2013 event: 6-m wind bias (PX/ACM2/USGS) (d) 4-6 Oct 2013 event: 6-m wind bias (PX/ACM2/MODIS)



(e) 13-15 May 2014 event: 6-m wind bias (PX/ACM2/USGS) (f) 13-15 May 2014 event: 6-m wind bias (PX/ACM2/MODIS)

Figure 4.43: As in **Fig. 4.41**, but for the PX LSM.

2013 case. One of the remarkable characteristics of this event, especially its first phase, was the development of the jump-like feature and wind reversal above station WCK. It has emerged that only the LSMs that employed relatively larger z_0 values (e.g., PX) were able to capture this feature, which the observations indicate was prominent and persistent. The smoother the terrain's lee side, the faster and more uniform the flow that developed there was, preventing the WCK jump from forming and resulting in faster than observed winds farther down the slope.

Sensitivity to the infusion of random noise was assessed using the WRF model's SKEB technique. Except for the February 2013 episode's first phase, the influence of this kind of perturbation was not particularly large. We also investigated other potentially relevant model configuration variations. Model sensitivity to horizontal diffusion was found to be low. The "topo_wind" approach, which attempts to account for unresolved terrain features, is not superior to a more careful and accurate assignment of surface roughness values. The landuse database may matter, as it is the source of landuse assignments and roughness values.

CHAPTER 5

Gust forecasting for the SDG&E mesonet

In this chapter, we examine the ability of the WRF model to provide skillful gust predictions for the SDG&E network. Numerical models of the present type do not resolve short period (~3 sec) gusts, because they cannot directly capture the turbulent motions that these wind bursts represent. This is true even if the model employs a time step on the order of a few seconds, as model filters act to suppress variations with time scales of less than several minutes. Because of this, it is most sensible to compare model outputs to observed *sustained* winds that are subsequently converted to gusts in some fashion. As a consequence, we need to find a reasonable gust parameterization.

5.1 Gust speed prediction in operations

There are at least two, ostensibly distinct strategies being employed by operational centers for the purposes of gust forecasting, both representing an augmentation to the sustained wind predicted by the model. The first utilizes information from the overlying boundary layer to determine the gust. In the case of NCEP models[1], the augmentation to the anemometer-level wind (V_s) is a weighted function of the fastest wind within the PBL. The formula for the anemometer-level wind gust speed (V_G) is

$$V_G = V_s + max\left[f(z) \times (V(z) - V_s)\right] \ , \tag{5.1}$$

where $V(z)$ is the wind speed at height z within the PBL, and $f(z)$ is a function of height that decreases from 1.0 to 0.5 at 1 km AGL, remaining at 0.5 to the top of the

---

[1] http://ruc.noaa.gov/rr/RAP_var_diagnosis.html#gust

PBL. *Brasseur* [2001] has presented a more sophisticated version of this strategy.

The second approach only directly utilizes surface layer information. As of Cycle 40, the European Center for Medium-range Weather Forecasts (ECMWF) estimates gusts by augmenting the anemometer-level wind by a function of the friction velocity, u_* :

$$V_G = V_s + 7.71 u_* \,, \tag{5.2}$$

[This "EC gust" equation utilizes the studies of *Panofsky et al.* [1977] and *Beljaars* [1987] and would be modified in the presence of deep convection.] The friction velocity is a shear stress rewritten in terms of velocity [cf. *Lumley and Panofsky*, 1964], and is computed in the WRF model using the wind speed (V_a) at the height of the lowest model level (Z_a), the surface roughness z_0, a vertical stability-dependent function ψ, and von Karman's constant κ as:

$$u_* = \kappa \frac{V_a}{\ln \frac{Z_a}{z_0} - \psi} \,, \tag{5.3}$$

Since models also use the friction velocity to estimate the anemometer-level wind, V_s and u_* are not independent, which leaves the gust factor a function primarily of surface roughness and stability.

Details on the derivation of Eq. (**5.2**) are not readily available, but we note that a similar equation can be derived from *Burton et al.* [2011]'s equation (*2.42*) after making a few common assumptions. Rewritten in terms of the gust speed, the equation for a t-second gust is

$$V_G = V_s + 0.42 \sigma_u \ln \frac{3600}{t} \,, \tag{5.4}$$

where σ_u is the standard deviation of the wind. For the 3-s sampling interval employed in the SDG&E network, this becomes

$$V_G = V_s + 2.98 \sigma_u \,, \tag{5.5}$$

Close to the ground, the approximation $\sigma_u \sim 2.5 u_*$ is often employed [cf. *Burton et al.*, 2011], implying

$$V_G \approx V_s + 7.45 u_* \,, \tag{5.6}$$

which is reasonably similar to the equation used by the ECMWF. For neutral conditions, *Panofsky et al.* [1977] found $\sigma_u = 2.29\ u_*$ is more appropriate, which would lead to the approximation

$$V_G \approx V_s + 6.82u_* \ , \tag{5.7}$$

instead. Adjustments for different averaging intervals might also be applied [cf. *Wieringa*, 1973]. In a later subsection, we will employ the EC version (Eq. **5.2**), but should keep in mind that the multiplier applied to the friction velocity might require adjustment.

In the following analysis, the NCEP and EC gust estimation strategies are compared, along with a very simple approach that is motivated by the characteristics of the SDG&E observing network and the biases seen in even the most skillful simulations of the sustained wind in our WRF experiments. This simple technique is motivated in the next two subsections.

5.2 Gust factors in the SDG&E network

Instead of focusing on the gust speed, an alternative is to focus on the *gust factor* (GF), the ratio of the gust and the sustained wind. The GF should be a function of sampling interval, anemometer hardware and mounting height, anemometer exposure, and perhaps surface roughness as well [e.g., *Ashcroft*, 1994]. It may also be a function of the sustained wind itself, and vary among stations, and perhaps from event to event. As an example, *Wieringa* [1980] demonstrated that heterogeneity of the surface with respect to roughness may be why the GF varied with wind direction at a single station.

Figure 5.1 shows how the GF varies with the sustained wind at the SDG&E station Sill Hill (SILSD) over a period of about 1.5 years, representing over 42000 observations recorded at 10 min intervals. While we have seen this is a remarkable station, there is nothing particularly unusual about this figure; the decrease of variability with increasing sustained wind is very typical. However, the wind speed dependence of the variability is something of an illusion, for two reasons. First, when wind speeds are

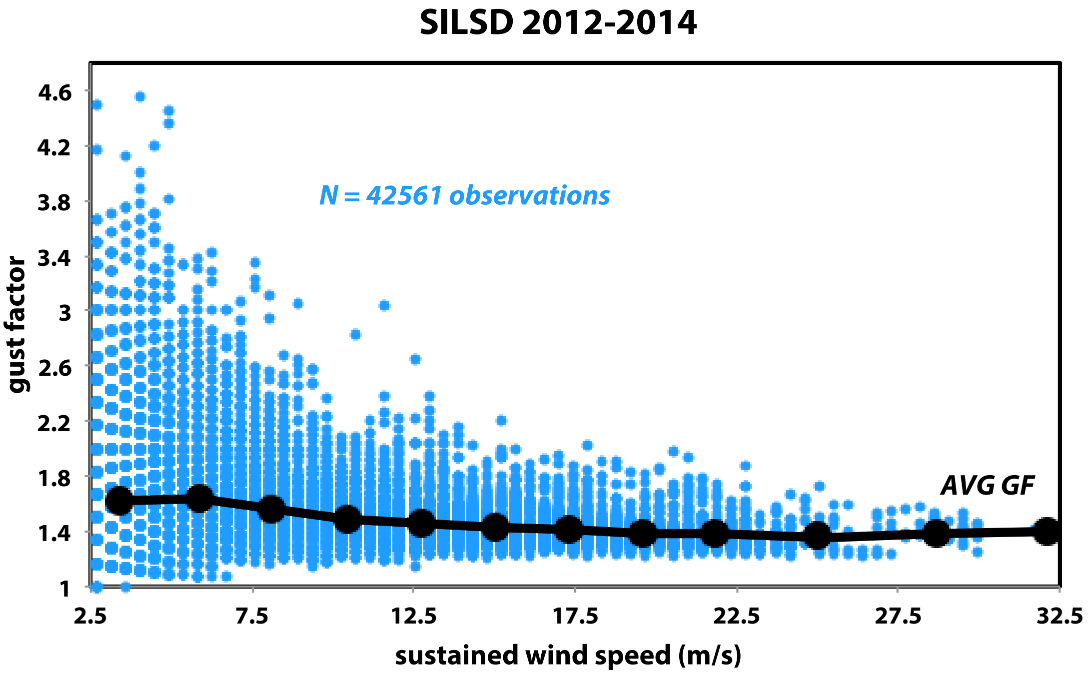


Figure 5.1: Gust factor (non-dimensional) calculated from 42561 observation times for the SDG&E station at Sill Hill (SIL) for observations collected during part sod 2012-2014, plotted against the sustained 10-min interval wind speed (m s^{-1}). Black dots denote the mean gust factor at a fixed sustained wind speed.

low, the gust factor variation is essentially unimportant because the gusts are also very likely to be weak. Second, the mean GF (large black dots) varies far less over the sustained wind speed range, decreasing from 1.6 when winds are slow to about 1.41 at the high end of the range.

The mean GF at SILSD is 1.6 over samples in which the sustained wind exceeded 2.7 m s^{-1}. The least-squares fit to the entire dataset yields GF = 1.51, with an R^2 of 0.95 (**Fig. 5.2**). (Owing to "flattening" of the frequency dimension, scatterplots can exaggerate the variability about the regression line.) Somewhat similar, very low dispersion plots may be obtained from the other stations in the network, the principal difference being in the gust factor slope (not shown). As an example, **Fig. 5.3** presents the event

GF determined via regression for each SDG&E station during the 13-15 May 2014 event. SIL's GF of 1.47 is 14^{th} smallest among the 142 stations shown, and consistent with the historical data given its average wind speed of 18 m s^{-1} (**Fig. 5.1**). The largest GF (3.15) is at station Mt. Laguna (MLGSD), the highest elevation site in the network.

Averaged over the entire SDG&E network, however, we have found the GF to be nearly constant, with a value very close to 1.7 and virtually no dispersion (**Fig. 5.4 a**). At an interval of 10 min, there were 330 observation times between 0510 UTC on 13 May to 1200 UTC on the 15^{th}, inclusive. For each time, the sustained wind and gust observations were averaged over the 142-station network, and the results are shown in the figure. The R^2 of the fit is 0.997. Although GFs do vary with station, sustained wind speed, and time, *the network-averaged GF can be represented by a single number*, independent of the magnitude of the network-averaged wind.

This surprising result is not confined to Santa Ana wind events. **Figure 5.4 b** presents a composite of 10-min observations over a period of twelve consecutive months consisting of 51,940 instances. Again, each point represents wind and gust averaged over roughly 140 stations at a single instant in time. Obviously, this interval contains all kinds of weather, but at 1.70, the slope has remained essentially unchanged, and the R^2 (0.988) is still extremely high. At each station, GF may change with the magnitude of the sustained wind, type of weather, and time of day and year, but when these data are combined into the network average, the variance vanishes.

Why this result is obtained for this network is not entirely understood. It is noted that, unlike many other networks or combinations of networks, the SDG&E mesonet is very homogeneous with respect to instrumentation, age of facilities, anemometer mounting height and station siting philosophy. It was already seen that while GF can vary substantially with sustained wind (**Fig. 5.1**) at a single station, but even at that site the variability was already considerably reduced when the wind-gust pairs were considered in bulk (**Fig. 5.2**). The least-squares fit to the entire dataset yields GF = 1.51, with an R^2 of 0.95. The dispersion, as measured by the R^2 statistic, was already

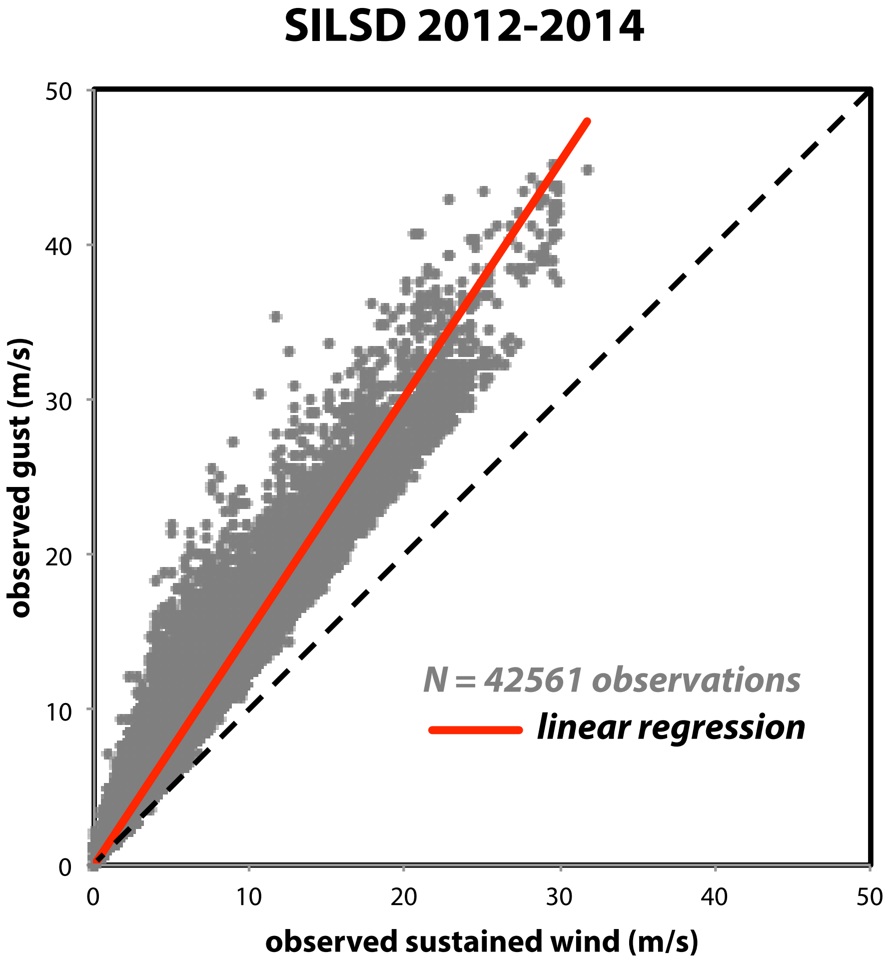


Figure 5.2: Scatterplot of network-averaged sustained wind vs. gust at station SIL during 2012-2014. 42561 observation times are plotted. The intercept-suppressed least--squares fit (red line) and the 1:1 line (dashed black) are shown for reference.

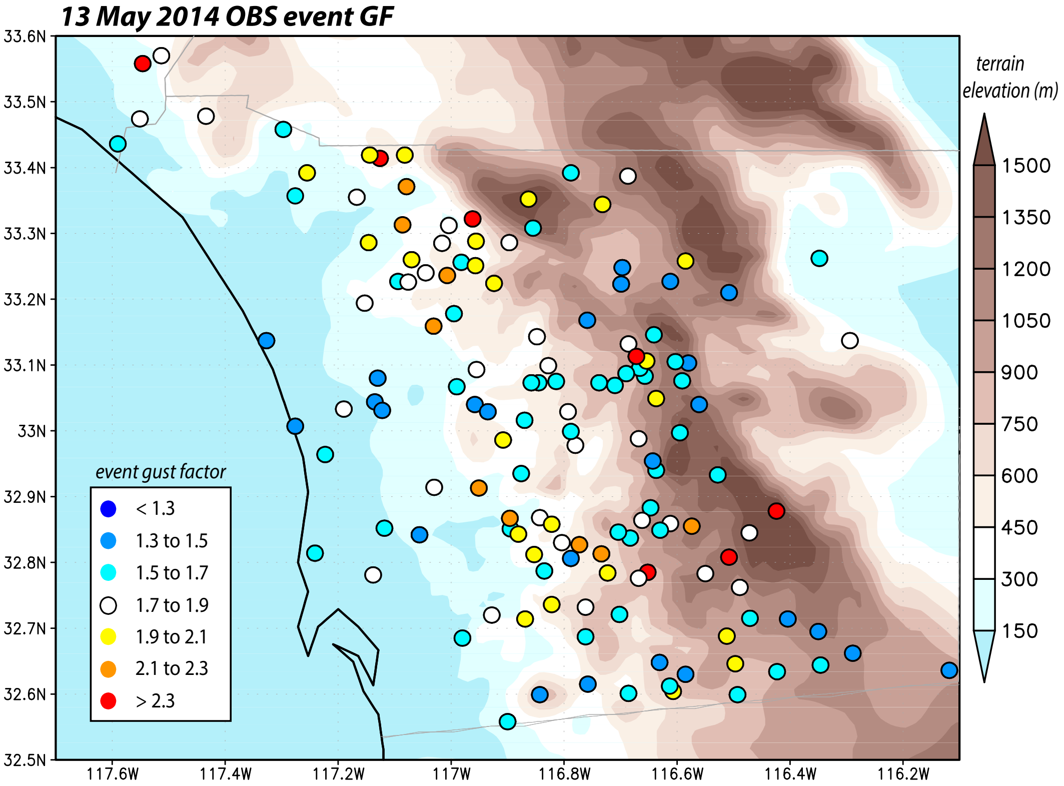


Figure 5.3: Similar to **Fig. 3.2**, but for the observed event-gust factor for the 13 -15 May 2014 event. The event gust factor is determined via regression for each SDG&E station during the event.

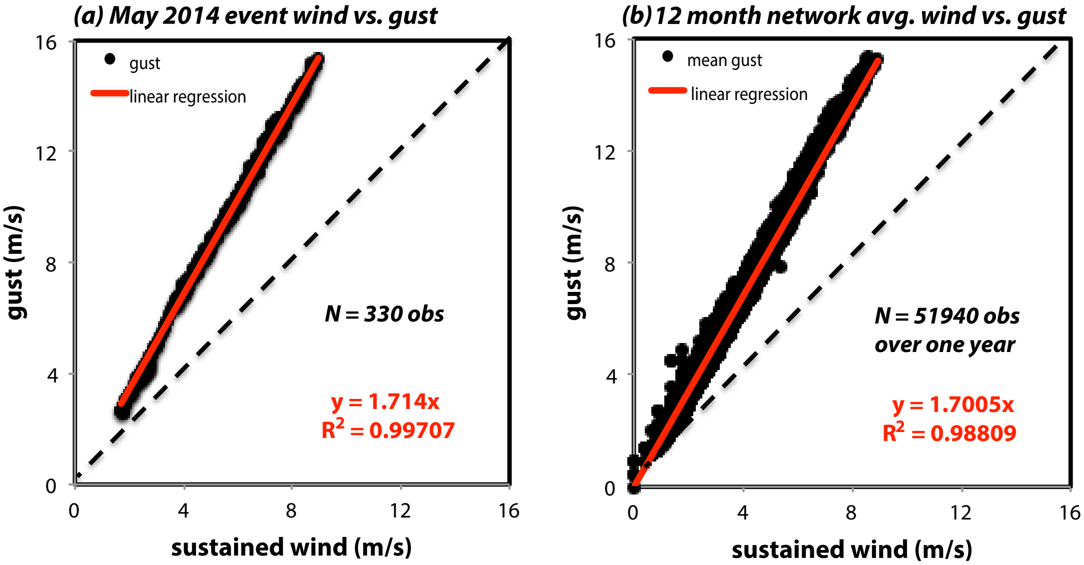


Figure 5.4: (a) Same as **Fig. 5.2**, but for the May 2014 event. The intercept-suppressed least-squares fit (red line) is shown with slope 1.714 and R^2=0.997; and (b) Same as (a), but for a composite of 10-min observations over a period of twelve consecutive months consisting of 51,940 instances. The 1:1 line (dashed black) is shown for reference.

low, and this is true at all sites in this mesonet. The network average is obtained by integrating over this collection of station datasets, each with relatively low dispersion. Furthermore, the network is nearly static – consisting of essentially the same hardware, placed at the same sites – over the period under examination. Given these factors, perhaps this result is not so surprising after all.

5.3 The network-averaged gust factor and systematic forecast bias

Whether or not we have adequately explained this result, we will take the SDG&E network-averaged GF as 1.7, and refer to it as "G". Now, it was already shown that, even for a single event, station GF can vary from the network average G, depending on location (e.g., **Fig. 5.3**). Each site contributes to the network composite, and G itself might be altered depending on its inclusion or exclusion. Furthermore, a given site's gust factor may vary from the network average value, at an instant of time, or integrated across all events, for a number of reasons. As an example, Eq. (**5-3**) above already shows that, for a given surface wind, the gust (and GF) should be larger over rougher terrain (larger z_0).

Many of the factors that lead to locally larger or smaller GFs are readily captured by the model, particularly when high resolution capable of representing fine-scale topography and variations in surface roughness and land cover is employed. We have already seen that wind speeds, in nature and in the simulations, vary considerably over relatively short distances, in large part owing to these influences. However, given that the *overall* (network-averaged) gust factor G is nearly insensitive to factors such as time of day, type of weather, wind direction, etc., we hypothesize that stations having individual GFs that vary significantly from the network average may, at least in part, represent the influence of very localized factors. These may include obstacles (buildings, trees, hills) that can act to weaken or intensify the wind at one location relative to another, depending on specific conditions. Furthermore, to the degree that these localized factors are *unresolvable*, even on a high-resolution grid, we anticipate finding that even a model that is properly configured overall will be more likely to have systematic biases at these stations.

Thus, other factors being equal, *we anticipate overpredicting the sustained wind at stations having GF > G, while underpredicting winds at stations with GF < G.* Additionally, the more the local GF deviates from the network average, the larger the bias is expected to be. This hypothesis is demonstrated in **Fig. 5.5**, with **Fig. 5.5 a** illustrating the standard case. The wind profile is assumed to be logarithmic with height and thus calm at height z = z_0. A parcel possessing faster horizontal velocity is transported downward by turbulent motions, and manifested at anemometer level as a gust U_max exceeding the sustained wind $\overline{U}$ at that height. If the gust factor for this station is comparable to the network average, we anticipate that a properly configured model will be able to represent the winds at this location without significant bias.

Despite best efforts regarding siting, however, some stations will experience at least very localized obstructions. For example, anemometers might have to be installed relatively close to landforms that might partially shield them, or be placed in an area with denser and/or taller vegetation than is representative of the grid cell in which it is found. In those cases, we anticipate that the sustained wind is slowed more than would be expected given the z_0 value assigned to the grid cell and the strength of the winds farther aloft. However, a parcel impulsively transported downward by turbulence would have less time to be influenced by the obstructions, and thus would appear *stronger* relative to $\overline{U}$, resulting in a larger station GF, as illustrated in **Fig. 5.5 b**. If these obstructions cannot be resolved on the model grid[2], or represented by the grid's roughness length, we anticipate overpredicting the wind at these stations with larger than average G.

Alternatively, some stations may be located with areas with landforms that serve to further accelerate the wind, including flow through favorably oriented canyons, near

---

[2] A good example of this is the aforementioned Mt. Laguna station (MLGSD), which is closely surrounded by trees. This specific site was selected to conduct comparisons with a pre-existing, similarly sheltered station nearby called MLGC1. B. D'Agostino and S. Vanderburg (personal communication).

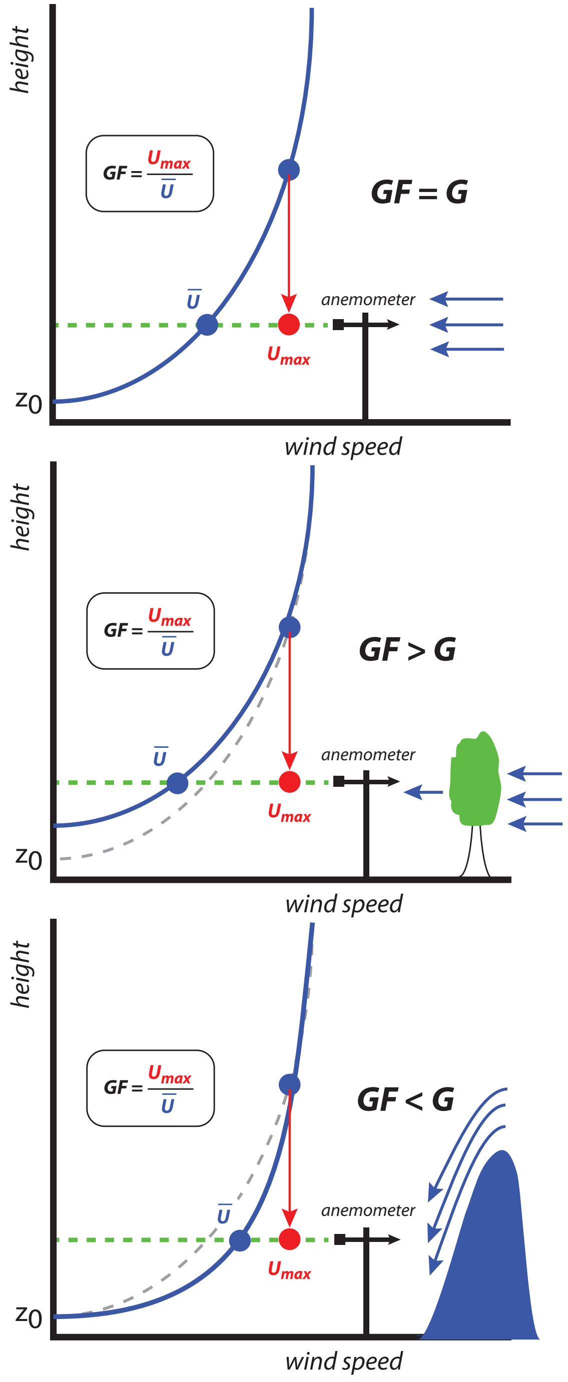


Figure 5.5: Illustration of wind bias concept, based on station GF relative to network average G: (a) standard case; (b) obstructed case; and (c) enhanced case.

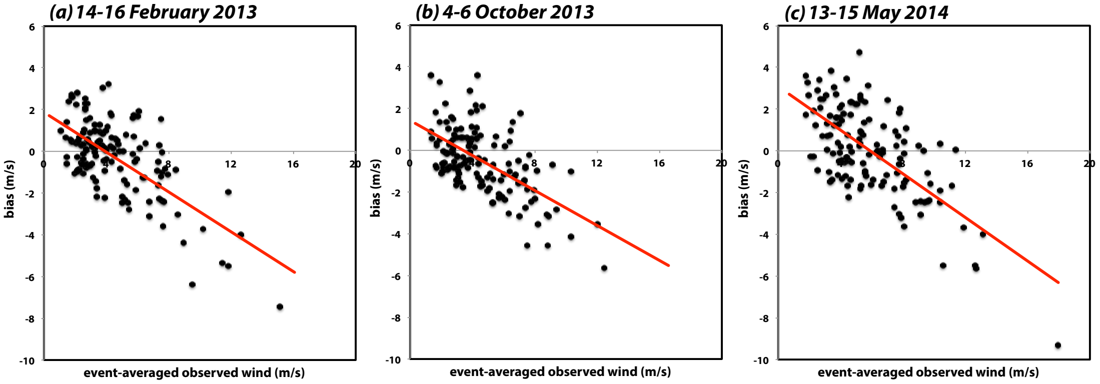


Figure 5.6: Scatterplot of event-mean observed wind vs. mean bias in the PX/ACM2 simulation for SDG&E stations for (a) the 14-16 February 2013 event; (b) the 4-6 October 2013 event; and (c) the 13-15 May 2014 event. A least-squares fit (red line) is shown on each figure for reference.

steep ravines, or over small hills. These may serve to enhance the sustained wind at anemometer level, such that a descending parcel has a relatively smaller speed excess over the mean flow there, resulting in GF < G (**Fig. 5.5 c**). If those amplifying features are unresolvable, we hypothesize that we will systematically underpredict the wind there.

We already showed in **Chapter 3** that while the control simulation (using the PX/ACM2 physics combination) for the 14-16 February 2013 event did a fine job of simulating sustained winds over the SDG&E domain overall, it remained that the model forecast error was negatively correlated with the observed wind (**Fig. 3.11 b; see also Fig. 5.6 a**). The model tended to underforecast sites that were relatively windy, while overpredicting locations with relatively weaker wind speeds through the event. In other words, the overall forecast bias was small, but the bias *itself* was biased. This did not occur for this event alone, as **Figs. 5.6 b and c** reveal that the same tendency occurred in the PX/ACM2 simulations of two other events, the 4-6 October 2013 and 13-15 May 2014 cases.

Our hypothesis is tested in **Fig. 5.7**, which compares the event-averaged forecast wind bias for each station in the SDG&E mesonet with its corresponding observed event GF for the same three Santa Ana episodes. Each plot is divided into quadrants, based on G (1.7), and the zero-bias line. As anticipated, there is a tendency for the model to exhibit positive bias at sites having gust factors exceeding the network average, while sites with GF < G are far more likely to be underforecasted. The R^2 for the October and May cases is ~0.52. Although the fit is poorer ($R^2 = 0.27$) for the February event, the sign of the bias is correctly identified for the majority of stations based on the location GF alone.

We believe this result suggests a very simple gust parameterization might be applied to properly configured (i.e., no network-average bias) model simulations of winds in the SDG&E mesonet, employing a uniform GF of about 1.7 (= G) to *all* stations at all times (except possibly during thunderstorms, which are rare in Southern California). This approach should at least partially remove the dependence of wind forecast bias on sustained wind speed. Stations with GF > G tend to be overpredicted already, so using a smaller GF value than justified from the observational record would work to mitigate the positive sustained wind forecast bias at those locations. Similarly, underpredicted stations generally have GF < G, so using a larger than observed GF helps correct for the underforecasted sustained wind. Certainly, a more sophisticated treatment could be designed[3] but we are encouraged that an attractively simple gust parameterization could be utilized with skillful sustained wind forecasts in this region.

5.4 Testing gust parameterizations

Figure 5.8 presents time series of the gust observations (black dots) averaged over the SDG&E mesonet for the three Santa Ana episodes. Also shown are three gust estimates (NCEP, EC and simple) derived from the model simulations made using the PX/ACM2 physics combination. The simple parameterization applies a constant gust factor of 1.7

---

[3] When sustained winds are lighter, larger GFs are probably appropriate, but the threat from weak gusts is not very substantial.

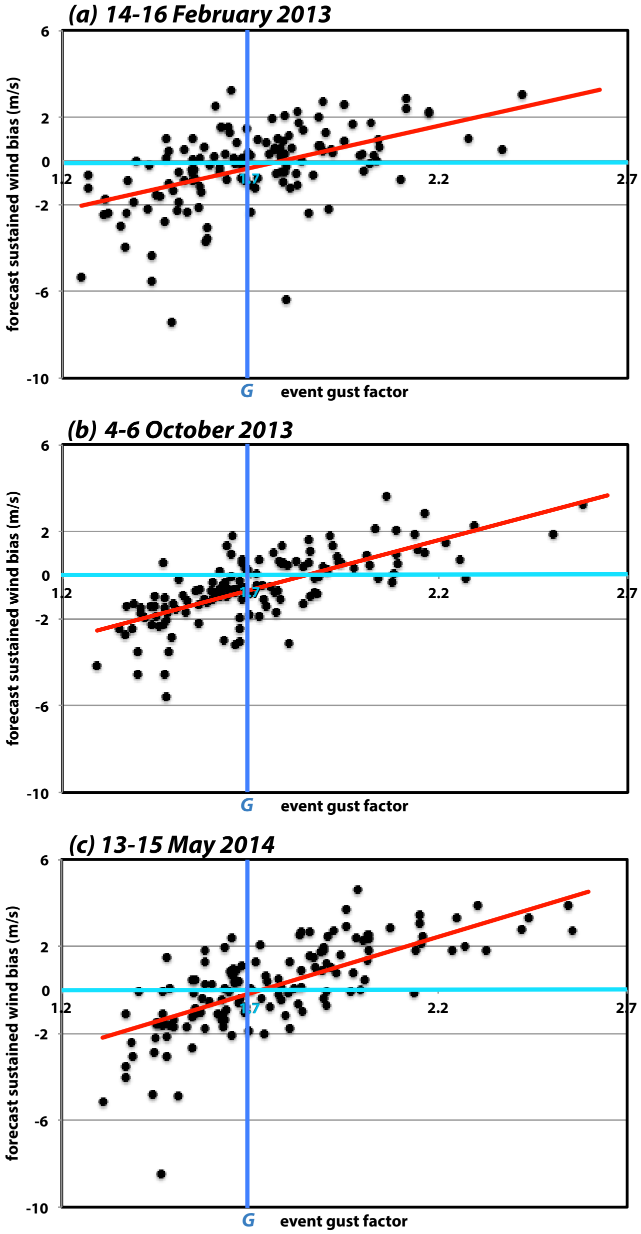


Figure 5.7: Scatterplot of event-mean sustained wind bias from the PX/ACM2 simulation vs. event GF for (a) the 14-16 February 2013 event; (b) the 4-6 October 2013 event; and (c) the 13-15 May 2014 event. The blue vertical line represents the SDG&E network gust factor average, G ~1.7. The cyan line represents the zero-bias line. The red line represents a linear least-squares fit, predicting model bias from station GF.

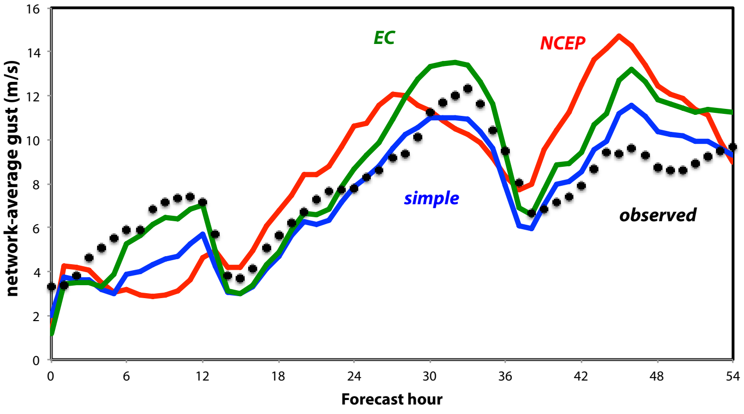
to the sustained wind predictions for each site. Naturally, all three strategies depend on the sustained wind predictions being correct.

The common feature among the three events is the NCEP algorithm overpredicts the gusts, but the more serious issue is it often peaks too early. The EC gust is also more likely to overpredict than underforecast. The EC and simple algorithms are obviously well correlated, so the EC's overprediction tendency might be remedied by training the friction velocity multiplier for this network. The simple algorithm does best for each case, having the smallest MAE averaged through each of the events (not shown).

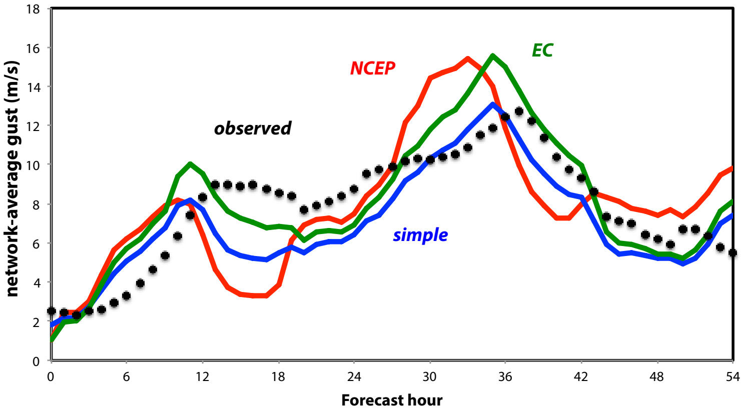
Finally, in **Fig. 5.9**, we examine how well the simple parameterization fares in predicting event mean gusts in the three Santa Ana wind events. Each point on these figures represents an SDG&E station. Considerable spread from the 1:1 line (dashed black) remains for the three events but, importantly, note that the gust forecasts are *unbiased*. Applying a constant GF to the sustained wind forecasts has taken predictions that were clearly biased (recall **Fig. 5.6**) and corrected for it. Naturally, bias correction can always be applied to raw model forecasts, and the simple algorithm constitutes such a correction. Additionally, more sophisticated correction procedures can always be crafted. In the present application, however, we believe we understand how the forecast bias originates and why the correction works.

5.5 Summary

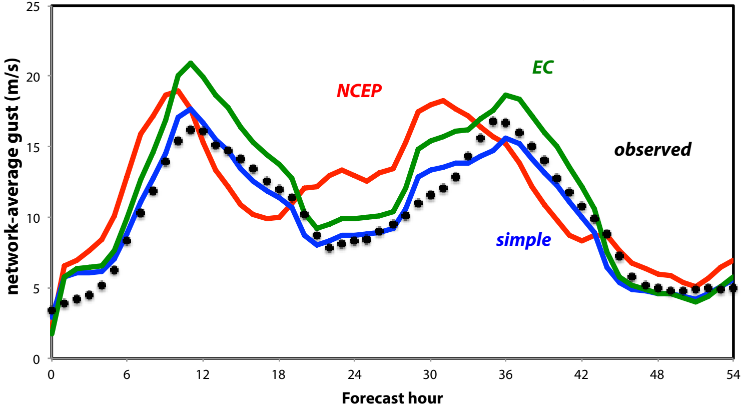
In fire-prone areas such as southern California, wind gusts are potentially hazardous to lives, properties, and public transportation since strong gusty winds can knock down trees and power lines, and spark and spread fires. Due to its chaotic nature and lack of adequate knowledge of its formation and behavior, it remains a day-to-day challenge to accurately predict wind gust speed, especially in the mountainous regions. In this chapter, we have examined the ability of the WRF model to provide skillful gust predictions for the SDG&E network. It is observed that each station tends to have its own



(a) Gust parameterization comparison: 14-16 Feb 2013



(b) Gust parameterization comparison: 4-6 Oct 2013



(c) Gust parameterization comparison: 13-15 May 2014

Figure 5.8: Time series of the gust observations (black dots) averaged over the SDG&E mesonet for (a) the 14-16 February 2013 event; (b) the 4-6 October 2013 event; and (c) the 13-15 May 2014 event. Also shown are the NCEP (red), EC (green), and simple (blue) gust estimates.

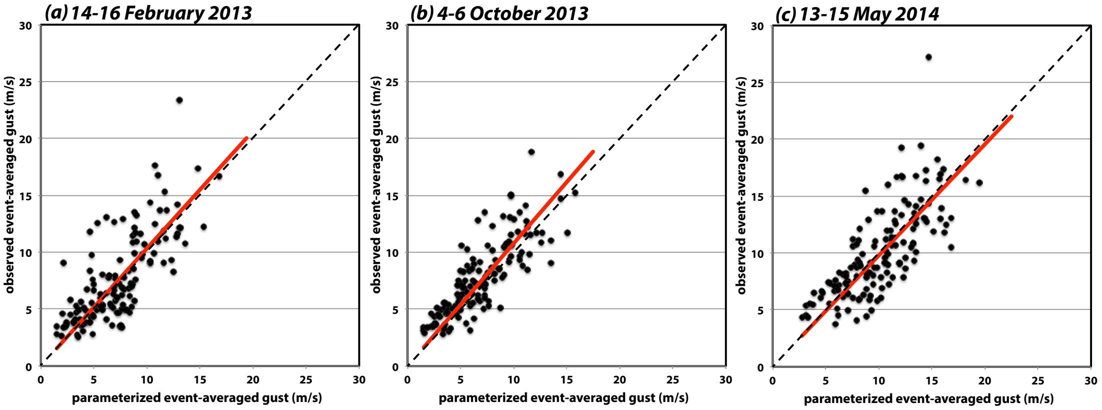


Figure 5.9: Scatterplot of observed event-averaged gust (vertical axis) vs. parameterized gust for the (a) Feb 2013; (b) Oct 2013; and (c) May 2014 events, based on event-averaged wind multiplied by a station-independent GF of 1.7. The least-squares fit (red line) and the 1:1 line (dashed black) are shown for reference.

event-averaged gust factor, and gust factors among stations vary greatly. However, to our surprise, we have discovered that the SDG&E network-averaged gust factor, which we termed G, is nearly constant with season, time, and event (apart from thunderstorm activity), with a value of about 1.7.

While this finding is not well-understood, we anticipated, and demonstrated, that stations with gust factors (GF) smaller than G were likely to be underpredicted in the model, while the winds at stations with GF greater than G were likely overpredicted. This was used to separate the forecast error into that which might still be rectified, by modifying surface characteristics and perhaps model physics, and that which was probably "unfixable" other than via ex post facto bias correction. Thus, we propose a simple gust parameterization, with a GF of 1.7, for all stations in the network, because the constant GF works to mitigate wind biases found at the more problematic stations.

Comparisons of the NCEP, ECMWF gust forecast techniques and our simple gust factor approach to observations reveals that the simple gust algorithm provides most reliable forecasts. However, it is cautioned that the constant network GF may reflect, and very likely depend on, the homogeneity of the SDG&E network, with respect to hardware, mounting height, sampling interval and siting philosophy, and therefore may not be applicable outside of the San Diego mesonet.

CHAPTER 6

Modeling live fuel moisture and greenness

The Santa Ana Wildfire Threat Index (SAWTI; **Chapter 1**) is heavily dependent on live fuel moisture (LFM) and greenness (G), two parts of the fuel moisture component (FMC) of Large Fire Potential (LFP). In this chapter, we describe our attempts to skillfully model these components using meteorological information.

6.1 Live fuel moisture

6.1.1 Problem and strategy

LFM is a measure of the moisture content of vegetation, which is important owing to its recognized role in the fire hazard [*Pyne et al.*, 1996, *Schoenberg et al.*, 2003]. Large fires often occur during periods of low LFM [*Dennison and Moritz*, 2009], with size increasing as LFM decreases [*Davis and Michaelsen*, 1995]. LFM is the ratio of the water weight in a particular sample to its dry weight, expressed as a percentage [*Pollet and Brown*, 2007]. Owing to its formulation, the LFM can exceed 100%, and often does during wet conditions.

While a variety of vegetation species are routinely sampled, the primary emphasis in Southern California is on chamise (*Adenostoma fasciculatum*), a common shrub of the chaparral family, owing to its sensitivity to drought [*Dennison and Moritz*, 2009, *Dennison and Roberts*, 2003]. Fuel moisture samples are usually taken twice a month by various fire agencies across southern California, using the technique described in *Countryman and Dean* [1979]. Unfortunately, the sampling sites are sparse and the sampling times are not coordinated between fire agencies. In addition, the equipment used to dry the samples is not standardized which can lead to inconsistencies in the data. The fact that the age of the sampled material may vary from site to site is another complicating factor, as when moisture is abundant, new growth[1] can contain much more water than older parts of the plant. It is understood there is a sizable amount of uncertainty in LFM measurements [*Weise et al.*, 1998].

Over the years, there have been a number of attempts to model LFM for various species, especially employing easily obtained or computed meteorological information [e.g., *López et al.*, 2002, *Viegas et al.*, 2001], although it is recognized that such information alone cannot fully describe the moisture content of live vegetation [*Fiorucci et al.*, 2007]. *Dennison and Roberts* [2003] evaluated the skill of models based on the Keetch-Byram Drought Index (KBDI) [*Keetch and Byram*, 1968] and the Cumulative Water Balance Index (CWBI) [*Dennison and Roberts*, 2003], as well as remotely sensed data, in reproducing LFM values sampled at sites in the Santa Monica mountains. A skillful relationship between CWBI and LFM for several species, including chamise, was demonstrated. *Qi et al.* [2012] used soil moisture sampled in the field along with remotely sensed data to model LFM values measured from species of oak and sagebrush. Although mean absolute errors were still sizable (±20% of LFM), soil moisture measured *in situ* emerged as the most useful LFM proxy.

The recognized dependence of chamise LFM on soil moisture availability [*Dennison and Roberts*, 2003] suggests that soil moisture could serve as a reasonable proxy for this species as well. However, sampling LFM is already labor-intensive and expanding the program to measure soil conditions at the same sites [as was done by *Qi et al.*, 2012] would add to that burden. Regarding Southern California chamise, *Dennison and Moritz* [2009] demonstrated the utility of using rainfall that, after all, is the source of soil moisture. We elected to pursue a different strategy, and test *whether a readily available, gridded soil moisture product could be used to reproduce historical*

---

[1] Plant growth of the current year.

LFM values. If completely successful, the resulting product would be of significant use in operational fire behavior forecasting, as the proxy could be applied to the enormous areas for which LFM sampling is unavailable. It would also be available to fill in temporal gaps in the observational record, as well as extend the historical record.

6.1.2 Data and methods

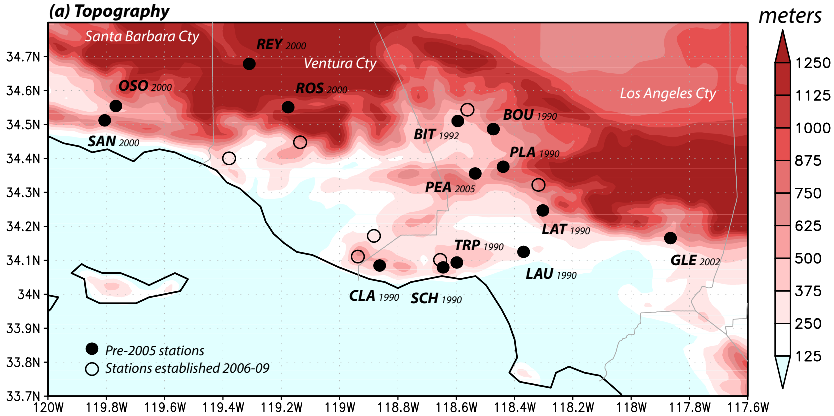
Historical live fuel moisture data was obtained from the National Fuel Moisture Database (NFMD) website[2]. For this study, we are focusing on LFM sampling sites located in the mountains located north and west of the city of Los Angeles, encompassing parts of Los Angeles, Ventura and Santa Barbara counties, at which chamise samples have been taken. Most of the sites are located on south-facing mountain slopes (**Fig. 6.1 a**), at elevations ranging from about 400 to 1200 m above mean sea level (**Table 6.1**). Only sites established prior to January 2006, indicated by the filled black circles, are considered (see **Table 6.1** for site information), as these have sufficient observations available to characterize the annual LFM cycle. Other stations exist, and those established between 2006 and 2009 are marked by the unfilled black circles. In the subset of sites where both new and old chamise is sampled, the new growth data are used.

For our soil moisture data source, we selected the North American Land Data Assimilation System (NLDAS) [*Mitchell et al.*, 2004, *Xia et al.*, 2012a,b] reanalysis product. The NLDAS project combines available observations and model output to provide surface and subsurface information over the conterminous United States at 1/8th degree spatial resolution. The NLDAS-2 model employs several LSMs, including the Noah model that is used in the operational weather forecasting models from the NCEP and is also a popular choice of WRF model users. While other available versions of the NLDAS-2 dataset have been considered for this work, we will confine our present analyses to the Noah product.

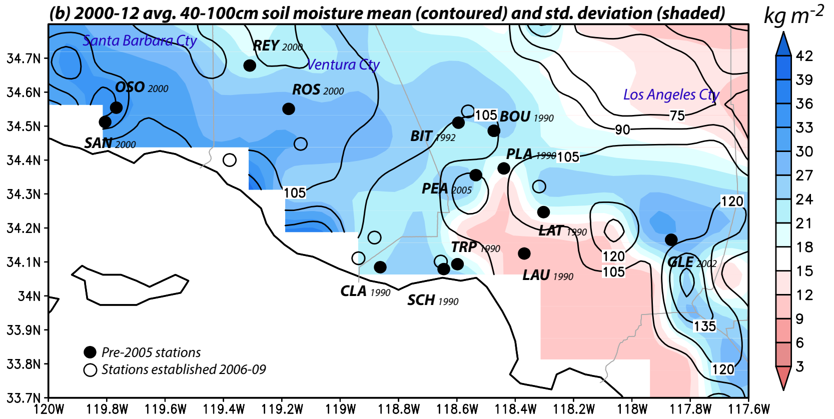
---

[2] http://www.wfas.net/nfmd/public/states_map.php?state=CA

(a) Topography



(b) 2000-12 avg. 40-100cm soil moisture mean (contoured) and std. deviation (shaded)



(c) Soil texture classification from the NLDAS/Noah dataset

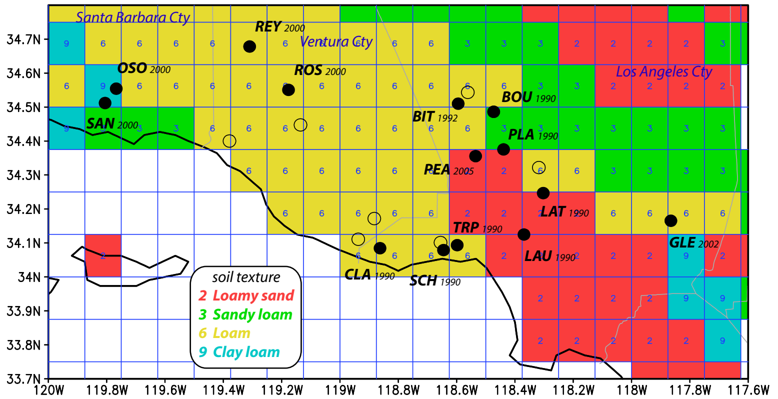


Figure 6.1: (a) Live Fuel Moisture surface site locations (black dots), with underlying topography shaded; (b) 2000-2012 averaged 40-100 cm layer soil moisture mean (contoured) and standard deviation (shaded); and (c) Soil texture classification over Southern California from the NLDAS/Noah dataset.

Table 6.1: Live fuel moisture sites elevation and the R-squares of the models.

Abbr.	Station	Elevation (m)	LFM annual model R^2	Soil moisture annual model R^2	Max partial corr. (lag)	Partial corr. at lag 29	Full model R^2
LAU	Laurel Cyn	398	0.57	0.57	0.81 (29)	0.81	0.85
LAT	La Tuna Cyn	610	0.61	0.54	0.69 (35)	0.68	0.82
PLA	Placerita Cyn	602	0.55	0.54	0.73 (36)	0.72	0.78
PEA	Peach Motorwy	579	0.63	0.47	0.54 (0)	0.48	0.74
PEA	Peach (mod)	579	0.63	0.55	0.66 (21)	0.66	0.82
GLE	Glendora Ridge	751	0.60	0.57	0.67 (38)	0.67	0.77
BIT	Bitter Cyn	532	0.48	0.46	0.68 (3)	0.64	0.70
BIT	Bitter Cyn (mod)	532	0.48	0.54	0.74 (32)	0.74	0.76
SCH	Schueren Rd	678	0.74	0.68	0.67 (12)	0.65	0.85
SAN	San Marcos	816	0.63	0.69	0.56 (19)	0.55	0.83
ROS	Rose Vly	1096	0.55	0.71	0.59 (18)	0.59	0.70
REY	Reyes Ck	1233	0.45	0.68	0.42 (28)	0.42	0.61

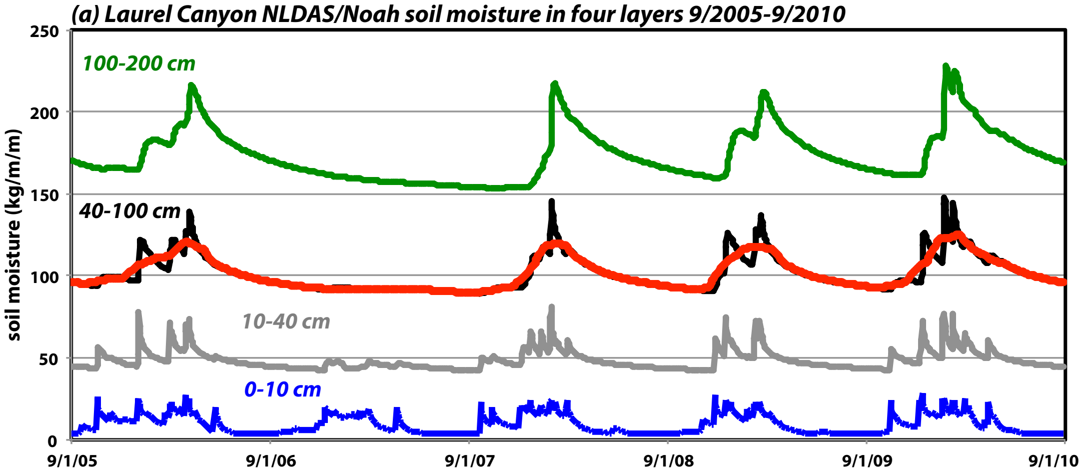
A few stations have been intentionally excluded from this analysis. The Bouquet Canyon (BOU on **Fig. 6.1 a**) record has several long gaps, including after the site was burned owing to the October 2007 wildfires. Upper Oso (OSO on **Fig. 6.1 a**) is one of the more infrequently sampled locations, and has several long stretches of missing data. This site is located very near San Marcos (SAN) although on the other side of the narrow coastal range; the relatively coarse NLDAS grid cannot differentiate between the two anyway. Clark Motorway (CLA) and Trippet Ranch (TRP) reside in the Santa Monica mountains near Malibu, and are highly correlated with a nearby station, Schueren Road (SCH), that is included. Subsequent to December 2011, the NLDAS soil moisture information at these three coastal sites has no value, a consequence of a change in NLDAS data sources and interpolation strategy (Youlong Xia, personal communication); as a consequence, our LFM modeling at SCH ceases on 1 January 2012. Templin Highway (not shown on **Fig. 6.1 a**) was established in 1990, and samples are being taken there at present, but data are missing for the years 1997 through 2012.

Chamise has a dual root system, “a broad, near-surface system to take advantage of light winter storms and a deep root system to tap deeper sources of water during summer drought" [*Henson et al.*, 1996] that can reach depths as large as 8 m [*Lambers et al.*, 2008]. The Noah LSM uses four soil layers, representing 0-10, 10-40, 40-100, and 100-200 cm below ground level. It is not precisely clear which of these best represents the shrub's root zone. We tested several layers and layer combinations before settling on the 40-100 cm layer for this analysis.

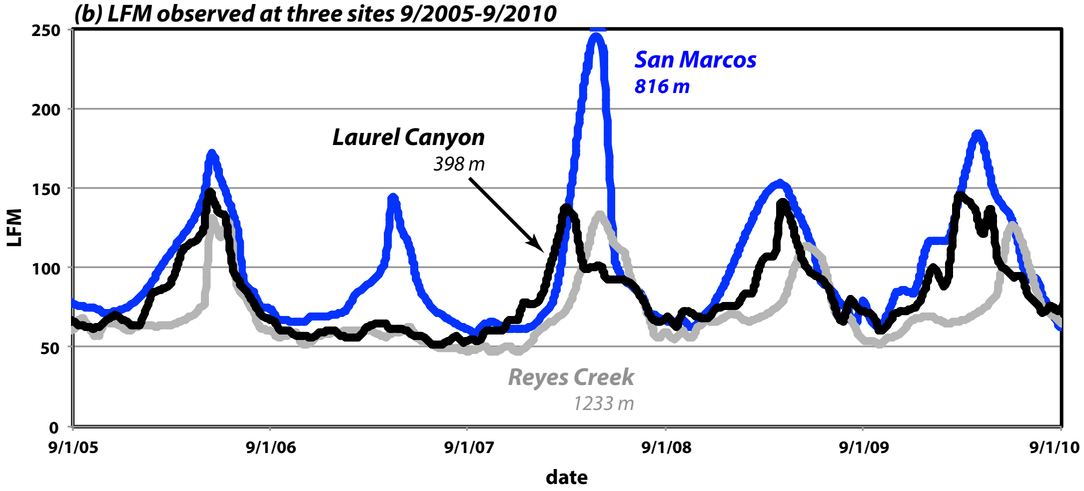
Both LFM and soil moisture data can be noisy, subject to sharp, but often short-lived spikes following precipitation events. In the case of LFM, the noisiness is exacerbated by infrequent sampling as well as the aforementioned inherent uncertainty. Our emphasis is on capturing relatively longer temporal variations of fuel moisture that might better characterize the seasonal fire hazard, rather than short-lived undulations in the LFM record. Thus, we view short-term variations in both series as distractions.

In this regard, the Noah 40-100 cm layer has some practical advantages. Unlike the shallower layers near the surface, the moisture in this soil layer is less subject to the high frequency pulses that immediately follow precipitation events. The temporal variation at this level is still sizable, but the time needed for moisture to pass through the top two layers acts as a natural filter. In contrast, the depth and thickness of the lowest level, 100-200 cm, renders variation in that layer to be small. This is illustrated for a single site, Laurel Canyon (LAU), in **Fig. 6.2 a**. Note only the more sizable and prolonged rainfall events propagate down to the lower two soil levels. Remaining short-period variations in the 40-100 cm record will be removed with smoothing, via application of a ±30 day unweighted filter (see red curve on **Fig. 6.2 a**).

Given the dual-root structure of chamise, a combination of both shallower and deeper soil moisture might represent an improvement over the approach outlined below. For the present, however, our goal is to create the simplest possible model that skillfully captures the temporal variation of LFM.



(a) Laurel Canyon NLDAS/Noah soil moisture in four layers 9/2005-9/2010



(b) LFM observed at three sites 9/2005-9/2010

Figure 6.2: (a) Site Laurel Canyon soil moisture in four layers September 2005-September 2010 from the NLDAS/Noah dataset; and (b) Observed LFM at Laurel Canyon, Reyes Creek and San Marcos September 2005-September 2010. The red curve represents smoothing via application of a ±30 day unweighted filter.

6.1.3 Analysis strategy

In general, LFM is influenced by a variety of factors, including moisture availability, evapotranspiration, and plant physiology [*Qi et al.*, 2012]. In the study area, nearly all of a typical year's precipitation falls in the non-summer months, so LFM values typically reach their lowest values in the fall, before the seasonal rains have commenced. For chamise, another control has to be air temperature, which to a certain extent, determines whether the shrub is active or dormant[3]. Due to dormancy, the vegetation at higher, cooler elevations may not take up moisture from the soil, even long after the winter rains have begun. The consequence of this is that the LFM of Southern California chamise has a pronounced annual cycle reflecting not only the temporal variation of rainfall, but also the influence of elevation.

This dependence can be seen in **Fig. 6.2 b**, which shows the temporal variation of LFM for three stations over a five-year period. The intermediate elevation station, San Marcos (816 m), is the moistest of the group, likely reflecting its coastal location above Santa Barbara (**Fig. 6.1 a**). Although this station's peak values vary substantially through the period, they tend to occur in the spring season, between March and May. In contrast, the lower (Laurel Canyon, 398 m) and higher (Reyes Creek, 1233 m) sites evinced earlier and later peaks, respectively, in three of the five years depicted.

Figure 6.3 a presents the LFM annual models for the ten stations selected for analysis, based on a day-of-year (DOY) measure that for convenience starts on September 1^{st}. These cycles were determined in two ways, both involving a "time function" (TF) consisting of four terms – a sine and cosine with a period of 12 months, and its first harmonic (6 month period) – a combination capable of identifying oddly shaped annual cycles[4]. In one application, the time function is fitted to LFM data interpolated using

---

[3] The time of year when chamise begins to emerge from dormancy is dependent on temperature; however, after the reproduction cycle is over, the plant will begin to enter dormancy regardless of temperature.

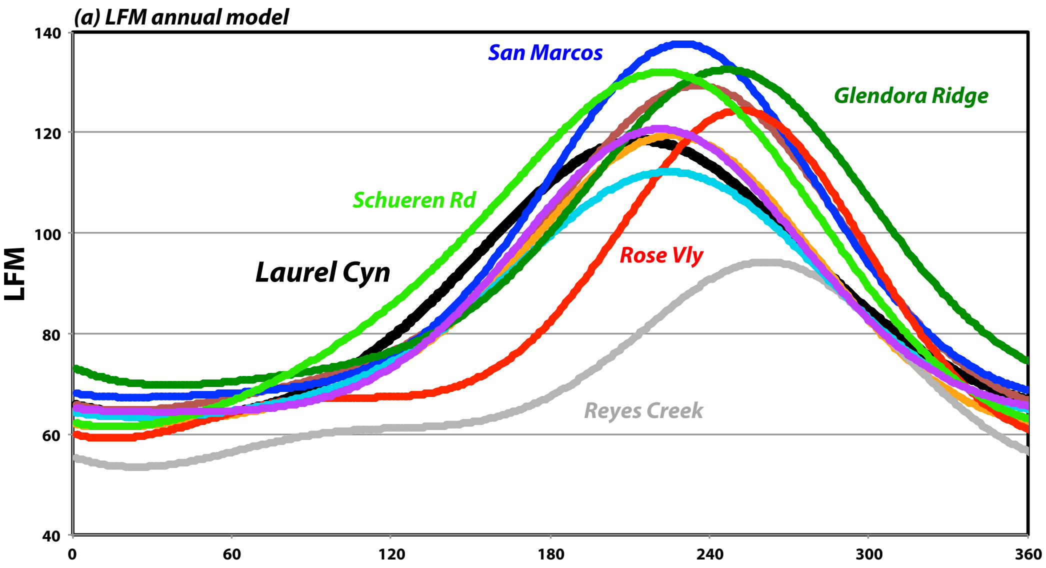
[4] While statistically significant for most stations, the 6-month sine and cosine terms contribute only trivially to the time function, and could be excluded with relatively little loss of skill. However, this shorter period function helps refine the shape of the annual cycle the most at locations like Reyes Creek. We want to use equations with the same functional form at all stations.

splines to daily frequency, representing a truncated periodogram spectral analysis but with coefficients determined via linear regression. The second regresses the four-term time function on LFM based on the original, irregularly spaced data. For sufficiently long series without serious data gaps, the two approaches yield nearly identical results, as anticipated, and this is one reason why we restricted our analysis to more established sites. A log transform is always applied to LFM data, as this helps to compress this positive definite, open-ended series.

It is clear that LFM annual cycles vary among the sites with respect to amplitude and phase. As was suggested by **Fig. 6.2 b**, lower elevation stations tended to reach peak LFM earliest in the season. Laurel Canyon's peak is on DOY 214, April 2^{nd} for a non-leap year. Reyes Creek, the highest elevation site sampled, has a long, slow increase through the wet season, peaking at DOY 261 (May 19^{th}). The dependence on peak DOY with elevation is imperfect but strong (**Fig. 6.4**).

The long-term average and standard deviation of the 40-100cm layer soil moisture (hereafter dubbed "SM") in the NLDAS Noah-based reanalysis is shown in **Fig. 6.1 b**. The temporal variation is small in the urbanized area, including nearby LFM sites Laurel Canyon (LAU), La Tuna Cyn (LAT) and Placerita Canyon (PLA). Glendora Ridge (GLE) and Peach Motorway (PEA) are located near pronounced local maxima and minima of soil moisture, respectively. The relatively coarse resolution of the NLDAS dataset is clear, especially along the coastline.

SM has a markedly different annual cycle than LFM (**Fig. 6.3 b**). With respect to phase and amplitude, the ten locations separate into two very distinct clusters, determined in part by soil characteristics such as texture (**Fig. 6.1 c**), hydraulic conductivity, and field capacity. The three sites at the urban margin, LAU, LAT and PLA, share a relatively small amplitude variation with an early peak at DOY 172 (February 19^{th}). The remaining stations reach peak soil moisture within a few days of DOY 194 (March 13^{th}). Most of the stations share the same annual mean, with SAN, PEA and Bitter Canyon (BIT) emerging as different.



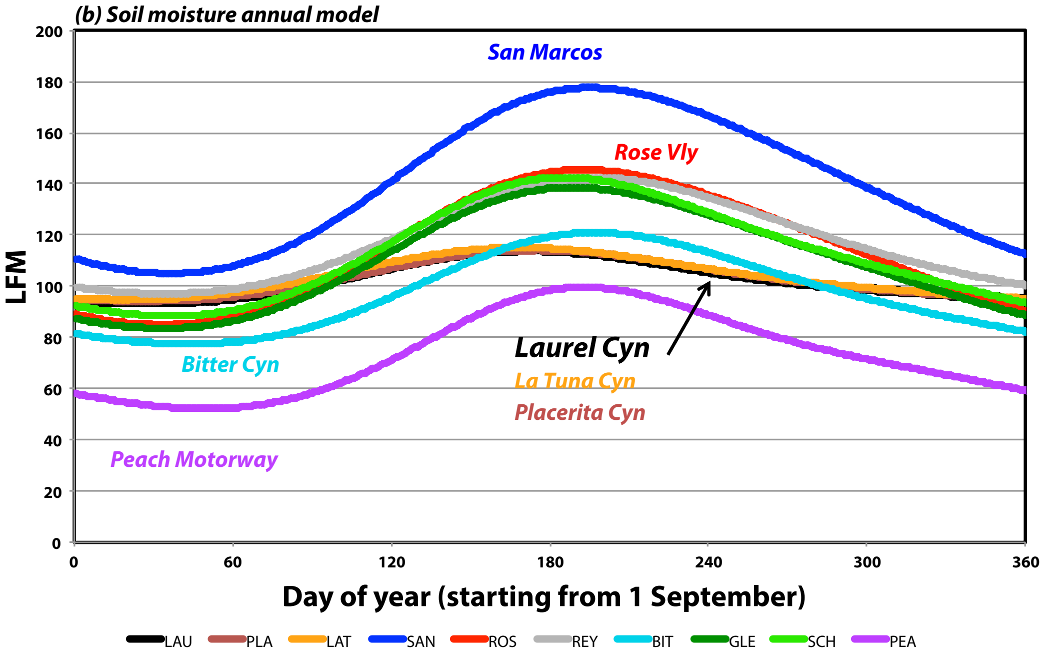


Figure 6.3: (a) LFM annual models for ten sites selected for analysis, based on a day-of-year measure that for convenience starts on September 1st; and (b) As in (a), but for the soil moisture annual model.

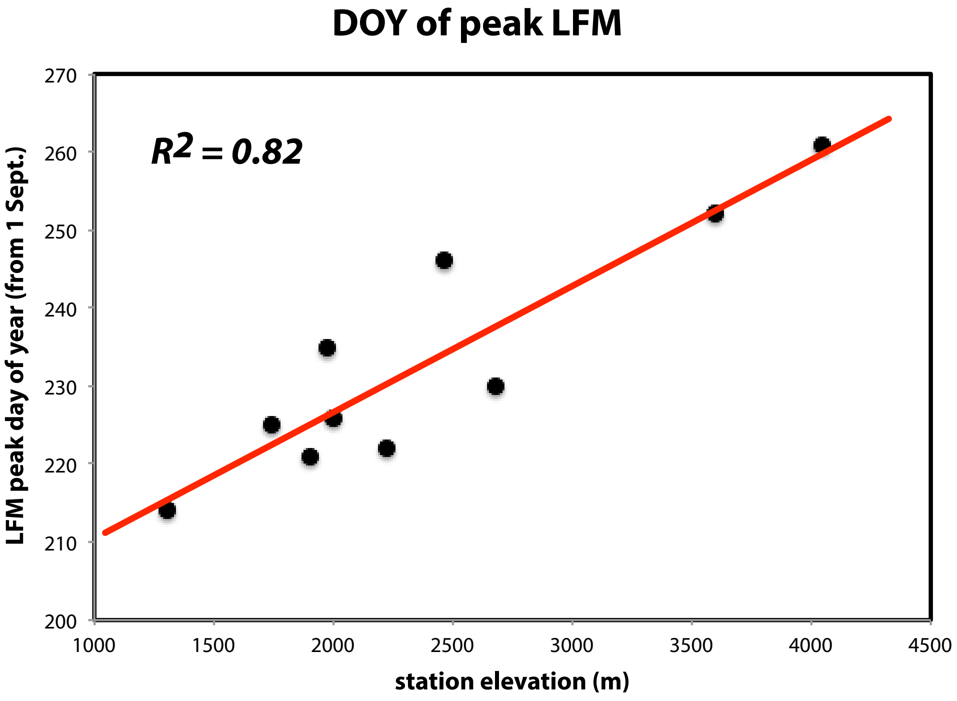


Figure 6.4: Scatterplot of the LFM peak day-of-year (starts on September 1^{st}) vs. station elevation (m). The least-squares fit is shown, with R^2=0.82.

As with LFM, the annual cycles were constructed from log-transformed SM values. Thus suppresses the natural increase of variability with magnitude, helping to keep the model fits homoscedastic. Models for each site were built using available LFM observations alone, and not interpolated data.

6.1.4 The modeling strategy, applied to Laurel Canyon

Our strategy is to predict the temporal variation of LFM using a time function in combination with SM. Specifically, we are concerned with SM's skill in capturing departures from a site's annual LFM cycle. After all, if a site deviates little from its established annual cycle, then climatology is by far the best forecast. As already shown in **Fig. 6.3 a**, climatology varies from station to station based on location, including elevation. Each station, therefore, has its own, unique time function.

This strategy is applied to the Laurel Canyon site in **Fig. 6.5**. Observed LFM (black dots) is compared to the time function (red curve) for the period extending roughly from February 2001 through June 2013 (**Fig. 6.5 a**). The TF captures a sizable 57% of the observed series'variance, which is the average for the 10 stations examined (**Table 6.1**). Removal of the clearly anomalous year of 2007 increases the coefficient of determination (R^2) to 0.7, but without altering the annual cycle shape very much.

The residuals from this fit (**Fig. 6.5 b**) reveal a number of discrepancies of variable length and magnitude. During the spring of 2002 (episode #1 on the figure), observed LFM values declined more rapidly than predicted by the time function, resulting in a several-month period of overprediction. The following year, LFM both ramped up more quickly during fall and declined more slowly in the spring (episode #2), creating an extended period with positive residuals. Episodes #3 and #5 involved near-peak spikes in LFM that occurred during otherwise more "normal" years, and episode #4 involved an early ramp-up followed by a more typical decline towards low warm-season readings.

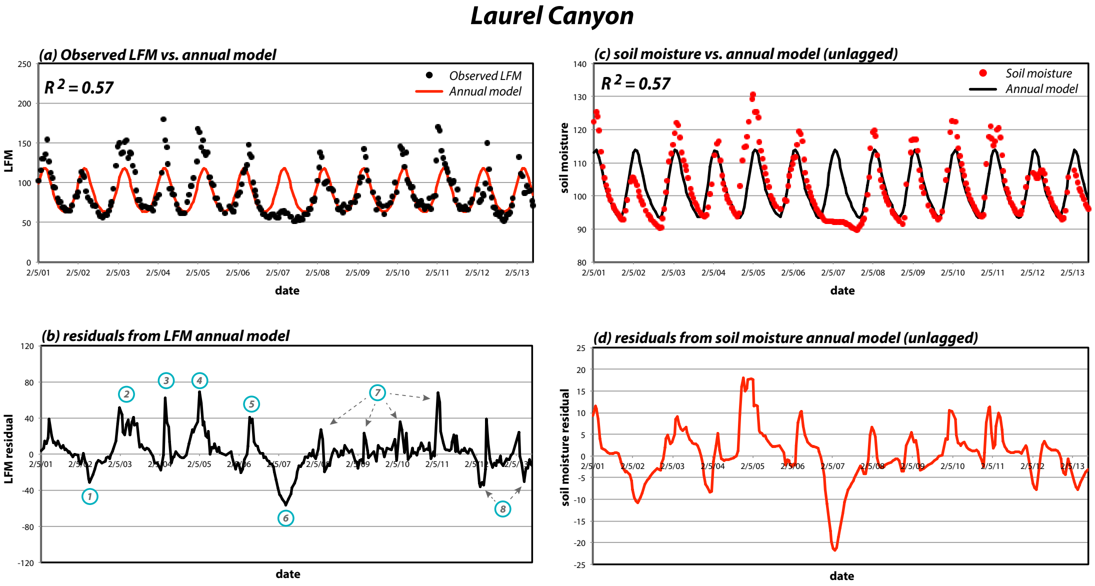


Figure 6.5: Time series of (a) the observed LFM (black dots) vs. the annual model predictions (red curve), R^2=0.57; (b) the residuals from LFM annual model; (c) the observed soil moisture (red dots) vs. the unlagged annual model prediction (black curve), R^2=0.57; and (d) the residuals from the unlagged soil moisture annual model at Site Laurel Canyon. Circled numbers are special episodes referred in the text.

LFM values at this station remained very low during the profound drought of 2007 (episode #6). During this period, LFM values remained well below *Dennison and Moritz* [2009] 's critical threshold value of roughly 79%, under which fire sizes could be significantly larger. The years 2008-11 were characterized by slightly wetter than anticipated fuel moistures, culminating in a more rapid than usual ramp-up in early 2011 (episode #7). The LFM rise was delayed in the winter of 2011, and the decline occurred earlier in the following year, representing the combined period marked episode #8.

As with LFM, the annual cycle for SM captures 57% of the total variance (**Fig. 6.5 c**). Residuals from this fit (**Fig. 6.5 d**) bear significant resemblance to the LFM annual model's departures, and direct comparison is facilitated by superimposing the two error series (**Figs. 6.6 a and d**). This comparison illustrates the partial correlation between the LFM and SM series, adjusted for the time function, which can be written as r(LFM, SM | TF). The partial correlation may be computed by separately regressing LFM and SM on the time function and then correlating the resulting residuals (r_L and r_S), as done here. The resulting squared correlation is 0.55 (r = 0.74), which represents an appreciable contribution of skill to the LFM prediction model that the TF itself could not provide.

While the congruence between the error series is substantial, it is also readily apparent there is a *systematic phase difference* between them. **Figs. 6.6 b** and **e** present the same data, after shifting the SM residuals later by 29 days. This is tantamount to relating LFM to the SM errors *lagged* by 29 days, so that an LFM residual now is informed by a soil moisture deviation from about a month prior[5]. This represents an even longer time interval between precipitation and the LFM response, as it takes time for water received at the surface to percolate downward to the 40-100 cm layer (**Fig. 6.2 a**). It was noted earlier that chamise has a network of shallow roots to take advantage

---

[5] This terminology can be a little confusing. The 29-day phase difference between SM and LFM is a lead or a lag, depending on viewpoint. The departure from SM climatology today is a leading indicator of LFM in the future, as it has skill in predicting the sign and magnitude of the future LFM anomaly. Our focus, however, is on LFM. Today's LFM discrepancy utilizes SM information from the past, meaning SM is a lagging indicator of LFM.

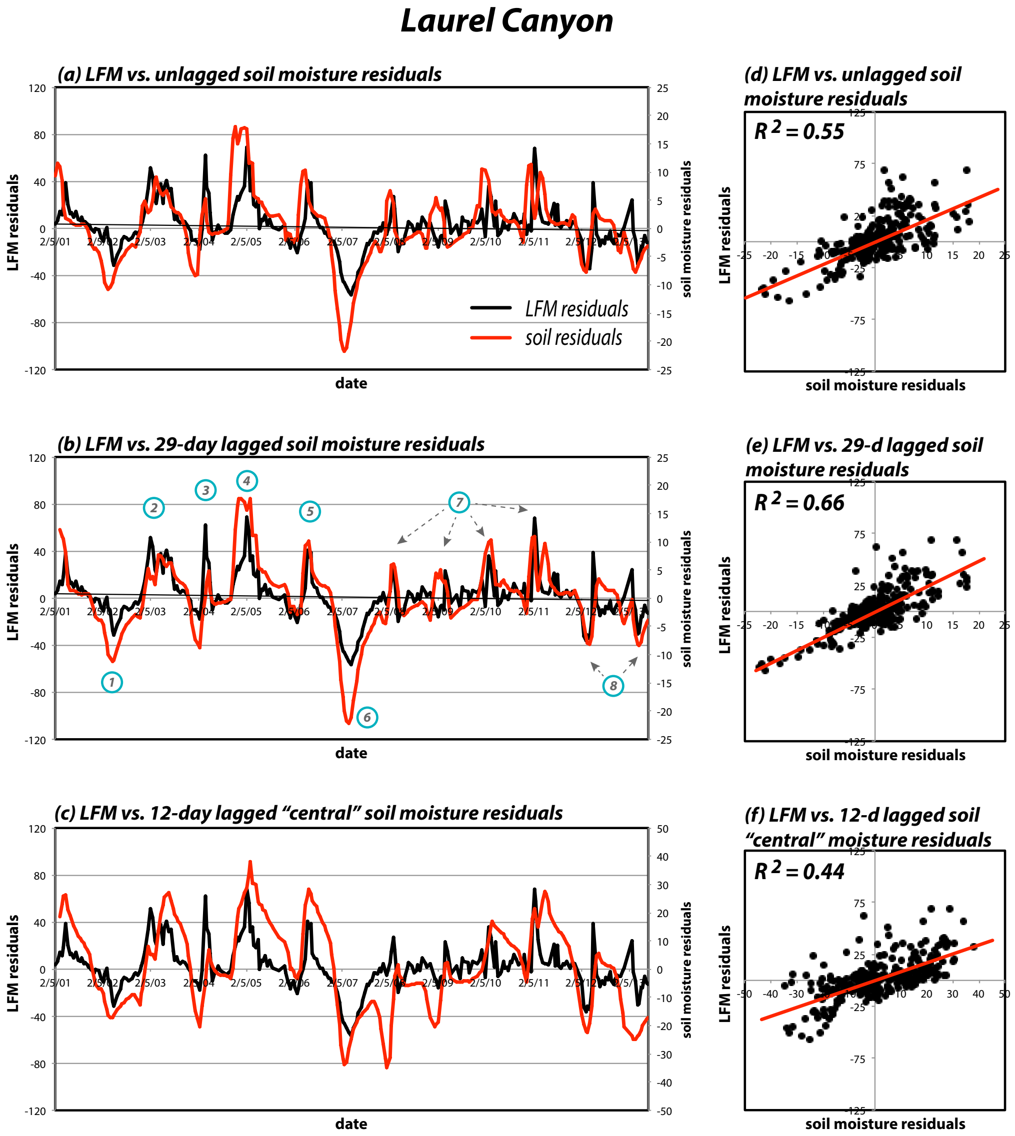


Figure 6.6: Time series of (a) the LFM residuals (black) vs. the unlagged soil moisture residuals (red); (b) the LFM vs. the 29-day lagged soil moisture residuals; and (c) the LFM vs. the 12-day lagged "central" soil moisture residuals at Site Laurel Canyon; and (d-f) are the LFM scatterplots of the LFM residuals vs. the soil moisture residuals of (a-c), with the leas-squares fitting line overlaid, and R^2s being 0.55, 0.66 and 0,44 respectively. Circled numbers are special episodes referred in the text.

of rains, but this result suggests that LFM is responding most directly and powerfully to SM changes farther below the surface, where the shrub's deeper taproots might be expected to reside.

After shifting the SM residual series, there is now a closer correspondence between these TF departures in terms of timing. The faster-than-expected decline of fuel moisture during the spring of 2002 (marked episode #1) corresponded with a more rapid than usual drop in SM, which had reached its own peak about 29 days earlier. LFM episodes #2-5 have significant, if imperfect, correspondences to SM variations, and episode #6 is dramatic in both series. (Naturally, the formal relationship between these series – the slope – will be determined by regression.) Even the small, higher frequency spikes during episodes #7 and 8 appear in the SM residuals. The squared partial correlation between LFM and M anomalies is increased to 0.66 ($r = 0.81$).

The prediction model for Laurel Canyon LFM, given by

$$\log(LFM) = intercept\ +\ TF +\ \log(SM) = \alpha\ +\ \beta_1 \cos(2\pi D/L)\ +\ \beta_2 \sin(2\pi D/L)\ + \\ \beta_3 \cos(4\pi D/L)\ +\ \beta_4 \sin(4\pi D/L)\ +\ \beta_5 \log(SM)\ + \epsilon, \tag{6.1}$$

where α is the intercept, β_i are the slope coefficients, D = September-based DOY, L is the length of year in days, and ϵ is the error, is shown along with observations in **Fig. 6.7**. The climatological TF is also displayed for reference. Overall, the prediction model is quite skillful relative to climatology (R^2=0.87 vs. 0.57). The rapid decline of LFM during episode #1 is very well captured, as are episode #4's early ramp-up, #2's delayed decline, and, especially, the dramatic drought of 2007 (episode #6). The sharp LFM peak during episode #3 is significantly underpredicted but the autumnal increase and springtime decline are well captured. The model does better with #5's sharp peak, and the periods marked #7 and 8 are also rather well handled. Scatterplots such as **Fig. 6.8** exaggerate the lack of fit (due to missing dimensions, such as frequency and time), but are useful for checking for heteroscedasticity. Results of the *White* [1980] test indicated heteroscedasticity is not a concern, which is due to the use of log transformations.

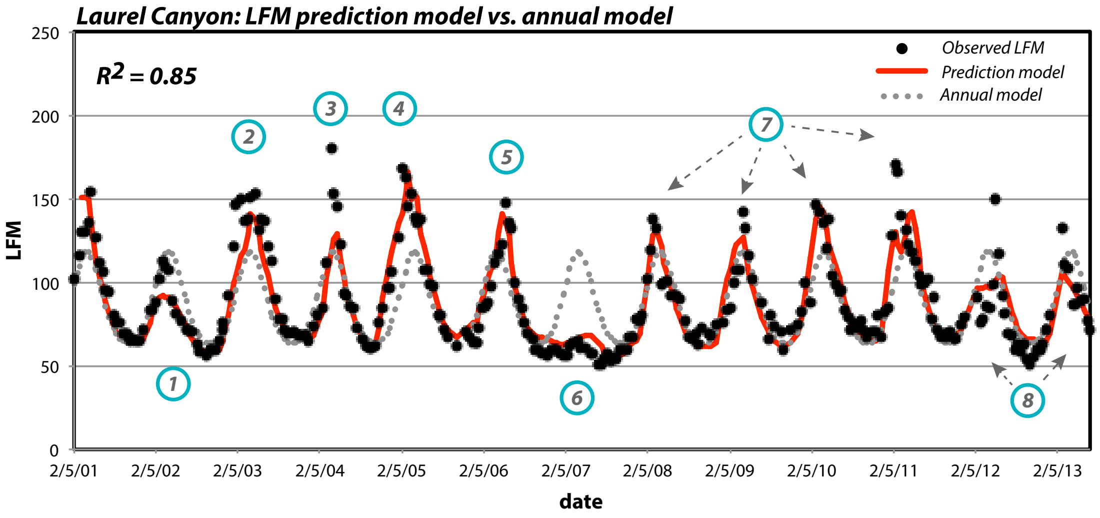


Figure 6.7: The time series of the LFM prediction model (red curve) vs. observations (black dots), with R^2 =0.85. The climatological time function (annual model, grey dotted line) is displayed for reference. Circled numbers are special episodes referred in the text.

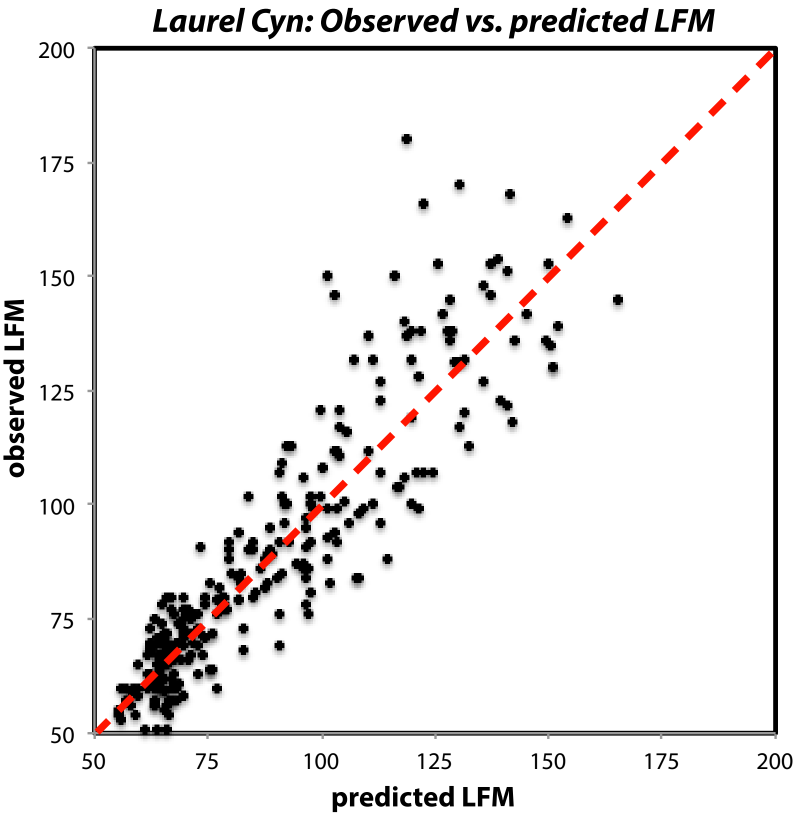


Figure 6.8: Scatterplot of the observed LFM vs. the predicted LFM at site Laurel Canyon.

6.1.5 Other sites

Table 6.1 reveals that our present SM and TF-based strategy worked best for Laurel Canyon. Other stations with model $R^2 \geq 0.80$ include La Tuna Canyon, Schueren Road (before January 2012) and San Marcos (**Figs. 6.9 c, g and h**). The poorest model was found for Reyes Creek, which has the most inland location. Overall, we believe the skillfulness of the fit at each station is high enough to provide useful information for operations, particularly for the timing of LFM up- and down-ramps.

There is a strong similarity among these time series, because they all share a fundamentally similar annual cycle. However, it is worth examining more closely how LFM and the prediction models behaved during the 2007 drought. As previously seen, LFM at LAU hardly changed during this period (**Fig. 6.9 a**), a behavior reproduced by the prediction model. A similar phenomenon is seen at Bitter Canyon and Reyes Creek (**Figs. 6.9 f and j**).

At other inland locations, such as PLA, LAT, PEA, and GLE, LFM values did manage to increase somewhat above the summertime minima, exceeding the Dennison-Moritz threshold in some cases. At the coastal sites, SCH and especially SAN, the year of 2007 did not appear particularly unusual. Their soil moistures were able to recover during the nominal wet season, consistent with their relatively higher fuel moisture readings. Taken together, these results indicate that spatial variations in LFM during the drought were associated with subtle differences in local soil moisture conditions, and these were faithfully captured in the NLDAS reanalysis.

6.1.6 Lag and location issues

For simplicity, all of the fits shown in **Fig. 6.10** and **Table 6.1** employed a 29-day lag with SM. This time interval proved optimal for Laurel Canyon, as seen in **Fig. 6.10 a**, which presents the partial correlation between the residual series r_L and r_S. This plot was constructed from daily SM data along with spline-interpolated LFM information.

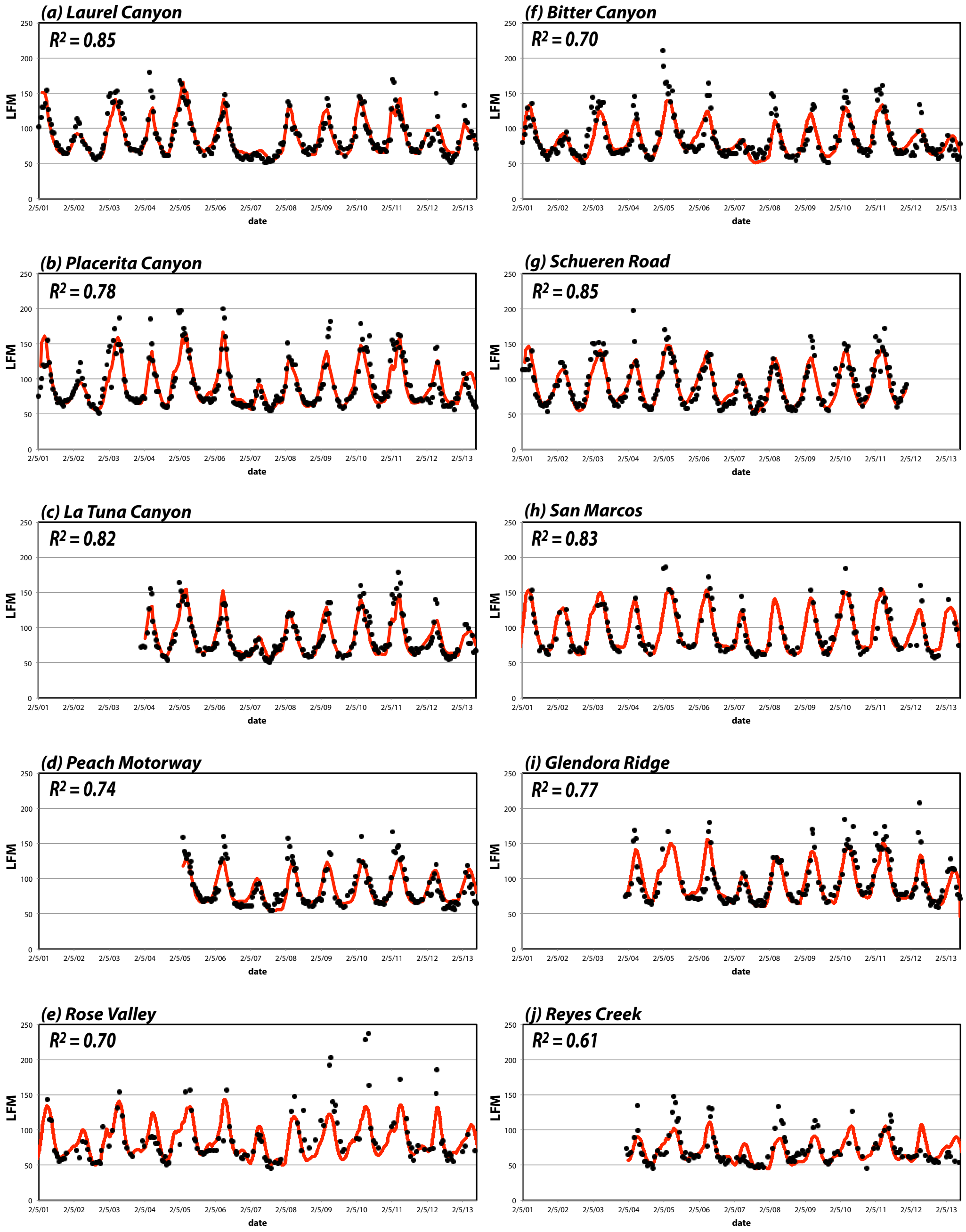


Figure 6.9: Time series of the predicted (red curve) and the observed (black dots) LFM of stations (a) Laurel Canyon; (b) Placerita Canyon; (c) La Tuna Canyon; (d) Peach Motorway; (e) Rose Valley; (f) Bitter Canyon; (g) Schueren Road; (h) San Marcos; (i) Glendora Ridge; and (j) Reyes Creek. The R^2s are labeled on each figure respectively.

Especially owing to the smoothing applied to the SM data, the variation with lag is relatively small, making the result less sensitive to the precise time interval employed.

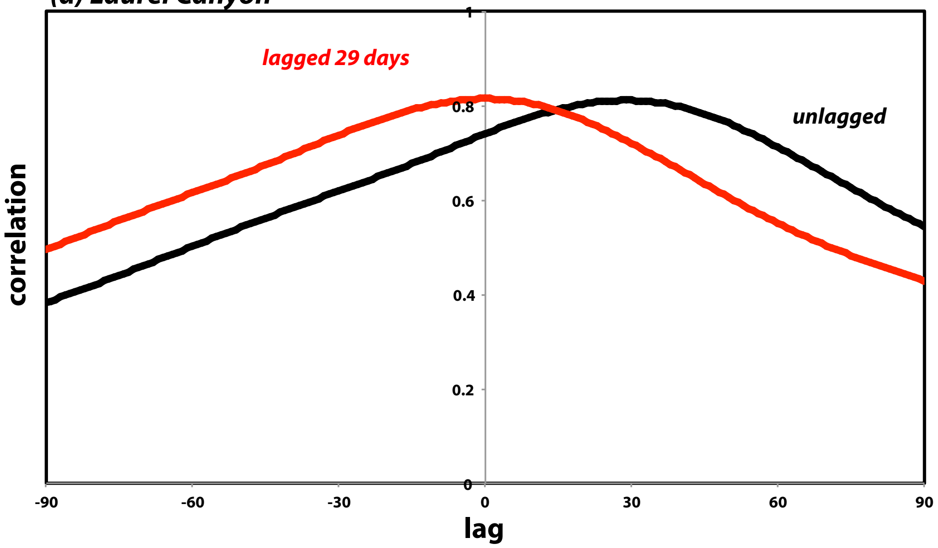
The optimal lag, however, is found to vary among the sites, ranging from 0 days at Peach Motorway to 38 days at Glendora Ridge (**Table 6.1**). **Figure 6.10 b** suggests there are three distinct behaviors relative to LAU exhibited: PLA, GLE and LAT, located in the eastern part of the study area, tend to have longer lags between SM and LFM, and the correlation drops off quickly for negative lags (LFM leading SM); SAN and ROS in the western part of the study area tend to have somewhat shorter lags; and SCH and BIT, located in between, had very short lags with a slow drop-off of correlation for negative values. The LFM-SM time lag also tended to increase with elevation (**Fig. 6.11**), separating into two distinct groupings. The more eastern stations, those in and near the major urban area and the region of sandy loam soils (**Fig. 6.1 c**) have longer lags than the remaining stations.

Note for positive lags (SM leading LFM), no station has a partial correlation higher than Laurel Canyon. As a consequence, perhaps this site should not determine the common network lag. However, we could certainly employ a different lag for each site, but in most cases the skill improvement is minor, with the possible exception of two locations.

The extremely very short LFM-SM lags for Peach Motorway and Bitter Canyon emerge as anomalies, and may represent issues associated with the NLDAS reanalysis'coarse resolution and/or the representativeness of NLDAS soil conditions for those sites. On the NLDAS grid, Peach is located very near a locally dry area (**Fig. 6.1 b**) representing the urbanized area of Santa Clarita, which is (likely inappropriately) impacting the Peach site. Bitter Canyon resides very close to a large reservoir, Castaic Lake, although that feature is not represented in the NLDAS land mask at all. Yet, using nearby Placerita Canyon's soil moisture information in the Peach and Bitter Canyon models result in improved fits at both locations (R^2 = 0.82 vs. 0.74 at Peach, and R^2 = 0.74 vs. 0.64 at Bitter, labeled as "mod" in **Table 6.1**), with lag times closer to the

Partial correlation between LFM and soil moisture vs. lag

(a) Laurel Canyon



(b) Additional stations

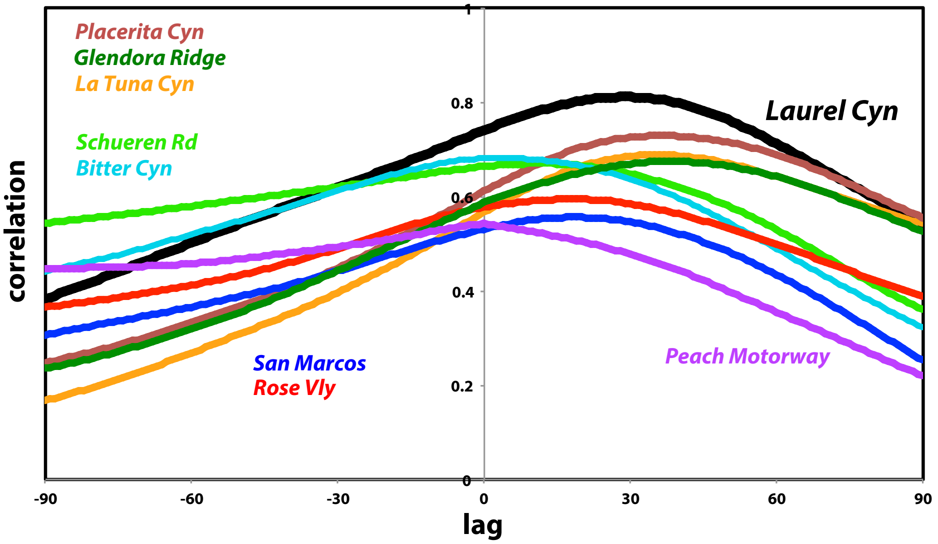


Figure 6.10: (a) Partial correlation between LFM and soil moisture vs. the time lag applied to the soil moisture for site Laurel Canyon (black), including correlation curve after 29 day shift (red); and (b) same as (a), but unshifted partial correlations for additional sites Placerita Canyon, Glendora Ridge, La Tuna Canyon, Schueren Road, Bitter Canyon, San Marcos, Rose Valley, and Peach Motorway.

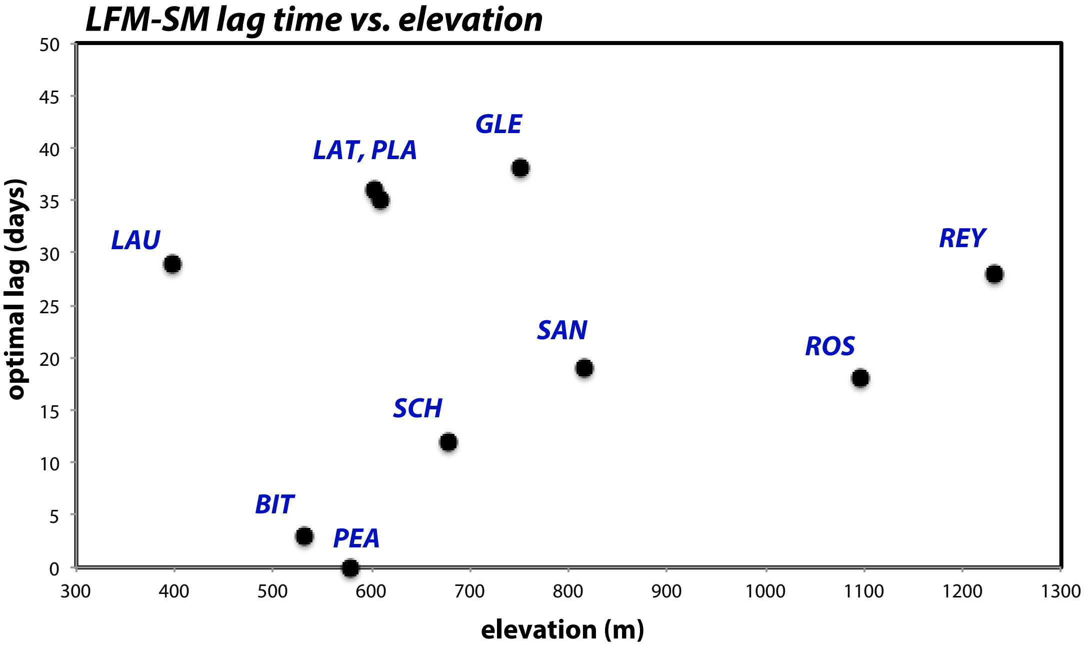


Figure 6.11: The optimal LFM-soil moisture lag (days) vs. the site elevation (m).

multi-site average. To a large degree, using Placerita SM means essentially adopting that site's soil characteristics in place of their NLDAS assignments.

It might be wondered how the Laurel Canyon fit would fare if soil moisture from another location were used in the prediction model. **Figure 6.6 c** replaces the site's SM information with that from a grid point located to the northwest, in the region of soil texture class 6 (loam; **Fig. 6.1 c**) that occupies a large (relatively unsampled) portion of the study area. While the gross features, representing the regional weather variations, have not changed much, the fit is clearly inferior (**Fig. 6.6 f**), even at the (shorter, 12 day) lag that maximizes the partial correlation. Finally, it might also be wondered how unsmoothed soil moisture might fare in these prediction models. At every site, the use of unsmoothed SM information results in diminished skill, mainly because SM acquires high frequency spikes that do not match up well with the noisy LFM data. This creates more of an appearance than a reality of reduced skill.

In conclusion, we feel we have outlined a viable strategy for reconstructing and (owing to the lag relationship between LFM and SM) making short-term forecasts for LFM in the mountainous region northwest of Los Angeles. Our approach is simple, and utilizes a gridded soil moisture product. Still, it is noted that it would be very useful to obtain a single regression equation to predict LFM from a time function and soil moisture data. Such an equation would have to incorporate elevation effects seen in **Figs. 6.2 b, 6.3 a** and **6.4**. Work on this aspect of the problem is ongoing.

6.2 Modeling greenness via NDVI

Following the onset of significant wetting rains, new grasses will begin to emerge in a process called "green-up". While the timing and duration of this process fluctuates from year to year, some degree of green-up usually occurs by December across southern California. During the green-up phase, grasses will begin to act as a heat sink, thereby preventing new ignitions and or significantly reducing the rate of spread among new fires. By late spring these grasses begin to cure with the curing phase normally completed by mid-June (Tom Rolinski, personal communication).

In practice, the greenness (G) part of FMC can be derived from satellite data, specifically the MODIS's Normalized Difference Vegetation Index (NDVI) product. The resolution of these data is at a resolution of 250 meters, and pixels consisting solely of grasslands can be identified and isolated. NDVI is defined by red and near-infrared (NIR) bands in the following equation:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \,, \tag{6.2}$$

where ρ_b= reflectance in band b [*Clinton et al.*, 2010]. It can be shown that NDVI values for Southern California grasslands generally range from about 0.25 (±0.05) to 0.75 (±0.05) for an average rainfall year. There is evidence that NDVI is affected by soil color [*Elmore et al.*, 2000], which may explain the NDVI variations (±0.05) seen among the selected Southern California grassland locations.

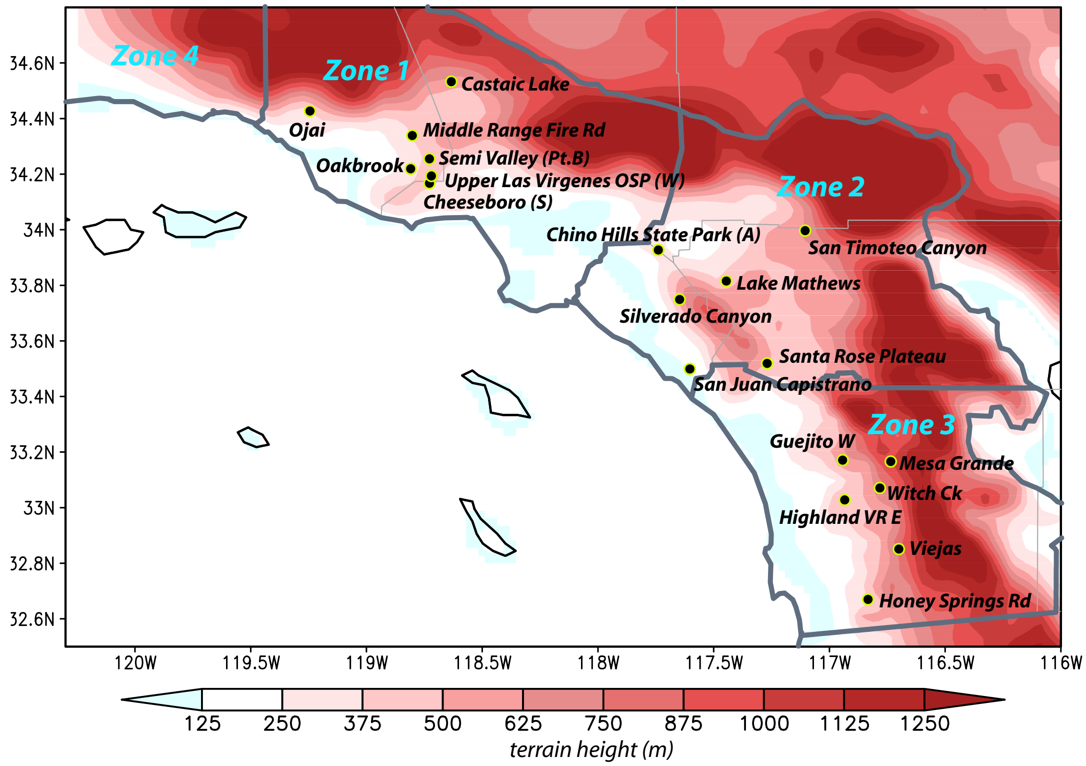


Figure 6.12: NDVI surface station locations (black dots), with underlying topography shaded and the four zone boundaries overlaid.

G is given a rating of 0-5 based on NDVI data, where 0 is green and 5 is fully cured. When applying the *White et al.* [1997] methodology to the general range of Southern California grasslands, green-up is estimated to have occurred when NDVI exceeds 0.50. However, our experience has suggested that this value can be closer to 0.55 for some sites. Therefore, NDVI values greater than or equal to 0.55 are assigned a value of 0, or green. Furthermore, NDVI values less than 0.35 are assigned a value of 5. This is because NDVI values are observed to be below 0.35 for all grassland sites during the dry season when grasses are known to be fully cured. A systematic relationship is suspected to exist between NDVI for Southern California grasslands and fire occurrence (Tom Rolinski, personal communication). For this reason, the transition between green and fully cured (or vice versa) grasses was given a rating of 1 to 4 in NDVI increments of 0.05 (**Table 6.2**).

Table 6.2: Greenness (G) values and their associated NDVI ranges.

NDVI	G
< 0.35	5 (Fully Cured)
$0.35 \leq \text{NDVI} < 0.40$	4
$0.40 \leq \text{NDVI} < 0.45$	3
$0.45 \leq \text{NDVI} < 0.50$	2
$0.50 \leq \text{NDVI} < 0.55$	1
$\text{NDVI} \geq 0.55$	0 (Green)

As with LFM, the literature on predicting NDVI is thin, and provides little guidance on how to model it using meteorological information. However, LFM values were only available for a relatively small number of locations, with the best sampling in the northern part of the Los Angeles basin we called Zone 1 (**Chapter 1**). In contrast, as a satellite-derived product, the spatial resolution of MODIS-derived NDVI information is much better. With more comprehensive coverage of the study area, we decided to develop a single prediction equation that could be applied to gridded model outputs as a diagnostic field. The analysis starts with biweekly NDVI data observed at the 19 stations shown in **Fig. 6.12**, interpolated to daily frequency using cubic splines. The data availability period was January 2004-June 2012.

Potential regressors we considered include: Julian day of the year, the sine and cosine functions of the Julian day; the accumulated summation of relative humidity (%) at 2 m AGL, surface precipitation (mm), and surface shortwave radiation flux (W/m^2) over 7-, 14-, and 30- day periods; the running average of 2 m relative humidity (%), precipitation (mm), and shortwave radiation flux (W/m^2) over 7-, 14-, and 30- day periods; annually accumulated precipitation (mm) since Sep 1st; the surface vegetation fraction (0-1); and the soil moisture (kg/m^3) and temperature (F) of various depths (0-10 cm, 10-40 cm, 40-100 cm, and 100-200 cm). As in the LFM model, our year dates start from September 1st for convenience, and annually accumulated values commence from that date. The surface and 2 meter variables aforementioned were retrieved from a 6 km resolution, 30-year historical WRF simulations initialized with the NNRP dataset [*Kistler et al.*, 2001]. For our soil moisture data source, we again used the NLDAS reanalysis product, although as mentioned previously, the coastal NLDAS soil moisture information near the coastline experienced some issue subsequent to December 2011. All the regressors have the same spatial length as the NDVI observations (January 2004-June 2012), and to get a universal equation over all the gridpoints in our domain, we merged the 19 stations, making both the observations and regressor data to be 58425 data points in length.

As for modeling LFM, our goal is to create the simplest skillful equation to capture the temporal variation of NDVI, both for ease of implementation and to avoid overfitting. First, we applied the "random forest" [*Breiman*, 2001] selection method to tag regressors with respect to their importance in explaining the NDVI variance. This was implemented by the "TreeBagger" package in MATLAB's statistics toolbox. From this process, the six most important regressors were found to be: the cosine function of the Julian day, annually accumulated precipitation, the 30-day running average of relative humidity, the surface vegetation fraction, the soil moisture of the 40-100cm depth, and the 14-day running average of precipitation.

Next, stepwise selection [*Efroymson*, 1960] was employed to further refine the selected independent variables as well as linearly regress the model. Stepwise selection is a one-step-at-a-time approach, and is based on t-tests of the individual parameters, which are known to be severely affected by multicollinearity (i.e., highly correlated predictors). During this process, the 14-day running average of precipitation regressor was removed from the model. With five independent variables retained (see **Table 6.3** for values), our prediction model for NDVI is given by

$$\begin{aligned} NDVI = \alpha \ + \ \beta_1 \cos(2\pi D/L) \ + \ \beta_2 \, annually\_accurain \ + \ \beta_3 \, 30day\_ave\_RH \ + \\ \beta_4 \, vegsfc + \ \beta_5 \, soilmo40\_100cm, \end{aligned} \tag{6.3}$$

where D = Julian Day starting at January 1^{st} and L is the length of year in days. The regressor “annually_accurain” is the September 1^{st} -based annually accumulated precipitation, “30day_ave_RH” is the 30-day running averaged to relative humidity, “vegsfc” is the surface vegetation fraction (0-1), and “soilm40_100cm” is the soil moisture content of the 40-100cm depth. The R^2 of the model is 0.73.

Table 6.3: Selected NDVI regressors.

Variable Number	Variable Name	Coefficients		Importance
		α	β	
1	cos(2π D/L)		0.11253592	11.5
2	year_accurain		1.44E-05	9.0
3	running_RH_30day	-0.314867	0.00355647	6.2
4	surface vegetation fraction (%)		0.911360168	5.9
5	soilmo40-100cm		0.002412815	5.5

We applied this model to the 19 sites in the three zones shown in **Fig. 6.12**. Observed NDVI (black dots) data are compared to the predicted NDVI (blue curve) values

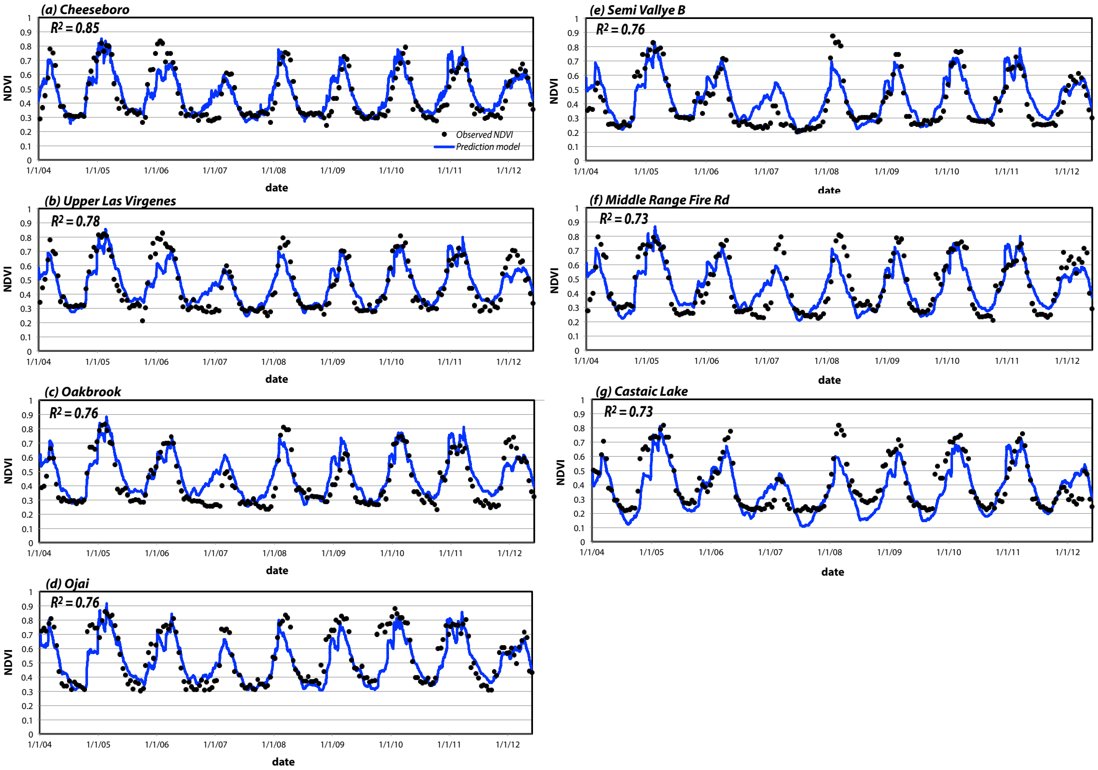


Figure 6.13: The observed (black dots) and predicted (blue contours) NDVI for stations in Zone 1.

from January 2004 through June 2012 (**Figs. 6.13** through **6.15**). The total variance captured in the model, R^2, being 0.79±0.06 for Zone 1 (**Fig. 6.13**), 0.69±0.1 for Zone 2 (**Fig. 6.14**), and 0.79±0.05 for Zone 3 (**Fig. 6.15**).

It is recognized that at some stations and times, the NDVI predictions are somewhat out of phase (i.e., the up and down ramps are too early or too late) with the observations (**Figs. 6.13 b, c, f , 6.14 a, c, f, 6.15 b, d, and f**), and the peaks are over- or underpredicted (**Figs. 6.13 e, f, g, 6.14 a, b, f, 6.15 a, c, and f**) at different locations and times. The marked drought year of 2007 is clearly a problem at some locations, especially in Zone 2, which accounts for that zone's relatively lower R^2 values. However, considering the fact that this is a universal and rather simple model developed for every model grid point in the Southern California area that only consists of five independent variables, we believe it has shown adequate skill overall.

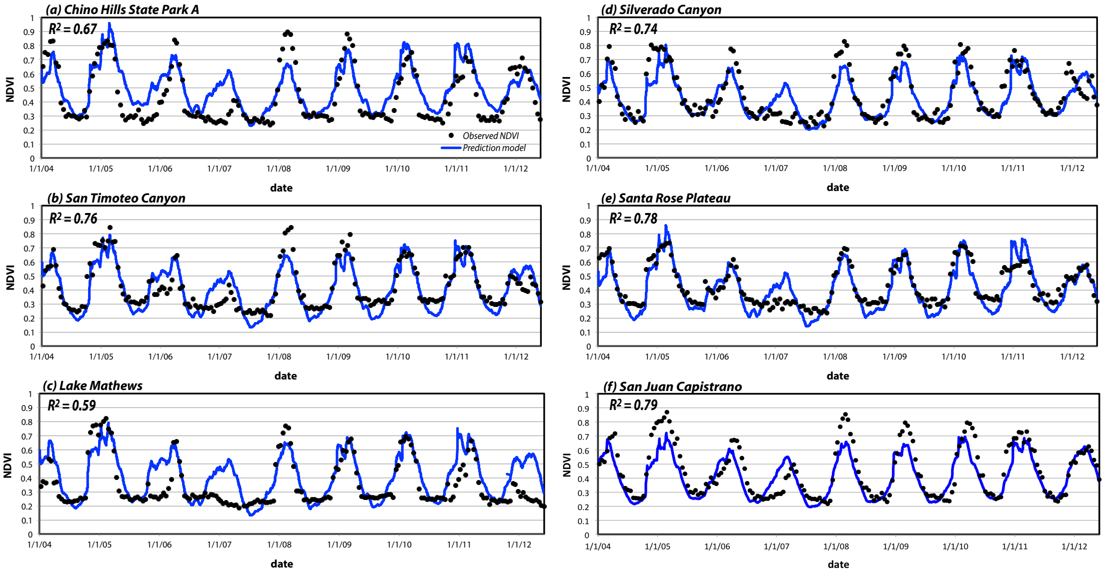


Figure 6.14: Same as **Fig. 6.13**, but for stations in Zone 2.

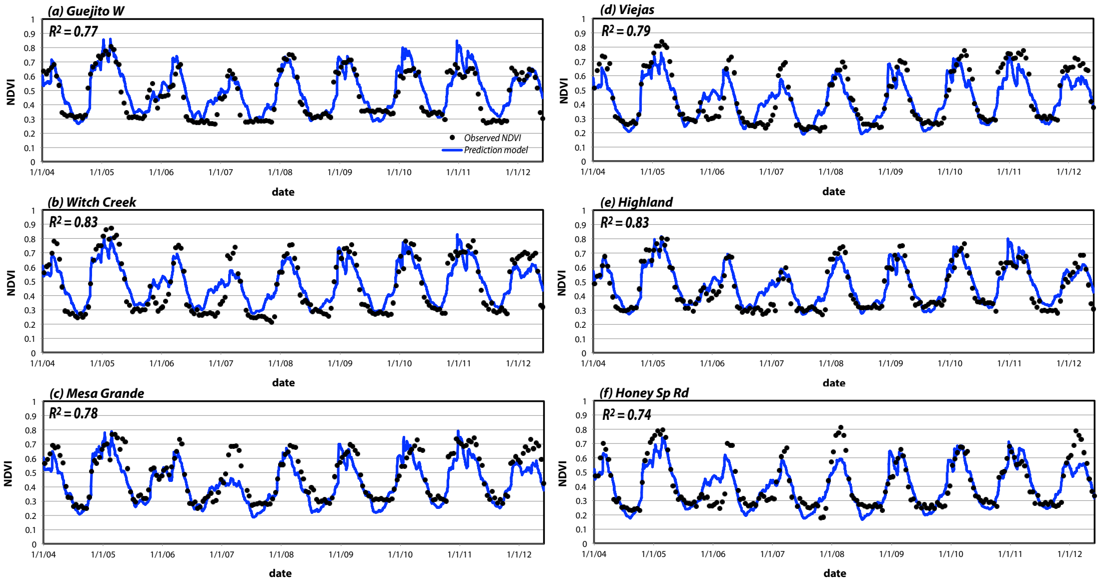


Figure 6.15: Same as **Fig. 6.13**, but for stations in Zone 3.

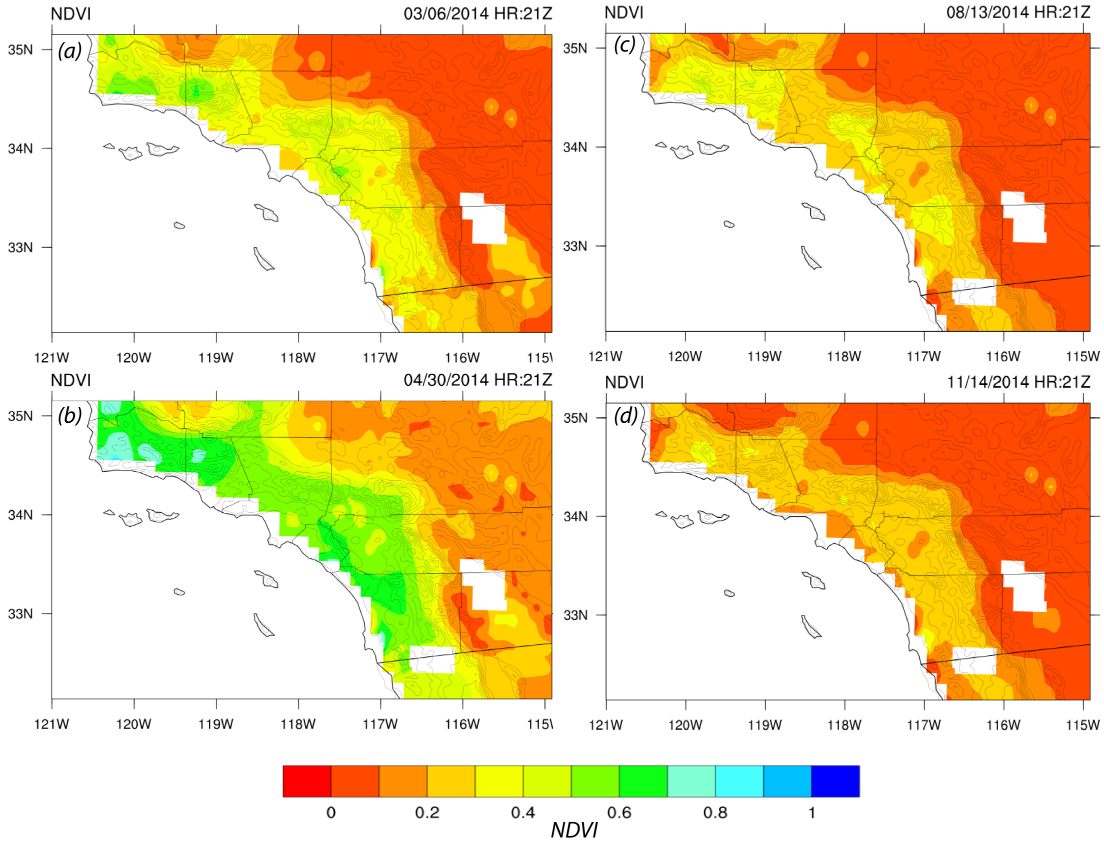


Figure 6.16: Predicted NDVI over the Southern California area at 2100UTC on (a) 6 March 2014; (b) 30 April 2014; (c) 13 August 2014; and (d) 14 November 2014. Courtesy of Dr. Scott Capps.

Currently, our NDVI model is being used operationally by the US Forest Service's Predictive Services unit, using model fields from daily 3 km WRF simulations initialized with the NAM Forecast System dataset. As shown in **Fig. 6.16**, the model developed is able to skillfully capture the annual NDVI variation such as the green-up (**Fig. 6.16 a**) and curing (**Fig. 6.16 c**) of annual grasses. For purposes of computing LFP, the NDVI model predictions are converted to G values using the thresholds presented in **Table 6.2** .

6.3 Summary

Plant moisture content plays an important role in determining the availability of natural fuels for wildfires in Mediterranean ecosystems. While dead fuel moisture variation mainly depends exclusively on the weather conditions, live herbaceous moisture content is relatively more difficult to predict due to its seasonality in response to plants' physiological and phenological processes such as spring flushing and fall curing as well as soil water availability. In this chapter, we have pursued a strategy to use a readily available, gridded soil moisture product, i.e., the NLDAS reanalysis dataset, to skillfully reconstruct the temporal variation of the historical LFM and NDVI observations, and to create the simplest models to predict the two essential factors that heavily influence SAWTI.

The LFM prediction model for each site we developed utilizes a time function in combination with soil moisture information lagged by 29 days, which was used to capture departures from the site's annual LFM cycle. The prediction model is simple yet skillfully predictive. The spatial variations in LFM during the drought were proven to be associated with subtle differences in local soil moisture conditions, which were faithfully captured in the NLDAS reanalysis. Work on obtaining a single regression equation to predict LFM based on a time function and soil moisture is ongoing.

Since more comprehensive coverage and higher resolution observation data are readily available for NDVI, we elected to develop a single prediction equation that could be easily implemented operationally. The model was obtained by performing the "random forest" and the "stepwise" selection methods, using meteorological data from a high-resolution WRF historical reconstructions and soil information form the NLDAS reanalysis dataset. The simple single NDVI model with merely five linear regressors has shown adequate skill in operationally forecast the NDVI annual variation.

The improved live fuel moisture model and NDVI model can work with other components of fuel moisture to help monitor fuel inflammability over southern California regions for fire danger assessment.

CHAPTER 7

Conclusions

This dissertation presents research work on forecasting the Santa Ana winds and gusts of Southern California and modeling two essential fuel moisture elements, i.e., the live fuel moisture and the greenness of grass, in the context of fire weather. With an introduction of the Santa Ana winds, the Santa Ana Wildfire Threat Index, the Ensemble Forecasting technique given in Chapter 1 and the description of the recently-installed, very high-density surface SDG&E mesonet, and the modeling and validation strategies in Chapter 2, the main results of this dissertation are summarized as follows:

In Chapter 3, a moderately strong and two-phase Santa Ana wind episode – the 14-16 February 2013 event – was examined, employing high-resolution WRF model simulations and utilizing the SDG&E observations for validation. The painstakingly selected model physics combination PX/ACM2 helped confirm that the downslope windstorm had different features during two phases, i.e., the hydraulic jump characteristic observed from the network in the first phase, and a normal downslope progression of winds during the second phase. The model simulations enabled us to view the downslope storm in a three-dimension perspective, which filled in gaps in the horizontal observation, and more importantly, provided the vertical structure of the wind field. In the vertical cross-section profiles, we have seen the evolution of the hydraulic jumps, and understood why winds can vary dramatically in speed and direction in a short distance. During the afternoon lull, the strong winds were not completely weakened, but resided near the mountain ridge, where no observations were available.

In Chapter 4, we further investigated two other Santa Ana episodes, i.e., the 4-6

October 2013 and the 13-15 May 2014 events, and the latter being a major Santa Ana event that had sparked several fires. We employed large physics ensembles composed of high-resolution simulations that involved an exhaustive examination of available model physical parameterizations. Model results are calibrated and validated against the SDG&E mesonet observations.

The ensembles demonstrated the wind speeds and structures are sensitive to model horizontal grid spacing, random perturbations, and especially, to the land surface schemes that determine the surface roughnesses employed in the simulations. The surface roughness length is important for winds because it determined how fast the near-surface winds were, how far downslope the offshore flow could remain close to the ground, and more importantly, the nature of the downslope flow. The PX LSM emerged as the best choice, owing to its unique treatment of surface roughness in the dominant shrubland landuse categories in the San Diego territory. Most other LSMs tended to overpredict winds due to their treating the surface is treated as being too smooth.

We further discovered that allowing other LSMs to use PX's roughness lengths improved their wind forecast skill to a large extent. A Santa Ana episode that possesses a jump-like feature, such as the February 2013 event's first phase, tends to be more sensitive to random perturbations than others due to its more turbulent nature. The "topo wind" approach was not proven to be superior to a more careful and accurate assignment of surface roughness values. The landuse database may matter. Model sensitivity to horizontal diffusion was found to be low.

Since mesoscale models are incapable of directly simulating the hazardous gusts, empirical and heuristic attempts have been made to estimate wind gusts by multiplying the resolved-scale sustained wind speed by a gust factor (a gust-to-wind ratio) empirically determined from available observations, or adding a temporally- and spatially-varying scalar value to the sustained wind, assuming a normal distribution of wind fluctuations. In Chapter 5, we have examined the ability of WRF model to provide skillful gust predictions and devised a simple gust parameterization that can skillfully anticipate extreme wind events in the complex topography such as San Diego county. Our simple gust factor approach outperformed the NCEP and ECMWF techniques via comparison. The improved gust forecast technique is of special interest in the context of routine weather combined with atmospheric humidity and fuel moisture information. However, whether the constant network gust factor depends on the homogeneity of the network and whether it is applicable outside of the SDG&E mesonet is unclear.

In Chapter 6, efforts were described to accurately predict live herbaceous moisture content as well as greenness annual and seasonally variability in southern California areas where large wildfires often occur when the fuels are dry and the Santa Ana winds are strong. Our results indicate a strong correlation between live fuel moisture of the new growth Chamise and soil moisture of 40-100cm depth layer with a lag ranging between 0-38 days depending on station elevations. A time function in combination with soil moisture information lagged by 29 days was shown to improve the prediction of the live fuel moisture annual and seasonal variations over a 12-year period for a set of ten stations. The average coefficient of determination (R^2) of the individual station models is as high as 0.77.

Regarding greenness, a single NDVI prediction equation composed of five regressors was obtained, using meteorological data from a high-resolution WRF historical reconstructions and soil information from the NLDAS reanalysis dataset. The average coefficient of determinations of the individual station is 0.79 in Zone1, 0.69 in Zone 2, and 0.79 in Zone 3. The model has shown adequate skill in operationally forecasting the NDVI annual variation and important deviations therefrom.

Bibliography

Ashcroft, J. (1994), The relationship between the gust ratio, terrain roughness, gust duration and the hourly mean wind speed, *Journal of Wind Engineering and Industrial Aerodynamics*, *53*(3), 331–355, doi:10.1016/0167-6105(94)90090-6.

Beljaars, A. C. M. (1987), The influence of sampling and filtering on measured wind gusts, *Journal of Atmospheric and Oceanic Technology*, *4*, 613–626, doi:10.1175/1520-0426(1987)004<0613:TIOSAF>2.0.CO;2.

Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer (2009), A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF Ensemble Prediction System, *Journal of Atmospheric Sciences*, *66*, 603–626, doi:10.1175/2008JAS2677.1.

Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder (2011), Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations, *Monthly Weather Review*, *139*, 1972–1995, doi:10.1175/2010MWR3595.1.

Bosart, L. F. (1975), SUNYA experimental results in forecasting daily temperature and precipitation, *Monthly Weather Review*, *103*, 1013–1020, doi:10.1175/1520-0493(1975)103<1013:SERIFD>2.0.CO;2.

Bradshaw, L. S., J. E. Deeming, R. E. Burgan, and J. D. Cohen (1984), *The 1978 national fire-danger rating system: Technical documentation*, General technical report INT, 44 pp., U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.

Brasseur, O. (2001), Development and application of a physical approach to estimating wind gusts, *Monthly Weather Review*, *129*, 5–25, doi:10.1175/1520-0493(2001)129<0005:DAAOAP>2.0.CO;2.

Breiman, L. (2001), Random forests, *Machine Learning*, *45*, 5–32, doi: 10.1023/A:1010933404324.

Buizza, R., M. Miller, and T. N. Palmer (1999), Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, *Quarterly Journal of the Royal Meteorological Society*, *125*, 2887–2908, doi:10.1002/qj.49712556006.

Buizza, R., P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu (2005), A comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems, *Monthly Weather Review*, *133*, 1076–1097, doi:10.1175/MWR2905.1.

Burton, T., N. Jenkins, D. Sharpe, and E. Bossanyi (2011), *Wind energy handbook, 2^{nd} edition*, 775 pp., John Wiley & Sons, Ltd.

Carlson, J. D., L. S. Bradshaw, R. M. Nelson Jr., R. R. Bensch, and R. Jabrzemski (2007), Application of the Nelson model to four timelag fuel classes using Oklahoma field observations: Model evaluation and comparison with National Fire Danger Rating System algorithms, *International Journal of Wildland Fire*, *16*, 204–216, doi:/10.1071/WF06073.

Chang, C.-H., and F. P. Schoenberg (2011), Testing separability in marked multidimensional point processes with covariates, *Annals of the Institute of Statistical Mathematics*, *63*(6), 1103–1122, doi:10.1007/s10463-010-0284-7.

Clinton, N. E., C. Potter, B. Crabtree, V. Genovese, P. Gross, and P. Gong (2010), Remote sensing–based time-series analysis of cheatgrass (Bromus Tectorum L.) phenology, *Journal of Environmental Quality*, *39*(3), 955–963, doi: 10.2134/jeq2009.0158.

Countryman, C., and W. Dean (1979), *Measuring moisture content in living Chaparral: A field user's manual*, General technical report PSW, 27 pp., U.S. Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station.

Davis, F. W., and J. Michaelsen (1995), Sensitivity of fire regime in chaparral ecosystems to climate change, in *Global Change and Mediterranean-type Ecosystems*, edited by J. M. Moreno and W. C. Oechel, pp. 435–456, Springer-Verlag, New York, NY.

Dennison, P. E., and M. A. Moritz (2009), Critical live fuel moisture in chaparral ecosystems: A threshold for fire activity and its relationship to antecedent precipitation, *International Journal of Wildland Fire*, *18*(8), 1021–1027, doi:10.1071/WF08055.

Dennison, P. E., and D. A. Roberts (2003), The effects of vegetation phenology on endmember selection and species mapping in Southern California chaparral, *Remote Sensing of Environment*, *87*, 295–309, doi:10.1175/MWR2905.1.

Doyle, J. D., and D. R. Durran (2004), NOWCAST: THE MAP ROOM: Recent developments in the theory of atmospheric rotors, *Bulletin of the American Meteorological Society*, *85*, 337–342, doi:10.1175/BAMS-85-3-337.

Durran, D. R. (1986), Another look at downslope windstorms. Part I: The development of analogs to supercritical flow in an infinitely deep, continuously stratified fluid., *Journal of Atmospheric Sciences*, *43*, 2527–2527, doi:10.1175/1520-0469(1986)043<2527:ALADWP>2.0.CO;2.

Durran, D. R. (2003), Downslope winds, in *Encyclopedia of Atmospheric Sciences*, edited by G. North, J. Pyle, and F. Zhang, pp. 644–650, Elsevier Science.

Efroymson, M. A. (1960), Multiple regression analysis, in *Mathematical methods for digital computers*, edited by A. Ralston and H. S. Wilf, 1, pp. 191–203, John Wiley & Sons, Ltd, New York, NY.

Elmore, A. J., J. F. Mustard, S. J. Manning, and D. B. Lobell (2000), Quantifying vegetation change in semiarid environments: Precision and accuracy of spectral mixture analysis and the normalized difference vegetation index, *Remote Sensing of Environment*, *73*(1), 87–102, doi:10.1016/S0034-4257(00)00100-0.

Epstein, E. S. (1969), Stochastic dynamic prediction, *Tellus*, *21*(6), 739–759, doi: 10.1111/j.2153-3490.1969.tb00483.x.

Fiorucci, P., F. Gaetani, A. Lanorte, and R. Lasaponara (2007), Dynamic fire danger mapping from satellite imagery and meteorological forecast data, *Earth Interactions*, *11*(7), 1–17, doi:10.1175/EI199.1.

Gyakum, J. R. (1986), Experiments in temperature and precipitation forecasting for Illinois, *Weather and Forecasting*, *1*, 77–88, doi:10.1175/1520-0434(1986)001<0077:EITAPF>2.0.CO;2.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: Basic concept, *Tellus*, *57*, 219–233, doi:10.1111/j.1600-0870.2005.00103.x.

Hamill, T. M., and C. Snyder (2000), A hybrid ensemble Kalman Filter 3D variational analysis scheme, *Monthly Weather Review*, *128*, 2905–2919, doi:10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Hamill, T. M., S. L. Mullen, C. Snyder, Z. Toth, and D. P. Baumhefner (2000), Ensemble forecasting in the short to medium range: Report from a workshop, *Bulletin of the American Meteorological Society*, *81*, 2653–2664, doi:10.1175/1520-0477(2000)081<2653:EFITST>2.3.CO;2.

Henson, P., D. Usner, and V. Kells (1996), *The natural history of Big Sur*, California Natural History Guides, University of California Press.

Huang, C., Y. L. Lin, M. L. Kaplan, and J. J. Charney (2009), Synoptic-scale and mesoscale environments conducive to forest fires during the October 2003 extreme fire event in Southern California, *Journal of Applied Meteorology and Climatology*, *48*(3), 553–579, doi:10.1175/2008JAMC1818.1.

Hughes, M., and A. Hall (2010), Local and synoptic mechanisms causing Southern California's Santa Ana winds, *Climate Dynamics*, *34*(6), 847–857, doi:10.1007/s00382-009-0650-4.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *Journal of Geophysical Research: Atmospheres*, *113*(D13103), doi:10.1029/2008JD009944.

Jackson, P. L., G. Mayr, and S. Vosper (2013), Dynamically-driven winds, in *Mountain weather research and forecasting*, edited by F. K. Chow, S. F. J. D. Wekker, and B. J. Snyder, Springer Atmospheric Sciences, pp. 121–218, Springer-Verlag, New York, NY.

Janjić, Z. I. (1994), The Step-Mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer, and turbulence closure schemes, *Monthly Weather Review*, *122*, 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Jiménez, P. A., and J. Dudhia (2012), Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model, *Journal of Applied Meteorology and Climatology*, *51*, 300–316, doi:10.1175/JAMC-D-11-084.1.

Kalnay, E. (2003), *Atmospheric modeling, data assimilation and predictability*, 364 pp., Cambridge University Press.

Keetch, J., and G. Byram (1968), *A drought index for forest fire control*, Research Paper SE, 35 pp., U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station.

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino (2001), The NCEP-NCAR 50-Year reanalysis: Monthly means CD-ROM and documentation., *Bulletin of the American Meteorological Society*, *82*, 247–268, doi: 10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Kożuchowski, K., and E. Żmudzka (2002), The warming in poland: the range and the seasonality of changes in air temperature during the second half of the 20^{th} century, *Miscellanea Geographica*, *10*, 103–111.

Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. Larow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran (2000), Multimodel ensemble forecasts for weather and seasonal climate, *Journal of Climate*, *13*, 4196–4216, doi:10.1175/1520-0442(2000)013<4196:MEFFWA>2.0.CO;2.

Lambers, H., F. Chapin, and T. Pons (2008), *Plant physiological ecology*, 605 pp., Springer-Verlag, New York, NY.

López, A. S., J. San-Miguel-Ayanz, and R. E. Burgan (2002), Integration of satellite sensor data, fuel type maps and meteorological observations for evaluation of forest fire risk at the pan-European scale, *International Journal of Remote Sensing*, *23*, 2713–2719, doi:10.1080/01431160110107761.

Lumley, J. L., and H. A. Panofsky (1964), *The structure of atmospheric turbulence*, no. 12 in Interscience Monographs and Texts in Physics and Astronomy, 239 pp., Interscience Publishers.

Mason, P. J., and D. J. Thomson (1992), Stochastic backscatter in large-eddy simulations of boundary layers, *Journal of Fluid Mechanics*, *242*, 51–78, doi: 10.1017/S0022112092002271.

Mitchell, K. E., D. Lohmann, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley,

D. P. Lettenmaier, C. H. Marshall, J. K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B. H. Ramsay, and A. A. Bailey (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *Journal of Geophysical Research: Atmospheres*, *109*, D07S90, doi:10.1029/2003JD003823.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis (1996), The ECMWF Ensemble Prediction System: Methodology and validation, *Quarterly Journal of the Royal Meteorological Society*, *122*(529), 73–119, doi:10.1002/qj.49712252905.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, *430*, 768–772, doi:10.1038/nature02771.

Nelson Jr, R. M. (2000), Prediction of diurnal change in 10-h fuel stick moisture content, *Canadian Journal of Forest Research*, *30*(7), 1071–1087, doi:10.1139/x00-032.

Oke, T. R. (1987), *Boundary layer climates*, 435 pp., Routledge.

Palmer, T. N., F. Molteni, R. Mureau, R. Buizza, P. Chapelet, and J. Tribbia (1993), Ensemble prediction, in *Proceedings of the ECMWF Seminar on Validation of Models over Europe*, vol. 1, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Panofsky, H. A., H. Tennekes, D. H. Lenschow, and J. C. Wyngaard (1977), The characteristics of turbulent velocity components in the surface layer under convective conditions, *Boundary-Layer Meteorology*, *11*, 355–361, doi:10.1007/BF02186086.

Pleim, J. E. (2007a), A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing, *Journal of Applied Meteorology and Climatology*, *46*, 1383–1395, doi:10.1175/JAM2539.1.

Pleim, J. E. (2007b), A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model, *Journal of Applied Meteorology and Climatology*, *46*, 1396–1409, doi:10.1175/JAM2534.1.

Pleim, J. E., and A. Xiu (1995), Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models., *Journal of Applied Meteorology*, *34*, 16–32, doi:10.1175/1520-0450-34.1.16.

Pollet, J., and A. Brown (2007), *Fuel moisture sampling guide*, 32 pp., U. S. Bureau of Land Management, Utah State Office, Salt Lake City, UT.

Pyne, S., P. Andrews, and R. Laven (1996), *Introduction to wildland fire, 2nd edition*, Wiley-Interscience Publication, 808 pp., John Wiley & Sons, Ltd.

Qi, Y., P. E. Dennison, J. Spencer, and D. Riaño (2012), Monitoring live fuel moisture using soil moisture and remote sensing proxies, *Fire Ecology*, *8*(3), 71–87, doi:10.4996/fireecology.0803071.

Raphael, M. (2003), The Santa Ana winds of California, *Earth Interactions*, *7*(8), 1–13, doi:10.1175/1087-3562(2003)007<0001:TSAWOC>2.0.CO;2.

Rothermel, R. C. (1972), *A mathematical model for predicting fire spread in wildland fuels*, Research Paper INT-115, 50 pp., U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.

Sanders, F. (1963), On subjective probability forecasting., *Journal of Applied Meteorology*, *2*, 191–201, doi:10.1175/1520-0450(1963)002<0191:OSPF>2.0.CO;2.

Sanders, F. (1973), Skill in forecasting daily temperature and precipitation: Some experimental results., *Bulletin of the American Meteorological Society*, *54*, 1171–1178, doi:10.1175/1520-0477(1973)054<1171:SIFDTA>2.0.CO;2.

Schamp, H. (1964), *Die Winde der Erde und ihre Namen*, 94 pp., Franz Steiner Verlag, Wiesbaden, Germany.

Schoenberg, F. P., R. Peng, Z. Huang, and P. Rundel (2003), Detection of non-linearities in the dependence of burn area on fuel age and climatic variables, *International Journal of Wildland Fire*, *12*(1), 1–6, doi:10.1071/WF02053.

Shutts, G. (2005), A kinetic energy backscatter algorithm for use in ensemble prediction systems, *Quarterly Journal of the Royal Meteorological Society*, *131*(612), 3079–3102, doi:10.1256/qj.04.106.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers (2008), A description of the Advanced Research WRF version 3, *Tech. rep.*, NCAR Technical Note, Boulder, CO.

Small, I. J. (1995), *Santa Ana winds and the fire outbreak of fall 1993*, NOAA Technical Memorandum, 30 pp., National Oceanic and Atmospheric Administration, National Weather Service Scientific Services Division, Western Region.

Sommers, W. T. (1978), LFM forecast variables related to Santa Ana wind occurrences, *Monthly Weather Review*, *106*, 1307–1316, doi:10.1175/1520-0493(1978)106<1307:LFVRTS>2.0.CO;2.

Stainforth, D. A., T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Piani, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe, and M. R. Allen (2005), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, *433*, 403–406, doi:10.1038/nature03301.

Steven Tracton, M., and E. Kalnay (1993), Operational ensemble prediction at the National Meteorological Center: Practical aspects, *Weather and Forecasting*, *8*, 379–400, doi:10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2.

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

Toth, Z., and E. Kalnay (1997), Ensemble forecasting at NCEP and the breeding method, *Monthly Weather Review*, *125*, 3297–3319, doi:10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

Verkaik, J. W. (2000), Evaluation of two gustiness models for exposure correction calculations, *Journal of Applied Meteorology*, *39*, 1613–1626, doi:10.1175/1520-0450(2000)039<1613:EOTGMF>2.0.CO;2.

Viegas, D., J. Piñol, M. Viegas, and R. Ogaya (2001), Estimating live fine fuels moisture content using meteorologically-based indices, *International Journal of Wildland Fire*, *10*(2), 223–240, doi:10.1071/WF01022.

Vosper, S. B. (2004), Inversion effects on mountain lee waves, *Quarterly Journal of the Royal Meteorological Society*, *130*, 1723–1748, doi:10.1256/qj.03.63.

Warner, T. T. (2011), *Numerical weather and climate prediction*, 550 pp., Cambridge University Press, New York, NY.

Weise, D. R., R. A. Hartford, and L. Mahaffey (1998), Assessing live fuel moisture for fire management applications, in *Fire in ecosystem management: Shifting the paradigm from suppression to prescription*, *Tall Timbers Fire Ecology Conference 20th Proceedings*, vol. 20, edited by T. L. Pruden and L. A. Brennan, pp. 49–55, Tall Timbers Research Station, Tallahassee, FL.

Westerling, A. L., D. R. Cayan, T. J. Brown, B. L. Hall, and L. G. Riddle (2004), Climate, Santa Ana winds and autumn wildfires in Southern California, *EOS Transactions*, *85*, 289–296, doi:10.1029/2004EO310001.

White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, *48*(4), 817–838, doi:10.2307/1912934.

White, M. A., P. E. Thornton, and S. W. Running (1997), A continental phenology model for monitoring vegetation responses to interannual climatic variability, *Global Biogeochemical Cycles*, *11*, 217–234, doi:10.1029/97GB00330.

Wieringa, J. (1973), Gust factors over open water and built-up country, *Boundary-Layer Meteorology*, *3*(4), 424–441, doi:10.1007/BF01034986.

Wieringa, J. (1976), An objective exposure correction method for average wind speeds measured at a sheltered location, *Quarterly Journal of the Royal Meteorological Society*, *102*(431), 241–253, doi:10.1002/qj.49710243119.

Wieringa, J. (1980), Representativeness of wind observations at airports, *Bulletin of the American Meteorological Society*, *61*, 962–971, doi:10.1175/1520-0477(1980)061<0962:ROWOAA>2.0.CO;2.

WMO (2010), Guide to meteorological instruments and methods of observation. WMO no. 8, Geneva, Switzerland.

Xia, Y., K. Mitchell, M. Ek, B. Cosgrove, J. Sheffield, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, Q. Duan, and D. Lohmann (2012a), Continental-scale water and energy flux analysis and validation for north american land data assimilation system project phase 2 (nldas-2): 2. validation of model-simulated streamflow, *Journal of Geophysical Research: Atmospheres*, *117*(D03110), doi:10.1029/2011JD016051.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, D. Lettenmaier, V. Koren, Q. Duan, K. Mo, Y. Fan, and D. Mocko (2012b), Continental-scale water and energy flux analysis and validation for the north american land data assimilation system project phase 2 (nldas-2): 1. intercomparison and application of model products, *Journal of Geophysical Research: Atmospheres*, *117*(D03109), doi:10.1029/2011JD016048.

Xiu, A., and J. E. Pleim (2001), Development of a land surface model. Part I: Application in a mesoscale meteorological model., *Journal of Applied Meteorology*, *40*, 192–209, doi:10.1175/1520-0450(2001)040<0192:DOALSM>2.0.CO;2.

Yoshino, M. M. (1975), *Climate in a Small Area: An Introduction to Local Meteorology*, 549 pp., University of Tokyo Press, Tokyo.