Downslope Windstorms of San Diego County. Part II: Physics Ensemble Analyses and Gust Forecasting

YANG CAO

Atmospheric Data Solutions, LLC, Santa Ana, California

ROBERT G. FOVELL

Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, New York

(Manuscript received 27 November 2017, in final form 30 January 2018)

ABSTRACT

The "Santa Ana" winds of Southern California represent a high-impact weather event because their dry, fast winds can significantly elevate the wildfire threat. This high-resolution numerical study of six events of moderate or greater strength employs physics parameterization and stochastic perturbation ensembles to determine the optimal model configuration for predicting winds in San Diego County, with verification performed against observations from the San Diego Gas and Electric (SDG&E) mesonet. Results demonstrate model physics can have a material effect on the strength, location, and timing of the winds, with the land surface model playing an outsized role via its specification of surface roughness lengths. Even when bias in the network-averaged sustained wind forecasts is minimized, systematic biases remain in that many stations are consistently over- or underforecasted. The argument is made that this is an "unavoidable" error that represents localized anemometer exposure issues revealed through the station gust factor. A very simple gust parameterization is proposed for the mesonet based on the discovery that the network-averaged gust factor is independent of weather conditions and results in unbiased forecasts of gusts at individual stations and the mesonet as a whole. Combined with atmospheric humidity and fuel moisture information, gust forecasts can help in the assessment of wildfire risks.

1. Introduction

We continue an investigation of gusty downslope winds in San Diego County, California, that occur during "Santa Ana" conditions, a cool-season weather pattern consisting of offshore flow emanating from the Great Basin and Mojave Desert that can be amplified by the terrain (cf. Raphael 2003; Conil and Hall 2006; Jones et al. 2010; Hughes and Hall 2010). The winds can be very dry, and sometimes hot, contributing to the fire danger in Southern California (e.g., Rolinski et al. 2016). The danger can be particularly acute in autumn, prior to the onset of winter rains (e.g., Westerling et al. 2004; Moritz et al. 2010). A striking example was the Santa Ana event of late October 2007 that was associated with multiple fires across Southern California, including the Witch Creek fire, which was sparked by power lines (cf. Cao and Fovell 2016; Fovell and Cao 2017).

This paper is the second in a series investigating the predictability of the gusty downslope windstorms of San Diego County. Part I of this study, Cao and Fovell (2016, hereafter CF16), focused on the moderately strong 14–16 February 2013 downslope windstorm event, in which near-surface wind gusts exceeding 40 m s^{-1} were recorded in a wildfire-prone mountainous backcountry area covered by a high-density San Diego Gas and Electric (SDG&E) mesonet. High-resolution simulations with the Advanced Research version of the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) revealed that the first phase of that event exhibited characteristics of a prominent hydraulic jump flow in part of the county, while the second phase showed a clear downslope progression of winds as the event wound down. Supporting evidence for our simulations and interpretations were found in the mesonet observations.

The success of CF16's simulations depended upon appropriate selections of the model physics and spatial resolution. Physical processes requiring parameterization

Corresponding author: Prof. Robert Fovell, rfovell@albany.edu

DOI: 10.1175/WAF-D-17-0177.1

in models like WRF include land surface and subsurface (soil) processes, mixing in the planetary boundary layer (PBL), radiative transfer, cloud microphysical processes, and the influence of subgrid turbulence and subgrid cloud activity. There are several viable options for each of these processes, resulting in a potentially enormous number of possible combinations, although some are undoubtedly better than others for this application. Using a ~50-member physics ensemble, CF16 demonstrated that most model configurations overpredicted the winds observed in the mesonet, with the forecast bias being most sensitive to the land surface model (LSM). However, even the most skillful configuration exhibited a wind-speed-dependent bias, which simultaneously overpredicted less windy sites and underpredicted windier locations. CF16 also showed that horizontal resolution profoundly influenced the spatial extent of the downslope flow, including controlling where the winds were strongest.

The first objective of this paper is to further investigate the sensitivity of the intensity, spatial extent, and structure of these windstorms to model physics, and to explain why some LSMs outperformed others. This will be done by examining additional Santa Ana events that have occurred since the deployment of the SDG&E mesonet. The second is to develop a strategy for parameterizing small-scale wind gusts, which cannot be resolved in mesoscale models. This is motivated by the fact that gusts can cause an enormous amount of damage to electrical infrastructure.

Gustiness is a remarkable feature of many downslope windstorms (e.g., Durran 2003; Jackson et al. 2013). As an example, the well-studied 11 January 1972 Boulder, Colorado, downslope windstorm had wind bursts as high as 60 m s^{-1} (e.g., Klemp and Lilly 1975). Empirical and heuristic attempts have been made to estimate wind gusts by multiplying the resolved-scale sustained wind speed by a gust factor [GF, the ratio of the peak wind speed of a given duration (gust) to the mean wind speed for a given averaging period] empirically determined from available observations (e.g., Mitsuta and Tsukamoto 1989) or adding a scalar value to the sustained wind, assuming a normal distribution of wind fluctuations (e.g., Wieringa 1973; Panofsky et al. 1977; Beljaars 1987). Brasseur (2001) pursued a gust parameterization based on physical considerations, reflecting boundary layer turbulence. In this work, we motivate a remarkably simple gust algorithm, which is shown to be skillful when applied to sustained wind forecasts for both individual SDG&E stations and the mesonet as a whole.

The organization of this manuscript is as follows. The available observations, model experimental design, and verification strategy are described in section 2. Model sensitivity to model physics, stochastic perturbations, and surface roughness is investigated in section 3. Section 4 presents an analysis of wind forecast bias for individual stations. A simple yet skillful gust parameterization for the SDG&E network is introduced in section 5, and the final section presents the summary.

2. Data and methods

a. Available observations

Wind observations are crucial for verifying and calibrating model forecasts of downslope windstorms. As in CF16 and Fovell and Cao (2017, hereafter FC17), we employed the dense, homogeneous, and high quality SDG&E surface observation network of (presently 158) stations sited in wind-prone areas, which commenced deployment in 2009 (see Fig. 1, and CF16's Fig. 1). SDG&E stations were purposefully sited in wind-prone areas, especially in the mountainous backcountry of San Diego County, and generally conform to the Remote Automated Weather Stations (RAWS) network standard with respect to anemometer mounting height (6.1 m) and sampling (3 s) and averaging (10 min) intervals. In the SDG&E network, the sustained wind is the temporal mean of the 3-s samples over each averaging interval, with the gust representing the highest wind speed sample in the interval. In contrast, RAWS gusts are not guaranteed to come from the same sequence of observations that was used to compute the sustained wind. Another difference is that SDG&E stations report every 10 min, while RAWS stations report once within each hour.

In CF16, we studied a moderately strong Santa Ana wind event of 14–16 February 2013 that was captured by the SDG&E network, employing a model with fine temporal and spatial resolution. Like many Santa Ana events, it spanned two days and evinced a pronounced diurnal cycle with a lull in the offshore winds in the late afternoon of the first day, possibly a response to boundary layer evolution (cf. Smith and Skyllingstad 2011). During the first phase, a hydraulic jump became visible in a vertical cross section oriented west–east across SDG&E station West Santa Ysabel (WSY), sited very near where the October 2007 Witch fire started. The jumplike feature persisted for several hours and was consistent with observations taken at stations arrayed along the west-facing slope, including sites WSY, Witch Creek (WCK), and Sunset Oaks (SSO). The second phase consisted of a marked westward downslope progression of winds with time as the overall winds amplified and waned.

In this study, six Santa Ana episodes (Table 1) were selected, adding to the February 2013 case two

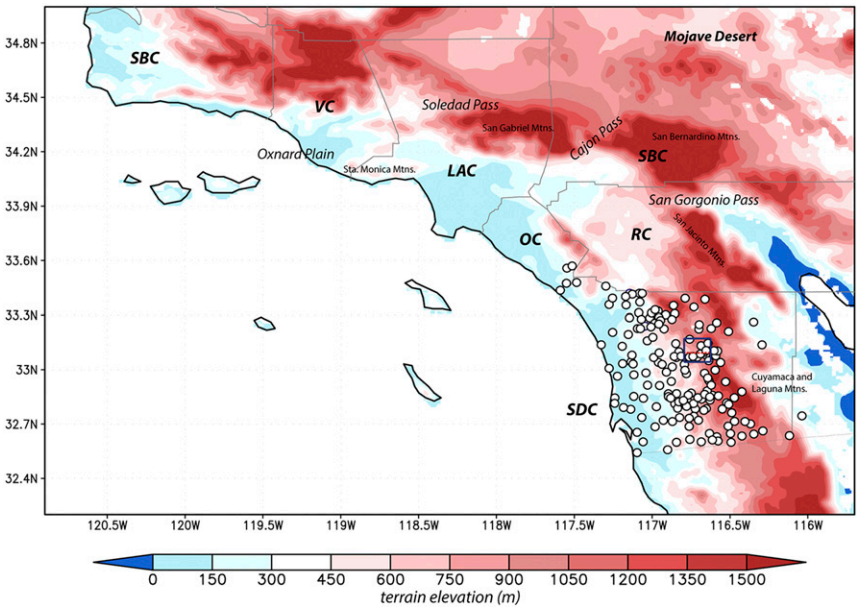

FIG. 1. Topography of Southern California (longitude on abscissa, latitude on ordinate), with selected place names. County outlines are in gray; identifiers are Santa Barbara (SBC), Ventura (VC), Los Angeles (LAC), Orange (OC), San Bernardino (SBC), and Riverside (RC) Counties. White dots denote SDG&E observational stations. The blue box highlights the Witch Creek region, and the red dashed line depicts the location of the vertical cross section across station WSY.

moderately strong events (in October 2013 and January 2015) and three strong events (in April 2014, May 2014, and February 2015). Multiphysics ensembles consisting of around 50 members each were constructed for the February 2013, October 2013, and May 2014 episodes, which may provide some insights into the optimal model configuration with respect to model physics that are applicable to other events of various strengths. Verifications employed the 135 SDG&E stations that were common to all six events and are shown in Fig. 1. All six events were used for the composite analyses presented in sections 4 and 5.

b. *Model setup and verification strategy*

As in CF16 and FC17, we employed WRF version 3.5 utilizing five domains telescoping to 667-m grid spacing that covered the highest terrain portion of the SDG&E network (Fig. 1). Domain 4 (2-km resolution) encompassed the entire SDG&E network (see CF16's Fig. 3), and all analyses were performed in this nest. This work was motivated by the need to anticipate winds that could impact electrical transmission lines, and so forecast data available in real time were used for the initialization and lateral boundary forcing. Specifically, all simulations were initialized with North American Mesoscale Forecast System (NAM) grids at either 1200 or 0600 UTC, with earlier initializations preferred for episodes with multiple peaks. All model integrations were 54 h long, starting prior to the onset of offshore winds, and including the bulk (if not the entirety) of the Santa Ana event. The model top was 10 hPa, with 50 layers (51 full-sigma vertical levels) employed, focusing the highest resolution in the lower troposphere in the usual fashion.

Physics ensembles consisting of variations of the LSM and PBL schemes were conducted to create many parallel realizations of these periods: 14–16 February 2013, 4–6 October 2013, and 13–15 May 2014 (Table 1). (The 14–16 February 2013 ensemble was introduced in CF16 and is examined in greater detail herein.) The five LSMs included the Noah land surface model (Chen and Dudhia 2001; Ek et al. 2003), the Noah model with multiparameterization options (NoahMP; Niu et al. 2011), the Rapid Update Cycle (RUC; Smirnova et al. 2000), the Pleim–Xiu (PX; Pleim and Xiu 1995; Xiu and Pleim 2001), and the thermal diffusion (TD; Skamarock et al. 2008) schemes. Ten PBL parameterizations were selected, those being the Yonsei University (YSU; Hong et al. 2006), Mellor–Yamada–Janjić (MYJ; Janjić 1994), quasi-normal scale elimination (QNSE; Sukoriansky et al. 2006), Mellor–Yamada–Nakanishi–Niino level 2.5 (MYNN2; Nakanishi and Niino 2004), Asymmetric Convection Model version 2 (ACM2; Pleim 2007a,b), Bougeault–Lacarrere (BouLac; Bougeault and Lacarrere 1989), Bretherton–Park (UW; Bretherton and Park 2009), total energy–mass flux (TEMF; Angevine et al.

TABLE 1. Santa Ana wind events studied and their model initialization time. Here, Y denotes the event was used to create a physics ensemble and N denotes the event was not used for creating a physics ensemble.

Event	Initialization time and date	Used in the physics ensemble?
14–16 Feb 2013	1200 UTC 14 Feb	Y
4–6 Oct 2013	0600 UTC 3 Oct	Y
29 Apr–1 May 2014	0600 UTC 29 Apr	N
13–15 May 2014	0600 UTC 13 May	Y
23–25 Jan 2015	1200 UTC 23 Jan	N
11–13 Feb 2015	1200 UTC 11 Feb	N

2010), Grenier–Bretherton–McCaa (GBM; Grenier and Bretherton 2001), and Medium-Range Forecast Model (MRF; Hong and Pan 1996) options. For each PBL scheme, the recommended and/or most frequently adopted surface-layer parameterization was employed. Some of the combinations were not workable, which left us with a total of 48 viable ensemble members. As in CF16, horizontal diffusion along the model surfaces was deactivated, and the MODIS land-use database was adopted. In contrast with Wilson and Fovell (2016), horizontal diffusion was not found to substantially influence our results or conclusions (cf. Cao 2015, CF16). As part of this work, the available topo_wind options were evaluated and determined to be unhelpful.

Unsurprisingly, given the weather associated with Santa Ana events, we found that the downslope windstorms were not very sensitive to the treatment of the microphysics or cumulus convection, and the influence of the radiation parameterization was also small. Therefore, all simulations examined herein employed the WRF single-moment 3-class microphysics scheme (Hong et al. 2004), a simple ice-bearing scheme suitable for mesoscale grid sizes, the Kain–Fritsch (Kain 2004) cumulus parameterization (in the 54- and 18-km domains only), and the Rapid Radiative Transfer Model for GCMs (RRTMG; Iacono et al. 2008) package for longwave and shortwave radiation.

For selected events (February 2013 and May 2014) and physics ensemble members (Noah–YSU and PX–ACM2), the sensitivity to random perturbations was assessed via ensembles created with WRF's stochastic kinetic energy backscatter scheme (SKEBS) option (Shutts 2005; Berner et al. 2011). This technique inserts random noise perturbations into the rotational horizontal wind components and the potential temperature field where and when turbulence is diagnosed. Each SKEBS ensemble consisted of one control run and 20 perturbed members, created by varying the random number seed used as input to the SKEBS procedure. SKEBS has several alterable parameters, and the standard or recommended values were adopted.

As discussed in CF16 and FC17, model-forecasted winds from 10 m above ground level (AGL) were adjusted using the stability-dependent logarithmic wind profile to the SDG&E anemometer height (6.1 m) before hourly verifications against observed sustained winds were conducted. Mean absolute error (MAE) and bias statistics were again used to assess how close pointwise model predictions $f_{j,i}$ were to their corresponding observations $y_{j,i}$. The MAE and bias are defined for station j and time i as

$$\mathrm{MAE}_{i,j} = |f_{j,i} - y_{j,i}| \tag{1}$$

and

$$\mathrm{bias}_{i,j} = (f_{j,i} - y_{j,i}). \tag{2}$$

CF16 reported that shifting the lowest model level downward to match the SDG&E anemometer height was not found to be sufficiently helpful to justify the increased computational cost (in the form of shorter time steps) for some physics combinations, and herein we employ the default value of roughly 27 m AGL. Particularly in the next section, our main emphasis will be on network- and/or event-averaged winds for the 135 SDG&E stations that are common to all six events.

3. Sensitivity tests: Model physics, stochastic perturbations, and surface roughnesses

In this section, we reexamine and extend CF16's physics ensemble for the February 2013 event, assess the sensitivity to stochastic perturbations and, with the aid of physics tests for two other episodes, assess the influence of surface roughness on the simulated airflows and forecast wind skill.

a. February 2013 event physics ensemble

CF16 selected the PX–ACM2 LSM–PBL combination because it minimized the forecast MAE and bias of the network-averaged sustained wind over multiple events and faithfully captured the temporal evolution and spatial variation of the flow during the February 2013 episode. In contrast, most other configurations systematically overpredicted the observed winds with respect to intensity (e.g., Fig. 17 in CF16) and/or offshore wind extent. The former is also illustrated in Fig. 2, which compares time series of network-averaged sustained wind for the February 2013 episode from the PX–ACM2 configuration with a more commonly used pair (Noah–YSU).

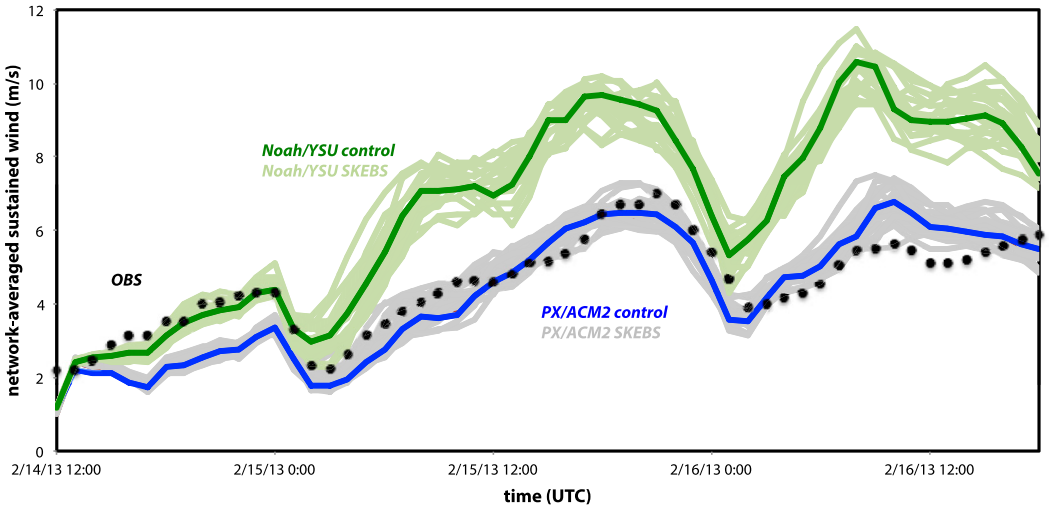

FIG. 2. Time series of network-averaged observed (black dots) and predicted (blue, PX–ACM2 control run; green, Noah–YSU control run) 6.1-m sustained winds ($m\,s^{-1}$) for the 14–16 Feb 2013 event. The gray and light green plumes reveal the ensemble spread created via SKEBS perturbations.

The PX–ACM2 reconstruction is clearly superior, at least after the first 12 h, and its event- and network-averaged sustained wind bias was $0.07\,m\,s^{-1}$, lower than Noah–YSU's $1.48\,m\,s^{-1}$.

Highlighted in CF16 (see their Fig. 6) was the persistent jumplike feature in the flow through the Santa Ysabel area (boxed area in Fig. 1) that appeared in the surface observations as weaker downslope winds (and, occasionally, reversed or upslope flow) at station WCK than at neighboring stations both up (WSY) and downhill (SSO). Figure 3 presents wind speeds averaged over a 4-h period, spanning the time of WSY's first peak gust (around 1800 UTC 15 February 2013) and WCK's first wind reversals. The flow can be presumed to be easterly (from right to left) nearly everywhere. The PX–ACM2 (Fig. 3a) and Noah–YSU (Fig. 3b) members both developed hydraulic jumps but positioned them differently. In the former, the jump was well formed and centered over WCK, and the winds were stronger both uphill and downhill of WCK, as was observed. The Noah–YSU simulation's jump was shifted somewhat downslope, resulting in higher winds at WCK than SSO, contrary to the observations. A jump was barely evident in the TD–YSU simulation as its temporally averaged winds were high at nearly all stations (Fig. 3c). The fourth configuration shown (Fig. 3d) will be discussed in section 3c.

Figure 4a shows the mean wind speed and potential temperature fields from the 48-member physics ensemble for the February 2013 event, computed for the same 4-h period as shown in Fig. 3. Owing to variations in position among the members, as well as its absence in some (such as TD–YSU), the jump feature was virtually absent in the ensemble average. The ensemble standard deviation was largest for both wind speed and potential temperature above WCK (Fig. 4c), where the jump was observed to occur. In contrast, the variation above WSY, where the highest winds in this cross section were recorded, was negligible.

Focusing on the model-diagnosed 10-m winds[1] averaged over the same 4-h interval (Fig. 5a), we note that the variation among the 48 ensemble members was quite small upwind of, and even beyond, the ridge, at least until the easterly flow passed station WSY. From that point downhill, the variation became quite substantial ($2–18\,m\,s^{-1}$). As suggested by the figure, only a subset of the physics ensemble members captured the weak winds associated with the jump observed at or very near WCK. Although not shown, we note that the region of largest variation on the lee side for the 4-h window around the peak of the *second* phase (during which time a jump did not appear in the observations) was found relatively farther down the slope, and there was slightly more spread in the flow on the east side of the ridge, perhaps reflecting the natural tendency for nonlinear simulations to diverge over time.

Figure 6a displays the event- and network-averaged 6.1-m forecast wind bias versus MAE for the February

[1] As this is a comparison among ensemble members, adjustment of these winds to the SDG&E anemometer height was not necessary.

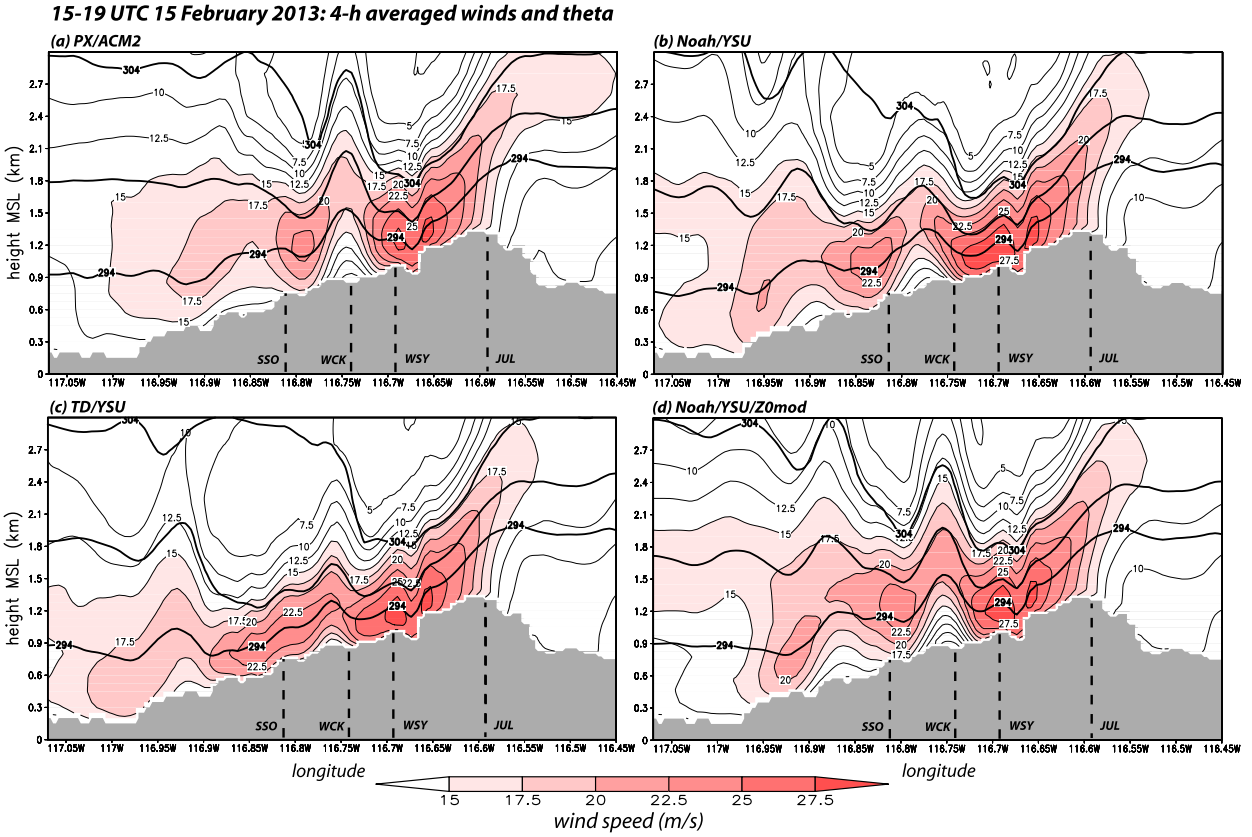

FIG. 3. Vertical cross sections of 4-h averaged horizontal wind speed (2.5 m s^{-1} contours and red shaded fields) and potential temperature (thick black 5-K contours) for the 1500–1900 UTC 15 Feb 2013 event, taken west–east across station WSY for four physics combinations: (a) PX–ACM2, (b) Noah–YSU, (c) TD–YSU, and (d) Noah–YSU–$z_0$mod.

2013 event, augmenting CF16's Fig. 17. (The quasi-linear relationship between these two metrics is anticipated.) Most physics ensemble members had a mean positive bias and also a larger MAE than the selected configuration, PX–ACM2, even following careful adjustments to the mesonet's 6.1-m anemometer mounting height. The result is clearly driven by the choice of LSM (differentiated by dot color) rather than the PBL scheme. The commonly employed Noah–YSU combination placed about average for this event, selecting TD resulted in most of the highest biases and mean absolute errors, and the RUC runs as a group were second best.

b. *Sensitivity to model random perturbations (February 2013 and May 2014 events)*

Prior to examining the physics ensembles for two other Santa Ana episodes, we will explore the effects of stochastic perturbations on wind speed patterns for the February 2013 and May 2014 events (including jump formation in the former).

Minimizing wind forecast errors entails determining wind speeds accurately at the right places and times. Although we are dealing with network-averaged quantities here, it might be anticipated that simulations that do not capture the jump, or do not position it correctly, will have relatively higher errors. It might be further anticipated that feature is also susceptible to noise or other perturbations. This is investigated using SKEBS-based ensembles, consisting of 20 members each for the Noah–YSU and PX–ACM2 configurations.

Figures 4b and 4d show that, by itself, applying perturbations to a single physics member, Noah–YSU, for the February 2013 event, generated qualitatively and even quantitatively comparable variations in wind speed and potential temperature to those produced by the physics ensemble during the jump period. Like the physics ensemble as a whole, members in the Noah–YSU SKEBS collection included runs in which the jump was very pronounced and others in which it failed to form (Fig. 5b), such that the feature disappeared from the perturbation ensemble mean (Fig. 4b). Therefore, at least part of the difference among the physics ensemble members was due to chance, at

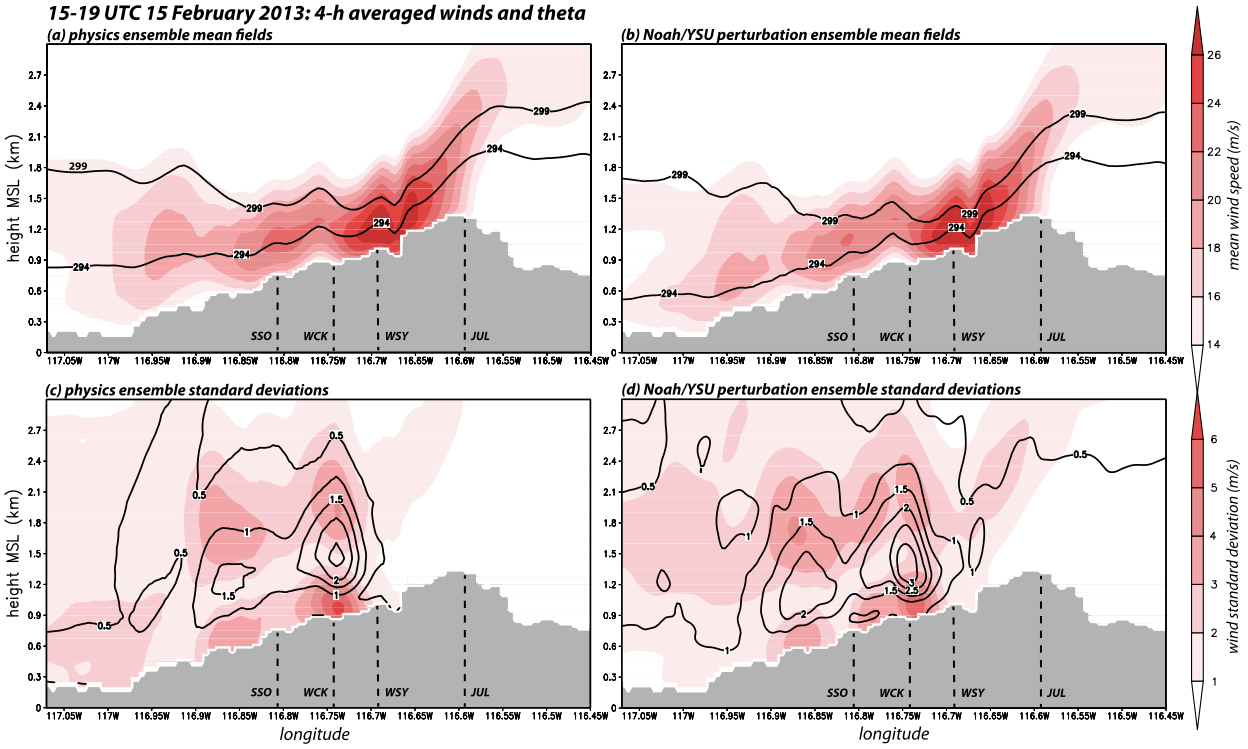

FIG. 4. As in Fig. 3, but for (a) the physics and (b) Noah–YSU perturbation ensembles' mean horizontal wind speed (shaded; m s^{-1}) and potential temperature (contoured; K) fields, and (c) the physics and (d) Noah–YSU perturbation ensembles' horizontal wind speed (shaded; m s^{-1}) and potential temperature standard deviations (contoured; K).

least as far as jump formation and its location are concerned.

Figures 5c and 5d show the variation of 10-m winds for the May 2014 event from its physics and Noah–YSU SKEBS ensembles, respectively, again averaged over the 4-h time period during which the winds on the lee slope were strongest. The May case did not produce a jump feature, either in the simulations or observations, and the variability provoked by the perturbations was rather small (Fig. 5d). The variation owing to the model physics (Fig. 5c) remained, which was still substantial at various places along the lee slope.

However, when averaged over the network, the sensitivity to SKEBS-generated perturbations is seen to have been relatively minor and was therefore swamped by the physics differences, especially the selection of the LSM. This is revealed for the February 2013 episode in Figs. 2 and 6a. All of the Noah–YSU reconstructions generated larger mesonet-average winds than any of the PX–ACM2 runs (including its SKEBS-perturbed members), at least after the first 12 h or so (Fig. 2). When plotted in bias versus MAE space (Fig. 6a), we see that the values for the Noah–YSU SKEBS runs (enclosed in the light blue ellipse) are rather far removed from all of the PX–ACM2 members (including its SKEBS runs enclosed in the light orange ellipse). Both figures also reveal that the stochastic perturbations excited less variability with the PX–ACM2 configuration.

c. *The influence of surface roughness*

Figure 7 summarizes the physics ensemble results with respect to LSM, aggregated over three events (February and October 2013 and May 2014) and representing around 150 simulations in total. The boxplots are based on physics ensemble MAEs and display their median and first and third quartiles, with the whiskers identifying the maximum and minimum values. A large spread in MAE, such as that seen for TD and NoahMP, indicates either more substantial sensitivity to the PBL scheme and/or greater skill variation among events; both are undesirable. Apart from a few outlying members, the Noah LSM had less variable (but relatively large) MAEs, while PX and RUC performed best in these trials. The red line depicts average biases for the LSM members, and it is seen that the bias was closest to zero with the PX LSM.

The boxplot, however, obscures an interesting and revealing finding. Figure 6 also shows results from the two other events (4–6 October 2013 and 13–15 May 2014) for

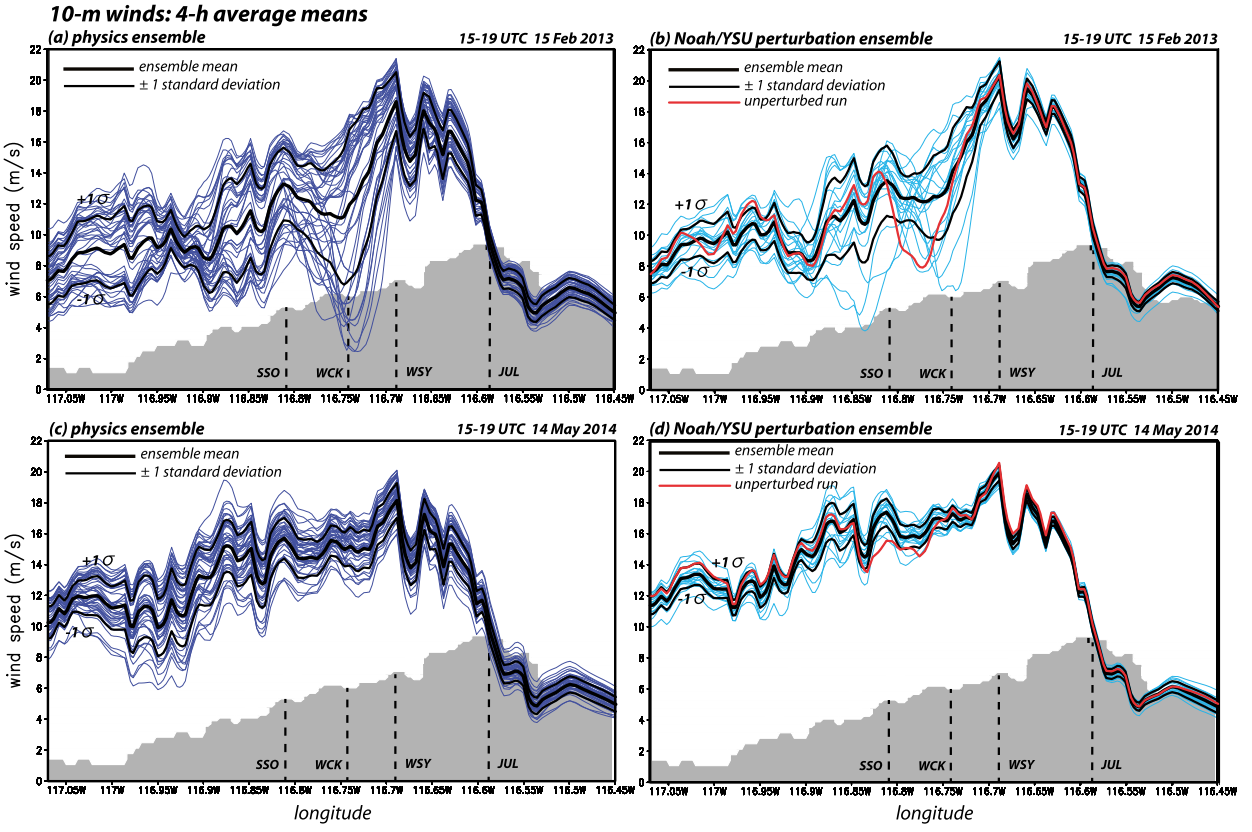

FIG. 5. Vertical cross sections taken west–east across WSY showing the ensemble mean (thick black line) and ±1 standard deviations (thin black lines) of the 4-h averaged horizontal 10-m wind speed (blue lines) for the first phase (1500–1900 UTC) of the 15 Feb 2013 event from the (a) physics and (b) Noah–YSU SKEBS ensembles. The same fields are also shown for the 1500–1900 UTC 14 May 2014 event from the (c) physics and (d) Noah/YSU SKEBS ensembles. The gray-shaded area depicts topography, shown for reference only. Approximate locations of some SDG&E stations are indicated.

which multiphysics ensembles were made. Consistent with the foregoing information, the errors clearly vary systematically with LSM, and the PX–ACM2 configuration remains one of the best choices in terms of minimizing errors. Note, however, that the ordering of LSMs with respect to skill shifts among the cases: for the October 2013 event, the Noah LSM clearly resulted in larger error than the other LSMs, with the Noah–YSU member actually being the *least* skillful of all.

At first glance, this may appear to contradict the information shown in Fig. 7. However, note that unlike NoahMP and TD, MAEs for the Noah subensemble were relatively uniform among the three events, and therefore the vertical extent of the Noah boxplot in Fig. 7 is small. In contrast, the TD and NoahMP MAEs varied more from event to event. TD, in particular, performed worse than Noah in the February and May cases but better for the October episode (Fig. 6).

This seasonal variation of forecast skill provided a clue that helped reveal why the PX LSM performed best among the five LSMs examined. The key difference lies in the surface roughness lengths z_0 applied to various land-use types, especially MODIS land-use categories 6 and 7 (closed and open shrublands[2]) that constitute a large (67.4%) fraction of the SDG&E network (see Table 2). Most LSMs start with roughness information provided in the LANDUSE.TBL and/or VEGPARM.TBL tables in the WRF Model, although some schemes subsequently make modifications. TD and NoahMP used the default MODIS roughness length values defined in the two tables, which vary abruptly from 0.01 m (November–April) to 0.06 and 0.05 m (May–October) for open and closed shrublands, respectively (Fig. 8; Table 2). The Noah LSM internally creates an annual cycle in z_0 with more gradual transitions that is also curiously shifted in phase from the default MODIS values, with the lowest roughness (0.01 m) in September

[2] "Closed shrubland" is characterized by dense foliage cover (70%–100%), while foliage for "open shrubland" is less dense (30%–70%) (https://en.wikipedia.org/wiki/Shrubland).

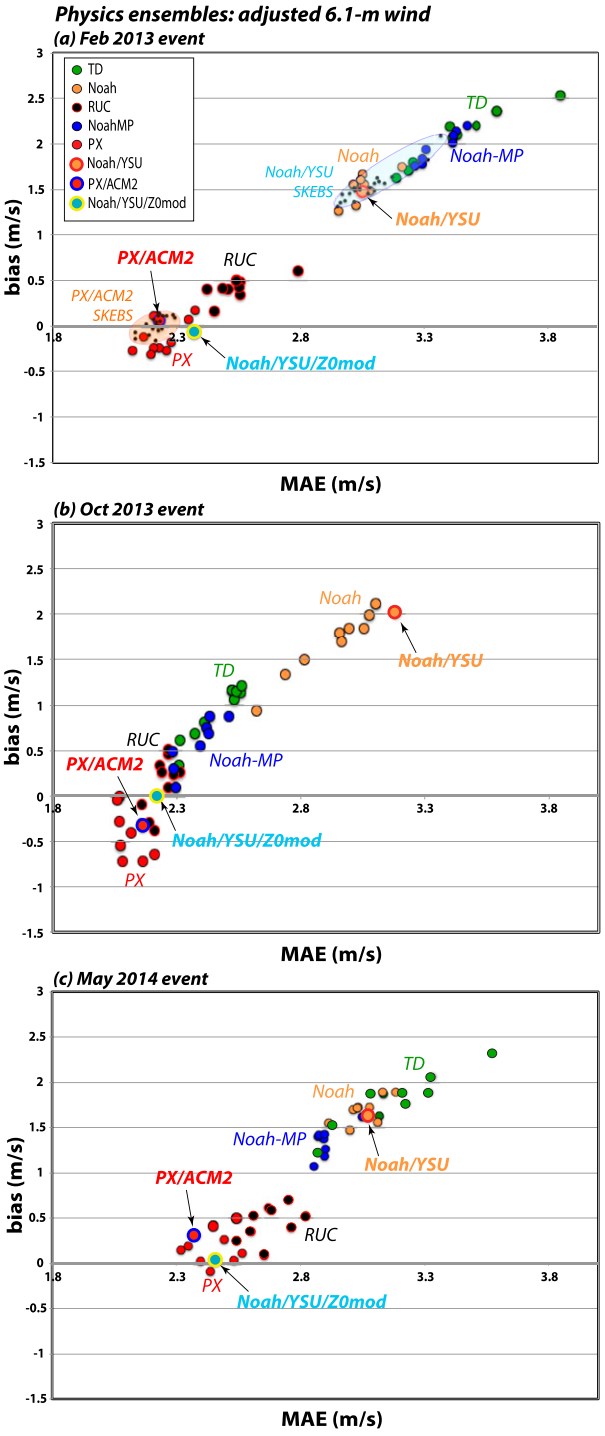

FIG. 6. Scatterplots of network- and event-averaged 6.1-m sustained wind bias vs MAE (both m s^{-1}) from the 48 physics ensemble members for the (a) 14–16 Feb 2013, (b) 4–6 Oct 2013, and (c) 13–15 May 2014 episodes, color coded by LSM. Small black dots in (a) show SKEBS perturbation ensemble members made for the PX–ACM2 and Noah–YSU configurations, respectively. Runs PX–ACM2, Noah–YSU, and Noah–YSU–$z_0$mod are marked. For members using the MYJ PBL scheme, the standard but cosmetic recalculation of the near-surface winds was removed (see CF16).

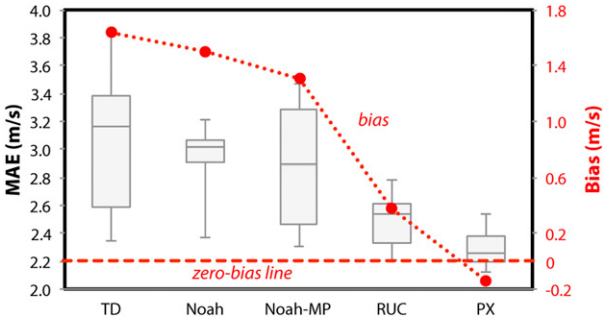

FIG. 7. Boxplots summarizing physics ensemble MAE (m s^{-1}) distributions stratified by LSM, displaying the median, first quartile, and third quartile values, with the whiskers identifying minima and maxima. The superposed red dotted line depicts their mean biases (m s^{-1}).

and the highest (0.05 m for closed shrublands and 0.06 m for open shrublands) in April. In contrast, PX and RUC assigned temporally constant and relatively high roughness length values of 0.15 and 0.1 m, respectively, to both shrubland categories. Note that the specified z_0 values used for a particular location are influenced by model grid interpolation and possibly other factors.

As a test, we modified the Noah LSM to utilize the same surface roughness values of each land-use type employed by PX (Table 2), which meant removing the seasonal cycle and increasing z_0 values for most categories, especially the open and closed shrublands that dominate the San Diego County landscape. The modified Noah simulations, dubbed "Noah–YSU–$z_0$mod" (Fig. 6), were very competitive with respect to event- and network-averaged bias and MAE, with near-to-zero biases and much smaller MAEs (comparable to PX's). Note that the revised Noah LSM also produced a temporally averaged airflow during the jump phase for the February 2013 event (Fig. 3d) that closely resembles that established by PX–ACM2 (Fig. 3a). During this time period, TD and NoahMP presumed the smoothest shrublands (z_0 = 0.01 m; Fig. 8) and were unable to capture the jump (as shown for the TD–YSU case in Fig. 3c). This leads us to the reasonable conclusion that surface roughness is a principal control of airflow on the lee slope, which may have implications for improving operational wind forecasts overall and at specific locations.

Thus, at least for our region of interest and present model configurations, wind forecast skill using the Noah LSM can be greatly improved by adopting roughness lengths employed by the PX scheme. As the Noah scheme has some advantages, especially with respect to 2-m temperature and dewpoint temperature forecast skills (not shown), the revised Noah LSM, Noah–YSU–$z_0$mod, emerges as one of the most skillful configurations

TABLE 2. Default roughness lengths (m) employed by land surface schemes for MODIS land-use (LU) categories occurring in the SDG&E network. Water areas of the 2-km nest are excluded.

MODIS LU index	Land fraction (%)	PX	Noah (February)	MODIS (winter)	Noah (October)	Noah (May)	MODIS (summer)	Type
1	8.2	1	0.5	0.5	0.5	0.5	0.5	Evergreen needleleaf forest
2	0.7	0.9	0.5	0.5	0.5	0.5	0.5	Evergreen broadleaf forest
5	11.1	1	0.3	0.2	0.23	0.5	0.5	Mixed forests
6	8.2	0.15	0.03	0.01	0.02	0.05	0.05	Closed shrublands
7	59.2	0.15	0.04	0.01	0.02	0.06	0.06	Open shrublands
8	0.7	0.25	0.05	0.01	0.06	0.05	0.05	Woody savannas
9	0.7	0.15	0.10	0.15	0.1	0.1	0.15	Savannas
10	3.0	0.07	0.1	0.1	0.2	0.1	0.12	Grasslands
13	8.2	0.8	0.5	0.8	0.5	0.5	0.8	Urban

of all. For simplicity, however, the PX–ACM2 combination remains the control model configuration for the remainder of this study. Finally, we note in passing that both versions of the topo_wind option have been found to reduce the high wind bias of the standard Noah–YSU configuration, but generally resulted in underforecasted winds (cf. Cao 2015).

4. Wind forecast bias analysis for individual stations

This section utilizes a composite of six Santa Ana events, adding the 29 April–1 May 2014, 23–25 January 2015, and 11–13 February 2015 events (Table 1) to the three considered for the physics ensemble. Again, each simulation is 54 h long and comparisons with observations were performed hourly, so the composite dataset consisted of 324 observation times, initialization times having been excluded.

Figure 9a shows average forecast wind biases for the composite at the 135 SDG&E stations common to all six events in rank order. As with the similar figure shown for the February 2013 event in CF16 (their Fig. 10d), the average bias is nearly zero among the stations, but systematic errors exist in that some locations are persistently either overpredicted (red) or underpredicted (blue) with respect to the event-averaged winds.[3] The spatial distribution of the composite biases (Fig. 10) reveals that some clustering is evident, but stations with a wide range of biases can be found in close proximity, such as in the region around WSY that is highlighted in the inset at the top-right corner. This suggests that very localized conditions (e.g., a single tree, a tiny hill), representing landforms not resolvable even on the 667-m grid utilizing the ~10-m USGS database (cf. CF16), may be responsible for much of these persistent biases. The locations of stations identified by name in Fig. 9a are also indicated in Fig. 10.

Station rankings for the observations and selected simulations are combined in Fig. 9b. The station ordering is not necessarily the same for each, as this plot is intended to compare the relative distributions of averaged wind speeds across the network. The observed distribution has a "hockey stick" shape with a blade composed of eight stations (6% of the network) with sustained winds greater than 9 m s^{-1}. The standard PX–ACM2 case (thick black line), with 667-m grid spacing over much of the network, compares quite well with the observations (red line), although its somewhat flatter shape and shorter blade again mean that there was some overprediction at sites with relatively slower wind speeds to go along with the windiest locations that were underforecasted. Shown for contrast are rankings from simulations made at 10-km grid spacing for PX–ACM2 and Noah–YSU that overpredicted the winds nearly everywhere, except for the few top wind locations. CF16 demonstrated that resolution influences the resolved shape of the topography, and overly smooth resolved terrain using a horizontal grid spacing larger than 2 km exaggerates the horizontal extent of the downslope winds (see their Fig. 16).

An analysis of the source of the forecast biases is presented in Fig. 11, in which each dot on the scatterplots represents an individual station. First of all, forecast wind bias is not a function (coefficient of variation $R^2 = 0.00$) of the forecast wind strength itself (Fig. 11a), at least when aggregated over the 324 hourly forecasts for each station. However, forecast bias is systematically related to (and negatively correlated with) the observed wind, with $R^2 = 0.45$ (Fig. 11b). As in CF16, which discussed the same issue with respect to the February 2013

[3] This is true for our topo_wind simulations as well (not shown), so that option does not mitigate this issue.

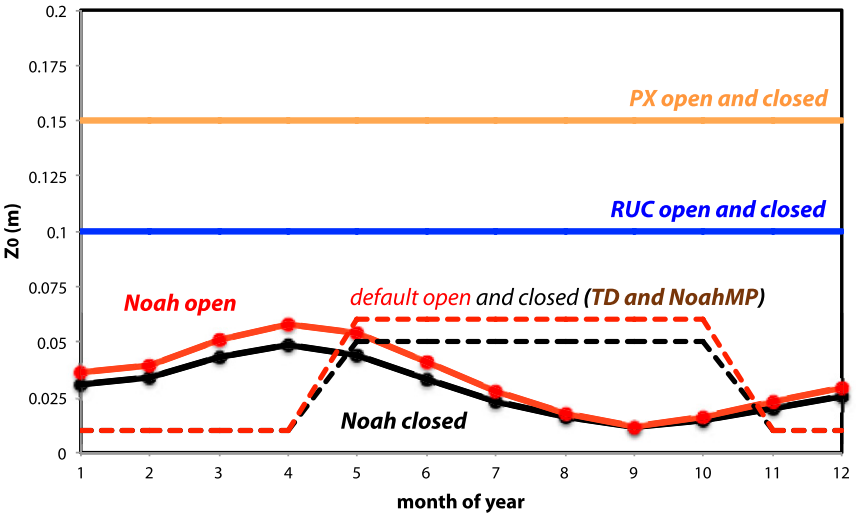

FIG. 8. Seasonal cycles of the default shrubland surface roughness length z_0 (m) for each LSM [red and black solid, Noah LSM open and closed shrublands; red and black dashed, MODIS default (TD and NoahMP LSMs) open and closed shrublands; blue, RUC LSM open and closed shrublands; and orange, PX LSM open and closed shrublands].

event (their Fig. 18), this relationship is partly determined by the windiest locations, and the R^2 drops to 0.25 (yet still statistically significant, as the simple linear correlation is $r = -0.5$) when the top eight highest mean wind stations are removed (not shown).

To improve forecast skill, we seek to identify and mitigate all possible errors. Beyond initialization issues and shortcomings in physical parameterizations, which can be sizable, some additional error sources can include factors such as unrepresentative surface roughnesses as well as a seasonal cycle in z_0 that appears questionable for this region. Very localized characteristics, including natural small-scale terrain features and artificial obstacles that can act to enhance or suppress windiness at a given location, would represent errors that may be inherently unavoidable and unfixable except perhaps through postprocessing. Here, we pursue the idea that as the network-averaged bias is nearly zero, most (but not all) of the fixable portion of the forecast bias has been addressed, and what remains is dominated by *unresolvable exposure issues*. We further posit that *information regarding local exposure can be derived from the* GF, which is defined as the gust divided by the sustained wind.

In this application, the GF was computed using composite (324 h) averages for both quantities, thereby representing a single value for each location characteristic of Santa Ana episodes. Station GF is moderately (and negatively) related to the average observed wind ($R^2 = 0.26$, $r = -0.51$) such that faster wind stations tend to have relatively smaller GFs (Fig. 11c). In contrast, GF is a fairly good predictor ($R^2 = 0.48$, $r = 0.69$) of forecast wind bias (Fig. 11d). The average GF among these stations is 1.78, indicated in Fig. 11d with a vertical blue line, and ranges from 1.29 [site Volcan Mountain (VCM); Fig. 10] to 2.55 [site Sunrise Highway (SRH); Fig. 10].

By way of explanation, we speculate that larger than average GFs indicate locations that are relatively more sheltered. As at least part of the gustiness represents higher-momentum parcels being brought closer to the surface by turbulent air motions, it is reasonable that gusts would be somewhat less influenced by a station's local conditions than the temporally averaged sustained winds. Therefore, anemometer exposure restrictions are anticipated to influence the sustained wind more than the gust, resulting in larger GFs. In contrast, stations with smaller GFs are presumed to have landform features (e.g., small hillsides, tiny valley wind corridors) that help enhance the sustained wind more than the gust, such that the ratio of gust to wind is reduced. As a consequence, we interpret GFs *both larger and smaller* than the all-station average as a proxy for the influence of localized and largely unresolvable conditions.[4]

While most of the stations cluster close to the least squares line, there is a subset of stations for which the bias is less well explained by the GF, especially those located well below the regression line (Fig. 11d). This is part of a tendency for sites that are assigned larger

[4] The fact that the mesonet is homogeneous with respect to instrumentation, mounting height, and data-processing characteristics may be crucial to this result.

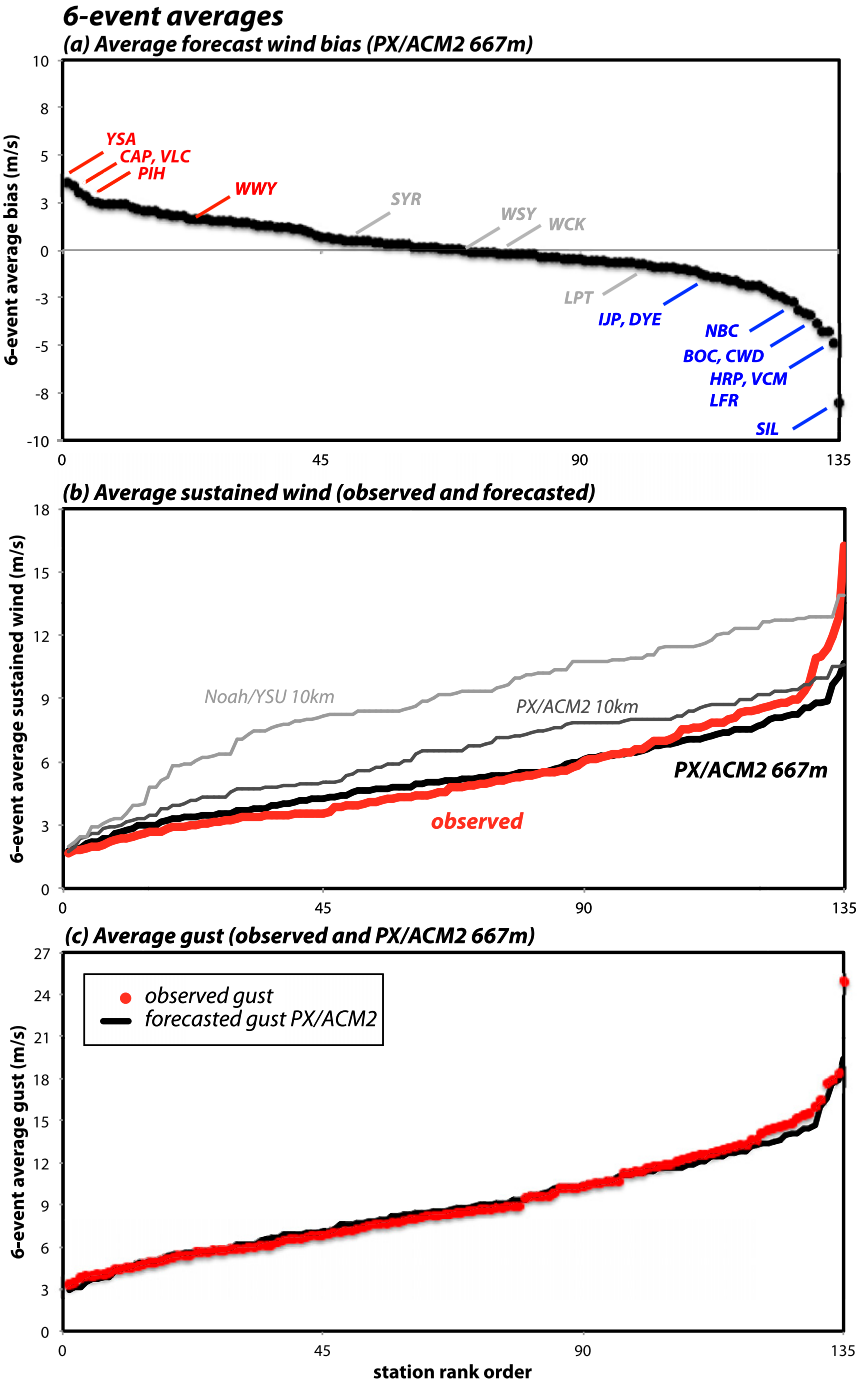

FIG. 9. (a) Sustained wind bias ($m\,s^{-1}$) for PX–ACM2 simulations, (b) sustained winds (red, observed; black, PX–ACM2 667 m; dark gray, PX–ACM2 10 km; and light gray, Noah–YSU 10 km; $m\,s^{-1}$), and (c) gusts (red, observed; black, forecasted using PX–ACM2; $m\,s^{-1}$), all averaged over six Santa Ana events (see Table 1), in station rank order with selected stations identified in (a), color coded by sign of the bias. Only the 135 stations available for all six events are included.

surface roughnesses to be more severely underpredicted (Fig. 11e). While the overall relationship between bias and z_0 is not large ($R^2 = 0.17$, $r = -0.41$), note that of the 17 stations with $z_0 > 0.7$ m, only two [Julian (JUL), which is forested, and El Monte (ELM), in an urbanized area; Fig. 10] do not have a negative bias. Indeed, this subset has an average forecast bias of $-3.4\,m\,s^{-1}$, while it is $+0.25\,m\,s^{-1}$ for the remaining 118 stations. This

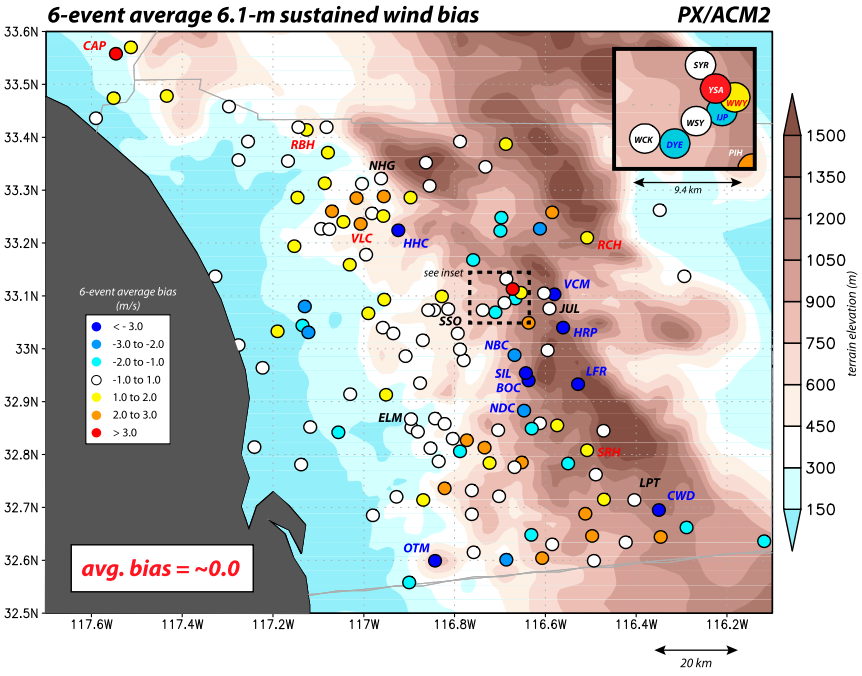

FIG. 10. Spatial distribution of six-event-average 6.1-m sustained wind bias ($m\,s^{-1}$), color coded as indicated. The average bias over the entire SDG&E network is around $0\,m\,s^{-1}$. Inset shows the subset of the 135 stations in the Santa Ysabel vicinity. The locations of stations mentioned by name in the text and/or figures are identified.

suggests that while network forecasts benefited from the PX scheme's assumptions regarding surface roughness, the highest z_0 assignments are probably excessive.

The 17 stations with roughness lengths exceeding 0.7 m are identified with red markers in Fig. 11d. These represent two basic land-use types: 11 are associated with forests (categories 1, 2, and 5 from Table 2) and six with urban land (category 13). Personal site inspections and discussions with a meteorologist at SDG&E (S. Vanderburg 2016–17, personal communications) have reinforced the idea that at least some of the forest assignments appear inappropriate or outdated, if only owing to the effect of extensive wildfires (e.g., the 2003 Cedar fire and the 2007 Witch and other nearby fires) that have significantly altered the landscape.

In postprocessing, we recomputed the winds for 10 forested locations (excluding JUL), presuming a more moderate (and empirically selected) surface roughness of 0.45 m, which is more comparable to those employed by the Noah LSM (Table 2) for these sites. This modification improved the 17-station subset's average bias from -3.4 to $-0.6\,m\,s^{-1}$ and raised the overall R^2 from 0.48 to 0.56 (Fig. 11f). Some of these stations remain outliers, and the names for the more obvious ones are marked in the figure. These include VCM, Otay Mountain (OTM), Sill Hill (SIL), and Lucky Five Ranch (LFR), which are situated on hilltops or very close to steep (and largely unresolved) ravines, which may play significant roles in their relatively large sustained winds. Further adjustments could be pursued, but the need is not apparent and the benefits are likely small. These z_0-adjusted forecasts are used in the next section.

The least squares fit shown in Fig. 11f (which includes the outliers) was used to predict forecast bias for the observed GF, and the residuals from this model are shown in Fig. 11g, plotted against the observed wind. In contrast to Fig. 11b, *no* signal of the observed wind remains after the bias explained by the station GF has been removed ($R^2 = 0.02$). This indicates to us that *the GF variation from the network mean serves as a useful proxy for local site exposure issues*, such that, after relatively minor adjustments to z_0 for a handful of sites, the systematic and persistent forecast biases seen in Figs. 9a and 11b essentially reflected unavoidable errors to be remedied in postprocessing.

5. A gust parameterization for the SDG&E network and its stations

Numerical models of the present type do not resolve short-period (~3 s) 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 will still act to suppress variations with time scales of less

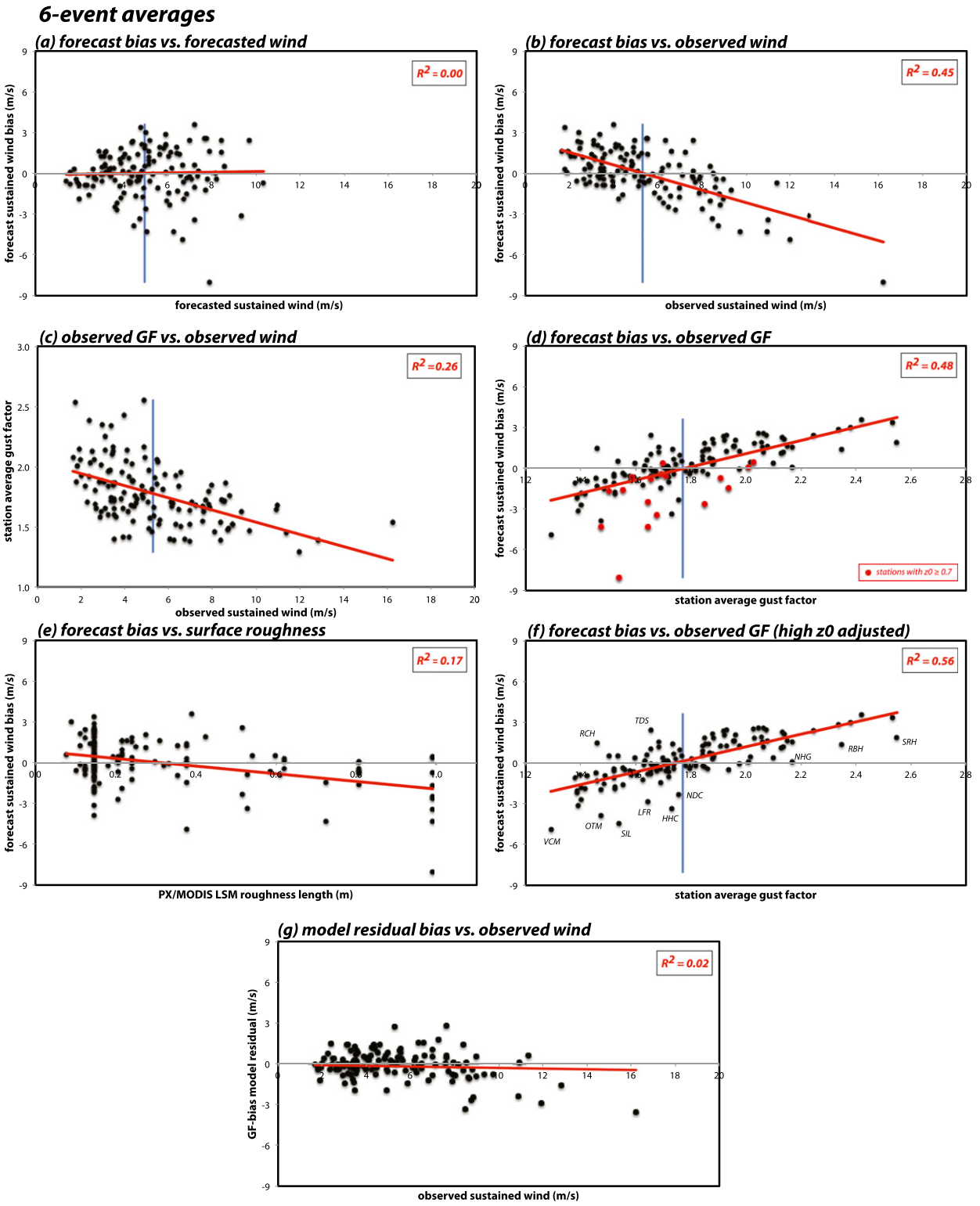

FIG. 11. Scatterplots of six-event-mean (a) forecast sustained wind vs sustained wind bias, (b) observed sustained wind vs sustained wind bias, (c) observed sustained wind vs observed GF, (d) observed station-averaged GF vs sustained wind bias, (e) PX–MODIS LSM roughness length vs sustained wind bias, (f) observed station average GF (after high z_0 station adjustment) vs sustained wind bias, and (g) observed sustained wind vs GF-bias model residuals for 135 SDG&E stations. Each dot represents a station. A least squares fit (red line) is shown in each panel for reference, with R^2 values indicated. The vertical blue lines represent network-averaged values. (Units are m s^{-1} for winds, biases, and residuals; m for roughness lengths; and GF is nondimensional.)

than several minutes. Because of this, it is most sensible to compare model outputs to observed sustained winds and construct a reasonable gust parameterization. In this section, we pursue a simple gust model for individual stations and the SDG&E mesonet as a whole.

a. Station wind and gust predictions

Figure 12a presents the relationship between observed sustained winds and gusts for the six-event composite dataset, each point representing an SDG&E station's values after averaging over the 324 hourly observations. The no-intercept least squares fit to the entire dataset yields a slope (or GF) of 1.7 with an R^2 of 0.89. This suggests that adequate forecasts of sustained winds could result in skillful predictions of observed gusts in a simple manner via a constant gust factor, that being 1.7 in the present example. The relationship between the composite-averaged forecasted versus observed sustained winds is nearly 1:1 (Fig. 12b), despite the previously discussed tendency to overforecast sites with slower winds and underforecast those with stronger average values. The systematic bias has mainly served to increase the scatter ($R^2 = 0.62$).

Using the average forecasted winds to predict the observed *gusts* instead (Fig. 12c) also results in a slope of very nearly 1.7, but with an even larger R^2 (0.75) than that associated with the relationship between forecasted and observed winds. The noticeable decrease in scatter occurred precisely because some of the systematic biases in the sustained wind forecasts have been mitigated via the tacit presumption of a uniform gust factor. Recall that we have hypothesized that GFs higher than the all-station average (1.7) indicate locations that are relatively more sheltered (resulting in model overprediction of the sustained winds), and that sites with smaller than average GFs have local features that contribute to underprediction. The unresolvable local exposure issues that made the forecasted winds either too high or low have less impact when those same winds are used to anticipate gusts, which are ostensibly less influenced by those issues.

As a consequence, *biased* forecasts for the observed sustained winds can be used to make *unbiased* estimates of the observed gusts. This is demonstrated in Figs. 12d and 9c, which now plot forecasted versus observed gusts. The gust predictions were made by multiplying the sustained wind prediction for each site by a single, constant GF of 1.7, as suggested by Fig. 12c. There are still errors that could possibly be decreased via revisions to terrain and/or land-use characteristics, but again at this point the benefits are likely too small to justify the effort.

b. Network-averaged wind and gust predictions

Next, we pursue a very similar strategy to make predictions of network-averaged gusts from network-averaged sustained wind forecasts. Each point in Fig. 13a now represents one of the 324 observation times from the six Santa Ana wind events. Remarkably, after averaging over the 135-station network, the relationship between the observed winds and gusts has virtually no dispersion, with an R^2 of 0.99 and a slope of 1.7. Each station has its own GF at each point in time, which can vary for a variety of reasons (see below), but it remains that when these 324 network-averaged wind and gust pairs are plotted, there is no scatter.

This curious finding is not just a characteristic of Santa Ana wind events, as illustrated in Fig. 13b, which shows network-averaged winds versus gusts every 10 min over 12 consecutive months. Again, each of the 51 940 points on the plot represents a wind and a gust averaged over 135 stations at a single instant in time. Obviously, this 1-yr period contains a wide variety of weather conditions and surface wind directions, and yet the slope (1.7) and the R^2 (0.99) have remained unchanged. At each station, the GF may shift with the magnitude of the sustained wind, type of weather, and time of day and year, but when these data are combined into network averages, the variation is found to vanish.

Why this result has been obtained is not entirely understood. [This finding was first reported in Fovell and Cao (2014), and Cao and Fovell (2015) and Gallagher (2016) have demonstrated that this very low scatter also occurs in other homogeneous observational networks, such as the Dugway Proving Ground Mesonet in Utah, the NOAA/Air Resources Laboratory Field Research Division Mesonet in Idaho, and the Delaware Environmental Observing System Mesonet.] It is noted that, unlike some other networks or combinations of networks, the SDG&E mesonet is nearly homogeneous with respect to instrumentation, age of facilities, anemometer mounting height, and station-siting philosophy. At a single station, the GF can vary substantially with the sustained wind (e.g., FC17, their Fig. 3), but even at one site the variability can be considerably reduced when the wind–gust pairs are considered in bulk (not shown). The dispersion, as measured by the R^2 statistic, was already low, and this is true at all sites in the mesonet. The network average is then obtained by integrating over this collection of station datasets, each with relatively low dispersion. Perhaps this is simply a manifestation of "central tendency," given that the network is nearly static, with neither

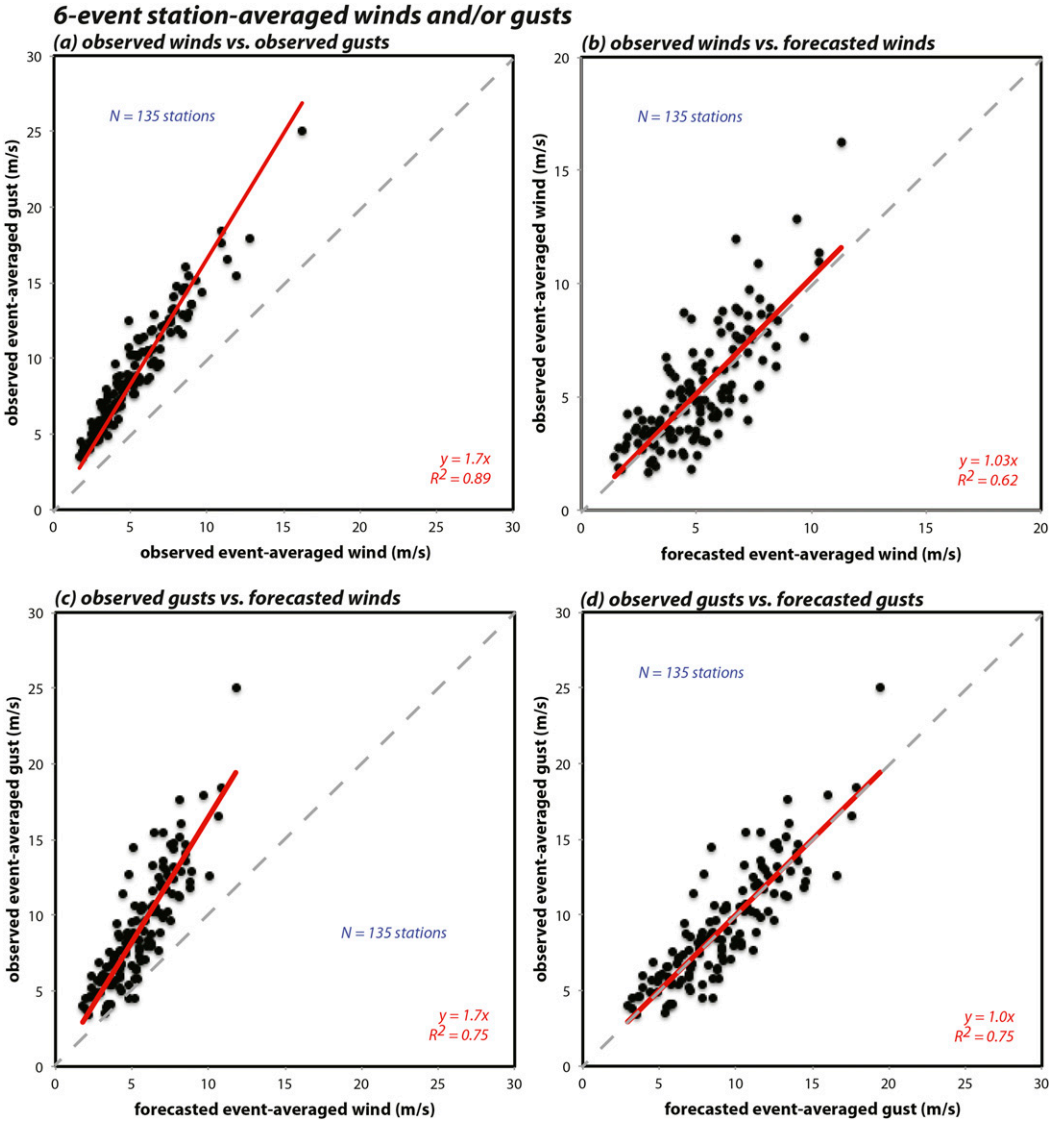

FIG. 12. Scatterplots of six-event-mean (a) observed event-averaged sustained wind vs observed event-averaged gust, (b) observed event-averaged sustained wind vs forecasted event-averaged sustained wind, (c) forecasted event-averaged sustained wind vs observed event-averaged gust, and (d) observed event-averaged gust vs forecasted event-averaged gust using a gust factor of 1.7 for the 135 SDG&E stations. Each dot represents a station. A zero-intercept least squares fit (red line) and the 1:1 line (dashed gray) are shown in each panel for reference, with slopes and R^2 values indicated. (All units are m s^{-1}.)

hardware nor station location changes over the period under examination.

In any case, the no-intercept least squares fit between the network-averaged observed and forecasted winds over the six Santa Ana events yielded a slope of 0.99, with a skillful R^2 of 0.80 (Fig. 13c). The PX–ACM2 configuration was already shown to consistently produce unbiased forecasts when averaged over the mesonet, which naturally averages out the previously discussed systematic biases. Applying a constant gust factor of 1.7 to these forecasted sustained winds to get predicted gusts yielded a scatterplot (Fig. 13d) that is visually comparable to Fig. 13c with respect to dispersion, possessing a slope of 1.0 and the R^2 being 0.82. Importantly, note that the network-averaged gust forecasts are unbiased.

Finally, we apply this constant GF concept to make predictions of network-averaged gusts for individual Santa Ana wind events. Figure 14 presents time series of the observed (dotted lines) and forecasted (solid lines) winds

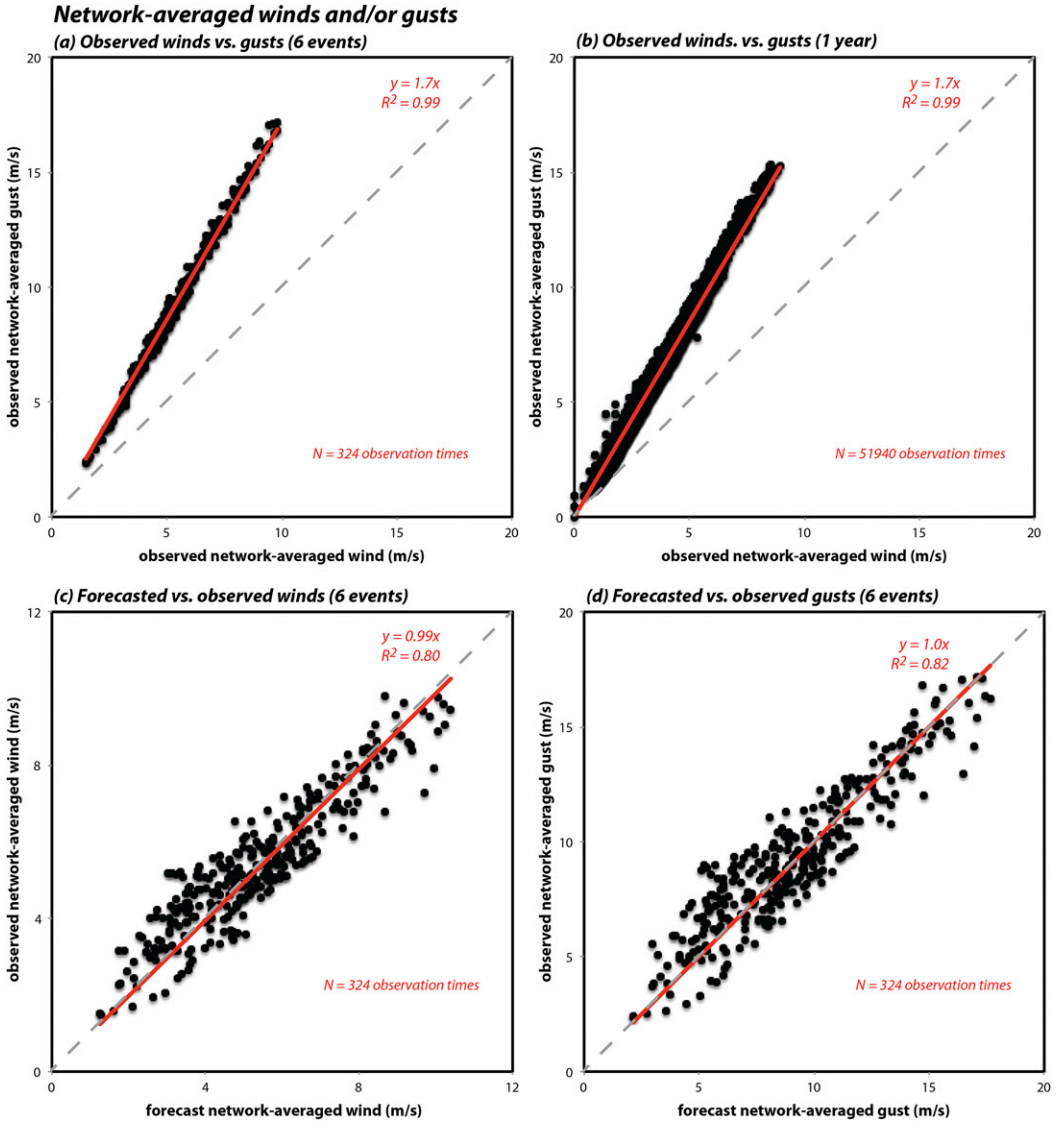

FIG. 13. Scatterplots of network-averaged (a) observed wind vs observed gust over the six events (324 observation times), (b) observed wind vs observed gust for 1 yr (51 940 observation times), (c) observed wind vs forecasted wind over the six events, and (d) observed gust vs forecasted gust over the six events. Each dot is a network average based on 135 SDG&E stations. A zero-intercept least squares fit (red line) and the 1:1 line (dashed gray) are shown in each panel for reference, with slopes and R^2 values indicated. (All units are $m\,s^{-1}$.)

(black) and gusts (red) over the SDG&E mesonet for the six Santa Ana episodes. Generally, the gust parameterization captures the amplitude and temporal evolution of the gusts for all six of the events quite well, especially the peaks. Obviously, performance depends upon the sustained wind predictions being correct in the first place. Gust overpredictions such as the first peak of the April 2014 event (Fig. 14a) and underpredictions such as the October 2013 (Fig. 14e), the February 2013 (Fig. 14f), and the January 2015 (Fig. 14d) events are consistent with the corresponding sustained wind biases. However, all can be considered to be rather skillful gust forecasts, especially for the May 2014 (Fig. 14b) and the February 2015 (Fig. 14c) episodes. Averaged over the six events, the SDG&E network forecasted gust bias is merely $-0.2\,m\,s^{-1}$, which is comparable to the averaged forecasted sustained wind bias ($\sim 0.0\,m\,s^{-1}$).

6. Discussion and summary

We seek to obtain skillful gust forecasts in San Diego County during Santa Ana wind events. These episodes are relatively frequent during the winter half-year (Raphael 2003; Jones et al. 2010) and the dry

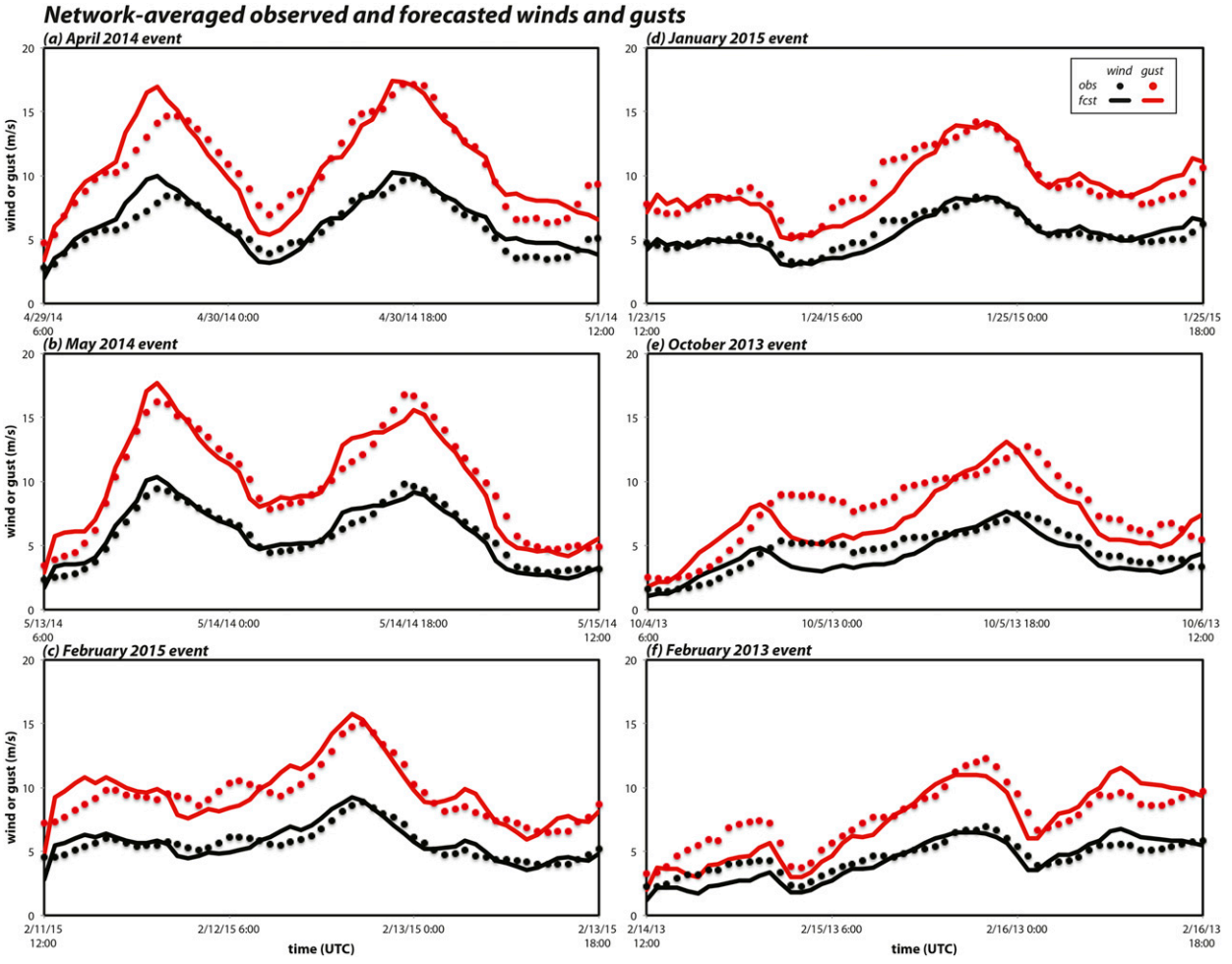

FIG. 14. Time series of network-averaged observed (black dots) and predicted (red curves) 6.1-m sustained winds (black; m s^{-1}) and gusts (red; m s^{-1}) for the (a) April 2014, (b) May 2014, (c) February 2015, (d) January 2015, (e) October 2013, and (f) February 2013 events (see Table 1). The GF used to forecast gusts is 1.7, based on the network-averaged wind vs the network-averaged gust.

(sometimes hot) winds contribute greatly to the fire hazard (Rolinski et al. 2016). As models of the present type cannot resolve gusts, this effort must start with making sure sustained wind predictions are as accurate as possible and then deducing gusts via an algorithm or parameterization of some kind. We attempted to partition the forecast sustained wind bias into ''fixable'' and ''unavoidable'' components. The fixable part may be addressed via the model configuration, including refinements of physical parameterizations, among many other things, leaving the unavoidable portion to be mitigated via postprocessing.

Utilizing hundreds of WRF simulations made for CF16, FC17, and herein, and the high-density SDG&E surface mesonet, we have demonstrated that the wind speeds and flow patterns during moderately strong to strong Santa Ana episodes are sensitive to horizontal resolution, stochastic perturbations (cf. Berner et al. 2011), and model physics selections, especially the land surface models (LSMs) that determine surface roughness. The Pleim–Xiu LSM scheme emerged as the best overall with respect to sustained wind forecast skill, having a nearly zero bias when averaged over the network and multiple events, largely because of its unique treatment of surface roughness in the shrubland categories that dominate the west-facing slopes in the SDG&E network. In particular, PX uses substantially larger z_0 values for those categories and holds them fixed throughout the year, while most WRF simulations resulted in positive wind biases as they treated the surface as being too smooth. Altering other LSMs such as Noah to mimic the PX roughness lengths improved their MAE and bias scores, confirming this is the dominant factor. This is important, because other land surface

treatments could provide superior temperature and humidity reconstructions (cf. Cao 2015), which are crucial within the context of Southern California fire weather predictions (e.g., Rolinski et al. 2016).

It is intuitive that increasing the surface roughness would slow down the winds, at least overall. It may also change the nature of the downslope flow, in ways that could be detectable even with limited (i.e., surface only) observational data, thanks to the relatively dense SDG&E mesonet. The hydraulic jump that occurred during the first phase of the February 2013 episode is one such example. The jump was consistently present only in simulations having relatively rougher surfaces, although it could occasionally emerge in configurations with lower z_0 values via stochastic forcing. The downslope extent of strong surface winds is another, related example. This was larger with LSMs specifying lower roughness lengths and in these cases contributed to the positive network-averaged wind biases.

Even a model configuration that yielded negligible bias when averaged over multiple events and stations was still found to possess systematic errors at individual sites. These errors were shown to be inversely correlated with the average observed wind: the model tended to overpredict locales with weaker winds while underpredicting speeds where measured winds were stronger. Forecast sustained wind speed bias was also seen to be correlated with, and proportional to, the observed gust factor (GF), the ratio of the observed sustained wind and gust. This is mainly true since the GF and sustained wind are (negatively) correlated: sites with larger GFs tend to have slower winds.

We interpreted the difference between a given station's GF and the network average (1.7 for the SDG&E mesonet as a whole, with virtually no scatter or weather dependence) as a measure of very localized anemometer exposure that cannot be captured even at reasonably high spatial resolution and, thus, an unavoidable component of the bias. Sites with obstacles or landforms that tend to slow the temporally averaged sustained wind more than the transient, impulsive gusts would result in GFs that are larger than the network average, and its winds would be more likely to be overpredicted. In contrast, stations having local features that help enhance the winds relative to the gusts would have lower GFs and be underpredicted. In practice, both were found to be the case, as after using the observed GF to predict sustained wind forecast bias, the remainder of the bias was independent of observed wind speeds.

The gust algorithm that emerged from this study is extremely simple: we multiply sustained wind forecasts at each site by the network average GF of 1.7. Gust factors vary in space and time, but gust forecasts made this way had more skill than the sustained wind predictions. This was because using a constant value partially mitigated the unavoidable local exposure bias discussed above. Applying a constant value of 1.7 to the network-averaged wind forecasts also proved to be beneficial, which we demonstrated for the six Santa Ana events under examination. Averaging the sustained wind predictions over the network already filtered out the exposure bias.

Certainly, a more sophisticated treatment of predicting gusts could be designed, but we are encouraged that an attractively simple gust parameterization could actually improve upon already skillful sustained wind forecasts in this region. That being said, we emphasize that the specific 1.7 factor is a consequence of SDG&E mesonet design factors including (and not limited to) mounting height and averaging intervals, and the optimal multiplier will certainly be different for networks with other characteristics. We also caution that our study area experiences little in the way of convective weather, and so a gust parameterization this simple may not work well in areas where thunderstorms are common.

Acknowledgments. This research was sponsored by the San Diego Gas and Electric Company. The authors thank SDG&E meteorologists Brian D'Agostino and Steven Vanderburg for their assistance, access to data, and useful discussions throughout this study, and three anonymous reviewers for their comments. Dr. Travis Wilson helped design many scripts involved in the model verification section, which employed the Model Evaluation Tools (MET) packages from the Development Testbed Center. The authors also 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.

REFERENCES

Angevine, W. M., H. Jiang, and T. Mauritsen, 2010: Performance of an eddy diffusivity–mass flux scheme for shallow cumulus boundary layers. *Mon. Wea. Rev.*, **138**, 2895–2912, https://doi.org/10.1175/2010MWR3142.1.

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

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. *Mon. Wea. Rev.*, **139**, 1972–1995, https://doi.org/10.1175/2010MWR3595.1.

Bougeault, P., and P. Lacarrere, 1989: Parameterization of orography-induced turbulence in a mesobeta-scale model. *Mon. Wea. Rev.*, **117**, 1872–1890, https://doi.org/10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.

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

Bretherton, C. S., and S. Park, 2009: A new moist turbulence parameterization in the Community Atmosphere Model. *J. Climate*, **22**, 3422–3448, https://doi.org/10.1175/2008JCLI2556.1.

Cao, Y., 2015: The Santa Ana winds of Southern California in the context of fire weather. Ph.D. thesis, University of California, Los Angeles, 173 pp., https://escholarship.org/uc/item/31p6863n.

——, and R. G. Fovell, 2015: Improving gust forecasting through sustained wind bias corrections. *16th Annual WRF Users' Workshop*, Boulder, CO, NCAR, http://www2.mmm.ucar.edu/wrf/users/workshops/WS2015/posters/p49.pdf.

——, and ——, 2016: Downslope windstorms of San Diego County. Part I: A case study. *Mon. Wea. Rev.*, **144**, 529–552, https://doi.org/10.1175/MWR-D-15-0147.1.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Conil, S., and A. Hall, 2006: Local regimes of atmospheric variability: A case study of Southern California. *J. Climate*, **19**, 4308–4325, https://doi.org/10.1175/JCLI3837.1.

Durran, D. R., 2003: Downslope winds. *Encyclopedia of Atmospheric Sciences*, G. North, J. Pyle, and F. Zhang, Eds., Elsevier, 644–650.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, https://doi.org/10.1029/2002JD003296.

Fovell, R. G., and Y. Cao, 2014: Wind and gust forecasting in complex terrain. *15th WRF Users Workshop*, Boulder, CO, NCAR, 5A.2, http://www2.mmm.ucar.edu/wrf/users/workshops/WS2014/ppts/5A.2.pdf.

——, and ——, 2017: The Santa Ana winds of Southern California: Winds, gusts, and the 2007 Witch fire. *Wind Struct.*, **24**, 529–564.

Gallagher, A. A., 2016: The network average gust factor, its measurement and environmental controls, and role in gust forecasting. M.S. thesis, Dept. of Atmospheric and Environmental Sciences, University at Albany, State University of New York, 139 pp.

Grenier, H., and C. S. Bretherton, 2001: A moist PBL parameterization for large-scale models and its application to subtropical cloud-topped marine boundary layers. *Mon. Wea. Rev.*, **129**, 357–377, https://doi.org/10.1175/1520-0493(2001)129<0357:AMPPFL>2.0.CO;2.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339, https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

——, J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120, https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318, https://doi.org/10.1175/MWR3199.1.

Hughes, M., and A. Hall, 2010: Local and synoptic mechanisms causing Southern California's Santa Ana winds. *Climate Dyn.*, **34**, 847–857, https://doi.org/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. *J. Geophys. Res.*, **113**, D13103, doi:10.1029/2008JD009944.

Jackson, P. L., G. Mayr, and S. Vosper, 2013: Dynamically-driven winds. *Mountain Weather Research and Forecasting*, F. K. Chow, S. F. J. De Wekker, and B. J. Snyder, Eds., Springer Atmospheric Sciences, Springer-Verlag, 121–218.

Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Jones, C., F. Fujioka, and L. M. V. Carvalho, 2010: Forecast skill of synoptic conditions associated with Santa Ana winds in Southern California. *Mon. Wea. Rev.*, **138**, 4528–4541, https://doi.org/10.1175/2010MWR3406.1.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Klemp, J. B., and D. R. Lilly, 1975: The dynamics of wave-induced downslope winds. *J. Atmos. Sci.*, **32**, 320–339, https://doi.org/10.1175/1520-0469(1975)032<0320:TDOWID>2.0.CO;2.

Mitsuta, Y., and O. Tsukamoto, 1989: Studies on spatial structure of wind gust. *J. Appl. Meteor.*, **28**, 1155–1160, https://doi.org/10.1175/1520-0450(1989)028<1155:SOSSOW>2.0.CO;2.

Moritz, M. A., T. J. Moody, M. A. Krawchuk, M. Hughes, and A. Hall, 2010: Spatial variation in extreme winds predicts large wildfire locations in chaparral ecosystems. *Geophys. Res. Lett.*, **37**, L04801, doi:10.1029/2009GL041735.

Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31, https://doi.org/10.1023/B:BOUN.0000020164.04146.98.

Niu, G.-Y., and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): I. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, D12109, https://doi.org/10.1029/2010JD015139.

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. *Bound.-Layer Meteor.*, **11**, 355–361, https://doi.org/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. *J. Appl. Meteor. Climatol.*, **46**, 1383–1395, https://doi.org/10.1175/JAM2539.1.

——, 2007b: A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model. *J. Appl. Meteor. Climatol.*, **46**, 1396–1409, https://doi.org/10.1175/JAM2534.1.

——, and A. Xiu, 1995: Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. *J. Appl. Meteor.*, **34**, 16–32, https://doi.org/10.1175/1520-0450-34.1.16.

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

Rolinski, T., S. B. Capps, R. G. Fovell, Y. Cao, B. J. D'Agostino, and S. Vanderburg, 2016: The Santa Ana wildfire threat index:

Methodology and operational implementation. *Wea. Forecasting*, **31**, 1881–1897, https://doi.org/10.1175/WAF-D-15-0141.1.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102, https://doi.org/10.1256/qj.04.106.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., http://dx.doi.org/10.5065/D68S4MVH.

Smirnova, T. G., J. M. Brown, and D. Kim, 2000: Parameterization of cold-season processes in the MAPS land-surface scheme. *J. Geophys. Res.*, **105**, 4077–4086, https://doi.org/10.1029/1999JD901047.

Smith, C. M., and E. D. Skyllingstad, 2011: Effects of inversion height and surface heat flux on downslope windstorms. *Mon. Wea. Rev.*, **139**, 3750–3764, https://doi.org/10.1175/2011MWR3619.1.

Sukoriansky, S., B. Galperin, and V. Perov, 2006: A quasi-normal scale elimination model of turbulence and its application to stably stratified flows. *Nonlinear Processes Geophys.*, **13**, 9–22, https://doi.org/10.5194/npg-13-9-2006.

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, Trans. Amer. Geophys. Union*, **85**, 289–296, https://doi.org/10.1029/2004EO310001.

Wieringa, J., 1973: Gust factors over open water and built-up country. *Bound.-Layer Meteor.*, **3**, 424–441, https://doi.org/10.1007/BF01034986.

Wilson, T. H., and R. G. Fovell, 2016: Modeling the evolution and life cycle of stable cold pools. *Wea. Forecasting*, **31**, 1753–1769, https://doi.org/10.1175/WAF-D-16-0108.1.

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