Statistical Modeling of Downslope Windstorms in Boulder, Colorado

ANDREW E. MERCER, MICHAEL B. RICHMAN, AND HOWARD B. BLUESTEIN

School of Meteorology, University of Oklahoma, Norman, Oklahoma

JOHN M. BROWN

NOAA/Earth System Research Laboratory/Global Systems Division, Boulder, Colorado

(Manuscript received 16 August 2007, in final form 29 April 2008)

ABSTRACT

Downslope windstorms are of major concern to those living in and around Boulder, Colorado, often striking with little warning, occasionally bringing clear-air wind gusts of 35–50 m s^{-1} or higher, and producing widespread damage. Historically, numerical models used for forecasting these events had lower than desired accuracy. This observation provides the motivation to study the potential for improving windstorm forecasting through the use of linear and nonlinear statistical modeling techniques with a perfect prog approach. A 10-yr mountain-windstorm dataset and a set of 18 predictors are used to train and test the models. For the linear model, a stepwise regression is applied. It is difficult to determine which predictor is the most important, although significance testing suggests that 700-hPa flow is selected often. The nonlinear techniques employed, feedforward neural networks (NN) and support vector regression (SVR), do not filter out predictors as the former uses a hidden layer to account for the nonlinearities in the data, whereas the latter fits a kernel function to the data to optimize prediction. The models are evaluated using root-mean-square error (RMSE) and median residuals. The SVR model has the lowest forecast errors, consistently, and is not prone to creating outlier forecasts. Stepwise linear regression (LR) yielded results that were accurate to within an RMSE of 8 m s^{-1}; whereas an NN had errors of 7–9 m s^{-1} and SVR had errors of 4–6 m s^{-1}. For SVR, 85% of the forecasts predicted maximum wind gusts with an RMSE of less than 6 m s^{-1} and all forecasts predicted wind gusts with an RMSE of below 12 m s^{-1}. The LR method performed slightly better in most evaluations than NNs; however, SVR was the optimal technique.

1. Introduction

Downslope windstorms are a common occurrence in cities that are located on the lee side of the Rocky Mountains, such as Boulder, Colorado (Fig. 1). According to recent data from the National Climatic Data Center, winds of 35–50 m s^{-1} or more are commonly observed near Boulder about 10 times each year, causing extensive tree and property damage. Despite these impacts, no modeling technique is used operationally that predicts these wind events accurately on a consistent basis. A previous study by Leptuch (2001) attempted to model mountain windstorms with a set of 18 predictors using linear models. Leptuch's linear modeling establishes the foundation for the present research.

According to Scheetz et al. (1976) and Brown (1986), two general classes of windstorms are observed to the lee of the Rocky Mountains. The first is the classic Chinook case, which is referred to in this study as a "prefrontal" windstorm. This type of storm is most common in late fall and winter and occurs with WNW to NW flow at 500 hPa ahead of a rapidly moving shortwave trough. These windstorms are accompanied often by the formation of a surface lee trough over the northern and central high plains (Fig. 2). The second class of windstorm is referred to as "postfrontal" because it is associated with strong cold advection behind a surface cold front that has crossed the Rockies from west to east. Associated with this cold front is a sharp, mobile trough or ejecting cutoff low aloft that couples with warm, low static-stability air over the high plains to produce strong cyclogenesis (Fig. 3). Because of the marked cold advection characteristic of these events, a postfrontal windstorm is sometimes referred to as a bora, after the strong, cold downslope wind that

Corresponding author address: Andrew E. Mercer, School of Meteorology, University of Oklahoma, 120 David L. Boren Blvd., No. 5632, Norman, OK 73072.
E-mail: amercer@rossby.metr.ou.edu

DOI: 10.1175/2008WAF2007067.1

© 2008 American Meteorological Society

FIG. 1. Topographical map of the Boulder area. Boulder as well as the three sounding sites (DEN, GJT, and LND) are indicated.

blows along the north and east shore of the Adriatic Sea.

Various theories have been proposed to explain strong downslope winds (see Durran 1990 for a review). It is now generally accepted that the explanation is tied to nonlinear steepening or overturning in large-amplitude vertically propagating gravity waves, triggered by passage of flow over the mountain ridge of the Continental Divide. The occurrence of vertically propagating waves and their steepening or breaking can occur as a result of a "mean-state critical layer," in which the component of the synoptic-scale flow normal to the mountain ridge changes sign above the top of the mountain ridge. However, this situation is rare for windstorms on the east side of the Rockies, occurring only every few years. A more common scenario is to have a westerly component to the flow at all levels from the mountaintop ridge through the lower stratosphere. Consequently, for the Boulder region, the normal component of flow has the same sign through this depth. Under these conditions, nonlinear wave steepening or wave breaking is favored by a decrease with height or, at most, a slow increase of the component of flow normal to the mountain ridge of the Continental Divide, and a decrease of static stability with height in the troposphere in the upstream flow that is able to cross this ridge, rather than being diverted northward into Wyoming. This wave steepening or breaking leads to a region of near-zero static stability aloft above the higher static stability air crossing the ridge. This permits the low-level air crossing the ridge to become dynamically analogous to hydraulic flow transitioning to supercritical as it crosses a weir or dam; therefore, it accelerates down the lee slope of the ridge.

One of the earlier theories (Klemp and Lilly 1975; also discussed in Durran 1990) was based on reflection of upward-propagating gravity wave energy at the tropopause, leading to constructive interference at the

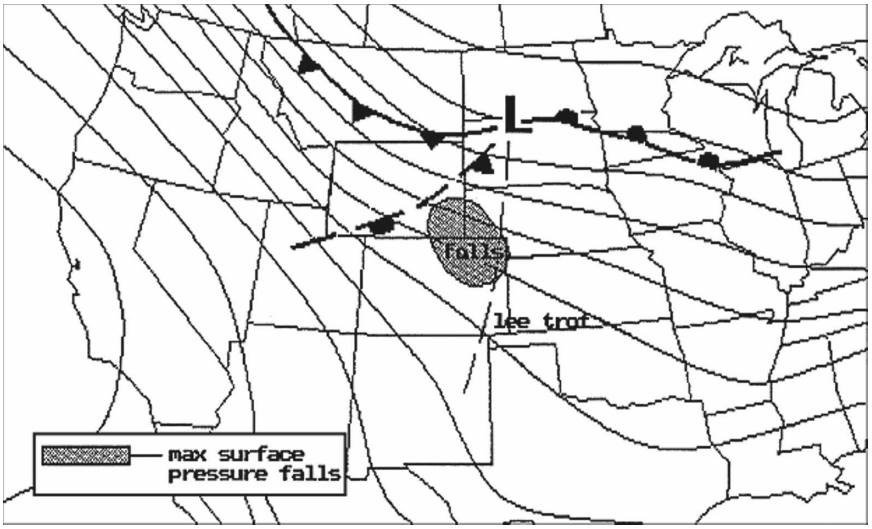

FIG. 2. Schematic composite chart for the typical synoptic pattern associated with a "prefrontal" windstorm at Boulder. Surface frontal positions are shown, along with the locations of the surface warm, cold, and Pacific occluded fronts. The shaded area labeled "falls" indicates the location of the largest 3-h falls in surface-station pressure, which typically has tracked eastward across southern WY. The typical position of a "lee trough" on the high plains is also indicated. Thin solid curved lines are 500-hPa height contours. [Adapted from Scheetz et al. (1976).]

surface in a suitably tuned atmosphere. Klemp and Lilly devised an algorithm to predict this reflection using as input forecast soundings from the old Limited-Area Fine Mesh (LFM) operational model. This was the first attempt to formulate an algorithm based on theory and intended for operational use by forecasters. Results were not very reliable; peak wind predictions were sometimes twice the observed. Hence, the method did not attain sufficient credibility with National Weather Service (NWS) forecasters for it to be used routinely. In part, this may have been due to the crude vertical resolution of the LFM-predicted soundings in the 1970s.

The period of exceptionally damaging downslope windstorms in Boulder during the late 1960s and 1970s led to a number of other attempts to devise physically based and statistical algorithms to provide guidance for the prediction of these storms. Scheetz et al. (1976) devised a pattern recognition algorithm based on synoptic typing and model-predicted changes in 500-hPa vorticity. Sangster (1977) developed a method using the 850- and 700-hPa geopotential heights at the upstream upper-air stations of Salt Lake City, Utah; Grand Junction, Colorado; and Lander Wyoming (site now moved to Riverton, Wyoming). Essentially, this method amounts to the computation of the geostrophic wind

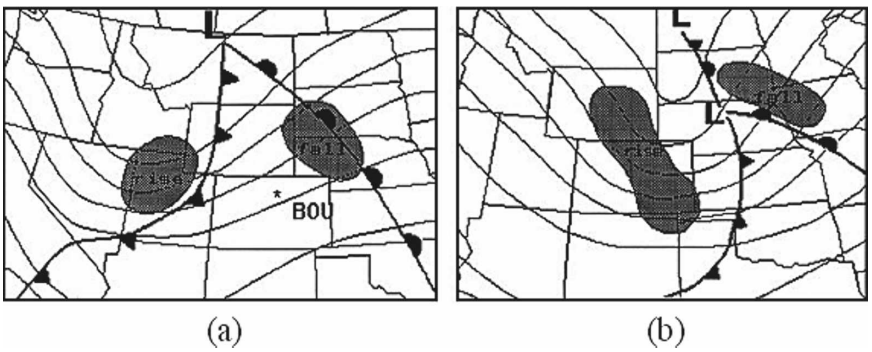

FIG. 3. As in Fig. 2, but for the "postfrontal" windstorm, showing typical synoptic evolution: (a) and (b) are about 12 h apart, with (a) prior to strong winds in Boulder (BOU, asterisk)] and (b) near or just after the strongest high wind in Boulder. Rise–fall couplet in surface-station pressure is shown, with the rise area west of the cold front.

component from about 300°. Results of Sangster (1977) were built upon by Brown (1986), who developed a decision tree that had, as an essential ingredient, the strength and direction of the [then LFM and Nested Grid Model (NGM)] model-predicted 700-hPa geostrophic wind. Additionally, this decision tree incorporated relationships between vertical wind shear, stability, and windstorm occurrence gleaned by the inspection of numerous windstorm cases. These relationships turned out to be consistent with the theoretical work of Durran (1986) that helped establish the hydraulic analog of strong downslope flow previously described. This decision tree for Boulder, along with a statistical method for Fort Collins developed by Weaver and Phillips (1990), were packaged for operational use by the Denver NWS Forecast Office (Rockwood et al. 1992). This package met with modest initial success and has been available for use on the Automated Weather Information Processing System (AWIPS) at the Denver office for a number of years.

It has been noted that the presence of a surface-based, shallow, cold, stable air mass left behind by an antecedent synoptic event can inhibit or delay the onset of strong downslope winds in Boulder and Fort Collins when conditions are otherwise favorable. This has been demonstrated through two-dimensional numerical simulations of windstorms by Lee et al. (1989), through insertion of a shallow cold pool downstream of a ridge. This inhibiting factor is (subjectively) included in both the Brown (1986) decision tree and the Weaver and Phillips (1990) statistical procedure [and incorporated in the Rockwood et al. (1992) package for NWS use].

The operational NWS numerical weather prediction (NWP) guidance has evolved over the years since these methods were developed and introduced. In particular, as models improved in horizontal resolution, their resolvable scales reached down into horizontal scales where mountain-induced gravity waves began to appear in model output. The appearance of mountain waves created difficulties for statistical methods such as that of Brown (1986) using pressure or height gradients and necessitated making adjustments to maintain the usability of these procedures. It was decided, at the outset of this study, that the best way to circumvent this problem (as well as related problems that come with frequent updates to the operational models) was to use the "perfect prog" approach, in which it is assumed that the forecast model is "perfect" and the dependent data sample used to derive the statistical technique is based solely on observations. Consistent with the perfect prog approach, we chose to use the actual rawinsondes, rather than, for example, the North American Regional Reanalysis dataset, to define the subsynoptic environment, since we wanted to capture aspects of the stratification and hodograph upstream of the Continental Divide that might be missed by the use of isobaric data at prescribed pressure levels. Given this general perfect prog approach, the prediction scheme described here takes 3D output fields from an operational NWP model forecast to construct forecast soundings at selected rawinsonde sites. For one or more of the statistical models, summarized in section 2, this scheme calculates the predictors from these soundings. For each statistical model, the predictors produce a forecast maximum wind gust over a specified 3-h time interval.

As yet, this prediction scheme has not been applied operationally, although we do present results of a small number of forecasts based on output from the operational Nonhydrostatic Mesoscale Model (NMM) dynamical core version of the Weather Research and Forecast model (WRF), which is used as the forecast component of the North American Mesoscale (NAM) modeling system run operationally at the National Centers for Environmental Prediction (NCEP).

In the next section we describe the predictors and the statistical procedures used. Most of these procedures are being applied to the downslope windstorm problem for the first time. In section 3, the results of the various models are presented, as well as results from forecast soundings. The work is summarized in section 4.

2. Methods

a. Data

This study makes use of NWS 0000 and 1200 UTC soundings from five sites: North Platte, Nebraska; Denver, Colorado; Grand Junction, Colorado; Lander, Wyoming; and Salt Lake City, Utah. The soundings are valid between 1 January 1969 and 31 December 1978. However, owing to the climatology of windstorms, no soundings between 15 May and 15 September are used.

The verification wind gust data are obtained from the (then) National Bureau of Standards (NBS), in Boulder. Following the extremely destructive storms of 7 January 1969 (Bergen and Murphy 1978) and 10–11 January 1972 (Doyle et al. 2000), there was an intensive ongoing effort to study Boulder windstorms. Strip-chart recording anemometers were set up at a number of private residences and schools, the NCAR Mesa Laboratory (NCAR-ML), and NBS. However, inspection of the strip charts by one of us (J. M. Brown) many years ago for locations other than NBS and NCAR-ML revealed many gaps, timing uncertainties, and other problems, particularly after 1975. These data were discarded

once it was realized that the effort to render them into usable form was prohibitive, given the resources available. The NCAR-ML site was atop the north tower of the Mesa Laboratory at an elevation of about 1880 m, 150 m or more in elevation above other sites in Boulder, and received many windstorms that did not affect the main residential sections of Boulder. It was therefore regarded as insufficiently representative to be used for purposes of this or the earlier studies of Brown (1986) and Rockwood et al. (1992). The NBS data collection is unique in that the Guard Force at the NBS site was charged with the responsibility to record by hand every 3 h (at 0000, 0300, 0600, . . . , 2100 UTC) the peak gust during the previous 3-h period, if this peak gust equaled or exceeded 20 kt (about 10 m s^{-1}). A paper copy of these tabulations was obtained (W. E. Sangster 1990, personal communication) and put into electronic form. (The original strip charts upon which these real-time tabulations are based had disappeared well before any of us were aware that the NBS tabulated data still existed; Sangster had only the tabulated sheets.) After conversion to electronic form, the data were scanned by eye several times for errors against the original tabulations and discrepancies were corrected in favor of the original tabulations. A further check was performed for the strongest storms by comparing the NBS data against NCAR-ML strip charts. In addition, for all wind events with gusts of 40 mi h^{-1} (18 m s^{-1}) or more, microfilm constant-pressure charts were inspected to determine if the wind gusts were consistent (based on synoptic experience) with the synoptic pattern (e.g., a strong wind was not associated with weak or easterly flow at 700 hPa). Using these latter two "sanity" checks, no clearly erroneous data were found. For these reasons, we have confidence that the verification data are free of errors that could seriously compromise the results reported in later sections.

The NBS site is located in a residential area of Boulder, 1.5–3 km to the east of the base of the Front Range foothills. The local terrain slopes gently upward to the SW–W before the foothills are abruptly encountered. A three-cup anemometer was located atop the northeast end of the main building on the site, roughly 15 m above ground level, at about 1650-m elevation. This anemometer is of unknown type, but one of us (J. M. Brown) recalls (ca. 1980) that it was of a generic rugged design, likely characterized by a distance constant rather larger than that typical of current designs.

The unique and valuable aspects of this dataset are that it is homogeneous over a 10-yr period, and it is representative of wind conditions where many people live. This homogeneity that is not tainted by secular changes in the environment (e.g., growth of vegetation, new construction) or instrumentation changes is rare. Such factors have the potential to complicate the analysis and the dataset.

Since we are interested in relating sounding properties upstream (west) of the Continental Divide to windstorms occurring just downstream, we chose to correlate the sounding data to Boulder gust occurrence near or after the launch time of the NWS rawinsonde balloons (typically about 1115 and 2315 UTC). Given the nature of the 3-h periods, for which we had peak gusts, we chose to relate the soundings to peak gusts occurring in seven different time periods: 0–12, 0–6, 6–12, 0–3, 3–6, 6–9, and 9–12 h after the nominal 0000 or 1200 UTC sounding times. These represent the time period over which the peak gust was observed. For example, in a 1200 UTC sounding, the 0–6-h peak gust would be the gust observed between 1200 and 1800 UTC. With these different combinations, prediction could take place over different time periods spanning the 12-h interval between soundings. Roughly 5% of daily wind data are missing and thus are not available for our study.

b. Parameters

The following is a list of the 18 predictors, each of which is followed by a description and brief definition. In the following discussions, low levels are classified as the layer between 3100 and 5600 m (roughly 700–500 hPa), midhigh levels are the layer between 5600 and 7400 m (roughly 500–400 hPa), and high levels are the layer between 5600 and 9400 m (roughly 500–300 hPa). Parameters that depend on the geostrophic wind at mandatory levels (700 hPa, 500 hPa, etc.) are determined by using the geostrophic wind obtained by fitting a plane to the geopotential height reported from the rawinsonde launches at Grand Junction (GJT; Fig. 1), Lander (LND; Fig. 1), and North Platte, and applying the geostrophic wind relation to the height gradient to obtain the geostrophic wind speed and direction at the center of the plane. [The reader will note that the geopotential height at Denver (DEN) is not used. During windstorms, the 700-hPa height at DEN is often lower than what would be anticipated from synoptic considerations, contaminating the geostrophic wind calculation. Since use of the geostrophic wind is only appropriate as a descriptor of the synoptic scale, we elected to choose a somewhat larger triangle of stations rather than suffer this contamination.] Parameters that depend on the observed wind hodograph or temperature stratification upstream of the Continental Divide [see sections 2b(5), 2b(9), 2b(11)–2b(14), and 2b(17)] are derived from the individual rawinsonde launches at GJT and LND and a mean of the two. This leads to

three sets of parameters for each of the two types of windstorms.

1) TEMPERATURE ADVECTION (K DAY^{-1})

The first variable calculated is the 700-hPa geostrophic temperature advection. Assuming thermal wind balance, it is estimated using the 700-hPa geostrophic wind and the 700-hPa temperature gradient. Positive values represent warm advection, and negative values cold advection. This parameter is used in classification between prefrontal and postfrontal windstorm types. The temperature advection at 700 hPa is important as a simple, approximate indicator of the sign of low-level quasigeostrophic vertical motion.

2) 700-HPA GEOSTROPHIC WIND DIRECTION (°)

The 700-hPa geostrophic wind direction is determined to be significant since 700 hPa is the pressure of the rawinsonde mandatory reporting level closest to mountaintop level; previous studies (Brown 1986, Rockwood et al. 1992) found that Boulder high-wind events rarely occurred with 700-hPa geostrophic wind direction outside the range of (240°, 350°).

3) 700-HPA GEOSTROPHIC WIND SPEED (M S^{-1})

Observational geostrophic wind data indicate that there is a minimum value of geostrophic wind speed at the ridgetop level below which windstorms do not occur, roughly 9 m s^{-1}.

4) DEN $\overline{U}$ (M S^{-1})

This wind component is an approximation to the flow component normal to the terrain west of the Continental Divide west of Boulder. It is derived from the Denver rawinsonde. The direction of $\overline{U}$ is typically westerly, but for the Boulder area it backs linearly with height (290° below 3100 m to 270° above 4200 m) due to the orientation of the terrain. Inspection of a topographic map (Fig. 1) reveals that the Front Range west of Boulder (crest approximately 3700–4200 m) is oriented N–S. However, when considering the state as a whole, the 3350-m (~11 000 ft) elevation contour west of the Continental Divide tends to be oriented NNE–SSW (Fig. 1). To account for this, we take the perpendicular used to compute $\overline{U}$ as backing linearly with height from 290° at and below 3100-m altitude to 270°at and above 4200 m.

5) LOCAL $\overline{U}$ (M S^{-1})

This parameter is computed identically to $\overline{U}$DEN, but for either GJT or LND, or the average of the two.

6) 700–500-HPA GEOSTROPHIC WIND DIRECTION SHEAR (°)

This parameter is defined as the direction of the 700-hPa geostrophic wind minus the direction of the 500-hPa geostrophic wind. If it is positive, it corresponds to cold-air advection, and if it is negative, it shows warm-air advection. Advection is significant as it is used to diagnose the synoptic type of the storm as either prefrontal or postfrontal (section 1 and Figs. 2 and 3). A postfrontal case is defined as one characterized by a sounding containing more than 15° of backing in this layer, while all other cases were deemed as prefrontal.

7) RATIO OF THE 700-HPA WIND SPEED TO THE 700-HPA GEOSTROPHIC WIND SPEED AT DENVER

This quantity diagnoses the degree of ageostrophy at DEN. The total wind is selected, as opposed to normal-to-barrier wind, so that the results give the departure from geostrophy in cases where the winds are both normal to the barrier and parallel to the barrier. It is expected to depart substantially from unity when mountain waves are present.

8) DIFFERENCE BETWEEN THE 700-HPA WIND DIRECTION AND THE 700-HPA GEOSTROPHIC WIND DIRECTION (°)

As with 2b(7), this parameter is used to determine whether the DEN sounding site is contaminated by mountain waves or other local wind perturbations.

9) MOUNTAINTOP RELATIVE HUMIDITY

Downslope winds can be augmented by convective outflow. Such an event may occur when evaporative cooling from precipitation falling from a foehn cloud intensifies the downslope wind, leading to strong surface wind gusts. One such event occurred on 3 July 1993, in Fort Collins, Colorado, which was a rare summertime windstorm event that was the result of a poorly forecast surface low over western Nebraska combined with snowfall at higher elevations west of Fort Collins (Cotton et al. 1995). This event resulted in the strongest wind gusts limited to regions of deep cloudiness associated with the snowfall, suggesting a link between the two. This type of event is very difficult to forecast using objective methods, and relative humidity in upstream soundings is chosen to account for these situations. High relative humidity can also effectively reduce the static stability at mountaintop (important mainly for prefrontal windstorms). This parameter is calculated by averaging the relative humidity from an

upstream sounding (GJT, LND, or the average of the two) between 3100- and 3700-m altitude.

10) CROSS-MOUNTAIN HEIGHT DIFFERENCE (M)

This quantity is the difference between the 850-hPa height (m) at GJT and that at DEN. The height difference between the two locations implies a pressure gradient force between the two sites. Based purely on synoptic reasoning, a large value implies high winds, whereas a small value implies weak synoptic flow and lighter winds.

11) STATIC-STABILITY RATIO

This parameter is the ratio of the square of the mean-layer Brunt–Väisälä parameter, N^2, of low levels to that at high levels. It is applied to diagnose the change with height of the static stability in the atmosphere. Decreasing static stability with height within the troposphere upstream of the mountain range, with high static stability near ridgetop level, is thought to be favorable for windstorm occurrence (Durran 1990). Thus, values of the static stability ratio greater than one are favorable for windstorms.

12) FROUDE HEIGHT (KM)

This parameter is the height at which the Froude number equals unity. This quantity represents the elevation above sea level on the upstream slope of the mountain above which the flow passes over the ridge instead of being blocked, or to go around it (Smith 1989), Overland and Bond 1995). The Froude height, H, is calculated by

$$\mathrm{Fr} = \int_{H}^{Z_{\mathrm{mountain}}} \frac{N}{\overline{U}}\, dz = 1, \tag{1}$$

where N is the Brunt–Väisälä frequency and $\overline{U}$ is the magnitude of the normal-to-mountain component of the wind for the upwind station (either GJT, LND, or a mean of the two). The upper limit on the integral, Z_{mountain}, is the elevation of the top of the mountain ridge, taken as 3700 m.

13) INTEGRATED SCORER PARAMETER (M^{-1})

This parameter is designed to determine the phase shift ϕ_l of vertically propagating waves between the Froude height and the tropopause, and is given by the following:

$$\phi_l = \frac{1}{2\pi} \int_{H}^{Z_{\mathrm{trop}}} l\, dz, \tag{2}$$

where l is the Scorer parameter (Scorer 1949) and represents the vertical wavenumber, which is defined as $2\pi/L_z$, with L_z being the vertical wavelength in meters, and Z_{trop} is the height of the tropopause in meters. The Froude height is chosen as a lower limit as it allows for more flexibility in the bottom surface and does not define a rigid value of 700 hPa [an improvement over Klemp and Lilly (1975)]. A phase shift of $n\pi$ radians represents the optimum phasing for a reflecting wave. The original equation for the Scorer parameter has the following two terms:

$$l^2 = \frac{N^2}{\overline{U}^2} - \left(\frac{1}{\overline{U}}\right) \frac{\partial^2 \overline{U}}{\partial z^2}, \tag{3}$$

where $\overline{U}$ is the magnitude of the local normal-to-mountain component of the wind and N is the Brunt–Väisälä frequency. In calculating the Scorer parameter, it can be shown that including the second term leads to significant noise in a small dataset; thus, this term was neglected. If the value of l^2 is determined to be negative over a thin layer, it is set it to zero. Negative values of the Scorer parameter are an artifact of small superadiabatic layers that are created by the splines used to fill missing sounding data.

14) CHARACTERISTIC IMPEDANCE RATIO

The characteristic impedance ratio (C.I.R.), is used to determine the joint vertical variation of cross-mountain flow and static stability. This parameter is first defined in Blumen (1985) and is given as

$$\mathrm{C.I.R.} = \left(\frac{36}{25}\right) \frac{\sum_{i=1}^{m} (N_i \overline{U}_i)_{\mathrm{low}}}{\sum_{i=1}^{n} (N_i \overline{U}_i)_{\mathrm{high}}}, \tag{4}$$

where $\overline{U}_i$ is the magnitude of the local normal-to-mountain component of the wind, n is the number of 100-m-thick vertical layers between 5600 and 9400 m, m is the number of such layers between 3100 and 5600 m, and N_i is the Brunt–Väisälä frequency. The coefficient 36/25 was determined empirically from Leptuch (2001). If the quantity $N_i \times \overline{U}_i$ tends not to increase with height (i.e., $\overline{U}_i$ commonly increases with height so N_i decreases with height), there is a greater likelihood of windstorms. The quantity i represents the height of the calculation of the C.I.R.

15) LOWEST TROPOPAUSE (KM)

This parameter shows the height of the lowest of the tropopause levels at DEN, GJT, and LND. Lower val-

ues of the lowest tropopause imply the presence of an upper-level trough directly overhead, which suggests weaker winds near the center of the trough. Large values of the lowest tropopause imply the presence of an upper-level ridge, with weak winds near the center of the ridge.

16) LOCAL TROPOPAUSE (KM)

This parameter is the tropopause height at the local sounding site (either GJT or LND depending on the dataset analyzed).

17) POSTFRONTAL PARAMETER (M S^{-1})

This parameter (PF), calculated using the following equation, is designed to measure forward wind shear in postfrontal situations:

$$\mathrm{Pf} = \overline{U}_{\mathrm{mountaintop}}\left(\frac{\overline{U}_{\mathrm{mountaintop}}}{\overline{U}_{\mathrm{midhigh}}}\right), \tag{5}$$

where $\overline{U}$ is the magnitude of the normal-to-mountain component of the wind and the subscripts "mountaintop" and "mid-high" refer to the 3700-m level and the average over (5600-, 7400-) m altitude. A commonly observed feature of postfrontal windstorms is strong cross-mountain flow at the mountaintop associated with cyclogenesis in the plains and weaker cross-mountain flow aloft with a near-zero or reverse cross-mountain thermal wind component. This latter feature is connected with an often-observed tendency for a cold pool at 500 hPa to pass directly over the region.

18) SANGSTER PARAMETER (M)

This parameter is approximately proportional to the component of the 850-hPa geostrophic wind from about 300°, upstream of Boulder. As such, it provides a measure of the overall strength of the low-level flow approaching the Continental Divide from the west. According to Sangster (1977), this parameter is defined as

$$Z_{850\mathrm{SLC}} + Z_{850\mathrm{GJT}} - 2Z_{850\mathrm{LND}}, \tag{6}$$

where $Z_{850\mathrm{SLC}}$, $Z_{850\mathrm{GJT}}$, and $Z_{850\mathrm{LND}}$ are the 850-hPa heights at SLC, GJT, and LND. As noted in section 1, Sangster (1977) found this quantity to be an important very short-term (0–12 h) predictor of high winds in Boulder. The NWS forecasters have found this parameter useful because it can be calculated readily in real time on an hourly basis from the elevations and observed altimeter settings at these METAR (routine aviation weather report) stations. It provides very short-range guidance probability of wind gusts exceeding 27 and 36 m s^{-1}, and continues to be used in a subjective fashion by NWS forecasters at the Boulder office.

c. Models

Three types of statistical models are selected for this study: stepwise linear regression, a feedforward neural network, and support vector regression. The final two models are included to determine if nonlinear statistical modeling of Boulder windstorms is capable of improving the forecasts of these events.

1) LINEAR REGRESSION MODEL

Multiple stepwise linear regression (LR) is used owing to the multivariate nature of the dataset. The prediction equation used in multiple LR (Wilks 1995) is

$$\hat{Y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i, \tag{7}$$

where x_i is each individual parameter and β_i is the coefficient for each predictor and is analogous with the slope. Each coefficient is calculated using

$$\beta_i = \frac{n\sum_{j=1}^{n} x_{ij}y_j - \sum_{j=1}^{n} x_{ij}\sum_{j=1}^{n} y_j}{n\sum_{j=1}^{n}(x_{ij})^2 - \left(\sum_{j=1}^{n} x_{ij}\right)^2}, \tag{8}$$

where β_i is each coefficient, n is the number of observations, x_{ij} is each predictor from each row, and y_j is the wind observation from each row. An example of the solution obtained by multiple LR is given in Fig. 4. In Fig. 4, a prefrontal case from GJT at 0000 UTC is selected for analysis. This figure shows a plot of the Sangster parameter versus the peak wind gusts over 0–12 h. A regression line is plotted as well, to visualize how LR would fit a line to the data.

2) SUPPORT VECTOR REGRESSION MODEL

The support vector regression (SVR) model is selected since it has been shown to offer promising results in other studies (Richman et al. 2005) as well as for its ability to model nonlinear input data. In SVR, the functional dependence of the dependent variable is based on a set of independent variables. Like other regression problems, the relationship between the independent and dependent variables is given by a deterministic function plus some additive noise. The goal is to find a functional form for the deterministic function that can correctly predict new cases that the SVR has not been presented with before. This can be achieved by training the SVR model on a sample set, that is, a training set,

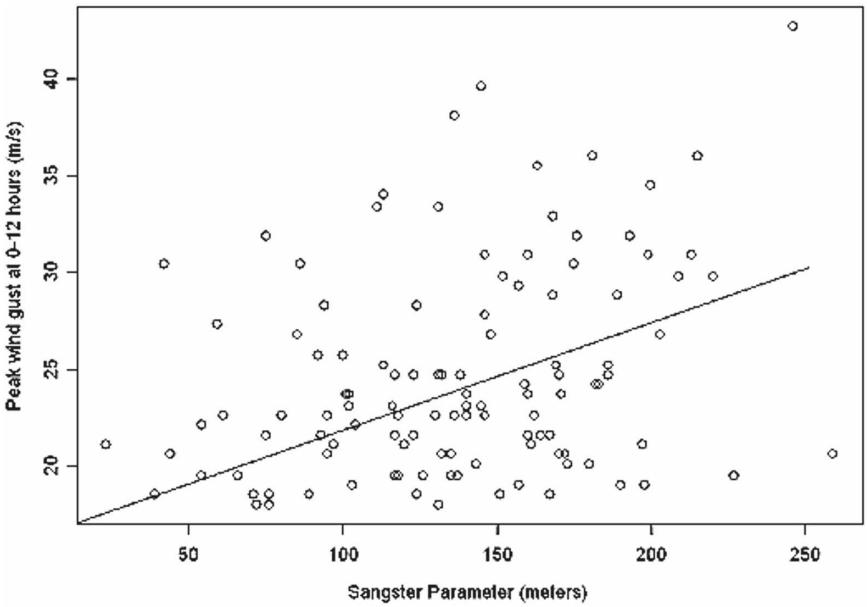

FIG. 4. A scatterplot of the Sangster parameter vs peak wind gusts for the prefrontal GJT 0000 UTC dataset with a hypothetical regression line.

a process that involves the sequential optimization of an error function.

Consider a dataset (x_j, y_j), where x_j is the input and y_j is the output. It is desired to find a function $f(x)$ that has ξ_j deviation from y_j for all training data. The goal of SVR is to minimize the difference between y_j and $f(x_j)$, such that

$$\begin{aligned} |f(x_j) - y_j| \le \xi_j \quad & \text{for every } x_j \\ \xi_j \ge 0 \quad & \text{for every } x_j \end{aligned}. \tag{9}$$

For a basic example, consider a linear case, with the following as our prediction function:

$$F = \{\mathbf{x} \in \Re \mapsto \langle \mathbf{w} \cdot \mathbf{x} \rangle + \mathbf{b} : \|w\| \le B\}, \tag{10}$$

where $B > 0$, $w = \Sigma_j \alpha_j x_j$, $\mathbf{b}$ represents an intercept parameter as with LR, and $\alpha_j \in \Re$. The notation given by $\mathbf{x} \in \Re \mapsto \langle \mathbf{w} \cdot \mathbf{x} \rangle$ indicates that the dot product $\langle \mathbf{w} \cdot \mathbf{x} \rangle$ exists for all real numbers $\mathbf{x}$. From Eq. 9, the prediction function becomes:

$$\begin{aligned} \left| \sum_j \alpha_j \langle x_j, x_i \rangle + \mathbf{b} - y_j \right| \le \xi_j \quad & \text{for every } x_j \\ \xi_j \ge 0 \quad & \text{for every } x_j \end{aligned}. \tag{11}$$

For our windstorm case, a nonlinear SVR methodology is employed. Due to this nonlinearity, the data were mapped into a higher dimensional space through the use of a kernel function:

$$k(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle_H, \tag{12}$$

where H represents the higher-dimensional Hilbert space and ϕ represents the feature map resulting from the kernel function. The new prediction equation for nonlinear SVR becomes

$$F = \{\mathbf{x} \in \Re \mapsto \langle \mathbf{w} \cdot \phi(\mathbf{x}) \rangle_H + \mathbf{b} : \|w\|_H \le B\}, \tag{13}$$

where $B > 0$, $w = \Sigma_j \alpha_j \phi(x_j)$, and $\alpha_j \in \Re$.

Since we wish to reduce the value of the slack variables ξ_i, we need to minimize three main quantities for SVR: $|\xi_i|$, $|b|$, and $\|w\|_H$. To solve this problem, we must perform quadratic programming to solve the following equation:

$$\begin{aligned} & \min(\boldsymbol{\alpha}^{\mathrm{T}} \mathbf{K} \boldsymbol{\alpha} + C \boldsymbol{\xi}^{\mathrm{T}} \boldsymbol{\xi} + b^2) \\ & \text{subject to } |\mathbf{K}\boldsymbol{\alpha} + b\mathbf{1} - \mathbf{y}| \le \boldsymbol{\xi}' \end{aligned} \tag{14}$$

where $C > 0$ is defined as the cost function, $\mathbf{K}_{ij} = k(x_j, x_i)$, and $\mathbf{y}$ is the vector of y_j's. After finding the optimal solution for this quadratic programming problem (α^*, b^*), the final version of the prediction equation becomes

$$f : \mathbf{x} \mapsto \sum_k \alpha_k k(\mathbf{x}_k, \mathbf{x}) + b^*, \tag{15}$$

where the vectors $\mathbf{x}_k$ for which values of α_k^* are nonzero are called support vectors.

Three common kernel functions are used with support vector regression, including

1) polynomial, $k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^{\mathrm{T}}\mathbf{y} + 1)^p$; (16)

2) radial basis function, $k(\mathbf{x}, \mathbf{y}) = e^{-1/2\sigma^2\|\mathbf{x}-\mathbf{y}\|^2}$; and (17)

3) tangent hyperbolic, $k(\mathbf{x}, \mathbf{y}) = \tanh(\beta_0\mathbf{x}^{\mathrm{T}}\mathbf{y} + \beta_1)$. (18)

For this study, the polynomial kernel function is chosen, as after several runs this kernel function is found to offer the most accurate predictions on the testing data. Figure 5 shows a sample of support vector machines from Richman et al. (2005) in which a set of binary data are optimally separated into two classes in feature space, represented by 1 and −1. These data are applied to a support vector machine in an attempt to determine the optimal solution for the modeling function, represented by the dashed line in Fig. 5. This figure is merely an illustration, as support vector regression does not classify data, but instead yields a modeled result that, in our case, is a wind speed in meters per second.

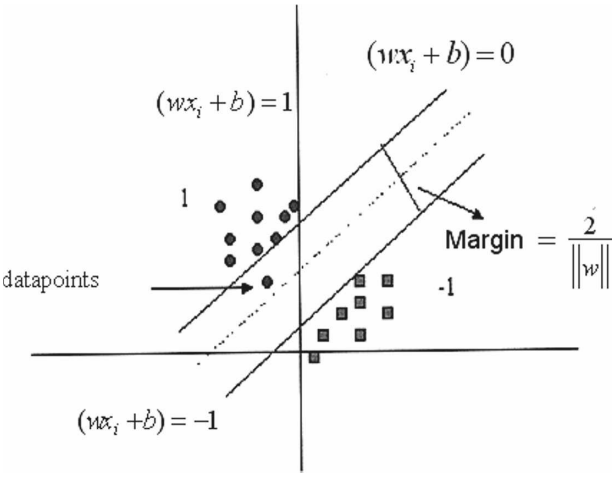

FIG. 5. A geometric representation of SVR. The dashed line is the optimal solution.

3) NEURAL NETWORKS MODEL

The final form of the model used in this study is the feed-forward neural network (Haykin 1999). This technique has been applied to wind forecasting prior to this study (Kretzschmar et al. 2004; Tebaldi et al. 2002), but never to the downslope windstorm problem. The neural network (NN) uses a nonlinear technique to attempt to fit a curve to a dataset, as opposed to fitting a line. The NN operates similarly to the human brain. In the brain, data are entered into the body through nerve endings; the data run through a network of nerves and end up at the brain. The brain then processes the nerve input and yields an output to the senses. Similarly, the NN enters an input dataset through the input layer, and then these data travel through neurons to one or more hidden layers where modeling takes place. The data exit through the output layer of the network as a modeled result.

The NN operates by using an activation function to introduce nonlinearity into the system, thereby attempting to model peak winds in a nonlinear way. The primary goal of the NN is to discover an optimal set of weighted neurons. The following equation shows the goal for the NN in function form:

$$y_j(n) = \varphi_j[v_j(n)], \tag{19}$$

where y_j is the output of the NN (in this case a predicted peak wind gust), φ_j is the activation function, and v_j is the input into the hidden layer, defined as

$$v_i = \sum_{i=0}^{m} w_{ji}(n)x_j(n), \tag{20}$$

where m is the number of input predictors (in this case 18), x_j is the input dataset, and w_{ji}'s are the initial weights selected at random. The activation function selected for this study is the hyperbolic tangent activation function, which is defined as

$$\varphi_j[x_j(n)] = \frac{1 - \exp^{-x_j}}{1 + \exp^{-x_j}}. \tag{21}$$

The goal of the feed-forward NN is to discover a value of w that will minimize the error:

$$\min[e_j = \sqrt{(y_j - d_j)^2}], \tag{22}$$

where e_j is the root-mean-square error in the model and d_j is the observed peak wind gust. According to Haykin (1999), in order to discover the optimal solution for the weights, an optimization technique must be used. In our work, this is the steepest descent method. This method converges on the optimal weights the fastest by calculating the gradient at each weight, thus determining optimal decrease or increase values. The following shows how this method works:

$$\Delta w_{ji}(n) = \eta\delta_j(n)y_j(n), \tag{23}$$

where $\Delta w_{ji}(n)$ is the change in the weights, η is a parameter called the learning rate that determines how quickly the system converges, δ_j is the gradient, and y_j is the output from the network. To put this result in terms of the activation function, take the derivative and multiply it by the error, which leads to the following:

$$\Delta w_{ji}(n) = \eta e_j(n)\varphi'[v_j(n)]y_j(n), \tag{24}$$

where $\varphi'[v_j(n)]$ is the derivative of the activation function. After adjusting the weights with the determined result from the NN, one epoch has been completed. Normally, several hundred epochs are required to determine an optimal solution.

d. Limitations of the methodologies

One complication with using NN and SVR is that many of the variables used in each calculation are not set to specific values. Multiple experiments of each set are required to determine the optimal solutions for different user-defined variables for each of these two models. For example, in this study, five hidden nodes are selected for the NN, with a learning rate of 0.3, a hyperbolic tangent activation function, and 1000 epochs. This procedure is found to be the most efficient for both computing resources and for modeling. In SVR, a cost function value C of 1 is selected, with a quadratic loss function and a polynomial kernel function. These values are determined after performing many experiments with varying values for the cost function, the kernel, and the quadratic loss function. One other concern is the variable nature of the NN, as weights are determined randomly and several model runs are needed to determine a mean of the modeled wind speeds, as well as to converge on an optimal solution. Figure 6 represents the NN selected for modeling peak wind speeds.

An additional difficulty involved with both NN and SVR is the inability to distinguish which parameters provide the most information about the gusts. However, unlike LR, these nonlinear techniques do not lend themselves to this type of analysis without some forward or backward parameter selection. To do this analysis properly, all parameter combinations would have to be tested, which is beyond the scope of this paper.

Since raw sounding data are used to calculate the 18 predictors, it is common to find missing values. A Cardinal cubic spline is implemented to account for the data gaps in the soundings. A layer resolution of 100 m is used for the spline, with a spline tension coefficient of 0.5. All data are separated into prefrontal and postfrontal cases, as described previously. Furthermore, all cases that had a 700-hPa wind direction that is not between 240° and 350° and a 700-hPa wind speed that is less than 8 m s^{-1} are removed (Brown 1986).

After obtaining data for each of the cases surviving the screening process, these are analyzed line by line in search of nonreal values, leaving roughly 1000 cases for each prefrontal dataset and 200 cases for each postfrontal dataset.

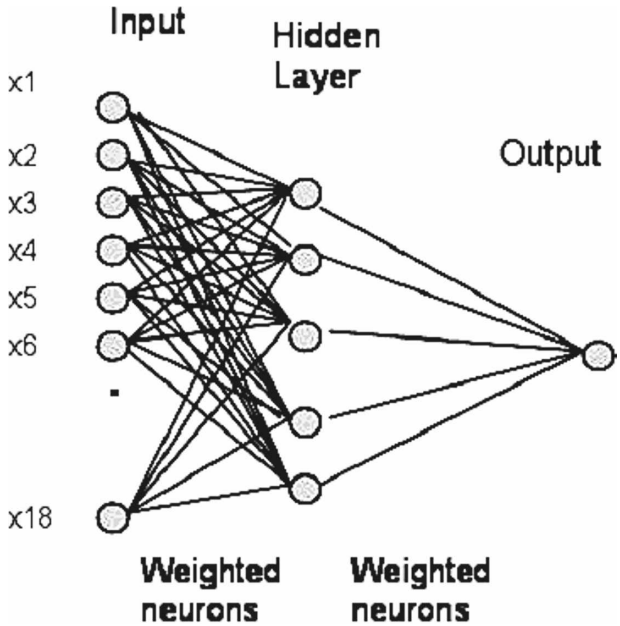

FIG. 6. The architecture of the NN selected for this study. The x values represent the 18 predictors; circles represent layers that show a run through the model, and lines represent weighted neurons in which the data are optimally weighted.

Another problem with the dataset is the abundance of 0 m s^{-1} wind gusts (see section 2a for discussion of the dataset). The lack of lower wind speed data proved troublesome when attempting to implement linear models on our data, since it leads to a large bias in the results as well as poor residual values. A threshold is selected to remove all wind data that had peak gusts that were below 18 m s^{-1}, classifying this threshold as a strong wind event. In doing this, we remove roughly 80% of our data, which makes it more difficult to model. Table 1 shows how many data points are used in each model.

3. Results

Three models are tested on these data: an LR model using the "efroysom" (Insightful Corporation 2002) technique, an SVR model, and an NN. Training and testing data are obtained by dividing the datasets into two halves. Each half is used in training, and the other half is used to test on the trained model. For example, the GJT prefrontal dataset for 0000 UTC has 230 data points. We divide the matrix of predictors for this case into two matrices of 115 points, train a model on the first half, and test on the second, then we reverse the process. The training data are selected at random using a random number program in Matlab (Mathworks

TABLE 1. Number of data points used in each model.

Dataset name	0–12 h	0–6 h	6–12 h	0–3 h	3–6 h	6–9 h	9–12 h
Prefrontal_gjt_0z	230	162	157	96	126	107	115
Prefrontal_gjt_12z	222	137	170	97	104	128	116
Postfrontal_gjt_0z	46	41	21	34	24	18	9
Postfrontal_gjt_12z	58	39	41	21	30	38	26
Prefrontal_lnd_0z	227	157	155	89	121	110	113
Prefrontal_lnd_12z	223	137	171	92	59	126	117
Postfrontal_lnd_0z	47	41	22	33	24	18	10
Postfrontal_lnd_12z	59	39	42	22	30	39	28
Prefrontal_lg_0z	232	162	159	95	59	36	115
Prefrontal_lg_12z	230	144	176	101	109	131	122
Postfrontal_lg_0z	48	42	23	34	25	19	10
Postfrontal_lg_12z	57	38	40	21	29	37	25

2002). Table 2 lists the models used, as well as a number that corresponds to each model, which will be referenced in the results.

a. Linear model

S-Plus (Insightful Corporation 2002) is used to create the LR models that are designed to forecast peak wind gusts. After implementing the LR, often only a few predictors remain. The use of stepwise prediction was included to remove highly correlated parameters that contain redundant information and linear dependence. Only two linear predictors were present in many model runs: the 700-hPa geostrophic wind speed [section 2b(3)] and the 700-hPa cross-mountain wind in Denver [section 2b(4)]. A complete summary of the number of times each predictor appears in a model is included in Table 3 (out of a possible 168 models; 24 training and testing sets multiplied by seven valid times). Other significant predictors include the Sangster parameter [section 2b(18)], the Froude height [section 2b(12)], 700-hPa temperature advection [section 2b(1)], and the 500–700-hPa geostrophic wind direction shear [section 2b(6)].

Another method used to determine the bias of each model is the median residual. Each residual is calculated as observed − predicted; thus, a negative value represents an overforecast and a positive value represents an underforecast. The median residual measures the bias of the model by showing the data point at the

TABLE 2. List of models and model numbers used the boxplots, beginning with Fig. 5.

Model name	Model No.
GJT prefrontal 0000 UTC, first half	1
GJT prefrontal 1200 UTC, first half	2
GJT postfrontal 0000 UTC, first half	3
GJT postfrontal 1200 UTC, first half	4
LND prefrontal 0000 UTC, first half	5
LND prefrontal 1200 UTC, first half	6
LND postfrontal 0000 UTC, first half	7
LND postfrontal 1200 UTC, first half	8
GJT and LND prefrontal 0000 UTC, first half	9
GJT and LND prefrontal 1200 UTC, first half	10
GJT and LND postfrontal 0000 UTC, first half	11
GJT and LND postfrontal 1200 UTC, first half	12
GJT prefrontal 0000 UTC, second half	13
GJT prefrontal 1200 UTC, second half	14
GJT postfrontal 0000 UTC, second half	15
GJT postfrontal 1200 UTC, second half	16
LND prefrontal 0000 UTC, second half	17
LND prefrontal 1200 UTC, second half	18
LND postfrontal 0000 UTC, second half	19
LND postfrontal 1200 UTC, second half	20
GJT and LND prefrontal 0000 UTC, second half	21
GJT and LND prefrontal 1200 UTC, second half	22
GJT and LND postfrontal 0000 UTC, second half	23
GJT and LND postfrontal 1200 UTC, second half	24

TABLE 3. Number of times each predictor appears in the LR model.

Predictor name	No.
Temp advection	35
700-hPa geostrophic wind direction	17
700-hPa geostrophic magnitude	64
700-hPa normal wind Denver	56
700-hPa normal wind local	29
700–500-hPa geostrophic wind shear	30
700-hPa wind to 700-hPa geostrophic wind ratio	20
700-hPa wind direction minus 700-hPa geostrophic wind	22
Relative humidity	18
Cross-mountain height difference	25
Static stability ratio	12
Froude height	29
Scorer parameter	15
Characteristic impedance ratio	14
Lowest tropopause level	22
Local tropopause level	20
Postfrontal parameter	29
Sangster parameter	33

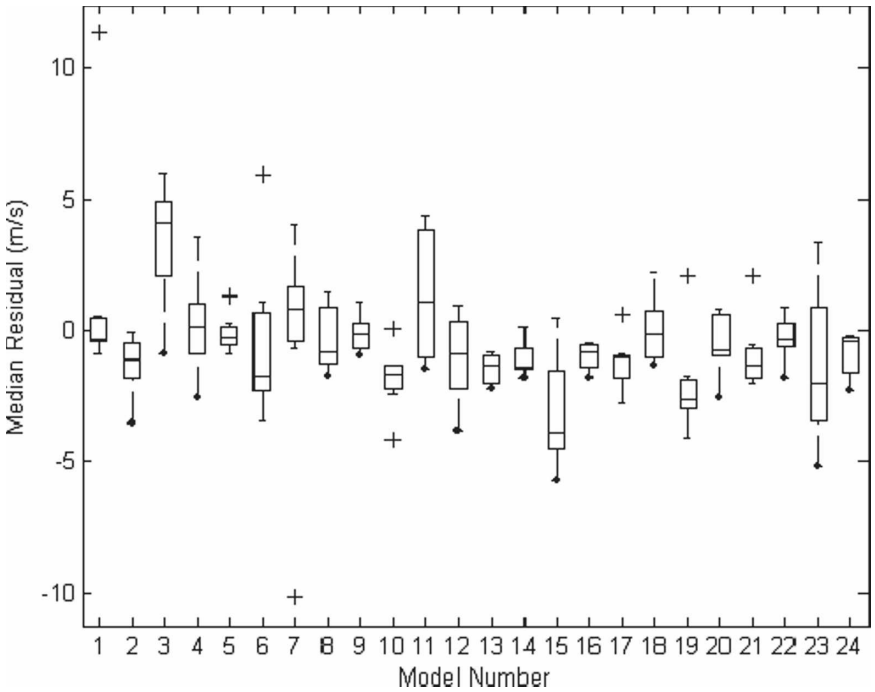

FIG. 7. The time-averaged boxplot of the median residual (m s^{-1}) vs model number for the LR models. Numbers along the horizontal axis represent the particular model that is listed in Table 2.

second quartile, which is a good measure of the central tendency of the dataset. The median residual is selected for this study because the residual vectors contained numerous outliers. The mean would include these outliers, so it would not provide a good interpretation of the central tendency of the models. Figure 7 shows boxplots for the median residuals for the seven time periods (e.g., 0–12, 0–6, 6–12, . . . , 9–12 h after 0000 or 1200 UTC). The center line in each boxplot gives the grand median over all of the seven time periods. Each model number corresponds to Table 2, which shows the model name versus the model number. Although the models seem to center around 0 m s^{-1}, there is a slight bias toward the negative (in 18 out of the 24 models), implying that, for LR, the models tend to overforecast the peak winds.

The parameter used to represent error is the root-mean-square error:

$$\mathrm{RMSE} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n}}, \qquad (25)$$

where Y is the observed peak wind gust and $\hat{Y}$ is the modeled, or where $Y - \hat{Y}$ is the residual winds (m s^{-1}). The models have a mean value of RMSE of 6.33 m s^{-1}. Boxplots of RMSE for every model are given in Fig. 8. Some problems with the LR model are observed when considering the RMSE values for a particular model. Some models have large outliers, such as the LND value for model 6 that contained an RMSE value of 21.15 m s^{-1}. In part, this is attributable to a few high correlations between the predictors (e.g., the Sangster parameter and 700-hPa geostrophic wind direction had a correlation of -0.708). These relatively few high correlations may have led to some degradation of the results for LR, but the results are improved upon using SVR.

b. *Support vector regression model*

We implement a Matlab (Mathworks 2002) program to use SVR (Trafalis et al. 2003) to model the wind data. Our SVR methodology requires training on half the data and testing on the other half.

Figure 9 indicates that the median residuals show little skewness and reside around ± 2 m s^{-1}, which implies that the models forecast with little bias. A very slight tendency toward underforecasting is observed, as more models seem to report a positive value of the median residual. However, this result is better than that obtained from LR (Fig. 7), as there are no large median residual values from SVR. Compared to LR (Fig. 7), the number of SVR median residuals (Fig. 9) having

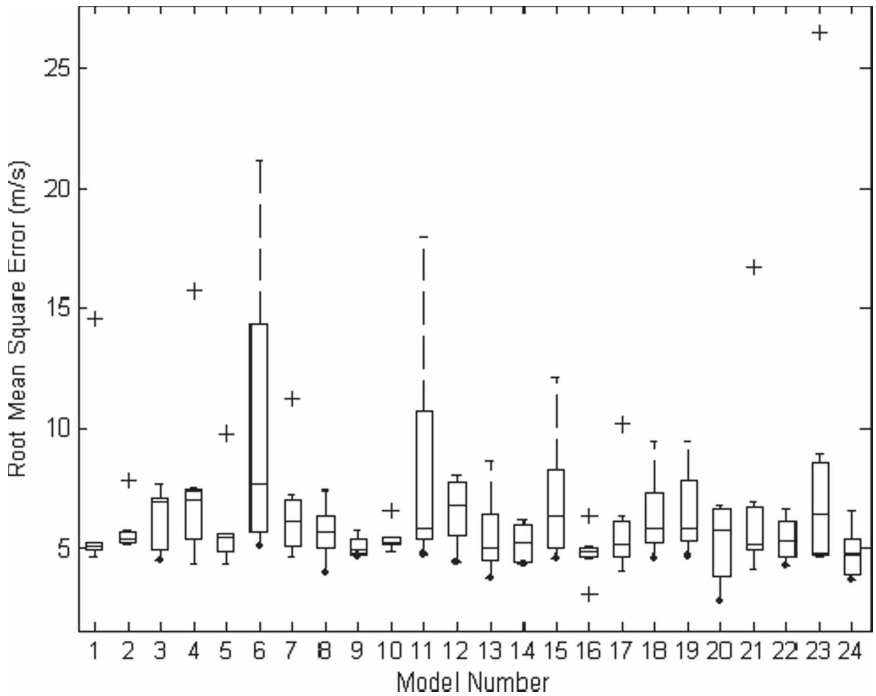

FIG. 8. Same as Fig. 7 but for RMSE instead of median residual.

grand medians exceeding ± 2.5 m s^{-1} is reduced from three to zero.

The mean value of RMSE for SVR (Fig. 10) is 5.22 m s^{-1}, which is lower than the 6.33 m s^{-1} that is obtained from LR (Fig. 8). Moreover, there are few significant outliers, with only one value (11.69 m s^{-1}) exceeding 10 m s^{-1}. In contrast, the LR RMSE (Fig. 8) has an outlier as large as 26.5 m s^{-1} and has numerous

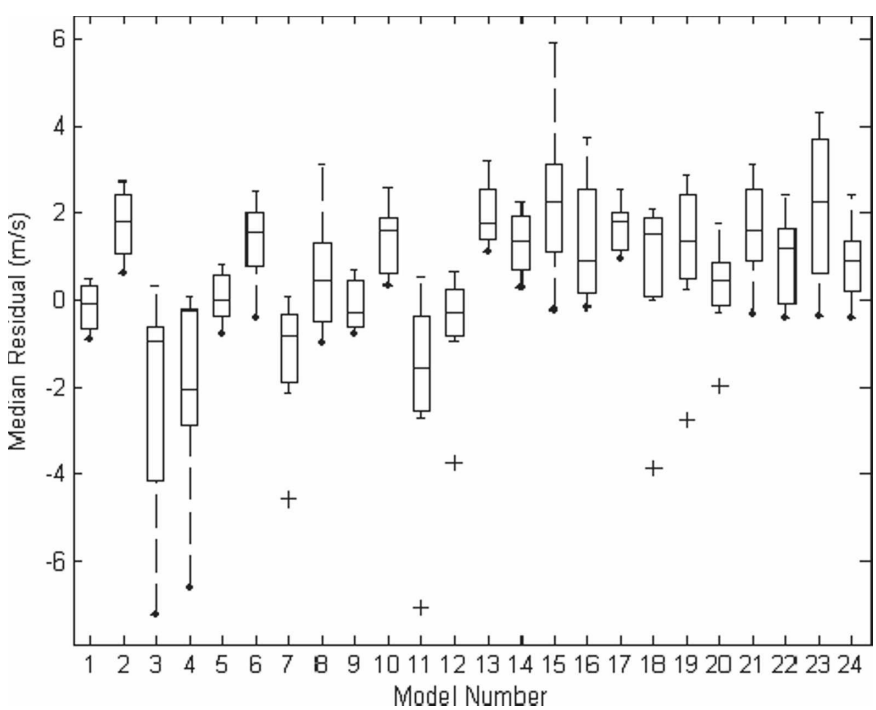

FIG. 9. Same as Fig. 7 but for SVR.

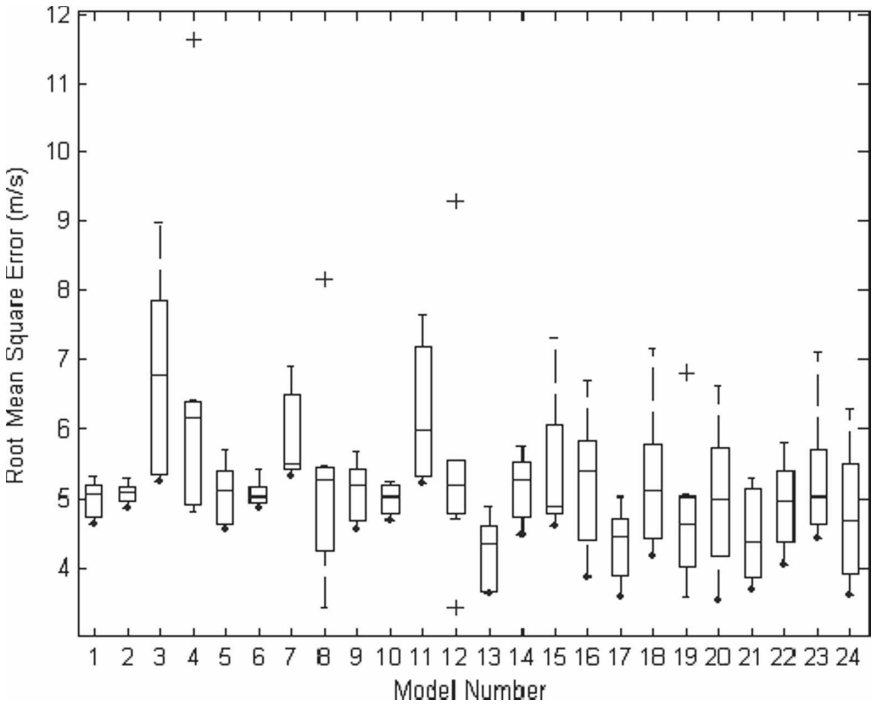

FIG. 10. Same as Fig. 8 but for SVR.

errors exceeding 10 m s^{-1}. Taken collectively, these results indicate that SVR is superior to multiple LR runs for modeling these data.

c. Neural networks

A feed-forward NN (Haykin 1999) is applied to the windstorm data using the "newff" command in Matlab (Mathworks, Inc. 2002). Figure 11 shows the median residual obtained from the NN from the 10 model runs, and Fig. 12 illustrates the RMSE results from the 10 runs. Figure 11 illustrates severe underforecasting for NN, with five models exceeding 2.5 m s^{-1}, compared to a single model exceeding 2.5 m s^{-1} for LR (Fig. 7) and no SVR models exceeding 2.5 m s^{-1} (Fig. 9). Furthermore, Fig. 11 indicates that the NN model overforecasts peak winds when using the second half of the dataset as a training set and the first half of the dataset as a testing set. Ironically, it underforecast peak winds when using the first half of the dataset as a training set and the second half as a testing dataset. The RMSE values center near 7 m s^{-1}, which is slightly larger than LR (Fig. 8) and considerably larger than for SVR (Fig. 10). Some extreme RMSE outliers are present in different NN model runs, including a 16.91 m s^{-1} RMSE. One result of note is that the values of RMSE for NN tended to be higher for a prefrontal case than for the postfrontal case (Fig. 12), and that NN and SVR seemed comparable when analyzing postfrontal cases.

d. Model comparison summary

In Fig. 13, a plot of RMSE versus the percentage of windstorm forecasts below given RMSE thresholds is shown for each of the three models. Analysis of this figure reveals that the SVR model (large dashed line) increases toward 100% accuracy faster than the other two models, while LR does better than neural networks. This finding is consistent with individual comparisons in sections 3b and 3c, which indicate more outliers exist in the NN than in either LR or SVR. Whereas, SVR is the superior model, LR was closer to SVR than to NN, although LR tended to produce extreme forecast outliers. If neural network networks are considered, perhaps additional architectures need to be investigated and techniques employed to minimize overfitting, such as simulated annealing (Laarhoven and Aarts 1987).

e. Real-world application

To test the success of this methodology, 10 recent windstorm cases were selected as independent testing data for the algorithms. These windstorms varied in intensity, timing, and type. Table 4 lists the 10 cases, as well as their peak gust, type, and time.

Once cases were selected, numerical model output from the 80-km NAM over the contiguous United States was used to create forecast soundings. Forecast

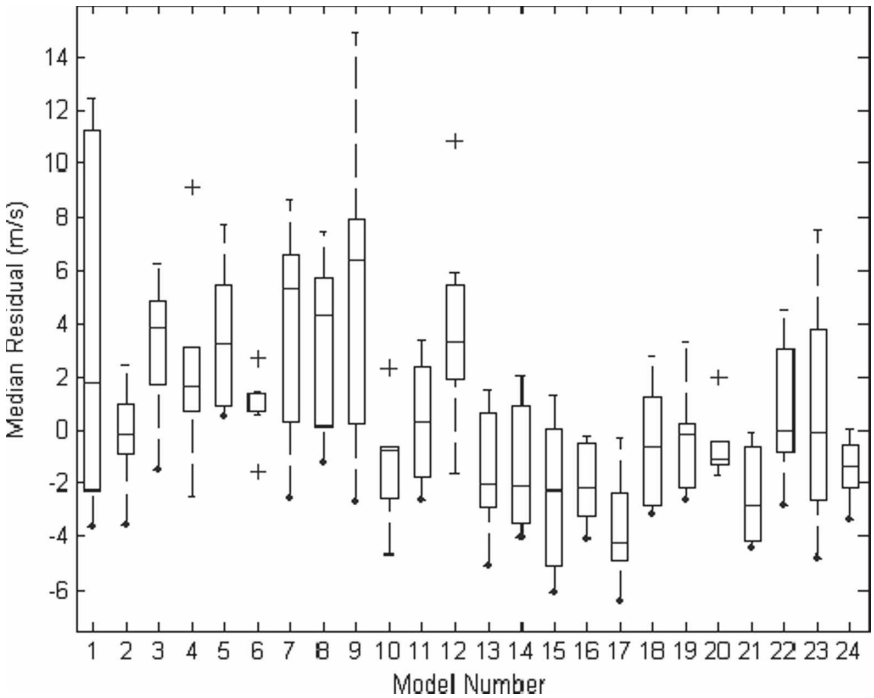

FIG. 11. Same as Fig. 7 but for NNs.

soundings were created based on model runs that were initialized 24 h prior to the windstorm event (24-h forecast soundings). Barnes interpolation to the five raob sites was done on these gridded data, and the output forecast soundings were processed to compute the different windstorm parameters. Based on output from the geostrophic wind direction shear parameter [section 2b(6)] for these 10 cases, they were classified as either

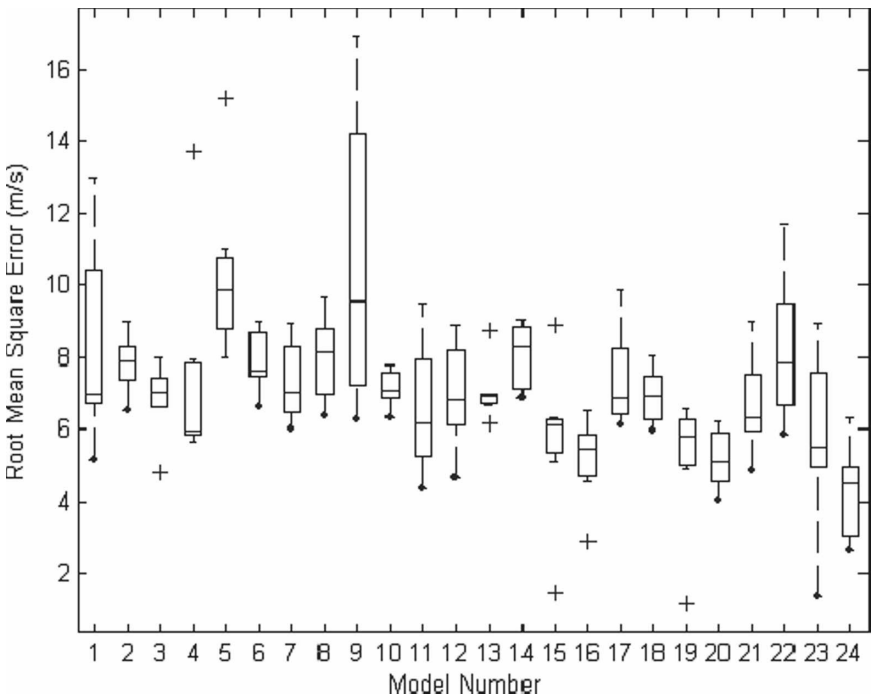

FIG. 12. Same as Fig. 8 but for NNs.

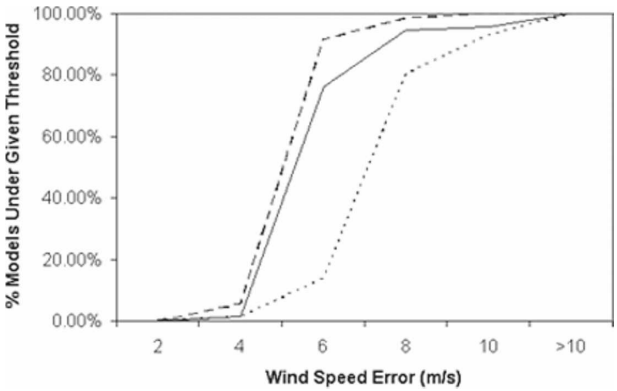

FIG. 13. Plot of RMSE vs the percentage of model runs below the given RMSE threshold for the LN (solid line), SVR (dashed), and NN (dotted) models. The x axis is divided into 0–2, 2–4, 4–6, 6–8, 8–10, and >10 m s^{-1}.

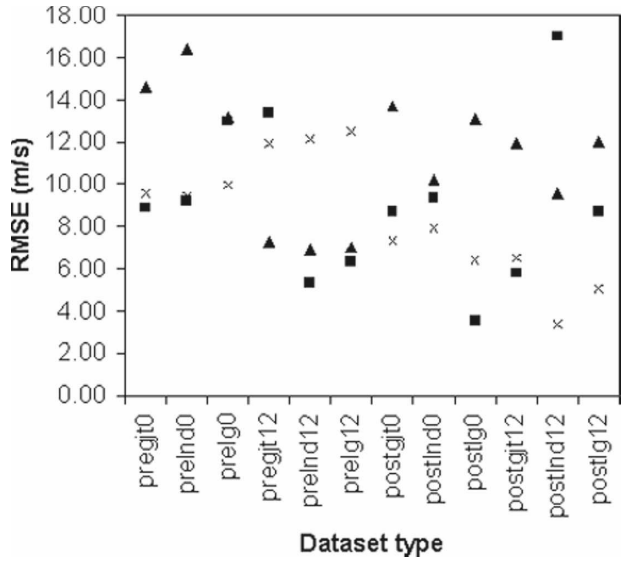

FIG. 14. Plot of RMSE results from the forecast sounding application of the methodology. The x axis contains the dataset type used in the training and testing of the model, where pre or post represent the windstorm type; GJT, LND, or LG mean the central sounding site(s) as with previous results; and 0 or 12 represent the valid time in UTC. Triangles represent LR, squares represent the NN, and x's represent the SVR results.

prefrontal or postfrontal. Finally, the statistical models were trained with the data valid from 1969 to 1978, and the 10 model output cases were tested. All model parameters for the NN and the SVR models remained the same as for the analysis in section 2d.

The results differed as a function of the statistical model. Overall, SVR was most accurate over the 12 datasets, with an average RMSE (Fig. 14) of 8.51 m s^{-1}, compared to 9.10 m s^{-1} (NN) and 11.32 m s^{-1} (LR). However, some biases in these model results existed. The SVR model performed best with postfrontal cases, with the lowest RMSE for the postfrontal LND parameter set valid at 1200 UTC (3.38 m s^{-1}). However, SVR did not predict accurately the large-magnitude wind gusts such as the 8 January 2007 case, as the error for all SVR predictions underforecast this event by 16 m s^{-1}. Overall, SVR had an RMSE of over 10 m s^{-1} for 3 out of the 12 datasets. The NN models were better at predicting the stronger cases, such as 8 January 2007, as the error for this case was only 1.5 m s^{-1}. However, the NN had a tendency to overforecast the weaker events such as 19 October 2006. As with SVR, 3 out of 12 datasets had RMSEs exceeding 10 m s^{-1}. The LR runs did not show biases in underforecasting or overforecasting these events; instead, all the events were predicted poorly (7 out of 12 datasets had RMSEs of over 10 m s^{-1}).

TABLE 4. List of cases used in the forecast sounding analysis.

Date	Time (UTC)	Observed peak gust (m s^{-1})	Type
16 Sep 2006	1200	24.2	Postfrontal
12 Oct 2006	0000	24.7	Prefrontal
19 Oct 2006	0000	26.2	Prefrontal
7 Nov 2006	1200	20.6	Prefrontal
20 Nov 2006	1200	20.1	Prefrontal
1 Dec 2006	0000	27.8	Postfrontal
14 Dec 2006	1200	37.6	Prefrontal
4 Jan 2007	1200	31.9	Prefrontal
8 Jan 2007	1200	41.7	Prefrontal
29 Jan 2007	0000	26.8	Prefrontal

Two possible sources of error exist for the SVR and NN techniques applied to these forecast soundings. First, the soundings are model output forecasts, which will have inherent errors not present in the observations (as were used in the development dataset). Underestimation and overestimation of parameters may have resulted from the use of model soundings in the algorithm (a drawback of the perfect prog approach). In addition, use of the model output at 80-km horizontal resolution (the only output readily available to us), instead of the model's native 12-km output, may have affected adversely the results since local terrain effects, apparent in the development dataset soundings and resolvable by the model, may have been eliminated in the coarsening of the output to 80 km. Second, the lack of significant cases with severe winds in the training set disadvantaged the SVR. Increasing the number of significant cases (larger than 40 m s^{-1}) in the training set (currently less than 1% of our training windstorm database) will provide a more robust dataset, improving SVR results for the large magnitude events such as 8 January 2007. Despite the large error with this testing

case, SVR outperformed both models, and was superior in predicting lower- and midrange-magnitude cases.

4. Conclusions

The purpose of this study is to create models to be used to forecast peak wind gusts during downslope windstorms in the Boulder, Colorado, area. In the LR model, stepwise regression is used. Implementing this model suggests certain predictors could have been ignored. It is difficult to see which predictors could be neglected, as each is used in a minimum of 12 models (Table 3); however, results indicate that 700-hPa geostrophic magnitude is selected most often. Both SVR and the feed-forward NN do not filter out any predictors, but instead fit a nonlinear function to all predictors, so that no important data are discarded. However, the LR analyses did not discover conclusively any subset of these 18 predictors that was significantly more accurate in forecasting the peak winds.

Through comparison of both RMSE and median residuals, it is found that SVR models performed the best. In the development dataset, 85% of the SVR forecasts predict maximum wind gusts with an RMSE of lower than 6 m s^{-1}, and all of the SVR forecasts predict wind gusts with an RMSE of lower than 12 m s^{-1}. The linear model forecasts wind gusts lower than 6 m s^{-1} 60% of the time, and lower than 12 m s^{-1} 95% of the time. For NN forecasts, wind gust errors are less than 6 m s^{-1} 20% of the time, while the RMSE is less than 12 m s^{-1} 90% of the time. The real-world application of 10 recent windstorms, based on model data, suggests that the SVR technique can be implemented into an operational setting with some additional tuning of the model.

Procedures developed in this work may be applied to other downslope windstorm prone areas as well, although the exact parameters will be unique for each region. Upstream rawinsonde data may not be available for other windstorm prone areas as well, hindering the application of these methods to, for example, windstorms to the lee of the Cascades of Oregon and Washington. Investigation of additional candidate predictors of peak wind gusts is important to improve the models further. Forward and backward selection of predictors can be applied to the nonlinear techniques in order to give more insight as to how the parameters interact with peak wind forecasting. Another possibility would be to test these results against real-time verification of explicit surface wind predictions in the WRF-NAM and Rapid Update Cycle (RUC) model runs to determine if the statistical methods are superior. Additionally, probabilities of windstorms could be computed and compared against current operational methods to determine which are optimal. If a larger dataset for Boulder downslope windstorms becomes available, stratification of the storms into strong and weak categories is possible. Strong and weak windstorms could then be classified, in which false alarm rates, hits, and probabilities of detection could be formulated. Finally, inclusion of more significant events in future training of SVR will allow for improved forecasting of strong windstorm cases.

Acknowledgments. Much of the work described herein was funded by COMET Grant S-98-93860 through UCAR. One of the authors (MR) was partly funded through NSF Grant ATM-0527934. We wish to thank Terry Simmons as well for providing data and assistance with this project and Peter A Leptuch for initiating the pilot study. We thank John Weaver for a careful review and an anonymous reviewer for helpful suggestions.

REFERENCES

Bergen, W. R., and A. H. Murphy, 1978: Potential economic and social value of short-range forecasts for Boulder windstorms. *Bull. Amer. Meteor. Soc.,* **59,** 29–44.

Blumen, W., 1985: Reflection of hydrostatic gravity waves in a stratified shear flow. Part I: Theory. *J. Atmos. Sci.,* **42,** 2255–2263.

Brown, J. M., 1986: A decision tree for forecasting downslope windstorms in Colorado. Preprints, *11th Conf. on Weather Forecasting and Analysis,* Kansas City, MO, Amer. Meteor. Soc., 83–88.

Cotton, W. R., J. F. Weaver, and B. A. Beitler, 1995: An unusual summertime downslope wind event in Fort Collins, Colorado, on 3 July 1993. *Wea. Forecasting,* **10,** 786–797.

Doyle, J. D., and Coauthors, 2000: An intercomparison of model-predicted wave breaking for the 11 January 1972 Boulder windstorm. *Mon. Wea. Rev.,* **128,** 901–914.

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. *J. Atmos. Sci.,* **43,** 2527–2543.

——, 1990: Mountain waves and downslope winds. *Atmospheric Processes over Complex Terrain, Meteor. Monogr.,* No. 45, Amer. Meteor. Soc., 59–81.

Haykin, S., 1999: *Neural Networks: A Comprehensive Foundation.* Pearson Education, 842 pp.

Insightful Corp., 2002: S-Plus. Version 6.1. University of Oklahoma.

Klemp, J. B., and D. K. Lilly, 1975: The dynamics of wave-induced downslope winds. *J. Atmos. Sci.,* **32,** 320–339.

Kretzschmar, R., P. Eckert, D. Cattani, and F. Eggimann, 2004: Neural network classifiers for local wind prediction. *J. Appl. Meteor.,* **43,** 727–738.

Laarhoven, van P.J.M., and E.H.L. Aarts, 1987: *Simulated Annealing: Theory and Application*. Reidel, 204 pp.

Lee, T. J., R. A. Pielke, R. C. Kessler, and J. Weaver, 1989: Influence of cold pools downstream of mountain barriers on

downslope winds and flushing. *Mon. Wea. Rev.,* **117,** 2041–2058.

Leptuch, P. A., 2001: Forecasting downslope windstorms in the vicinity of Boulder Colorado, an empirical statistical technique. M.S. thesis, School of Meteorology, University of Oklahoma, 200 pp. [Available from Bizzell Library, University of Oklahoma, Norman, OK 73019.]

Mathworks, 2002: Matlab. Version 6.5.0.180913a. University of Oklahoma.

Overland, J. E., and N. A. Bond, 1995: Observations and scale analysis of coastal wind jets. *Mon. Wea. Rev.,* **123,** 2934–2941.

Richman, M. B., B. Santosa, and T. B. Trafalis, 2005: Feature selection of radar-derived tornado attributes with support vector machines. Preprints, *Fourth Conf. on Artificial Intelligence Applications to Environmental Sciences,* San Diego, CA, Amer. Meteor. Soc., J5.1. [Available online at http://ams.confex.com/ams/pdfpapers/87991.pdf.]

Rockwood, A. A., J. M. Brown, J. F. Weaver, B. D. Jamison, and R. Holmes, 1992: An expert system for the prediction of downslope windstorms. *Proc. Fourth AES/CMOS Workshop on Operational Meteorology,* Whistler, BC, Canada, AES/CMOS, 352–357.

Sangster, W. E., 1977: An updated objective forecast technique for Colorado downslope winds. NOAA Tech. Memo. NWS CR-61, NWS Central Region Headquarters, Kansas City, MO, 24 pp.

Scheetz, V. R., J. F. Henz, and R. A. Maddox, 1976: Colorado severe downslope winds: a prediction technique. Final Rep., Contract 5-35421 to SDO/NWS/SDO/TDL from Geophysical Research and Development Corporation, 40 pp.

Scorer, R., 1949: Theory of waves in the lee of mountains. *Quart. J. Roy. Meteor. Soc.,* **75,** 41–56.

Smith, R. B., 1989: Hydrostatic airflow over mountains. *Advances in Geophysics,* Vol. 31, Academic Press, 1–41.

Tebaldi, C., D. Nychka, B. G. Brown, and B. Sharman, 2002: Flexible discriminant techniques for forecasting clear-air turbulence. *Environmetrics,* **13,** 859–878.

Trafalis, T. B., B. Santosa, and M. B. Richman, 2003: Prediction of rainfall from WSR-88D radar using kernel-based methods. *Int. J. Smart Eng. Syst. Design,* **5,** 429–438.

Weaver, J. F., and R. S. Phillips, 1990: An expert system application for forecasting severe downslope winds at Fort Collins, Colorado, USA. Preprints, *16th Conf. Severe Local Storms,* Kananaskis Park, AB, Canada, Amer. Meteor. Soc., 13–15.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.