



Electronic Journal of
SEVERE STORMS METEOROLOGY

On the Use of Indices and Parameters in Forecasting Severe Storms

CHARLES A. DOSWELL III

*Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma,
Norman, Oklahoma*

DAVID M. SCHULTZ

*Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma,
and NOAA/National Severe Storms Laboratory, Norman, Oklahoma*

(Submitted 11 May 2006; Accepted ?????)

ABSTRACT

This paper describes our concept of the proper (and improper) use of diagnostic variables in severe-storm forecasting. A framework for classification of diagnostic variables is developed, indicating the limitations of such variables and their suitability for operational diagnosis and forecasting. The utility of diagnostic variables as forecast parameters is discussed, in terms of what we consider to be the relevant issues in designing new diagnostic variables used for making weather forecasts. Finally, criteria required to determine whether a new diagnostic variable represents an effective forecast parameter are proposed. We argue that many diagnostic variables in widespread use in forecasting severe convective storms have not met these criteria for demonstrated utility as forecast parameters.

1. Introduction

Operational and research meteorologists often refer to diagnostic variables, such as convective available potential energy (*CAPE*) or the supercell composite parameter (*SCP*; Thompson et al. 2003), as forecast parameters. We contend that they are not necessarily forecast parameters; rather, they constitute a set of diagnostic variables. For most such variables, their forecast value generally has not been established via rigorous verification. That operational forecasters have used many of them for decades is not sufficient, in our opinion, to establish their capability as forecast parameters.

Furthermore, diagnostic variables can lead to faulty perceptions of the state of the atmosphere, owing to various issues tied to their computation

and representativeness. Although there is nothing inherently wrong with diagnostic variables, forecasters need to be aware of the limitations on their use. For example, one of the most common and well established among these is *CAPE* in various forms. Monteverdi et al. (2003) showed that for at least one type of severe weather forecasting (tornadoes in California), *CAPE* proved to be of little value in discriminating tornado cases from nontornadic cases when tested as a forecast parameter. The forecasts based on *CAPE* were verified against observed events, whereas another variable (0–1-km shear) appeared to be capable of making such discriminations.

Diagnostic variables have a long history in association with forecasting severe convection (Schaefer 1986; Johns and Doswell 1992). It is evident that the value of such variables is strongly associated with their capacity to summarize in a single number (or field variable)

Corresponding author address: Dr. Charles A. Doswell III,
The University of Oklahoma/CIMMS, 120 David L. Boren
Blvd, Suite 2100, Norman, OK 73072-7304.
E-mail: cdoswell@hoh.gcn.ou.edu

some characteristic of the severe storm environment. Rather than having to consider the full complexity of four-dimensional atmospheric data, it is often regarded as a benefit for forecasters working with hard deadlines to be able to distill that complexity into a single variable.

Forecasters and researchers generally acknowledge that any single diagnostic variable considered in isolation has little forecast value. Nevertheless, in our experience, we have seen instances where forecasters, often under forecast deadline pressure, will make forecast decisions based heavily, if not primarily, on a single diagnostic variable. It has been our observation that forecasters are most prone to rely heavily on a single diagnostic variable in the context of determining the general likelihood of severe weather. For example, when *CAPE* or strong vertical wind shear is found to be absent at the diagnosis time, the likelihood of severe convective weather subsequently is sometimes dismissed as unlikely. An important concern is that most of the widely used diagnostic variables have not been validated as proper forecast parameters (to be defined below). As diagnostic variables, they can be useful in assessing quantitatively the state of the atmosphere at the time of their calculation, but their capability to inform forecasters about weather in the future can be quite limited, at best.

One purpose of this paper is to propose a classification scheme for diagnostic variables in use for severe weather forecasting, in order to understand their characteristics and limitations. Another goal is to address the issue of what it takes to validate the value of a variable as a proper forecast parameter. We are not seeking to discourage the use of diagnostic variables, per se, in forecasting severe convective storms. Rather, we seek to develop a perspective through which their value to forecasting can be maximized.

2. Diagnostic variables

What specifically do we mean by a *diagnostic variable*? A diagnostic variable is some quantity, valid at a specific instant in time, that either is a basic observed variable (e.g., pressure, temperature, wind, humidity) or can be calculated from those variables. The relationship between diagnosis and prognosis, as developed by Doswell (1986), can be expressed mathematically as follows. Let Φ be an n -dimensional vector of atmospheric state variables

$$\Phi = (\phi_1, \phi_2, \dots, \phi_n).$$

In general, this vector is a function of its location in space, denoted by the position vector \mathbf{X} , and is known at some particular time, $t = t_0$, such that

$$\Phi = \Phi(\mathbf{X}, t = t_0) \equiv \Phi_0.$$

A basic principle of numerical weather prediction (NWP) is that the information about the state of the atmosphere at some initial time can be used to make a forecast of the state of the atmosphere at some future time in the following way:

$$\Phi(\mathbf{X}, t = t_0 + \delta t) = \Phi(\mathbf{X}, t = t_0) + \delta t \left. \frac{\partial \Phi}{\partial t} \right|_{t=t_0}.$$

Thus, the state of the atmosphere at some future time is the sum of its state at the initial time (the diagnosis) and the product of the time step, δt , with the local time rate of change of Φ , where

$$\frac{\partial \Phi}{\partial t}(\mathbf{X}, t) = f[\Phi(\mathbf{X}, t)].$$

That is, the time trend of the atmospheric state variables at any point in space at a particular time is a function of the current spatial distribution of those state variables. In NWP, this time trend is expressed in the form of a set of governing equations used for a particular NWP model. Thus, although it is strictly true that a forecast can be calculated from the state variable vector Φ_0 at the initial time, we exclude from consideration as diagnostic variables the explicit calculation of local time tendencies for atmospheric state variables via this NWP-like process. This does not, as we will show, exclude as diagnostic variables the estimation of time tendencies by other means.

Establishing a quantitative understanding of the state of the atmosphere at a given time using diagnostic variables is possible. Qualitative understanding, in turn, can be substantially enhanced by this quantitative analysis. For example, a forecaster can simply look at the plotted winds on a surface chart and recognize zones of strong rotation (vorticity) and convergence, but without a quantitative evaluation of the vorticity and convergence values, a forecaster's sense of the strength of those variables is obviously subjective. A quantitative assessment is generally preferred over subjective assessment—in forecasting, numerical values of variables often matter to the forecast decision.

We define a diagnostic variable being used as a *forecast parameter* as one that allows a forecaster to make an accurate weather forecast based on the *current* values of that variable. That is, it is expected that, for a proper forecast

parameter, there should be a time-lagged correlation between the parameter and the weather being forecast. This is to be distinguished from using a diagnostic variable calculated from a forecast of the state variables (say, from NWP model gridded forecast fields) valid at some future time to make a forecast of the weather for that valid time. How accurately an NWP model forecasts that diagnostic variable is a very different issue from how accurate a forecast can be using only the current distribution of that variable. A diagnostic variable might be very good at discriminating between different weather events when known at the time of the event, but still be ineffective as a forecast parameter because its values *before* the events occur make no directly useful statement about the future weather. Obviously, diagnostic variables change with time, and indeed, a description of the time tendency for some diagnostic variable can be an important diagnostic variable in its own right and may even be valuable as a proper forecast parameter. The accuracy of forecasts made using a diagnostic variable as a forecast parameter will tend to decay with time. Some forecast parameters might permit reasonably accurate forecasts only for a short period very close to the time of their diagnosis, whereas others might show high lagged correlation with the weather for a relatively long time.

Therefore, what we are concerned with herein is a forecast by a human forecaster that makes a statement about the future weather (say, the occurrence of severe storms), rather than a statement about the distribution of atmospheric state variables. For a variable to be useful in making a weather forecast by a human, it must be shown rigorously that there is some measure of forecast skill when the current values of the variable are used to make that weather forecast. This will be developed in detail in what follows.

When calculating diagnostic variables, forecasters unaware of the caveats associated with the particular variable being analyzed can be misled. What are the unique sensitivities in those calculations? How much confidence can you put in the numbers and how they might change over time? Each diagnostic variable has its own story and, if diagnostic variables are to be used properly, forecasters need to be aware of each such story.

In order to discuss diagnostic variables and their limitations, we propose a classification scheme for diagnostic variables.

a. Simple observed variables

Simple observed variables are those measured by meteorological instruments. Strictly speaking, most modern meteorological instruments make electronic measurements (e.g., resistance or capacitance) that are calibrated to provide readouts of the meteorological variables (e.g., pressure, temperature, dewpoint, wind direction and speed).

b. Simple calculated variables

Calculated variables are not observed directly but are computed from the raw measurements using relatively simple conversion formulae. Calculated variables typically involve combinations of two or more observed variables, but they are not combined in an arbitrary fashion. Rather, they are combined in ways having a physical basis. Such calculated variables are usually sought because they have some valuable physical property, such as being conserved under certain reasonable assumptions. For example, at the surface, the temperature and dewpoint temperature are the common observed variables. However, for reasons discussed in Sanders and Doswell (1995), mixing ratio and potential temperature are conserved variables that incorporate the pressure observations, as well as the temperature and dewpoint. The formulae for calculating the mixing ratio and potential temperature are determined by the laws governing the physical properties of air parcels under the assumption of adiabatic flow.

c. Derivatives or integrals (spatial or temporal) of simple observed or calculated variables

Time and space derivatives and integrals of the observed or deduced variables form the next class of diagnostic variables. In effect, these diagnostic quantities allow estimates of the terms in formulae that might arise in mathematical descriptions of atmospheric structure. An important caveat in the calculation of such diagnostic variables is the inevitable truncation error that arises from the limited temporal and spatial resolution of our meteorological information, whether it be observed or modeled. If we were to calculate, say, the Eulerian time rate of change of the 500-hPa height at some location where soundings are taken, the true instantaneous value is not known, simply because soundings generally are taken at 12-h intervals. Of course, estimating that same variable at some point other

than where a sounding is launched¹ is problematic because of the small number of locations where soundings are made. The usual approach is to make estimates using available information; this typically involves making assumptions about the behavior of the fields where we have no information. Realizing the limitations of those calculations is critical. Just because the calculations were performed by computer is no guarantee they are even remotely accurate.

d. Combined variables

Next is the class of *combined* diagnostic variables. Many ways exist to take two or more diagnostic variables and combine them in some way that might be more useful for some specific purpose than the raw observations or simple derivatives and integrals of those variables. Moisture flux convergence (*MFC*), discussed in detail by Banacos and Schultz (2005), is an example of a combined variable. The formulation of *MFC* can vary from one application to another, but it is often formulated as the finite difference calculation of the quantity:

$$MFC = -\nabla_h \cdot (r\mathbf{V}_h) = \underbrace{r \nabla_h \cdot \mathbf{V}_h}_1 + \underbrace{\mathbf{V}_h \cdot \nabla_h r}_2$$

where r is the mixing ratio, ∇_h is the horizontal gradient operator, and \mathbf{V}_h is the horizontal wind vector. This calculation involves a calculated variable (the mixing ratio) and a spatial derivative of an observed variable (the horizontal wind vector). *MFC* is often calculated using the relatively dense surface observations, and, as used in forecasting severe storms, *MFC* is purported to show where ascent is occurring in the presence of surface moisture. Banacos and Schultz (2005) emphasize four points about *MFC* relevant to this discussion.

1. The scientific justification for using *MFC* as a forecast tool for convective initiation is inadequate.
2. The putative value of *MFC* as a forecast variable has never been firmly established by a careful statistical verification study. Its popularity is based almost entirely on anecdotal evidence and heuristic arguments.
3. Term 1 on the rhs of the *MFC* equation—associated with horizontal diver-

gence—is characteristically *much* larger than term 2—that associated with moisture advection. The *MFC* field tends to look very much like that of the horizontal divergence field alone on mesoscale and smaller scales.

4. *MFC* is an inadequate tool as a forecast parameter because it *combines* two of the three ingredients required for deep, moist convection (e.g., Johns and Doswell 1992). The two components of the major term in this combination, the mixing ratio and the divergence field, can evolve quasi-independently. What this variable shows is where those two main components overlap at the time of the analysis. Thus, *MFC* is first and foremost a diagnostic variable.

CAPE is another example of a combined variable; its calculation involves a vertical integral of multiple state variables and incorporates a number of assumptions (e.g., Doswell and Rasmussen 1994; Doswell and Markowski 2004). Nevertheless, that calculation can be represented by its most basic components: large *CAPE* is generally observed where low-level moisture is found in the presence of conditionally unstable lapse rates in the lower mid-troposphere. As with *MFC*, these constituent fields can evolve quasi-independently and can be superimposed by differential advection processes. Prior to that superposition, the air streams carrying conditionally unstable lapse rates and low-level moisture (typically at different levels in the atmosphere advecting variables in different directions and at different speeds) have not yet interacted and so little or no *CAPE* is found. The presence of *CAPE* indicates when its constituents overlap, but an absence of *CAPE* prior to that superpositioning cannot be used to infer large *CAPE* will not be present in the future. *CAPE* depicts where moisture and conditional instability are already superimposed, not where they will (or will not) be superimposed in the future. A comparable statement can be made about *MFC*. This aspect of combined variables is an important element for understanding the difference between a diagnostic variable and a proper forecast parameter.

e. Indices

Indices, the final class of diagnostic variables, can be broken down into two distinct subclasses: indices based on physically based formulae and indices representing more or less arbi-

¹ In certain situations, the displacement of the sounding instrument from its original launch location has to be accounted for in the calculations.

trary combinations of diagnostic variables. This is a complex topic and is the subject of a wholly separate section (section 4, below). Before we discuss indices, however, we consider several issues associated with diagnostic variables that affect their utility in diagnosing the current state of the atmosphere.

3. Issues affecting the suitability of diagnostic variables

All diagnostic variables are subject to error, but not all are equally error-prone. Errors can be decomposed into *measurement* errors, which have a number of sources that are generally instrument-dependent (Brock and Richardson 2001; §1.1.3), and *sampling* errors (associated with the finite number of observations). Measurement and sampling errors are associated with much of the *volatility* of diagnostic variables. Here we use volatility to mean that the variable can vary considerably over both time and space as a result of sensitivity to both measurement and sampling errors. Some diagnostic variables are inherently more volatile than others.

Consider *CAPE* as an example of a combined variable, and compare it to the Showalter index. Both of these purport to be variables pertinent to forecasting severe convective weather, and both are based on simple parcel theory. The Showalter index is determined by the temperature difference between a hypothetical parcel lifted from 850 hPa to 500 hPa (a calculated diagnostic variable) and that at 500 hPa (an observed diagnostic variable). The Showalter index was originally proposed because it uses only three observed variables: 500-hPa and 850-hPa temperatures, and 850-hPa dewpoint temperature. At the time, manual processing of soundings meant that the mandatory pressure-level data were available more quickly than the rest of the sounding. Calculating an index requiring only mandatory-level data was advantageous because it gave forecasters a quick look at the convective instability before the whole sounding came in later.

The temperature of a parcel lifted from 850 to 500 hPa depends strongly on the 850-hPa temperature–dewpoint spread. When that spread is large, unless the 850–500-hPa lapse rate is very nearly dry adiabatic, the parcel is likely to arrive at 500 hPa colder than the observed 500-hPa temperature because it will have followed a dry adiabat for most of its ascent. Conversely, when the 850-hPa temperature–dewpoint spread is small, the parcel will ascend mostly along a

moist adiabat and is likely to be warmer than the observed 500-hPa temperature, unless the lapse

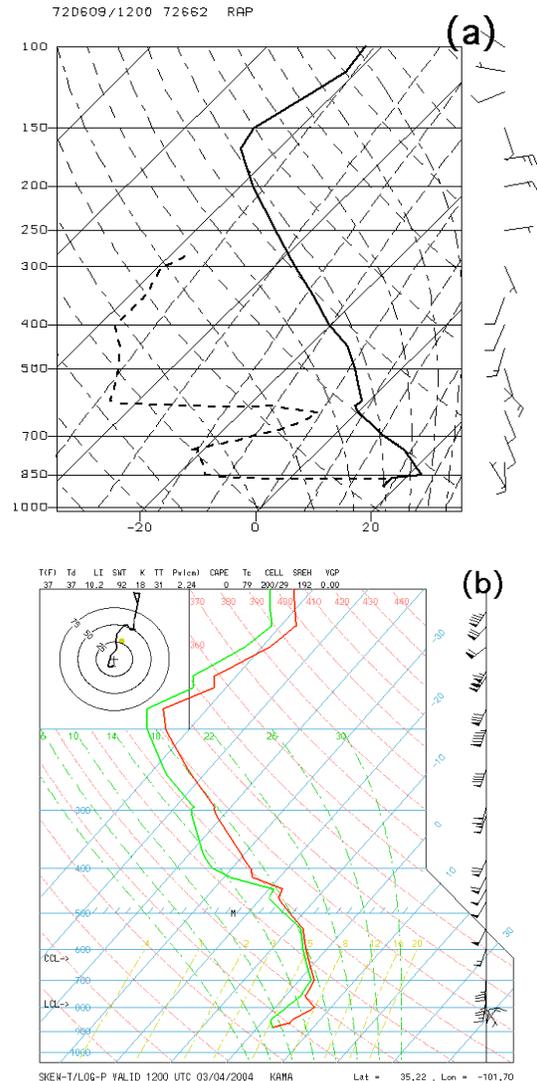


Figure 1. Rawinsonde plots on a Skew- T -log p diagram for (a) Rapid City, SD on 09 June 1972, and (b) Amarillo, TX on 04 March 2004. See text for discussion.

rate is less than moist adiabatic. So far, so good, but consider the example shown in Fig. 1a. If, as in this case, the low-level moisture cuts off just below 850, the Showalter index may be calculated correctly, but its implications about the convective instability of the atmosphere can be misleading. This example was taken on the morning of a devastating flash flood in the vicinity that evening—the calculated Showalter index would be nominally indicative of little or no

chance for thunderstorms simply because of the low dewpoint temperature at 850 hPa.

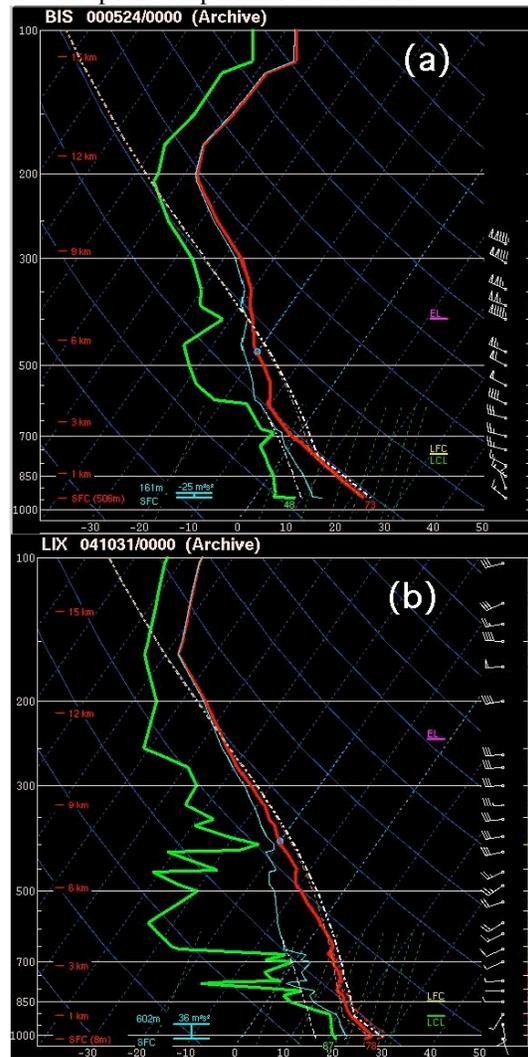


Figure 2. Soundings with very nearly equal *CAPE* values, but showing distinctly different vertical distributions of *CAPE*—(a) for Bismarck, ND at 00 UTC on 24 May 2000, and (b) for Slidell, LA at 00 UTC on 31 October 2004.

Indices can be dramatically affected when soundings rise through precipitation (as in Fig. 1b), or from errors that just happen to fall at mandatory pressure levels, thereby rendering calculation of parameters such as the Showalter index, unrepresentative or even meaningless. This sensitivity to a small number of observations occurs in part because the Showalter index involves a differential quantity, and derivatives generally are much more sensitive to errors than are the basic quantities being differenced.

On the other hand, the *CAPE* calculation is a vertical integral that uses many more point measurements. By virtue of being an integral quantity, rather than a derivative, calculation of *CAPE* is inherently less sensitive to small differences that might arise from measurement errors. Unfortunately, that apparently useful property of integration renders its values non-unique—that is, you can get the same *CAPE* value from distinctly different vertical distributions of a parcel’s thermal buoyancy (Fig. 2). It is likely that one would interpret these soundings differently in terms of the weather forecasts, but considering only the *CAPE* values without seeing the soundings would permit no such discrimination.

Moreover, as discussed in Brooks et al. (1994), finding a *representative* sounding can be something of a challenge, owing to undersampled variability in space and time. This undersampled variability also affects any interpretation of the Showalter index, of course. At least in the short term, little can be done about measurement errors and undersampling. However, when considering how to interpret diagnostic variables, it seems quite unlikely to be able to define a diagnostic variable that distills the relatively rich complexity of a complete sounding into a single number that isn’t vulnerable to being unrepresentative under some circumstances. In fact, we assert that *no* single number can replace the value of a forecaster simply looking at the soundings, as well as looking at diverse diagnostic variable computations based on those soundings. Any obvious errors in the sounding, or any characteristics that would render diagnostic variable computations misleading, will be apparent to someone experienced in sounding interpretation.

Further, we suggest that this principle applies to *any* diagnostic variable, not just those based on soundings. The diagnostic variable calculations still add value beyond what our hypothetical forecaster can see simply by viewing the data on a display, but using *only* the diagnostic variables as a substitute for considering all the information in the data is inherently risky in making weather forecasts.

Volatility is associated with the specific way in which the observations are used to construct the fields of a diagnostic variable. Consider the horizontal divergence, expressed in conventional meteorological notation as:

$$\text{div}_h = \nabla_h \cdot \mathbf{V}_h = \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right),$$

where (u,v) is the Cartesian coordinate (x,y) form of the horizontal velocity vector. When expressed in terms of natural coordinates—along and normal to the horizontal wind vector—Panofsky (1964, p. 33 ff.) has shown that the horizontal divergence is the difference between two relatively large numbers, so the calculation can be quite sensitive to small changes in the winds. Note that this sensitivity is not present when calculating the vertical component of the vorticity, ζ , where

$$\zeta = \mathbf{k} \cdot \nabla \times \mathbf{V} = \left(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x} \right),$$

and where \mathbf{k} is the unit vector in the vertical—the two terms in this calculation have no characteristic tendency to cancel one another.

Further, a divergence (or vertical vorticity) calculation is dependent on the resolution of the data. To show this, assume a characteristic magnitude for the difference in wind velocity between two points (δV , which is the same order of magnitude as the wind velocity itself, V), separated by a characteristic distance scale L , so that

$$\text{div}_h \sim \frac{\delta V}{L},$$

which is a simple scaling law for the divergence (or vorticity) magnitude. The order of δV turns out to be not very scale dependent, typically of order $1\text{--}10 \text{ m s}^{-1}$. This characteristic difference rarely reaches 100 m s^{-1} or becomes as small 0.1 m s^{-1} . Thus, over a fairly wide range of scales, the divergence tends to depend most strongly on the distance between sample points.²

If divergence is calculated from a sparse network of points, such as rawinsonde sites ($L \sim 400 \text{ km}$), a rough order of magnitude for the divergence, according to the above scaling rule, is about $(1\text{--}10 \text{ m s}^{-1}) \div 400 \text{ km} = 2.5 \times (10^{-6}\text{--}10^{-5}) \text{ s}^{-1}$. In contrast, for the network of surface observations ($L \sim 100 \text{ km}$), the simple scaling rule gives a value of $1 \times (10^{-5}\text{--}10^{-4}) \text{ s}^{-1}$, which is four

times larger. According to this simple scaling law, horizontal divergence is inversely proportional to the scale of the data resolution over a wide range of scales.

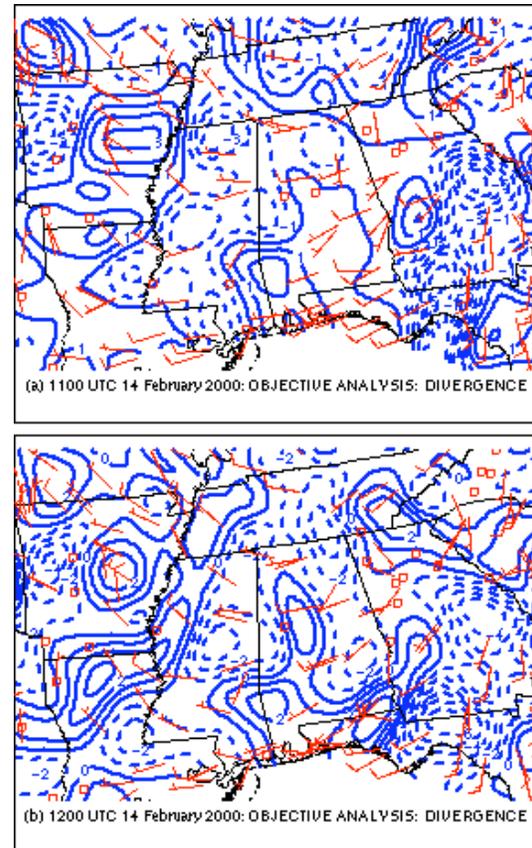


Figure 3. Objective analyses of surface divergence at (a) 1100 UTC and (b) 1200 UTC, illustrating the volatility of the details. Solid contours are nonnegative divergence, starting at zero, and dashed contours are negative values. Conventional wind barbs show the wind observations used in the calculations.

Like the Showalter index, divergence calculations can be volatile. This volatility also follows from the simple scaling law above: if winds are only known to within 2 m s^{-1} , then a divergence estimate based on a station separation of 100 km is known only to within $2 \times 10^{-5} \text{ s}^{-1}$. The fields tend to be noisy and behave somewhat erratically over time as a result of undersampled variability in the wind field and measurement errors, although the basic shape of the fields might be fairly consistent from one time to the next (Fig. 3).

To some extent, this volatility can be reduced by heavy smoothing that limits the spatial and/or

² Note that at *synoptic* scales (where $L \sim 1000 \text{ km}$), because the wind is nearly geostrophic, the characteristic divergence magnitude is an order of magnitude less than this simple scaling law predicts because the geostrophic wind calculated on an f -plane in many meteorological coordinate systems is nondivergent (see Doswell 1988). This can be accounted for by changing the scaling law to

$$\text{div}_h \sim Ro \frac{V}{L}$$

as shown by Haltiner and Williams (1980, p. 57), where Ro is the Rossby number and is of order 10^{-1} on synoptic scales, but dynamics-based scale analysis is beyond the scope of this paper. Again, synoptic-scale vertical vorticity scaling is not subject to this consideration.

temporal scales of the features retained to those that can be depicted reliably in the analysis (Doswell 1977). The consistency of the basic shape of the fields, combined with the volatility of the details, is a direct indication of the sensitivity of the details in the field to small changes in the wind.

Having considered several of the issues that confront users of diagnostic variables, we are now prepared to consider the topic of indices.

4. Indices

Indices have a long history in severe storms forecasting that perhaps began with the Showalter index (*SI*) and has continued through a growing plethora of constructs, including, among many others, the lifted index (*LI*), SWEAT index, bulk Richardson number (*BRN*), energy-helicity index (*EHI*), *SCP*, significant tornado parameter (*STP*), and enhanced stretching potential (*ESP*; J. Davies 2005, personal communication). These and selected other indices are listed in Table 1. Indices also have been applied to forecasts other than severe convective weather. One example is the Garcia (1994) method for forecasting snowfall. See Wetzel and Martin (2001) and Schultz et al. (2002) for a discussion of the scientific integrity of the Garcia (1994) method and other approaches.

Some indices are associated with a physical argument. The early stability indices (e.g., *SI*, *LI*) were based on simple parcel theory, although we have suggested some issues with their use as diagnostic variables previously, to say nothing of their use as forecast variables. The dimensionless *BRN* is at least related to the true Richardson number, but the actual physical significance of any Richardson number to the physics of deep moist convection is unclear, as the original intent of the Richardson number was to address topics in turbulence theory (Tennekes and Lumley 1972, p. 98 ff.).

Many of these indices, including the SWEAT index, *EHI*, *SCP*, and *STP*, have combined variables in ways that have no physical rationale. In other words, the process of forming sums, products, and ratios has not been done in accordance with a formula originating in the mathematics describing a physical process. Rather, the mathematical expression for the index is more or less arbitrary. Why a sum of two variables divided by a third? Why not one variable raised to a power defined by a second variable multiplied by a third?

As we have mentioned for the case of simple diagnostic variables calculated from observations (cf. section 2b), combining two or more variables in a way that has a physical basis affects the interpretation and use of the resulting variable. If a variable is conserved during certain physical processes, for example, that is quite relevant to its application in diagnosis or forecasting.

At issue is whether or not the variable can be related to physical principles. Examples of diagnostic variables based on physical principles might be something like the static stability time tendency, potential vorticity, or energy dissipation rate. Combining two or more variables in an arbitrary way leaves open many questions and makes it difficult to relate the variable to any physical understanding of the process. The individual diagnostic variables used to form an index may have physical relevance to the problem at hand, but when specific formula combining them is unphysical, this can be problematic.

Furthermore, the physical dimensions of these indices may or may not make any physical sense. For example, the ratio of *CAPE* to shear (a ratio used in one form or another for several indices within Table 1) has dimensions of $\text{J kg}^{-1} \text{s}$, the product of energy per unit mass and time, whereas the *product* of *CAPE* and shear has dimensions of $\text{J kg}^{-1} \text{s}^{-1}$. The former has no obvious physical interpretation, whereas the latter has dimensions of energy per unit mass per unit time, which at least can be related to terms in an energy budget.

a. Example of representative problems with indices: EHI

In order to demonstrate these issues using one of these indices, we use the *EHI* as a representative example, although any of the above-listed indices would possess similar problems. As described in Rasmussen and Blanchard (1998, p. 1154), “This index [using the 0–1 km layer] is used operationally for supercell and tornado forecasting, with values larger than 1.0 indicating a potential for supercells, and $EHI > 2.0$ indicating a large probability of supercells.” Our problems with *EHI* center around five issues: combination of ingredients, arbitrary construction, choice of scaling constant, ambiguous physical meaning, and lack of proper validation.

First, *EHI* is a combination of two separate variables that may not even be collocated during the event and can evolve separately, as discussed in section 3. For instance, although large *CAPE*

Index Name	Formula	Reference
Showalter index	$SI = T_{500} - T_{lp(850 \text{ hPa})}$	Showalter (1947)
Lifted index	$LI = T_{lp(\text{fcst surface})} - T_{500}$	Galway (1956)
K-index	$K = (T_{850} - T_{500}) + D_{850} - (T_{700} - D_{700})$	George (1960, pp. 407–415)
Convective available potential energy	$CAPE = \int_{LFC}^{EL} (\alpha_{lp} - \alpha) dp$	Glickman (2000, p. 176)
Vertical totals	$VT = T_{850} - T_{500}$	Miller (1972)
Cross totals	$CT = D_{850} - T_{500}$	Miller (1972)
Total totals	$TT = VT + CT$	Miller (1972)
SWEAT index	$SWEAT = 20(TT - 49^\circ\text{C}) + 12D_{850} + 2V_{850} + V_{500} + 125[\sin(\Delta V_{500-850}) + 0.2]$	Miller (1972)
Bulk Richardson number	$BRN = \frac{CAPE}{\frac{1}{2}(\bar{U} - U_0)^2}$	Weisman and Klemp (1982)
Storm-relative helicity	$SRH = - \int_{z_0}^z \mathbf{k} \cdot \left[(\mathbf{V}_h - \mathbf{C}) \times \frac{\partial \mathbf{V}_h}{\partial z} \right] dz$	Davies-Jones et al. (1990)
Energy-helicity index	$EHI = \frac{(CAPE)(SRH)}{160,000}$	Hart and Korotky (1991)
Supercell composite parameter	$SCP = \left(\frac{MUCAPE}{1000 \text{ J kg}^{-1}} \right) \left(\frac{SRH_{0-3 \text{ km}}}{100 \text{ m}^2 \text{ s}^{-2}} \right) \left(\frac{\bar{U} - U_0}{40 \text{ m s}^{-1}} \right)$	Thompson et al. (2003)
Significant tornado parameter	$STP = \left(\frac{MLCAPE}{1000 \text{ J kg}^{-1}} \right) \left(\frac{SHR_{0-6 \text{ km}}}{20 \text{ ms}^{-2}} \right) \left(\frac{SRH_{0-1 \text{ km}}}{100 \text{ m}^2 \text{ s}^{-2}} \right) \left(\frac{(2000 \text{ m} - MLLCL)}{1500 \text{ m}} \right)$	Thompson et al. (2003)
Enhanced stretching potential	$ESP = \left(\frac{\partial T}{\partial z} \Big _{0-2 \text{ km}} - 7^\circ\text{C km}^{-1} \right) \left(\frac{MLCAPE_{3 \text{ km}}}{50 \text{ J kg}^{-1}} \right)$	Davies (2005 – personal communication)

Table 1. A selection of indices commonly used in the United States for severe storm forecasting. In the formulae, T denotes a temperature and D denotes a dewpoint temperature in $^\circ\text{C}$, with a subscript indicating at what mandatory pressure level (in hPa) this value is to be taken from; α denotes the specific volume and a subscript lp denotes a value associated with a lifted parcel; LFC stands for a lifted parcel’s level of free convection and EL stands for its equilibrium level. For the Lifted index, the lifted parcel is for a surface parcel with *forecast* properties at a representative time of day. For the *SWEAT* index, V denotes a wind speed (in knots), and ΔV denotes a wind direction difference (in degrees). For the Bulk Richardson number, \bar{U} denotes the density-weighted speed of the mean vector wind in the layer 0–6 km, and U_0 denotes the speed of the mean vector wind in the layer from the surface to 500 m—the quantity $(\bar{U} - U_0)$ is sometimes referred to as the “*BRN* shear”. For the storm-relative helicity, \mathbf{C} denotes the storm motion vector. See Thompson et al. (2003) for an explanation of symbols used for the *SCP* and *STP* calculations.

and large *SRH* are both pertinent to supercell forecasting, they need not be precisely collocated in space—many severe storm forecasters believe *CAPE* and *SRH* need simply be in proximity to

each other, perhaps with some overlap. Therefore, a variable based on the combination of these two variables may not adequately reflect the true potential for storms.

Further, *EHI* is the combination of two ingredients in an unphysical, arbitrary fashion. Can it

be shown that the formation of a supercell depends in some well-defined physical way on the product of *CAPE* and *SRH*? Although we cannot exclude such a possibility in the future, as of this writing, this product has not been shown to be physically pertinent, in the sense of appearing in some physically-relevant mathematical formula. Schultz et al. (2002) made a similar argument about the lack of scientific justification in reference to the PVQ parameter defined and proposed by Wetzel and Martin (2001).

Consider the following argument. Imagine a hypothetical world where the *EHI*'s two constituent parameters were all that is needed to forecast supercell tornadoes perfectly. Then picture a two-dimensional (2D) phase space of *SRH* and *CAPE* in which some irregular region within this phase space was associated with the occurrence of supercell tornadoes. By assumption, such a scenario would represent a perfect forecasting tool: if the values of *SRH* and *CAPE*, fall within this presumably complex region of 2D phase space, supercell tornadoes *always* occur. Anywhere outside of this region, supercell tornadoes *never* occur. Use of the *EHI* compresses the information within this hypothetical 2D space into a single number, eliminating our ability to apply the information in this phase space.

Furthermore, *EHI* is scaled somewhat arbitrarily by 160,000, so that whether the value of the *EHI* is 1 or 100 is similarly arbitrary. *EHI*'s scaling constants are associated with a "standard" value for *CAPE* of 1000 J kg^{-1} and for *SRH* of $160 \text{ m}^2 \text{ s}^{-2}$ —note that these units are actually equivalent and so *EHI* has dimensions of $(\text{J kg}^{-1} = \text{m}^2 \text{ s}^{-2})^2$, which has little obvious physical interpretation. For indices of this sort, the scaling constants are determined by what the developer felt to be typical values for the input variable. What is typical for some variable associated with severe weather in one part of the world may not be typical elsewhere. Hence, this can lead to incorrect interpretations of the index when used outside of the region in which it was developed (Tudurí and Ramis 1997). It is reasonable, we believe, to ask that a forecast parameter's utility and interpretation should not vary around the world.

Moreover, consider our hypothetical world again, in which the constituent parameters for *EHI* can be used to construct a 2D phase space that includes some complex region wherein supercell tornadoes always occur. It is conceivable that *some* transformation of the *EHI*'s constituent variables (*CAPE* and *SRH*) would convert the

complex region into a simple one—say, a circle. However, it seems highly unlikely that the existing scaling and unphysical construction of *EHI* could correspond to such a transformation.

This leads to the next point: how does a forecaster interpret *EHI* values? Put another way, does a situation where $EHI = 2$ mean that supercells are twice as likely as a situation where $EHI = 1$? If the *CAPE* doubles, doubling the *EHI* (assume *SRH* remains unchanged), does this imply twice the chance of supercells? There might be some empirical way to determine the significance of these values, but there is no physical rationale for interpreting them. Is the relationship between *EHI* and severe weather linear or nonlinear? How does the lagged correlation between *EHI* and the observed weather vary as a function of the lag time? Such questions ought to be of concern to any severe weather forecaster, but are certainly not readily answerable by the method *EHI* was constructed.

Finally, how has the forecast value of *EHI* been validated? Rasmussen (2003) described his evaluation of *EHI* for three classes of observed proximity soundings: supercells with significant tornadoes (those rated F2 and greater), supercells without significant tornadoes (no tornadoes reported, or only F0 or F1 tornadoes), and nonsupercell convection, defined in Rasmussen and Blanchard (1998). The weather events occurred anywhere from three hours before to six hours after the nominal 00 UTC sounding time. Rasmussen (2003) found, "only 25% of [proximity soundings from supercells without significant tornadoes] had $EHI > 0.5$, whereas nearly 2/3 of the [proximity soundings from supercells with significant tornadoes] had values this large." From the number of soundings in each dataset from Rasmussen and Blanchard (1998), this means about 30 events occurred in each category—a relatively small sample size. Taken together, these results indicate that when a supercell occurs with an $EHI > 0.5$ there is about a 50% chance of it producing a significant tornado. Observe that this figure amounts to a conditional probability, where the condition is the presence of a supercell.

Rasmussen and Blanchard used a contingency table and scatter diagrams (see section 5) to evaluate the diagnostic potential for *EHI* and several other candidate variables. But recall that their study evaluated indices from proximity soundings on the basis of observed events occurring in a time period around the sounding's nominal time. Therefore, it is not truly an analysis of the forecast potential for the variables con-

sidered—it is instead directed at a related problem: How well do *diagnosed* values of the indices discriminate among the observed events? This sort of analysis does not consider the topic of the lagged correlation between the proposed forecast parameter and the forecast severe weather events.

b. Pros and cons regarding the use of indices

We have shown, using the example of *EHI*, what types of problems can arise with the use of indices constructed in an arbitrary, unphysical way. We recognize that there are advantages as well as drawbacks to their use in severe weather forecasting.

As already noted, many severe weather forecasters already recognize the risks in relying on a single forecast parameter. Certainly most would never consider using a single variable to determine their forecast, but our experience suggests that *some* forecasters might be tempted to do so, perhaps because the use of some diagnostic variable (such as *CAPE*) is so widespread (e.g., the “disengaged” forecasters studied by Pliske et al. 2004). The pervasive use and ready availability of diagnostic variables is a trap for the unwary. In situations where the time pressure becomes intense, some might be inclined to do so in the interest of making a quick decision. Such a practice is antithetical to good forecasting, in general.

In a related line of reasoning, the press of time can encourage severe weather forecasters’ use of indices and other diagnostic variables to obtain a quick look at the data for the purpose of identifying the hot spots upon to focus more attention on in a diagnosis. By itself, this is a reasonable strategy. However, for reasons we’ve already mentioned, a serious forecaster is not likely to get much forecasting help from such a cursory consideration of atmospheric structure. There is nothing inherently wrong with a quick look, unless the forecaster limits the diagnosis to that. For all the reasons we have described, a conscientious weather forecaster should always try to find the time to do a comprehensive analysis of the data. We believe that using ingredients-based forecasting methods (e.g., Doswell et al. 1996) is a scientifically sound way to keep the diagnosis within practical time limitations in an operational forecasting environment.

It also can be argued from a purely utilitarian perspective that if a forecast parameter works successfully in forecasting, no matter *how* it was derived, it seems unreasonable to ask forecasters to cease using it. We don’t disagree with this at

all, especially when proven forecast parameters based on physical arguments are either unavailable or demonstrably inferior to a nonphysically constructed variable. If it can be shown rigorously that an arbitrarily constructed index does indeed have forecast utility (see the following section), we do not advocate ignoring its proven value to the challenge of weather forecasting – unless a physically-based forecast parameter is known to be superior for forecasting. Generally, physical reasoning is always preferable for the construction of diagnostic variables and indices, owing to the relative ease with which such forecast parameters can be interpreted and applied anywhere in the world.

5. Evaluation of forecast utility for a candidate prognostic variable

Because many diagnostic variables with potential forecast utility have never been tested rigorously as forecast parameters in their own right, we develop herein a general description of what we believe are the requirements that a proper forecast parameter would have to meet. As already discussed, diagnostic variables have their own specific purposes, but a diagnostic variable might also have the capability to make a reasonably accurate and perhaps even skillful prediction of the weather at some future time. See Murphy (1993) for a discussion of the difference between accuracy and skill.

One way to conduct a rigorous assessment of a variable as a forecast parameter is to use a developmental data set to form a classic 2 x 2 contingency table, the standard verification table when considering a dichotomous (yes/no) forecast for some dichotomous event (Wilks 2006, p. 260 ff). One example is given by Monteverdi et al. (2003) for tornadoes in California (see their Table 3 and Fig. 8). To create such a table for a potential forecast parameter, begin with choosing a threshold value for the candidate variable—forecast “yes” if the variable is at or above the threshold, and “no” if the variable is below the threshold. An assessment of the accuracy of the forecasts using the developmental dataset would be done by filling in the contingency table using that threshold. Optimizing the choice for the threshold value of the variable using the so-called Relative (or Receiver) Operating Characteristic curves associated with signal detection theory is possible; see Wilks (2006, p. 294 ff.) for more information.

The accuracy of the forecasts using that threshold can be assessed using standard meth-

ods based on the contingency table. The *skill* of the forecasts is determined by comparing the accuracy of the proposed forecasts based on use of that variable against the accuracy of some standard forecasting method (e.g., climatology or persistence, or some other forecast scheme, such as Model Output Statistics). If the forecast scheme using the proposed diagnostic variable shows statistically significant skill in comparison with some standard method, then it can be considered a useful forecast parameter.

To do a thorough assessment, however, another dataset is needed that is completely independent of the developmental dataset—in other words, a wholly different set of cases than those used for development and testing of the threshold values for the variable. If the results using the verification dataset are comparable to those found from the developmental data, confidence in the use of the variable as a forecast variable is correspondingly high. If there is a statistically significant difference between the results from the two datasets, then perhaps a larger sample is needed, but in any case, confidence in the forecast value of the variable (and its associated threshold value) is correspondingly low.

Two concerns often arise when a forecast parameter is proposed. First, many assessments of potential forecast parameter are done with a small number of cases, perhaps as few as a single example. Knowledge of how to determine an appropriate sample size is outside the scope of this paper; see the discussion of hypothesis testing by Wilks (2006, Ch. 5). It is incumbent on the developer of some forecast parameter to provide a reasonably thorough test using a sample of enough cases to give the parameter a rigorous test. Second, many attempts to validate the utility of some variable as a forecast parameter make the logical mistake of considering only values of the parameter when forecast events are known to have occurred. Diagnostic variables are of little use in forecasting until they can be shown to discriminate successfully between events and nonevents. Correct predictions of nonevents are not inevitably easy (e.g., Doswell et al. 2002) although in forecasting severe storms (relatively rare events), many nonevents are obvious.

Further, some diagnostic variables may have some value as forecast parameters, but it needs to be shown just how far in advance of the event they exhibit forecast accuracy and/or skill. That is, the accuracy of any diagnostic variable used as a forecast parameter is likely to increase as the time lag between the diagnosis and the event

decreases but this may not necessarily be a simple relationship. Contingency tables and full assessments as described above would have to be developed for a variety of diagnosis times relative to the beginning of the forecast events—say, 12, 6, 3, and 1 h before the actual events begin. Alternatively, an analysis of the time-lagged correlation between the forecast parameter and the observed weather could be carried out at a variety of lag times. However it is done, the accuracy of a candidate variable as a forecast parameter should be known as a function of time before the event. Having knowledge of forecast accuracy as a function of lead time is obviously important when using a diagnostic variable as a forecast parameter.

Another way to verify the potential of a forecast variable would be to construct a multidimensional scatter diagram (say, for the case of two dimensions, *CAPE* and shear) in which both events and nonevents are plotted with respect to observed values for the diagnostic variables. Using this plot, a probability of occurrence of the weather event as a function of its location in the scatter diagram could be developed, perhaps facilitated by the use of kernel density estimation methods (e.g., Ramsay and Doswell 2005). Examples of how this type of verification could be done are found in Rasmussen (2003) and Brooks et al. (2003), although this method would need to be accomplished with diagnostic variable values prior to the event, rather than proximity data.

Of course, the preceding does not present the *only* ways to assess the effectiveness of a proposed forecast variable. Many other possible methods could be used, but it does represent the level of rigor we believe is necessary before asserting that a diagnostic variable has real value as a true forecast parameter.

6. Conclusions

In our experience, many severe weather forecasters and researchers are seeking a “magic bullet” when they offer up yet another combined variable or index for consideration, whether or not they realize it. If forecasting were so simple as to be capable of being done effectively using some single variable or combination of variables, then the need for human forecasters effectively vanishes. There may be other reasons for the demise of human forecasters, but distilling the complex atmosphere with its nonlinear, possibly chaotic, interactions into an all-encompassing variable seems improbable. Any forecaster seeking to find such a variable is not only unlikely to

be successful, but, if success were achieved, the need for a human forecasting that event vanishes! Should advances in the science of severe convective storms ever produce such a forecast parameter, or should NWP models become near-perfect in terms of forecasting severe convection, then the need for human forecasters will indeed disappear, whatever our wishes might be. But it is our belief that this is not very likely to happen soon. Even if such an unlikely development ever occurs, in the interim, it remains incumbent on forecasters to use the information at their disposal as effectively as possible.

Acknowledgments. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce. We would like to thank our reviewers for their careful and thoughtful reviews and helpful suggestions, which have improved our presentation.

REFERENCES

- Banacos, P. C., and D. M. Schultz, 2005: The use of moisture flux convergence in forecasting convective initiation: Historical and operational perspectives. *Wea. Forecasting*, **20**, 351–366.
- Brock, F. V., and S. J. Richardson, 2001: *Meteorological Measurement Systems*. Oxford University, 290 pp.
- Brooks, H. E., C. A. Doswell III, and J. Cooper, 1994: On the environments of tornadic and non-tornadic mesocyclones. *Wea. Forecasting*, **10**, 606–618.
- Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94.
- Davies-Jones, R., D. Burgess, and M. Foster, 1990: Test of helicity as a tornado forecast parameter. Preprints, *16th Conf. Severe Local Storms*, Kananaskis Park, AB, Canada, Amer. Meteor. Soc., 588–592.
- Doswell, C. A. III, 1977: Obtaining meteorologically significant surface divergence fields through the filtering property of objective analysis. *Mon. Wea. Rev.*, **105**, 885–892.
- Doswell, C. A. III, 1986: Short range forecasting. *Mesoscale Meteorology and Forecasting*, P. Ray, Ed., Amer. Meteor. Soc., 689–719.
- Doswell, C. A. III, 1988: Comments on “An improved technique for computing the horizontal pressure-gradient force at the earth’s surface.” *Mon. Wea. Rev.*, **116**, 1251–1254.
- Doswell, C. A. III, 2004: Weather forecasting by humans – Heuristics and decision-making. *Wea. Forecasting*, **19**, 1115–1126.
- Doswell, C. A. III, and E. N. Rasmussen, 1994: The effect of neglecting the virtual temperature correction on CAPE calculations. *Wea. Forecasting*, **9**, 625–629.
- Doswell, C. A. III, and P. M. Markowski, 2004: Is buoyancy a relative quantity? *Mon. Wea. Rev.*, **132**, 853–863.
- Doswell, C. A. III, H. E. Brooks, and R. A. Maddox, 1996: Flash flood forecasting: An ingredients-based methodology. *Wea. Forecasting*, **11**, 560–581.
- Doswell, C. A. III, D. V. Baker and C. A. Liles 2002: Recognition of negative mesoscale factors for severe weather potential: A case study. *Wea. Forecasting*, **17**, 937–954.
- Galway, J. G., 1956: The lifted index as a predictor of latent instability. *Bull. Amer. Meteor. Soc.*, **43**, 528–529.
- Garcia, C., Jr., 1994: Forecasting snowfall using mixing ratios on an isentropic surface—An empirical study. NOAA Tech. Memo. NWS CR-105, PB 94-188760 NOAA/NWS, 31 pp. [Available from NOAA/National Weather Service Central Region Headquarters, Kansas City, MO 64106-2897.]
- George, J. J., 1960: *Weather Forecasting for Aeronautics*, Academic Press, 673 pp.
- Glickman, T. S., Ed., 2000: *Glossary of Meteorology*. 2nd ed., Amer. Meteor. Soc., 855 pp.
- Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology*. 2d ed., John Wiley & Sons, 477 pp.
- Hart, J. A., and W. Korotky, 1991: The SHARP workstation v1.50 users guide. National Weather Service, NOAA, U.S. Department of Commerce, 30 pp. [Available from NWS Eastern Region Headquarters, 630 Johnson Ave., Bohemia, NY 11716.]
- Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612.
- Miller, R. C., 1972: Notes on analysis and severe storm forecasting procedures of the Air Force Global Weather Central. Tech. Report 200(R), Headquarters, Air Weather Service, Scott Air Force Base, IL 62225, 190 pp.
- Monteverdi, J. P., C. A. Doswell III, and G. S. Lipari, 2003: Shear parameter thresholds for

- forecasting tornadic thunderstorms in northern and central California. *Wea. Forecasting*, **18**, 357–370.
- Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281–293.
- Panofsky, H., 1964: *Introduction to Dynamic Meteorology*. Pennsylvania State University, 243 pp.
- Pliske, R. M., B. Crandall, and G. Klein, 2004: Competence in weather forecasting. *Psychological Investigations of Competence in Decision Making*, K. Smith, J. Shanteau, and P. Johnson, Eds., Cambridge University Press, 40–68.
- Ramsay, H., and C. A. Doswell III, 2005: A sensitivity study of hodograph-based methods for estimating supercell motion. *Wea. Forecasting*, **20**, 954–970.
- Rasmussen, E. N., 2003: Refined supercell and tornado forecast parameters. *Wea. Forecasting*, **18**, 530–535.
- Rasmussen, E. N., and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164.
- Roebber, P. J., D. M. Schultz, and R. Romero, 2002: Synoptic regulation of the 3 May 1999 tornado outbreak. *Wea. Forecasting*, **17**, 399–429.
- Sanders, F., and C. A. Doswell III, 1995: A case for detailed surface analysis. *Bull. Amer. Meteor. Soc.*, **76**, 505–521.
- Schaefer, J. T., 1986: Severe thunderstorm forecasting: A historical perspective. *Wea. Forecasting*, **1**, 164–189.
- Schultz, D. M., J. V. Cortinas Jr., and C. A. Doswell III, 2002: Comments on “An operational ingredients-based methodology for forecasting midlatitude winter season precipitation.” *Wea. Forecasting*, **17**, 160–167.
- Showalter, A. K., 1947: A stability index for forecasting thunderstorms. *Bull. Amer. Meteor. Soc.*, **34**, 250–252.
- Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261.
- Tennekes, H., and J. L. Lumley, 1972: *A First Course in Turbulence*, MIT Press, 300 pp.
- Tudurí, E., and C. Ramis, 1997: The environments of significant convective events in the western Mediterranean. *Wea. Forecasting*, **12**, 294–306.
- Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520.
- Wetzel, S. W., and J. E. Martin, 2001: An operational ingredients-based methodology for forecasting midlatitude winter season precipitation. *Wea. Forecasting*, **16**, 156–167.
- Wetzel, S. W., and J. E. Martin, 2002: Reply. *Wea. Forecasting*, **17**, 168–171.
- Wilks, 2006: *Statistical Methods in the Atmospheric Sciences*, 2d ed., Academic Press, 627 pp.

REVIEWERS' COMMENTS: