On Moist Instability*

STEVEN C. SHERWOOD
Universities Space Research Association, Seabrook, Maryland

13 March 2000 and 1 May 2000

ABSTRACT

An argument is made that the concepts of conditional instability and conditional symmetric instability need to be revisited. Confusion in the profession has led to two extant definitions of conditional instability that are superficially similar but fundamentally inconsistent. Only one definition corresponds to a real instability. Further, spillover of this confusion into the analysis of slantwise instability mechanisms appears to be producing inappropriate diagnosis of conditional symmetric instability. Concepts from hydrodynamic stability theory are helpful in discussing the situation.

Conditional instability (CI) has traditionally been defined (and will be defined here) as the occurrence of an atmospheric lapse rate that is steeper than a moist adiabat but less steep than a dry one. More recently, the CI label has been attached to states in which an upward-moving air parcel can become positively buoyant. In this note I argue that the inconsistency between these two definitions runs deeper than is commonly understood, and mirrors confusion in the way many of us in the profession conceive of atmospheric stability itself. Further, in light of the recent paper by Schultz and Schumacher (1999, hereafter SS) who reviewed the use of upright moist instabilities and their ''symmetric'' counterparts (instabilities to slantwise perturbations), it is clear that the confusion is probably causing inappropriate diagnosis of conditional symmetric instability (CSI) even beyond the problems that SS already noted. This long-standing confusion therefore requires attention and remediation.

It will be useful to begin with a brief digression into fundamental concepts. To speak of instability it is essential to distinguish between the *basic state* and one or more *perturbations* to which that state is unstable. Unstable perturbations grow exponentially in time, drawing energy from the basic state. Thus instability requires some form of stored energy (as implicitly acknowledged in meteorology by the oft-mentioned concept of ''release'' of various forms of buoyant instability, which presumably means release of potential energy stored in the basic state). Stored energy does not guarantee instability, however; for example, the earth's potential energy could be lowered if Mount Everest were toppled into the Indian Ocean, but this does not lead us to speak of the planet as being ''unstable.'' The existence of instability in any useful sense also requires that there be some perturbation, whose occurrence is inevitable (or at least likely), capable of gaining access to the stored energy. In the atmosphere the term convection is typically reserved only for overturning motion that draws upon available *potential* (rather than kinetic) energy in an atmospheric column or slab, and we will so reserve it here.

It is simple to show that the atmosphere is gravitationally unstable to infinitesimal vertical perturbations of any horizontal scale if

$$d\theta/dz < 0 \qquad (1)$$

in the basic state. In baroclinic, rotating environments, the symmetric form of this instability also applies, with the derivative evaluated on a surface of constant absolute angular momentum rather than in the vertical direction. However, convection is observed to occur in atmospheres where neither the upright nor slantwise condition is met, a circumstance clearly due to moisture behavior.

It has therefore become standard practice to recognize two simple forms of *moist* instability, labeled potential instability (PI) and conditional instability, respectively, in which the variables θ_e and θ_{es}, respectively, are substituted for θ in (1). Thus CI is indicated for lapse rates between moist and dry adiabatic as defined earlier, and supposedly applies if the atmosphere can be brought to saturation somewhere by some means. As we will see below these terminologies stray from classical stability concepts, which can be blamed at least partly on the

* This note is the first of two in this issue that discuss concepts of moist instability. The other, by Schultz et al. (2000), appears next.

Corresponding author address: S. Sherwood, NASA Goddard Space Flight Center, Mail Code 916, Greenbelt, MD 20771.
E-mail: ssherwood@alum.mit.edu

failure of classical theory to tackle the tough problem of instability in a time-varying basic state. The central concern of this note is to explore in what sense (if any) CI and PI are really instabilities, and what this implies for their analogously defined symmetric counterparts CSI and PSI.[1]

One approach to understanding CI and PI is to consider them as possible examples of what is known in fluid mechanics as *subcritical* instabilities, which means that the basic state is unstable only to perturbations above a critical amplitude (Drazin and Reid 1981). In the moist-convective case, the critical amplitude is that required to saturate the atmosphere. Note that in this concept a single perturbation perseveres to become a growing disturbance of the same scale.

An alternative instability concept is to consider the sequence of events leading to convection as *modification of the basic state itself*—by what might be called a ''preconditioning'' perturbation—until it possesses ordinary instability. This view, popular in meteorology, is more appropriate than the subcritical instability concept if the expected ''preconditioner'' is qualitatively different from the perturbation that will actually consume the available energy (e.g., of different horizontal scale). In this case the original state is simply stable, but happens to lie nearby to an unstable one in state space. I shall call such states ''near unstable.'' Clearly, potential instability is an example of such near instability: if a whole layer possessing PI is lifted (adiabatically) far enough, it becomes unstable to convective-scale perturbations within the layer. Such a lifting preconditioner can readily occur in the real atmosphere.

But which of these two instability concepts includes the traditional, lapse-rate definition of CI? Neither one. We cannot classify CI as a near instability, since the required preconditioner would have to be isothermal moistening (i.e., increase in water vapor mixing ratio) of the basic state in order to raise its θ_e to θ_{es} with no other change—not a realistic event in the real atmosphere. Adiabatic lifting does *not* work as a preconditioner for CI since lifting of an unsaturated air mass does not conserve its θ_{es}. Similarly, we cannot classify CI as a subcritical instability either, since the decrease of (nonconserved) θ_{es} with height does not guarantee the existence of any finite perturbation capable of extracting potential energy from the basic state (a moisture-free atmosphere could meet this criterion, for example, yet be absolutely stable to any vertical air displacement). In fact lapse-rate CI is not an instability in any meaningful sense.

Despite this, the traditional definition of CI is a useful concept when viewed from a different perspective. When faced with a stability problem, one ideally wants

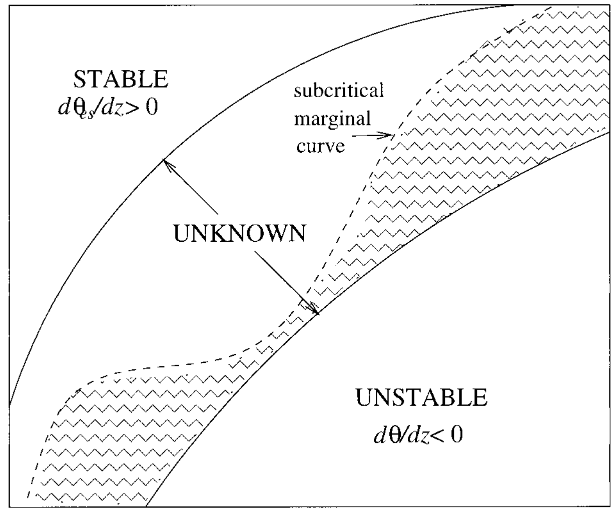

FIG. 1. Schematic diagram of state space at some location in the atmosphere. Solid curves are necessary and sufficient conditions for instability, defined by moist and dry adiabats, respectively. Dashed line shows the marginal stability curve for subcritical instability of a lofted near-surface parcel. The traditional lapse-rate definition of CI encompasses the entire region between the solid curves; subcritical instability occurs in the filled area only.

to establish the ''marginal stability curve'' (or surface) in state space that exactly separates stable basic states from unstable ones. But if this is impossible, the next best that one can do is to bracket the marginal curve with a *necessary* condition for instability on one side and a *sufficient* condition on the other, dividing state space into regions of stability, instability, and unknown stability (see Fig. 1). The choice of these conditions is arbitrary, but the ''dry'' condition (1) clearly serves as a simple choice of sufficient condition for moist instability. Likewise, if θ_{es} increases with height, instability is clearly ruled out, making decrease of θ_{es} with height a good choice for a necessary condition. Thus, lapse-rate CI corresponds to the middle region in Fig. 1 and can correctly be viewed as a statement of *uncertainty* about instability. From the forecaster's perspective, this usefully distinguishes cases needing further scrutiny from those where instability can quickly be confirmed or ruled out.

The problem is that, too often, CI is not thought of as a statement of uncertainty but as a real instability, even by those who stick to the lapse-rate definition. The confusion is propagated in many popular textbooks, which give this definition and immediately afterward begin to discuss conditions for the ''release'' of CI. The common interpretation of lapse-rate CI in terms of a quantitative ''condition'' for instability is fundamentally specious, since the required condition is a fictitious, saturated parcel that is different from any air actually measured. One may just as well posit ''conditional'' instability in a sounding if a hypothetical parcel 10 K warmer than any observed could have been buoyant in that sounding! In fact neither of the words conditional

[1] Schultz and Schumacher (1999) complain that the use of the symmetric terms in the literature has not always been consistent with that of their upright counterparts.

instability properly describes what is really meant by the lapse-rate definition (regrettable, but these things do happen). *Lapse-rate CI is ignorant of any specific mechanism for initiating a disturbance* (see again Fig. 1). Any actual mechanism must possess a marginal stability curve, which would lie within the CI region bracketed by the two adiabats. The figure shows only one such curve but any number of additional curves corresponding to other types of instabilities (e.g., PI) could be envisioned, too.

By now it must be clear that low-level air parcels often do possess characteristics enabling them to attain positive buoyancy if sufficiently lifted, corresponding to a genuine subcritical instability, but that the traditional lapse-rate formula for CI is incorrect for establishing this. This is widely recognized in practice. Forecasters and researchers who want to evaluate atmospheric stability nearly always bypass lapse-rate CI in favor of convective available potential energy (CAPE), lifted index, or some other variable that essentially compares low-level θ_e with higher-level θ_{es}. This ''parcel theory'' calculation truly does correspond to a test for subcritical instability, with the perturbation being forced, finite-amplitude uplift on the convective scale (though details of how best to perform this CAPE or CAPE-like calculation are still debated).

Since this uplifted parcel buoyancy situation is commonly associated with the term conditional instability, some have advocated *redefining* the CI label accordingly. Thus, more recent use of the term CI in the literature often refers to a state in which a lifted parcel *will* achieve positive buoyancy according to some calculation (e.g., Xu and Emanuel 1989), whereas the traditional definition only implies that instability *may or may not* exist. This has now led to the inclusion of a second definition of CI, based on the existence of available potential energy, in the new version of the *Glossary of Meteorology* (American Meteorological Society 2000). The two definitions are far from equivalent, however, since the lapse-rate ''instability'' criterion is met nearly everywhere in the low to midtroposphere, including subtropical areas that experience very little deep convection and hardly ever have (e.g.) CAPE. Any argument about which definition of CI is ''better'' is not likely to be productive since each has value and the preference will largely depend on one's philosophy.

The ingredients-based methodology espoused by SS and others, in which convection is said to require lifting and/or moisture in addition to nominal instability, is an effort to elevate lapse-rate CI into a true stability test by crudely identifying portions of the CI region that are likely to lie on the stable side of the marginal stability curve. This does not really solve the problem, since the additional ingredients necessary for convection depend crucially on which definition of CI is being used. The unfortunate labeling of one of the ingredients as an ''instability,'' when (classically speaking) instability is actually what the ingredients are meant to *diagnose*, adds confusion. The point (made by SS and others) bears repeating here that convection and instability are not the same thing either, and that subcritical instability (i.e., CAPE) does not automatically lead to convection.

This finally brings us to conditional symmetric instability. While the lapse rate criterion has basically been discarded for diagnosing upright instability, it typically continues to be used to establish CSI (except when the PSI criterion is erroneously used; see footnote 1), and is even suggested for use by SS. I believe it is a serious mistake to use a lapse-rate version of CSI to try to diagnose slantwise instability mechanisms, since in cases where a researcher is getting sufficiently ambitious to be doing this a posteriori, the mere statement of possibility inherent in the lapse-rate definition is simply too weak to be of any scientific value. Only the lack of lapse-rate CSI would be interesting. Instead, a slantwise version of CAPE (SCAPE) would seem necessary to establish a meaningful *instability* of any scientific value for helping to understand convection and its organization.

Schultz and Schmacher (1999) point out that SCAPE is not often used since it appears to be small or zero most of the time, and is not found to correlate well with the observed features thought to indicate slantwise convection. But it must also be pointed out that, despite the important role of CAPE in regulating upright deep convection, CAPE bears no simple correspondence to convective activity in observations; one would not expect the relation between SCAPE and slantwise convection to be any less subtle. If a slantwise subcritical instability is important, there must be some variant of SCAPE that can be used to diagnose it. One possibility is that significant SCAPE occurs more often than is believed, due to improper accounting of physical effects (such as precipitation) in its calculation. Perhaps cloud-resolving models with good microphysics can be used to discover a proper SCAPE-like criterion. It would be surprising if slantwise subcritical instability were of no importance, since lofted-parcel instability appears to be a very important mechanism for upright convection. For example, cloud-scale perturbations are often observed to initiate disturbances before large-scale lifting can properly convert PI into ordinary instability even when PI is present; also, in the Tropics, lofted-parcel instability is widely recognized to be more relevant than PI.

In summary, the traditional, lapse-rate definition of conditional instability does not correspond to an instability in any sense, and so is not aptly named, but does correspond to a statement of uncertainty that is useful in a forecasting context. The slantwise extension of this traditional lapse-rate definition has been used inappropriately in studies of slantwise convective mechanisms as if it indicated a real instability, producing misleading scientific discourse. I see no useful purpose for this slantwise extension. However, a very important subcritical or lofted-parcel instability certainly does exist—and is often called CI—but (as is commonly recognized) requires a calculation different from the traditional definition of CI. Finally, in contrast to CI, PI can be unambiguously understood as a straightforward near instability, but is not recognized by experts to be highly useful on its own in actual practice. Several suggestions seem to follow from this discussion:

1) Further confusion could be mitigated if separate terminology were adopted for the separate concepts of uncertainty based on lapse rate (now traditionally labeled CI), and actual parcel (subcritical) instability. At least one of these could use another name besides conditional instability.
2) Some legitimate index (e.g., SCAPE or a modification thereof) is essential in order to establish any meaningful slantwise parcel instability for understanding convective behavior; the traditional definition (the symmetric analog to the traditional definition of CI) is inadequate. Whether to call this type of instability CSI or not depends somewhat on suggestion 1 above.
3) Those studying convective mechanisms in models or observations may wish to pay close attention to any apparent convective ''trigger'' (a local precursor to convection) to see whether it seems to function as a ''preconditioning perturbation'' or whether the trigger itself grows into an unstable disturbance. This may not be easy, especially since multiple instability mechanisms may often act in concert as noted by SS, but should help to focus the line of inquiry and clarify how unstable growth occurs.

Acknowledgments. This note grew out of a brief series of stimulating discussions on this topic with Chuck Doswell, Dave Schultz, and Kerry Emanuel. The author thanks each of them for sharing their time and their views.

REFERENCES

American Meteorological Society, 2000: *Glossary of Meteorology.* 2d ed. American Meteorological Society, 855 pp.

Drazin, P. G., and W. H. Reid, 1981: *Hydrodynamic Stability.* Cambridge University Press, 527 pp.

Schultz, D. M., and P. N. Schumacher, 1999: The use and misuse of conditional symmetric instability. *Mon. Wea. Rev.,* **127,** 2709–2732.

Schultz, D. M., P. N. Schumacher, and C. A. Doswell III, 2000: The intricacies of instabilities. *Mon. Wea. Rev.,* **128,** 4143–4148.

Xu, K. M., and K. A. Emanuel, 1989: Is the tropical atmosphere conditionally unstable? *Mon. Wea. Rev.,* **117,** 1471–1479.