THE DEPICTION OF DROUGHT

A Commentary

BY KELLY T. REDMOND

This issue of *BAMS* contains several articles that emphasize how we describe the status of drought. Interest in this topic has risen because of a combination of natural and human factors. The Southwest–southern Great Plains drought of 1995–96 led to the establishment of the Western Drought Coordination Council (WDCC) by the Western Governors Association. The subsequent National Drought Policy Commission (NDPC) expanded on WDCC recommendations distilled from the western experience. Both groups emphasized climate monitoring as a necessity. In May 1999, drought in an area of tremendous climatic significance—Washington, D.C.—led to the establishment of the Drought Monitor. A small cadre of climatologists, originally in the West—the Climate Prediction Center, the National Climatic Data Center (NCDC), and the National Drought Mitigation Center—were involved in every step in this sequence, and participation has gradually grown since.

The Drought Monitor is both a product and an activity. An extended e-mail "conversation" takes place for about 2–3 days each week, and the results are assimilated by the "author of the week" into a consensus product in the form of a national map. During a typical week about 15–20 out of a total distribution of around 140 individuals participate, mostly from affected areas. The article by Svoboda et al. captures the flavor of the Drought Monitor very well. The discussions, ranging widely as conditions and issues unfold, have proven to be a rich source of provocative thought. It is from these and numerous earlier WDCC discussions, and other experiences over the prior 15 years, that the personal perspective has been acquired to offer the observations and comments that follow.

AFFILIATION: REDMOND—Western Regional Climate Center, Desert Research Institute, Reno, Nevada
CORRESPONDING AUTHOR: Dr. Kelly T. Redmond, Western Regional Climate Center, Desert Research Institute, 2215 Raggio Parkway, Reno, NV 89512-1095
E-mail: krwrcc@dri.edu

DEFINITIONAL ISSUES. In depicting the status of drought, we are almost immediately confronted with the question, "What is meant by drought?" The problem of defining drought is longstanding and has never been resolved to the satisfaction of all. This fact alone surely tells us something. Drought means many different things to many different audiences. Early definitions focused on purely meteorological or hydrological causes. In my opinion, the preferred definition would be that which has the most universal range of application, the one that works in the largest number of circumstances.

Most concepts of drought involve a water balance. This implies that both supply and demand must be considered, as well as the question of whether there is "enough" (and, enough for what?). Thus, through time I have come to favor a simple definition; that is, insufficient water to meet needs. This covers a broad range of situations, from an asteroid to a Pacific Ocean beach, if need be. By intention, it highlights the importance of both the supply and the demand sides of the issue. From this standpoint, a system is "in drought" when supply does not meet demand. Such could occur from diminished supply with relatively constant demand, or from relatively constant supply and increased demand. Supply involves mainly the physical climate system (biological feedbacks complicate this), but demand very heavily involves the biological world, which includes humans and other living things. The demand side is much more subject to manipulation by humans than the supply side, and some of this manipulation involves responses to past (or even anticipated) supply variations. It is mostly because of this demand side and the properties and causes of its variability that I feel we cannot escape consideration of impacts in how we characterize and describe drought.

This approach in effect defines drought in terms of its impacts. Population growth from hamlet to metropolis, and attendant water demand growth, could thus turn a dry spell into a drought. Like the tree falling in the forest, does drought occur if there is no human to record or experience it? In my view, the concept extends to vegetation and ecosystems, when insufficient water to meet needs necessitates adjustments at individual and community levels. If we take this approach, to properly track drought status we need access to data on both the physical hydroclimatic system and on all the various types of impacts.

MAKING INDICES. Like twilight, drought creeps stealthily into existence (with a few exceptions), unannounced and often not obvious. A set of early warning indicators is needed to alert us that societal adjustments may be needed and that other biological systems of concern face similar circumstances. In developing an index to assess drought status, we quickly discover that we cannot escape from the issue of defining drought itself.

The history of this endeavor is nicely recapped by Richard Heim in this issue. Over the years, a large number of such indices have been proposed and used. Many of these, suitable only for very restricted and specific applications, suffer from a kind of misplaced exactitude for use as general indicators, recalling what economist Kenneth Boulding once described as "suboptimization:" finding the best way to do something that should probably not be done at all.

There is often a temptation to proceed directly from datasets to the formulation of indices, with minimal attention to the larger issues surrounding the usage of indices. This reflects a kind of fascination with the numbers and numerics to the exclusion of more fundamental considerations. What is the purpose of an index? Who is the audience? What do they want it to tell them? These questions would be better asked at the start of the process rather than as an afterthought. What properties and behaviors would we like indices to exhibit? A short list assembled for a previous drought conference is included in Table 1. These questions are seldom brought up. In a very interesting article John Keyantash and John Dracup use six such criteria, some of which overlap those in Table 1, to evaluate a number of indices (many covered in the article by Richard Heim) commonly used to describe drought. The outcome of their evaluation seems to reflect reasonably well the real world experience with the Drought Monitor and earlier related activities.

The development of the Standardized Precipitation Index (SPI) furnishes an example of how helpful an orderly and thorough process can be for later users. It was designed at the Colorado Climate Center by T. McKee, N. Doesken, and J. Kleist, specifically for two reasons: (a) to recognize and even *emphasize* that accumulated precipitation can be simultaneously in excess and deficit on different timescales, and (b) for practical purposes of answering the five questions most commonly asked by water managers. For any one of many timescales, these are as follows.

1) What is the absolute amount of precipitation that has fallen (in units, e.g., millimeters or inches)?
2) What is the departure from average (in units, e.g., millimeters or inches)?
3) What is the departure in percentage terms?
4) What is the value in frequency space (percentile)?
5) What is a single number that best encompasses this information (e.g., the SPI, which is closely akin to a departure in standard deviations)?

What makes the SPI unusual, in addition to the strong user orientation that guided its definition, was the *very* extensive and careful testing and diagnosis of its properties—using nearly 1200 U.S. century-long Historical Climatology Network stations—prior to its formal introduction. As it happens, the SPI (a precipitation-only index) shows the highest correlation coefficients, mostly 0.80–0.90 or more, with the Palmer Index (which has a very convoluted algorithm) at timescales of 6 to 12 months, typically with a maximum around 9 months. This forms the intrinsic timescale of the Palmer Index. Such high correlations left a message from this exercise: temperature (in the Palmer family) adds very little information, so leave it out. The SPI is conceptually very simple; it measures transformed precipitation at a specific timescale. A fruitful exercise would be to perform similar comparisons with other indices to obtain their intrinsic timescales.

Having followed the origins of several drought indices, and even participated in the development of a few, I find that the process often calls to mind the old adage about laws and sausage: two things one would rather not want to watch being made. A typical sequence is as follows. Someone hatches an idea for a new index; an algorithm and code are created; requisite observational input data (time series) are assembled, usually at considerable travail and effort, often based on different periods of record, not always homogeneous, and many times exhibiting unusual statistical properties. (For example, all three of the latter problems plague streamflow and storage records from artificially managed river systems.) In many agency and operational settings, workplace pressures do not favor a thorough examination of the index behavior in a variety of real-world situations. Issues such as characteristic timescales, histogram shapes, return frequency, seasonal or geographic or hydrologic basin dependencies, response properties to known inputs, and behavior for limiting cases are often ignored or casually investigated.

Table 1. Desirable properties of indices (Redmond 1991).

1) Detailed understanding of caveats, limitations, assumptions should not be critical to proper interpretation for indices in wide public use. Indices should not be too complex.
2) Indices should not be overly simplified (e.g., "Colorado statewide precipitation" lumps too many things together).
3) Indices should offer improved information over raw data values.
4) For routine practical usage, historical time series of data must be readily available, recent values must be quickly computable, and both must be compatible (homogeneous record).
5) It is helpful if social and economic impacts are proportional to the value of the indices.
6) Indices values should be open-ended. Unprecedented behavior yields unprecedented values.
7) Normalization to background climate, in nondimensional units, greatly facilitates spatial comparisons across very different settings.
8) Statistical properties and sensitivities should be thoroughly evaluated before operational usage.
9) Subindices, component indices, or other spin-offs help debug or explain unusual, alarming, or otherwise interesting behavior.
10) Measures of placement within the historical context are invaluable and frequently requested, typically as percentiles. The goal should be a 50–100-year perspective.

EVALUATING INDICES. Because of such pressures, indices are usually implemented without a complete evaluation of either their temporal statistical properties or relation to climate impact information (e.g., soil moisture, crop conditions, or economic or societal indicators). A number of drought indices go into usage without description in reviewed publications, and fewer yet receive an exhaustive evaluation after escaping into widespread usage. The very widely used Palmer Index family was introduced in 1965, but not thoroughly documented until the excellent critique by William Alley in 1984. (Actually, Palmer himself did not intend or foresee significant usage beyond the Great Plains.) Going a step further, 26 years after the Palmer Index was invented, Ned Guttman at NCDC supplied climate "shocks" (one time step impulses in otherwise strictly climatological time series) and found bizarre nonphysical "ringing" patterns for years after the impulse. Such intensive evaluations, especially once an index is in use, take a great deal of effort and are difficult to justify in the operational environment of an agency trying to fulfill a mission that only partly depends on climate. The evaluations often involve activities that such agencies rarely seek: follow up on a past activity, with risk of an unfavorable outcome.

How do we evaluate the quality, accuracy, and utility of an index? As a beginning to evaluation, indices can be compared to each other (as in the article by Heim). For example, new indices are frequently compared with Palmer values for the same time sequence (e.g., the various *N*-month SPI series). However, good correlations indicate similarity of behavior, which means they can at least mimic each other, but does either match some indicator of reality? What serves as "ground truth?" What if there are many ground truths to choose from (which is almost always the case)? The apparently simple act of choosing such indicators turns out to be very difficult indeed. The Palmer Index is often the standard of comparison simply because it is familiar, widely available, and frequently quoted; it is an index that to some extent is used because it is used.

Some index problems that come to light during operational evaluation arise from the flaws in the input data rather than intrinsic flaws in the index itself. Most drought indices are heavily driven by precipitation, and some also include temperature and other elements. It is self evident that an index can be no better than the quality of its input data: bad input produces bad output. An index should not be blamed when the data quality is the main problem. As a case in point, several prominent indices use the widely available monthly data for the 344 climate divisions in the contiguous United States. If correctly computed, these values are extremely useful for many applications. However, rapidly formed estimations based on limited provisional or "temporary" data at the start of a new month are sometimes inaccurate. These can show up as conspicuous blotches on a map and make an index appear bad. Unfortunately, to confound things, in the West and elsewhere there can occasionally be large *real* differences between adjoining divisions. Inhomogeneities in the data time series can lead to similar effects. In the West, sampling periods and record lengths are often quite different between valleys and nearby mountains, influencing frequency distributions when climate regimes (wet or dry decades) are present. Stationarity of aggregated (multiple station) climate records cannot always be assumed when there are large climatology differences in small spatial domains (up to factors of 10–30 for a given month).

DEPICTION DILEMMAS. Complicating the description of drought even more are regional effects and differences, such as in the West. Prime among these is the presence of natural lags like those in the hydrologic system, wherein mountain precipitation does not become snowmelt for 3–9 months. Large differences in annual precipitation totals and in seasonality (the shape of the annual cycle) can be found only a few miles apart. Places where winter is the wet season can be within sight of places where winter is the dry season. Such elevational effects are found nearly everywhere in this region. Droughts in wet locations (the Pacific Northwest) can seem quite wet; droughts in dry locations (the desert Southwest) may appear barely different from normal aridity. In some regions (e.g., the West Coast, or at high elevations throughout the West), the most significant droughts are winter events. The effects of a dry winter cannot usually be made up until the next winter. In the meantime, can a drought in effect "hibernate," only to quickly reappear the next spring, if the winter was dry? The Drought Monitor discussions have also illustrated rapidly developing "flash droughts" in the Great Plains where vegetation quickly depletes recent moisture in shallow root zones while the deeper soil has remained dry all along.

Water supply regions for agriculture, human consumption, recreation, power, navigation, fish, and wildlife can be hundreds or thousands of miles from demand regions. As a result, severe and substantial drought impacts in one location can occur as a result of climate events far away, or long ago, or both. The 2001 energy crisis in California was greatly abetted by drought conditions in the Pacific Northwest and southwest Canada. One quickly realizes that portrayal of long distance and long lag connections, and other factors just mentioned, on a single map, in a form simple enough for use by newspapers and political audiences, always has limitations no matter how carefully crafted. Furthermore, although the Drought Monitor strives to use "D0–D4" to express approximate return values for "the" situation, there are in reality *many* situations at once—a multitude of return values in terms of both climate and impacts, from rare to common, at any one time or location, in addition to the distance and time connections.

Another kind of simplification pertains to the use of labels, such as the common-language terms that are mapped onto the "Dx" categories of the Drought Monitor, a practice not limited to this product. For example, the Palmer Index is a continuous metric to which numerical ranges have been assigned subjective labels ("incipient," "mild," "moderate," "severe," and "extreme"). Labels sometimes are assigned without strong rationale, often by other parties not involved in the design or subsequent computation. Thereafter the labels can easily take on a life of their own, used with little close questioning or research on how these terms are perceived by different audiences. A numerical drought index should not be judged on the basis of the subjective labels applied to it. Further, the Palmer Index is usually formally labeled as the Palmer *Drought* Index, a factor that alone surely facilitates usage. Labels are useful, and indeed needed, for some audiences, but the impacts associated with a given numerical value are usually variable across sectors and applications. Also there is debate as to whether numerical metrics should be considered as "drought indices" or simply as "climate indices," with implied freedom to describe neutral and excessive water availability as well.

It is not likely that there is an easy way out of this depiction dilemma. Drought is a many-headed creature, and its full description requires an equally diverse menagerie of indices and indicators.

IMPACT INFORMATION. The effect of drought varies widely, and the same physical climate sequence can easily produce very different impacts to a wide variety of economic sectors, ranging from none to extreme. In essence, as with rainbows, each person experiences their own drought. If a drought or climate index is to be adequately evaluated, through correspondence with observed effects, then (homogeneous) time series of such impact information need to be available to properly determine relationships. This impact information is typically difficult to find, often much more so than the climate information against which it will be compared. A knowledgeable and well-connected coordinating presence is usually needed to link information from the climate world with that from the impact world. In some states this role is invested in the state climate office, but many states lack the resources to support such expertise.

The national infrastructure for providing impact information has many flaws and is very uneven. The smaller spatial scales, where almost all impacts are ultimately felt, are not sampled or reported adequately, and in some cases, at all. This critical lack of attention to climate needs at local, state, and regional scales has been highlighted in a number of recent workshops and reports from the National Academy of Sciences and other organizations, and should be a high priority for future development as the nation proceeds to build a comprehensive national climate services program. Drought monitoring is encompassed under such an umbrella—probably the only umbrella needed for drought.

REFERENCES

Redmond, K. T., 1991: Climate monitoring and indices. *Proc. Seminar on Drought Management and Planning,* D. A. Wilhite, D. A. Wood, and P. A. Kay, Eds, Lincoln, NE, University of Nebraska, Lincoln, 29–33.