J11.1 DEVELOPMENT OF A STATISTICAL VALIDATION METHODOLOGY FOR FIRE WEATHER INDICES

Brian E. Potter
USDA Forest Service, East Lansing, Michigan

Scott Goodrick
USDA Forest Service, Athens, Georgia

Tim Brown
Desert Research Institute, Reno, Nevada

1. INTRODUCTION

Fire managers and forecasters must have tools, such as fire indices, to summarize large amounts of complex information. These tools allow them to identify and plan for periods of elevated risk and/or wildfire potential. This need was once met using simple measures like relative humidity or maximum daily temperature (e.g., Gisborne, 1936) to describe fire weather, and with increasing decision-support requirements over time, eventually led to more complex systems like the National Fire Danger Rating System (NFDRS). When there was no readily available index for some fire or fire-weather characteristic, managers and forecasters borrowed what they could find from related fields. For example, the K-Index and Lifted Index, designed for prediction of severe weather linked to thunderstorms, were used to indicate stability before the development of the Haines Index. High values of these non-fire indices resulted from high moisture as well as high instability, and as such were not ideal for fire weather forecasting.

Before any index is used operationally it should undergo rigorous testing, evaluation and validation to fully document its relationship to fire activity, and any associated limitations. In reality, there is cultural, political and economic pressure to provide as much information as possible for fire management and fire suppression decisions. Because of this pressure, researchers have often developed fire-weather indices – and forecasters and managers have often used them – without rigorous testing, evaluation and validation.

Across the nation various land management agencies – public and private, federal and state – currently employ a range of fire indices that rely on some form of atmospheric input. These indices include, among others, NFDRS, the Canadian Fire Weather Index (CFWI), the Keech-Byram Drought Index (KBDI), the Haines Index (HI), the Fosberg Fire Weather Index (FFWI), the Davis Stability Index, the Lifted Index, Convective Available Potential Energy (CAPE), the Ventilation Index, the Lavdas Dispersion Index, and the Fire Potential Index. Of the eleven indices listed here, only five (CFWI, KBDI, HI, FFWI, and Davis Stability Index) were originally and primarily concerned with the weather component of the fire environment (fuels and topography being the other components.) And of these five, only the HI was developed and tested, and then subjected to some scientific scrutiny by being published complete in a peer-reviewed process. Even though it was initially tested and reviewed, there is growing evidence that the HI does not work equally well in all regions. Only now, after nearly 25 years, is the NFDRS beginning to be tested against actual fire occurrence data (Andrews, et al. 2003)

We propose that the use of untested fire indices is unscientific, as well as potentially unsafe and unwise. It can lead to a false sense of confidence, flawed decisions, economic loss, property damage, and injury or fatalities. In this paper, we lay out a road map to establish standards for the validation of fire-weather indices.

2. PHILOSOPHY

This is not an appropriate forum for a lengthy discussion of the method. In a simple sense, the steps in the scientific method include observation, formulation of a hypothesis based on the observations, testing the hypothesis with observational data, and acceptance or rejection of the hypothesis based on results of the tests. A typical fire weather index develops as follows. Someone observes or believes there is an

* - *Corresponding author address:* Dr. Brian E. Potter, North Central Research Station, 1407 S. Harrison Road, Suite 220, East Lansing, MI 48823; email: bpotter@fs.fed.us

association between atmospheric property **A** and fire property **F**. Based on this, they propose the hypothesis "**A** and **F** are correlated" or "**A** contributes to **F**." They then collect some data, apply statistical tests, and find **A** either does or does not correlate with **F**. The investigator, or someone else, then proposes that **A** can be used to predict large values of **F**, and so **A** can serve as an index for a particular purpose such as prediction, assessment, or decision-making.

This breaks down for 2 basic reasons. First, fire-atmosphere interaction is incredibly complex and a simple correlation in no way guarantees a causal relationship (in fact, a strong apparent association may be the result of one or more unknown variables). Second, without an explanation of the physical reasons that **A** affects **F**, one cannot determine when it is and is not appropriate to use **A** as an index for **F**. One could as easily claim that rustling leaves on trees cause earthquakes.

In reality, the problems with fire indices are more profound than this. They often result from regional, short-term data samples that are then extrapolated to all times and other regions. For example, Haines (1988) used a data set that included many North Carolina fires to derive his Index. Werth and Werth (1998) showed that the HI works well for some western regions, like Idaho, but very poorly for other regions like southern Nevada. Davis (1969) noted that most fires in the Louisiana, Arkansas, Tennessee, Mississippi, and Alabama region occurred when conditions were "absolutely unstable." Based on his definition of "absolutely unstable" and Holzworth (1972)'s calculations of mixing heights, most of the continental United States is "absolutely unstable" for most of the summer.

To improve the reliability and utility of indices, we propose that if existing indices were at least tested and measured using common data and common tests as much as possible, their relative merits could be assessed, perhaps along with their regional or temporal idiosyncrasies.

4. DATA

We first suggest the creation of a standard index validation data base. This would comprise appropriate fire and meteorological data. The fire data should include both regional data from multiple years and data from specific, individual fires. Regional, multi-year data allow for examination of index performance on a regional and/or seasonal basis, as well as relationships between indices and the number of fires by size class. Using a large database should reduce the impact of suppression, slope, nonhomogeneous fuels, etc. Data from individual, specific fires allows examination of index association and value with respect to daily fire growth rates and/or behavior characteristics. This task will involve an extensive search of historical fires, covering the entire U.S. and perhaps Canada. A sufficient number of fires, days, or events must be found to ensure statistically robust results. In creating this database, we will look at Wildland Fire Use (WFU) fires, national and state records, and specific fire case histories, conferring with fire danger/behavior specialists when necessary. (Data sources may include Hardy and Bunnell, 1999; the NIFMIDS database; and DOI Form-1202 reports, available through the Shared Application Computer System, SACS.) WFU fires are especially relevant for this project because their characteristics are less impacted by human control efforts, and more directly a response to the atmosphere and fuels.

Meteorological data must match the fire data in time and location. The database could contain nearby surface and upper data, or simply time and location coordinates for the fires. Because of the breadth of meteorological data, it is impractical to include all surface observations, upper air data, satellite data, radar observations, model data, etc. that may exist. Many of these are available from other sources other than fire agencies and need not be retabulated for index validation.

Different indices claim to do different things. If an index claims to predict fire starts, it should be examined for days both with and without starts. If the index claims to predict only fire spread, then it should be examined for high and low spread days, but cannot be evaluated using days without any fires at all. An individual testing an index should take great pains to explain what the index claims to do, how the person chose to reflect that in the data analysis, and their rationale or justification for that choice. In the end their results may not be universally applicable, but it will be clear when and how they apply.

5. STATISTICAL METHODS

Some fire-weather indices are discrete (e.g., HI, Davis Stability Index) while others are continuous (CFWI). Some are theoretically unbounded (temperature) but many are bounded over small ranges (HI) or broad ranges (CFWI). Because of these variations, different indices must be tested with different statistical methods. All index and fire data should first be subjected to exploratory graphical and numerical methods to identify patterns, outliers, and other notable data features (e.g., Wilks, 1995.) Once these properties are properly assessed, then more formal statistical procedures can be applied. Discrete indices will then be examined through contingency tables, logistic regression (Hosmer and Lemeshow, 1989) and randomization cross-validation (Efron, 1982) procedures. Continuous indices will be examined utilizing contingency tables, least absolute deviation regression (LAD; Barrodale and Roberts, 1973) fitted models, and cross-validation. Note that we intend to incorporate more "non-classical" methods (e.g., LAD and randomization) to account for invalid statistical assumptions given no formal randomized experimental design, and the possibilities of outliers and influence values in small samples. In the end, all indices will be summarized according to these models, including statistical uncertainty (e.g., confidence intervals and probability values).

The results of these steps should be the expression of each index's performance in a limited set of statistical measures, always derived from the same fires.

6. CONCLUDING REMARKS

Operational indices on which lives and great expense depend must be tested and proven according to high standards. To date, this has seldom been true. We propose the creation of a general framework for index testing, and the development of a uniform data base to use in such testing. This would not solve all index validation and use concerns for all indices, but it would serve as a start and may lead to more rigorous guidelines in the future.

One concern we recognize is the potential overuse of a limited set of "good" fire data. Because quality fire data can be hard to find, any readily accessible and expansive data set will create a temptation to use it for index development or calibration. But once that is done, the same data cannot be used for validation of the same index. We propose that, if possible, two data bases be produced for general distribution and use. One could be designated for use in calibration and development, if a researcher so desired. The other would be used solely for validation.

5. REFERENCES

Andrews, P.L, D.O. Loftsgaarden, and L.S. Bradshaw, 2003. Evaluation of fire danger rating indexes using logistic regression and percentile analysis. *Int. J. Wildland Fire*, **12**:213-226.

Barrodale, I., and F.D.K. Roberts, 1973. An improved algorithm for discrete L1 linear approximation. *Society for Industrial and Applied Mathematics Journal on Numerical Analysis*, **10**:839-848.

Davis, R.T., 1969. Atmospheric stability forecast and fire control. *Fire Control Notes*, **30**:3-4,15.

Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. Society for Industrial and Applied Mathematics.

Gisborne, H.T., 1936. Measuring Fire Weather and Forest Inflammability. USDA Circular No. 398, Washington, D.C.

Haines, D.A., 1988. A lower atmosphere severity index for wildland fires. *National Weather Digest*, **13**:23-27.

Hardy, C.C., D.L. Bunnell, 1999: Course-scale spatial data for wildland fire and fuel management [Online] (1999, November). Prescribed Fire and Fire Effects Research Work Unit, Rocky Mountain Research Station (producer). Available: www.fs.fed.us/fire/fuelman [1999, December].

Holzworth, G.C., 1972. Mixing heights, wind speeds, and potential for urban air pollution throughout the contiguous United States. Pub. No. AP-101. Washington, DC: U.S. Environmental Protection Agency, Office of Air Programs: 3,34.

Hosmer, D.W., and S. Lemeshow, 1989. Applied Logistic Regression. John Wiley and Sons.

Werth, J. and P. Werth, 1998. Haines Index climatology for the western United States. *Fire Management Notes*, **58**:8-17.

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press.