J4.2 MODELING SEASONAL VARIATION IN THE TOTAL PROBABILITY OF WILDFIRES

Randall P. Benson*
South Dakota School of Mines and Technology, Rapid City, South Dakota

1. INTRODUCTION

Wildfires are commonplace in the Black Hills and in many parts of the United States each summer. They occur naturally by lightning and by various human causes such as escaped campfires, debris burning, and arson. The risk of wildfire ignition depends on an assortment of factors ranging from various forest fuel conditions, weather and climate conditions, and the day of the week and time of year (Carter and Rolph 1975, Nickey and Chapman 1979, Walker *et al.* 1979).

The combustion process, whether initiated by a lightning strike or caused by humans, occurs as fuel elements in the forest are "cooked" out of the fuel by intense heating. This process is known as pyrolysis, which is largely dependent on the temperature of the fuels (Susott 1984) and immediately precedes ignition. In fact, lower temperatures under which pyrolysis occurs leads to char formation while higher temperatures generated from the ignition source leads to flaming combustion (Kilzer and Broido 1965). Energy released from the ignition source must be transferred into the fuel to initiate pyrolysis. Because of the temperature and fuel duration requirements needed to begin pyrolysis, lightning-ignited fires are distinguished from propagating fires. A heat pulse of energy generates spontaneous ignition like that occurring from a lightning strike and either ignition occurs or not. In propagating fires, a steadily increasing or constant heat source is observed for longer durations, which may be followed by a pulse or increase in heat. Once again, ignition may or may not ensue (Cox 1995). It is assumed for this study that lightning-caused wildfires ignite under different fuel and weather conditions than do human-caused wildfires at least to some degree.

* *Corresponding author address:* Randall P. Benson, South Dakota School of Mines and Technology, Department of Atmospheric Science, 501 East Saint Joseph Street, Rapid City, South Dakota, 57701, email: randall.benson@sdsmt.edu

A definition of forest fire risk is given in the 2002 Glossary of Forest Fire Management Terms (http://wlapwww.gov.bc.ca/esd/fire_mgmt_gloss_2002.pdf) by "the probability or chance of fire starting determined by the presence and activities of causative agents." Estimating forest fire risk involves identifying the possible sources of ignition and the factors that allowed ignition to occur. Chuvieco and Congalton (1989) stated that fire risk is "the union of two components: fire hazard and fire ignition." Once the hazards and causes have been identified, a mathematical expression can be developed to quantify the forest fire risk. A fire risk expression may be developed to quantify forest fire risk on short time scales (days) or much longer time scales (months and seasons).

A short-term fire risk assessment tool should be based upon natural elements that may potentially change very rapidly. Specifically, the factors that affect the moisture content of forest vegetation are critical to understanding the risk that that vegetation has to burning (Dudley 2003). Fine forest fuels such as grasses, leaves, and needles may experience sudden changes in their moisture content. Contrastingly, forest fuels such as medium-sized sticks and branches and larger logs undergo moisture changes over a longer period of time on the scales of weeks and months (Brown and Davis 1973). Climate and weather aspects may combine to enhance vegetation drying. For example, long-term drought conditions combined with hot and dry weather spanning several days or a week may be sufficient to allow a wildfire to burn easily once ignition occurs. Because it is difficult to know the actual vegetation or fuel moisture content, it is often estimated through the use of meteorological variables (Viegas *et al.* 2000), or it is approximated through the use of remotely sensed data and techniques (Paltridge and Barber 1988, Lopez *et al.* 1991, Illera *et al.* 1996).

This research explores the concept of developing a single index or mathematical measure to quantify the daily risk of either a lightning or human-caused wildfire. A single equation denoting the total probability of a wildfire occurring is developed for the typical wildfire season in the Black Hills (March – October).

Multiple logistic regression is used to determine the daily probability of a lightning and/or human-caused wildfire to occur in the Black Hills National Forest (Figure 1). The candidate predictors are a combination of weather, fuels, National Fire Danger Rating Systems (NFDRS) fire danger indices, and periodic calendar functions. In the equation development it was assumed that lightning and human-caused fires occur independently of each other.

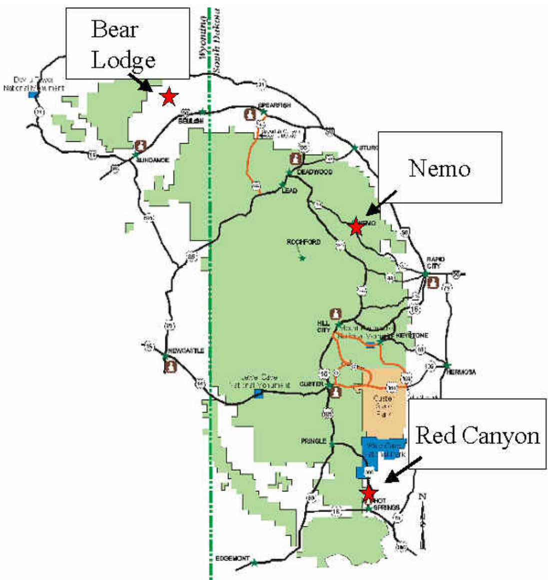

Figure 1. The Black Hills National Forest of Wyoming (left) and South Dakota (right) indicating the name and location of the RAWS stations used in this study.

During the period 1994-2003, lightning caused approximately 60% of all wildfires annually (Figure 2) in the Black Hills of South Dakota and Wyoming (Andrews and Bradshaw 1997). The mountain and plains topography act as an elevated heat and moisture source allowing convection to occur easily during the warm season. The Hills are dome shaped with a northwest to southeast oriented axis and exhibit an elevation change from the adjacent plains to the tallest peak of approximately 4,000 feet. Lightning activity (Vaisala, Inc.) begins in March and continues into October (Figure 3) with the most significant rainfall occurring in May and June (Dennis *et al.* 1966, Kuo and Orville 1972).

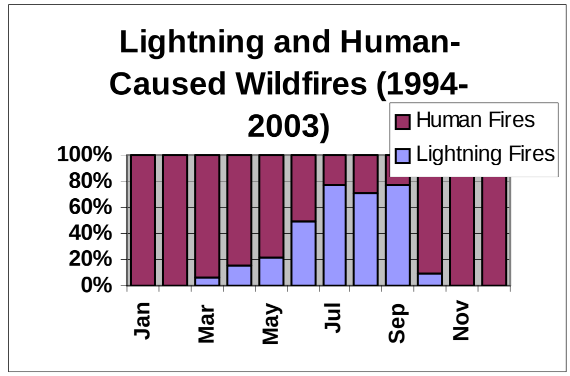

Figure 2. The annual distribution of human and lightning-caused wildfires in the Black Hills National Forest (Andrews and Bradshaw 1997).

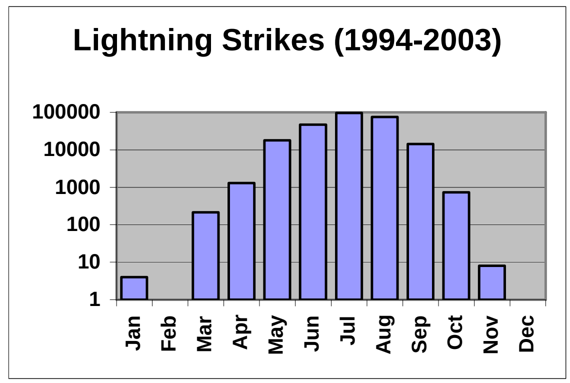

Figure 3. Lightning strikes in the Black Hills region by month (Vaisala, Inc.).

The ability to predict lightning occurrence and whether a lightning strike will generate a forest fire remains a complex daily challenge. Lightning flash rates of both in-cloud and ground flashes typically vary widely from one storm to the next although Peckham *et al.* (1984) did find that the average ground flash rate was higher in multicell storms than in isolated cells in Florida.

2. PREDICTAND, PREDICTORS, METHODS

A statistical tool to estimate the daily probability of one or more lightning strikes occurring is used in this study as a predictor in the probability calculation of a lightning-caused wildfire. Multiple logistic regression is used to formulate the equations to predict a lightning strike and is based upon 45 prospective thermodynamic predictors (Bolton 1980) derived from the daily 1200 UTC upper air radiosonde in Rapid City.

Equation predictors for the probability of a lightning fire (PLF) are comprised of daily 1 pm local time NFDRS calculated indices including the following: Spread Component (SC), Burning Index (BI), Energy Release Component (ERC), Ignition Component (IC), and the Keetch Byram Drought Index (KBDI). In addition, dead fuel moisture timelag terms are utilized as predictors. A fuel's timelag is proportional to its diameter and is commonly defined as the time it takes a fuel particle to reach two-thirds of its way to equilibrium with its local environment. Dead fuels in NFDRS fall into four groups or classes: 1-hr, 10-hr, 100-hr, and 1000-hr fuel moisture timelag fuels (Deeming *et al.* 1977). The herbaceous fuel moisture (FMH) value, which is the calculated value of the approximate moisture content of live herbaceous vegetation expressed as an oven dry weight of the original sample, is used as another predictor.

The meteorological predictors used in the logistic regression equations are taken from the daily 1 pm local time observations of temperature, relative humidity, wind speed and wind direction. The maximum and minimum temperature and relative humidity values are taken from the previous 24-hr period ending at 1 pm local time. The amount of rainfall in the previous 24-hr period and the duration of rainfall (recorded automatically by a Remote Automated Weather Station or RAWS) are also used as candidate predictors. Previous day lightning strike activity is used as a binary predictor in the equation set.

The logistic regression equations were calculated for periods of two months beginning in March and concluding in October. The predictand or response variable (lightning or human-caused wildfire occurring in the Black Hills National Forest) was used in the equation development to gauge the relationship of fire activity amongst three RAWS stations located in the Northern, Central, and Southern Hills (Bear Lodge, Nemo, and Red Canyon, Figure 1). In addition, a separate set of monthly equations was developed using the average of the three RAWS stations fuels, weather, and NFDRS indices. For the response variable, a binary "1" was used to designate at least one fire occurring daily and a "0" was used to indicate no fires occurring for the response variable.

A backward stepwise logistic regression technique was used to identify which of the possible predictors related most significantly to lightning or human-caused wildfires. A statistical software program (MINITAB) was used to formulate the binary logistic regression equations. This method utilizes the p-value (Pearson 1900), which is also known as the rejection level, to determine which of the predictors should be eliminated. The p-value is a probability ranging from 0 to 1 and indicates by a small value that the relationship between the predictor and the predictand is not likely to be a coincidence and that the predictor may have statistical significance. In the backward stepwise regression technique, the predictor with the largest p-value was deleted first and this procedure was continued until only the smallest p-values (< .05) remained. The final logistic regression equation (Wilks 1995) takes the following form,

$$y = \frac{e^{\left(b_0 + b_1 x_1 + \ldots + b_k x_k\right)}}{1 + e^{\left(b_0 + b_1 x_1 + \ldots + b_k x_k\right)}}$$

where y is the predicted probability of occurrence, b_0 is the intercept, b_k are the coefficients for the predictors, x_k and k is the number of predictors.

Nearly 40% of all the wildfires occurring in the Hills from the period 1994-2003 are caused by human activity (Figure 2). The majority of these wildfires are caused by escaped campfires (Andrews and Bradshaw 1997). It becomes obvious after analyzing the human-caused fires that there is a weekly (Figure 4) and seasonal influence (Figure 5) to the number of ignitions. Martell *et al.* (1989) used periodic functions of Julian calendar dates to incorporate seasonal variation into logistic regression equations that also used two of the six components of the Canadian Forest Fire Danger Rating System (CFFDRS), the Fine Fuel Moisture Code (FFMC), and the Build-Up Index (BUI). These indexes measure the moisture content of the upper-most layers of the forest floor of both live and dead fuels. The authors determined early in their study that human-caused wildfire exhibited seasonal variation. To account for this seasonality, they incorporated trigonometric sine and cosine functions to simulate periodicity within an assumed 140-day fire season.

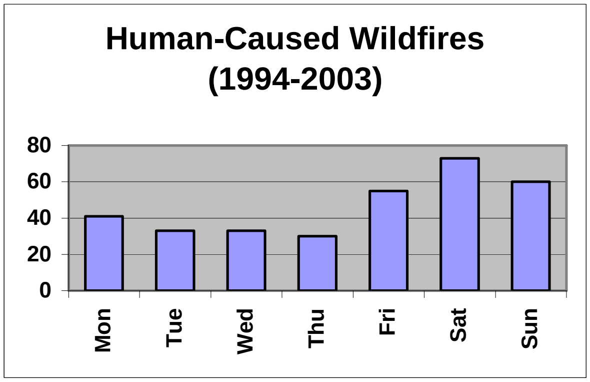

Figure 4. The human-caused fires occurring in the Black Hills National Forest by day of the week (Andrews and Bradshaw 1997).

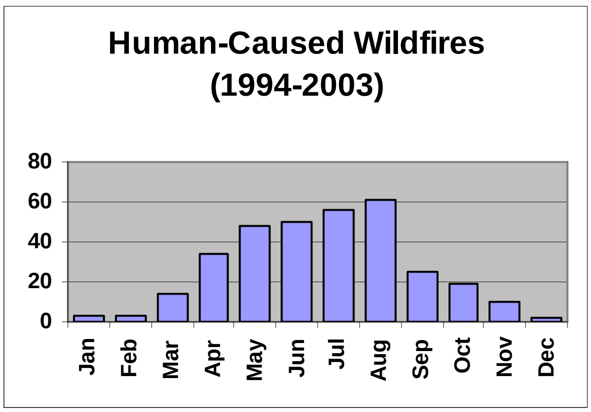

Figure 5. Human-caused wildfires occurring in the Black Hills National Forest by month (Andrews and Bradshaw 1997).

The predictors used to formulate the equations in this research for the Probability of Human-Caused Wildfire (PHF) were the daily 1 pm local time observations of temperature, relative humidity, wind speed, and wind direction from the three RAWS stations in the northern, central, and southern Hills. Additionally, a Julian day function was added to account for fires occurring in clusters at various times of the year. In the ten-year database (1994-2003) of human-caused fires, a binary "1" was used when there was a fire recorded in two or more of the Julian days on record. Otherwise, a "0" was recorded. The day of the week was also used as a predictor for human-caused fires. The trigonometric sine function was used with the day of the week in a seven-day period and of the form,

$$y = a * \sin[b(x + c)] + d$$

where *a* is the amplitude, *b* is the period, *x* is the fractional day of the seven-day week, c is the horizontal phase shift and d is the vertical phase shift.

Assuming independence of lightning-caused and human-caused fires, we may observe the probability of a wildfire of any type is the following,

$$P_{AF} = 1 - (1 - P_{HF})(1 - P_{LF})$$

where P_{AF} is the probability of all fires, P_{HF} is the probability of human fires, and P_{LF} is the probability of a lightning fire.

The model equation used to obtain P_{AF} for each RAWS location may be run operationally, given the probability of a lightning strike, several persistence variables, weather and fuels data, and forecast daily NFDRS output indices. For each RAWS station a probability is given for each type of wildfire. Then, the calculated probabilities are normalized according to their ranked historical values to determine a subjectively designed adjective rating for the likelihood of fire at that location. The adjective rating was created to resemble that of the NFDRS daily rating for convenience and are designated by ranked historical probabilities for each RAWS as follows: Extremely High, Very High, High, Moderate, Low, and Extremely Low.

3. RESULTS, DISCUSSION, CONCLUSIONS

As a first check in verifying the predictability of the equations, the daily calculated historical probabilities were ranked for each RAWS station (and the three-station average) and were separated and normalized into quartiles (1, 2, 3 and 4) based on the maximum and minimum probabilities for each two-monthly calendar period. Quartiles were used for the sorted historical probabilities for all fires, lightning fires, and human-caused fires and for each bi-monthly calendar period. A percentage of days with fire compared to the total number of days is shown in column 4 of Figure 6 for each quartile for the months of Jul-Aug as an average for each RAWS station (including the three-station average) for each wildfire category (all fires, lightning fires, and human-caused fires).

The results shown in Figure 6 indicate that for quartile 1 the outcomes indicate an expected result. When the calculated probability is high for a specific wildfire type (quartile 1), there is frequently a high occurrence of observed wildfire. The exception occurs at times when there is insufficient data to formulate a complete analysis (quartile 1 MAY-JUN in Figure 6 for LFD). It is worth noting that for the months of JUL-AUG, which corresponds to the peak in lightning activity (Figure 2) for the Hills (Vaisala, Inc.) , the percentage of calculated lightning fire days to actual lightning fire days is the highest for any of the time periods.

Figure 7 indicates that averaging the daily fuels, weather, and NFDRS indices may be more useful in predicting both all-fire days and lightning-fire days. It is reasonable to assume that because the RAWS stations are in close proximity to each other, that many of the fuels and weather conditions were often similar. The reason for the large percentage of calculated human-caused fire days compared to actual human-caused fire days at NEMO is unclear. Perhaps the high percentage is attributed to its central location in the Black Hills (Figure 1).

Conceptually, the method of being able to predict wildfire activity from historical wildfires, predicted lightning activity, current observations of fuels and weather data and with the inclusion of variables attributable to human-caused wildfires seems plausible. With the assumption that human and lightning-caused wildfires occur independently of each other and that there is at least some differences of the mechanisms of ignition between each wildfire type, it appears possible that a single statistical index could be used to estimate the daily likelihood of wildfire. Future work related to this study will be to explore the use of the Poisson probability distribution function to determine PAF, PHF, and PLF for multiple numbers of fires.

RAWS SUMMARY			
MAY-JUN	AVG AFD/TOTD	MIN	MAX
1	0.672	0.588	0.833
2	0.473	0.440	0.511
3	0.324	0.288	0.370
4	0.111	0.094	0.130
MAY-JUN	HFD/TOTD	MIN	MAX
1	0.527	0.500	0.577
2	0.358	0.327	0.405
3	0.254	0.206	0.303
4	0.072	0.063	0.081
MAY-JUN	LFD/TOTD	MIN	MAX
1	0.406	0.000	1.000
2	0.485	0.400	0.625
3	0.254	0.194	0.290
4	0.052	0.045	0.060
JUL-AUG	AVG AFD/TOTD	MIN	MAX
1	0.756	0.664	0.857
2	0.591	0.538	0.614
3	0.358	0.314	0.389
4	0.196	0.158	0.216
JUL-AUG	HFD/TOTD	MIN	MAX
1	0.611	0.333	1.000
2	0.439	0.392	0.500
3	0.249	0.117	0.336
4	0.065	0.032	0.079
JUL-AUG	LFD/TOTD	MIN	MAX
1	0.740	0.671	0.917
2	0.565	0.500	0.642
3	0.290	0.245	0.322
4	0.170	0.140	0.181
SEP-OCT	AVG AFD/TOTD	MIN	MAX
1	0.690	0.542	0.765
2	0.395	0.355	0.455
3	0.235	0.198	0.261
4	0.069	0.065	0.075
SEP-OCT	HFD/TOTD	MIN	MAX
1	0.756	0.600	0.875
2	0.382	0.200	0.500
3	0.244	0.179	0.303
4	0.051	0.049	0.055
SEP-OCT	LFD/TOTD	MIN	MAX
1	0.715	0.529	0.846
2	0.300	0.286	0.344
3	0.165	0.135	0.217
4	0.026	0.021	0.029

Figure 6. The percentage of days summary of each wildfire type separated by quartiles for each of the three RAWS stations including the RAWS station's average. Column 1 is the quartile for the percentage range. Column 2 is the average (AVG) percentage of all fire days (AFD) for each RAWS and the average RAWS compared to all the days for that quartile. Column 3 is the calculated minimum percentage value for each quartile of fire days compared to all days for each RAWS and the average RAWS. Column 4 is the calculated maximum percentage of fire days for each quartile. HFD = human fire days. LFD = lightning fire days. TOTD = total days.

Upper Quartiles RAWS Averages		
AFD/TOTD	**PAF 3 AVG**	**0.619**
HFD/TOTD	**PHF NEM**	**0.568**
LFD/TOTD	**PLF 3 AVG**	**0.601**

Figure 7. An average of the top two quartiles for March through October of the calculated percentage of wildfire days versus the actual number of wildfire days for the best-performing RAWS and/or the three-station average. AFD = All fire days. HFD = Human Fire Days. LFD = Lightning Fire Days. TOTD

4. ACKNOWLEDGEMENTS

I would like to thank Dr. Frank Matejcik from the Department of Industrial Engineering at the South Dakota School of Mines and Technology for his help in the statistical equation development. Also, I would like to express my appreciation for the help of Matt Bunkers, Science and Operations Officer for the National Oceanic and Atmospheric Administration's National Weather Service office in Rapid City for his help in building the lightning prediction tool. Lastly, I would like to thank Don Latham (retired USFS) for his help in the methodology used in this study. The position of fire meteorologist is funded by the State of South Dakota.

5. REFERENCES

Andrews K., and L.S. Bradshaw, 1997: "Fires: Fire information retrieval and evaluation system – a program for fire danger rating analysis." USDA Forest Service, Intermountain Research Station General Technical Report INT-367. 64 pp. (Ogden, UT)

Bolton, D., 1980: The computation of equivalent potential temperature. *Mon. Wea. Rev.,* **108,** 1046-1053.

Brown, A.A., and K.P. Davis, 1973: "Forest Fire: Control and Use." McGraw-Hill, New York.

Carter, G.M., and J.E. Rolph, 1975: New York City fire alarm prediction models. II. Alarm rates. R-1215-NYC. The New York City Rand Institute, New York.

Chuvieco, E., and R.G. Congalton, 1989: Application of remote sensing and geographic information systems to forest fire hazard mapping. *Remote Sensing of the Environment* 29, pp. 147-159.

Cox, G., 1995: "Combustion Fundamentals of Fire." Academic Press, London.

Deeming, J.E., R.E. Burgan, and J.D. Cohen, 1978: The National Fire-Danger Rating System - 1978. USDA, For. Serv. Gen. Tech. Rep. INT-39. Intermt. For. and Range Exp. Sta., Ogden, UT. 63 p.

Dennis, A.S., H.E. Hart, J.B. Niederauer and C. Lundy, 1966: Upper air climatology of Rapid City, South Dakota. Rept. 66-3, Inst. Atmos. Sci., South Dakota School of Mines and Technology, 40 pp.

Dudley, M., 2003: "Wildland Fire Danger Estimation and Mapping – The Role of Remote Sensing Data." World Scientific Publishing Co., Singapore.

Illera, P., A. Fernandez, A. Calle, and J.L. Casanova, 1996: Temporal evolution of the NDVI as an indicator of forest fire danger. *International Journal of Remote Sensing*, **17**, 1093–1105.

Kilzer, F.J., and A. Broido, 1965: Speculations on the nature of cellulose pyrolysis. *Pyrodynamics*, **2,** 151-163.

Kuo, J. T., and H.D. Orville, 1972: A radar climatology of summertime convective clouds in the Black Hills. *Journal of Applied Meteorology,* **12**, 359-368.

Lopez, S., F. Gonzalez-Alonso, R. Llop, and J.M. Cuevas, 1991: An evaluation of the utility of NOAA-AVHRR images for monitoring forest fires risk in Spain. *International Journal of Remote Sensing*, **12**, 1841–1851.

Martell, D.L., E. Bevilacqua, and B.J. Stocks, 1989: Modeling seasonal variation in daily people-caused forest fire occurrence. *Can. J. For. Res.,* **19**, 1555-1563.

Nickey, B.B., and C.B Chapman, 1979: Evaluating fire prevention effectiveness through a probability model. *Fire Technol.***, 15,** 291-306.

Paltridge, G.W., and J. Barber, 1988: Monitoring grasslands dryness and forest fire potential in Australia with NOAA/AVHRR data. *Remote Sensing of Environment*, **25**, 381–394.

Pearson, K., 1900: On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*, **50**, 157–175.

Peckham, D.W., M.A. Uman, and C.E. Wilcox, Jr., 1984: Lightning phenomenology in the Tampa Bay Area. *J. Geophys. Res.,* **89**, 11, 789-805.

Susott, R.A., 1984: "Heat of Preignition of Three Woody Fuels Used in Wildfire Modeling Research." Res. Note INT-342. USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah.

Vaisala, Incorporated. Remote Sensing Division, Tucson, AZ, U.S.A.

Viegas, X., G. Bovio, A. Ferreira, A. Nosenzo, and B. Sol, 2000: Comparative study of various methods of forest fire danger evaluation in southern Europe. *International Journal of Wildland Fire*, **9**, 235–246.

Walker, W.E., J.M. Chaiken, and E.J. Ignall, (*Editors*), 1979: Fire department deployment analysis: a public policy case study – the Rand fire project. North Holland, New York.

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