The Presentation of Precipitation Information in Television Broadcasts: What is Normal?

Thomas Hagen[1,2]

Justin Glisan[1]

Anthony R. Lupo*[1]

Eric Aldrich[1,3]

Patrick E. Guinan[1,4]

Patrick S. Market[1]

Neil I. Fox[1]

[1]Department of Soil, Environmental, and Atmospheric Sciences
302 Anheuser Busch Natural Resources Building
University of Missouri-Columbia
Columbia, MO 65211

[2]KSNT Channel 27
6835 NW Hwy 24
Topeka, KS 66618

[3]KOMU - TV 8
Highway 63 South
Columbia, MO 65201

[4]Missouri Climate Center
1 – 74 Agriculture Building
University of Missouri - Columbia
Columbia, MO 65211

Submitted to:

National Weather Digest

November, 2007 revised: May, 2008

* *Corresponding Author Address: Dr. Anthony R. Lupo, Department of Soil, Environmental, and Atmospheric Sciences, 302 E Anheuser Busch Natural Resources Building, University of Missouri - Columbia, Columbia, MO 65211, E-mail: LupoA@missouri.edu*

Abstract

In a typical weather broadcast, observed precipitation information such as the daily amount that fell and the accumulated monthly total are shown and compared to the mean monthly average or “normal” precipitation. Such information, however, may not adequately describe whether or not that particular month is fairly typical for the time of year or truly an unusual occurrence. Here it is shown that monthly average precipitation may not be representative of the typical value for a particular month at all. Thus it is suggested that the presentation of precipitation information can be augmented with elementary statistical information in order to give a more meaningful presentation of precipitation information without the need to explain the basis of such statistical information. A study of the climatological behavior of monthly precipitation values over a 118-year period for Columbia, Missouri is performed in order to provide the rationale for displaying "typical" precipitation ranges.

1. Introduction

Recently, two papers (Lupo et al., 2003; Holder et al., 2006) have demonstrated that statistical information could be used to augment daily temperature readings (e.g., maximum and minimum temperature) and be presented successfully in a three-to-five minute television weather segment. Both of these articles showed that simple information about the distribution and variability of daily temperature information could be included in a television graphic that uses the Tukey box plot (Tukey, 1977) as its basis. These papers suggested that daily maximum and minimum temperatures could be compared, not only to the daily mean, but to some range of temperatures that would be considered as “typical” for that date, month, or season. Lupo et al. (2003) define a range for “typical” daily temperatures using values within one standard deviation from the 30-year normal (σ = 68%; rounded by Lupo et al. 2003 to a “70% range”), whereas Holder et al. (2006) defined their range using a smaller value than one standard deviation (50% range). Each paper showed that temperature data at their respective locations was normally (or near normally) distributed about the mean. Naturally, these graphics should be constructed such that they are appealing to the general public as well (see Holder et al., 2006 for an example).

Precipitation information would be much more difficult to incorporate using the same basis as that for temperature, since precipitation does not occur daily. Additionally, (monthly) precipitation data does not have a normal distribution (e.g. Stephenson et al. 1999; Guinan, 2004), may accumulate on only a handful of days within a month, and is generally tabulated monthly and then compared to a monthly averaged value. Television meteorologists make this comparison of the monthly accumulated precipitation to the monthly mean (or a fraction of the monthly mean by dividing the monthly total by the number of days in the month and assuming an equal amount should fall daily) using simple subtraction. While this information may be meaningful to users, it does not provide more detail, or accuracy, regarding the statistical or historical context.

Generally, television meteorologists only comment on whether or not the month has been wetter or drier than normal, defined as the arithmetic mean. In the mid-west and plains states, viewers are very interested in this type of information since these areas have a strong agriculturally-based economy. Also, most television weather broadcasts are centered on the information that the broadest segment of the viewing public may be most interested in on a day-to-day basis, which is mainly temperature and precipitation information and forecasts.

In the last 30 years, there has been a dramatic increase in the attention paid to weather and climate information such as severe weather (Del Genio et al. 2007), El Niño and La Niña events, and climate change, including global climate change (e.g., Changnon and Kunkel, 1999; Kunkel et al., 1999). This has resulted in more than a five-fold increase in the television coverage of weather related events over that time period (e.g., Ungar, 1999). As such, a broader segment of the public is interested in weather broadcasts and has become increasingly "weather-savvy" with regard to the information presented to them.

Thus, this short paper has two simple objectives. The first objective is to perform a short statistical study by examining monthly precipitation data, their means and distribution, and variations for the Columbia, Missouri area. The second objective is to demonstrate how such information can be incorporated into weather broadcasts in order to provide the public with a more informative presentation.

2. Data and Methodology

a. Data

The data used in this brief study are the monthly precipitation records for Columbia, Missouri from January 1890 to December 2007. These data were obtained from the Missouri Climate Center. This temporal period was chosen since this period provided a larger data set for the statistical analysis. Also, this 118-year period provides for a continuous record for the precipitation information, and, as such, there was no need to artificially fill in missing precipitation data. In this study, data from the Columbia region were chosen since the results found here could be incorporated into local weather broadcasts. Also, precipitation observations have been taken from the locale named above over the duration of the last 38-year period. Only one change in the instrumentation was made in 1996, when the Automated Surface Observation Station (ASOS) instrumentation was installed. However, the airport did change location around 1970. This station moved approximately 22 km south-southeastward, but there are no indications that this move resulted in significant changes in the precipitation climatology (not shown). Before 1970, precipitation records were synthesized from records taken at the Columbia Municipal Airport (1930 – 1969) and an observation station in Columbia, MO (1890-1951). Each of the previous stations was located within the city limits. Lastly, the precipitation data and the calculations used in this study carried units of inches since that is still the standard unit for precipitation measurements in the United States and is still the standard unit for precipitation used in weather broadcasts.

b. Methods

The initial step required that the monthly precipitation information for the 118 year period, 1890 – 2007, was used to generate the statistics discussed here. Based on these monthly precipitation records, the average monthly precipitation values were generated for each month of the year, along with the standard deviations (Table 1). These were then tested in order to determine whether or not a normal or Gaussian distribution would best represent these data. Fig. 1 shows a histogram generated for one month and binned in 0.5 inch intervals. Inspection would reveal that this distribution is not Gaussian, and thus another distribution needed to be fit to the data. The observed distribution was tested in order to determine if it followed a standard normal distribution at a statistically significant level using the chi-square goodness-of-fit test (e.g., Neter et al., 1988). The statistical testing reveals that the observed distribution is not the same as (different from) the normal distribution at the 99% confidence level.

Comparing, however, the observed data to a gamma distribution for statistical testing reveals that the distributions are similar at the 95% confidence level. A gamma distribution was fitted using the standard parameters for the gamma probability density function following Guinan (2004) (see also Stephenson et al. 1999):

$$f(x,\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \qquad (1)$$

where x represents a bin value, and α and β are parameters needed to calculate the distribution and are as follows:

$$\alpha = \frac{1}{4A}\left[1+\left(1+\frac{4A}{3}\right)^{0.5}\right] \qquad (2a)$$

$$A = \ln(\bar{x}) - \sum_{i=1}^{n} \frac{\ln(x_i)}{n} \qquad (2b)$$

$$\beta = \frac{\bar{x}}{\alpha} \qquad (2c),$$

where α controls the shape of the distribution (shape parameter), and β is the scale parameter. The parameter α controls whether the distribution takes on the form of a skewed normal distribution, or appears more like the tail of a normal distribution (see Fig. 1).

Table 2 shows the monthly values of α and β for our data set. These demonstrate that these values do vary throughout the year. Values of α (β) are relatively low (high) during the wettest months (April and May) when the distribution would be expected to extend father along the abscissa. Our values here are not consistent with Guinan (2004) who used gamma distributions for two week periods to examine drought, but our values for α are more consistent with Stephenson et al. (1999). They studied the frequency and amounts of monsoonal precipitation events in India. Our winter month values for α and β were consistent with a study of precipitation frequencies from the University of Bergen in Norway[1].

Clearly, the Tukey box-plot (e.g., Fig. 2 – adapted from Lupo et al. 2003) which is typically based on normally distributed data would not be the proper basis for a precipitation graphic, and further discussion below will illustrate this point. Another strategy is proposed here and bases the graphic on dividing the observed data into quintiles, and the new graphic is described in the next section. Using observed data allows for yearly adjustment of the values in Table 3 if necessary. Additionally, the gamma distributions were generated and tested versus the observed distribution using commercially available software such as spreadsheets and mathematical / statistical software.

3. Results and Graphical Depiction

a. Results and Discussion

Figure 3a shows the proposed template for a new graphic. The monthly precipitation data for any month could be classified as within the range of “normal” (20% of the data, the third quintile – Table 3). The second and fourth quintiles would represent below and above average rainfall for that month, respectively. The first and fifth quintiles could represent extremely dry and wet conditions, respectively. Finally, a record wet month could also be represented as extremely wet. Two strategies could be used in representing this kind of information and these will be described below. Note also that the graphic could be color coded in a way that would be consistent with the color of land or the vegetation during dry or wet conditions, or colors that are associated with, for example, water (extreme wet). This suggested color scheme also corresponds roughly to the natural color spectrum, but alternative schemes could be experimented with and then implemented to better convey the information (e.g., Brewer, 2005). Additionally, the graphic could be made more viewer-friendly by using a rain gauge image in a similar manner that Holder et al. (2006) show the use of a thermometer in their temperature graphic (Fig. 4).

Probability Density Functions (PDF) for the entire 118-year period and for each month were tested as opposed to PDFs for individual days or seasons since precipitation is archived as monthly totals. Since precipitation does not occur every day, the daily precipitation information (for individual calendar days) over a 30-year climatic averaging period would represent a small sample from which it might be difficult to obtain statistically meaningful results. Similar to temperature anomalies, precipitation anomalies for each month can be viewed as being produced by the cumulative impact of a set of random synoptic-scale disturbances guided by quasi-randomly generated large-scale flow regimes which may be strongly influenced by the underlying surface of the earth (e.g., Lupo et al., 2007 and references therein).

However, summer-season flow regimes may possess different kinematic and dynamic characteristics from winter-season flow regimes over North America. Data sets which are produced by two different forcing regimes and which also produce clearly differing distributions (if analyzed separately) are called “mixed distributions” or “mixture distributions” (e.g., Wilks, 2006). In order to minimize this problem of differing seasonal flow regimes, the data sets were analyzed monthly over the 118 year period. This same assumption does not preclude the use of seasonal statistical results instead of, or in addition to, the monthly data sets.

In displaying precipitation information, it would be useful to show some measure that represents a typical range of monthly values or the typical variability for a particular month. Most television broadcasts only show the monthly accumulated precipitation and compare this to the monthly mean total or a portion of this value that depends on how much of the month has already past. However, when examining the gamma distributions for each month, it is revealed that the mean value is commonly near the top of the third quintile (compare Table 1 and 3), or in the case of October – December (see Table 3) is at the top of this quintile or even into the fourth quintile. Then, the implication of comparing observed precipitation to the arithmetic mean is that in the fall, most months will appear to the viewer to be dry (drier-than-normal), or conversely, an "average" month is actually one of the wetter months historically. Thus, comparing monthly precipitation to an average value is misleading.

Some television broadcasts show the record monthly high and low precipitation amounts, but typically only when the month approaches a record. These represent in a statistical sense (and loosely in a physical sense) the absolute range of the precipitation amounts that may be expected for a given location for a given month. Here, the graphic we chose (Figs. 3 and 4) is based on variability. Lupo et al. (2003) chose to use standard deviation (σ), which represents a measure of absolute variability in a data set (in their case, the 30-year daily temperature anomalies). For data in a set that are normally distributed, σ can be used to construct an interval (range) about the mean for which approximately 68% of the data points in a particular set of data should reside.

By using quintiles, we followed the method used by Guinan (2004), and the frequency of extreme events (1st and 5th quintile) will be similar to that of Lupo et al. (2003) or Holder et al. (2006), or would occur less than 50% of the time. Using January 2005 (Figs. 4 and 5) as an example would show that 5.94 inches of rain fell at the Columbia, Missouri, regional airport. This is shown in comparison to the monthly mean 1.85 inches and clearly depicts that January 2005 was extremely wet (occurring only 20% of the time, or once every 5 years), the range of which is 2.48 - 6.87 inches. This demonstrates also that January in this region of the country is usually quite dry. Over the course of a year, the astute viewer / weather observer would realize that winter in this region is the driest time of the year (Tables 1 and 3).

b. Application

In the previous section (3a), it was stated that two strategies could be used when applying this graphic. The first is to simply accumulate the daily precipitation data and the regular viewer would watch the monthly precipitation increase versus the historical record. This method would be simple to apply and would not require daily maintenance of the graphic. The second strategy would involve identifying each of the quintile values on the right hand side of the plot and dividing these by the number of days in a month, and then tallying up the rain daily as well as "stretching" out (re-calculating) the numbers on the ordinate of the graph with each passing day. The latter is recommended, since this would account for the fact that months could become progressively drier or wetter during the month, or as shown in Fig. 6 for March 2005, the month stays extremely dry in spite of precipitation occurring on two different occasions. Fig. 6a shows the graphic as it would appear on 10 March, and as no precipitation occurred during the next ten days, Fig. 6b shows the graphic as it would have appeared on 20 March. Finally, Fig. 6c shows the graphic as it would have appeared for the end of that month. Note that the numbers on the right side of the graphic grew progressively larger as the month progressed. Additionally, Fig. 7 shows an example during a month in which precipitation categories changed from dry month early to a record wet month by mid-month.

Such information could be presented as in Fig. 6 and 7 without the need to explain to the general public the concept of standard deviations, quintiles, and other statistical concepts. A similar display to these figures showing seasonal or annual means could be considered as well for the presentation of this kind of information.

4. Summary and Conclusions

In this study, the statistical properties of the 118-year record (1890 - 2007) of monthly precipitation observations for Columbia, Missouri, were examined with the goal of providing more information about the representativeness of observed monthly precipitation amounts with respect to climatological mean in television weather broadcasts. The data used in this study were obtained from the Missouri Climate Center and analyzed using standard statistical techniques. A 118-year period was chosen because the record for this time period has been continuous and is long enough to present a reasonable sample size for each month.

In general it was found that maximum and minimum monthly precipitation in the 118-year period occurs during the warm and cold seasons, respectively. For each month, the data were fitted to a gamma distribution which provided a better match than a normal distribution as was used for the mean temperature values by Lupo et al. (2003) and Holder et al. (2006). Quintiles were then used as a measure of variability and these were applied to the monthly samples. The third quintile was considered “typical”, while the second and fourth quintiles were considered dry and moist, respectively. Together, these three quintiles constituted 60% of the observations. The first and fifth quintiles were considered extremely dry and moist, respectively.

This information was incorporated into routine television weather broadcasts at KOMU and KSNT, the NBC affiliates in Columbia, Missouri and Topeka, Kansas, respectively, beginning with the fall of 2007. Meteorologists and weather broadcasters created graphics that used the monthly climatological values of precipitation. Thus, the viewer will not only have seen how observed precipitation compared to that which is typical rather than the mean, but how representative these observations were within the historical context for this region. In an era when weather information is presented more and more often, information regarding a typical range for precipitation can be used to separate out unusual monthly accumulations of precipitation from those that are more typical.

5. Acknowledgments

The authors would like to acknowledge each of the reviewers for their helpful review comments (Dr. Troy Kimmel, University of Texas at Austin; Dr. S. Jeffery Underwood, Nevada State Climatologist; and Dr. Richard W. Dixon, Texas State University). Their contributions greatly strengthened this manuscript.

6. References

Brewer, C., 2005: *Designing Better Maps: A Guide for GIS users*. ESRI press, 203 pp.

Changnon, S.A., and K.E. Kunkel, 1999: Rapidly expanding uses of climate data and information in agriculture and water resources: Causes and characteristics of new applications. *Bull. Amer. Meteor. Soc.*, **80**, 821 - 83.

Del Genio, A.D., M-S. Yao, and J. Jonas, 2007: Will moist convection be stronger in a warmer climate? *Geophy. Res. Lett.*, **34,** L16703, doi: 10.1029/2007GL030525.

Guinan, P.E., 2004: A seasonally adjusted index for projecting agricultural drought. *University of Missouri, Department of Soil, Environmental, and Atmospheric Science Ph.D. Dissertation, 144 pp.*

Holder, C., R. Boyles, P. Robinson, S. Raman, and G. Fishel, 2006: Calculating a daily normal temperature range that reflects daily temperature variability. *Bull. Amer. Meteor. Soc.*, **87,** 769–774.

Kunkel, K.E., R.A. Pielke Jr., and S.A. Changnon, 1999: Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: A review. *Bull. Amer. Meteor. Soc.*, **80**, 1077 - 1098.

Lupo, A.R., Kelsey, E.P., D.K. Weitlich, I.I. Mokhov, F.A. Akyuz, Guinan, P.E., J.E. Woolard, 2007: Interannual and interdecadal variability in the predominant Pacific Region SST anomaly patterns and their impact on a local climate. *Atmosfera*, **20,** 171- 196.

Lupo, A.R., E.P. Kelsey, E.A. McCoy, C.E. Halcomb, E. Aldrich, S.N. Allen, A. Akyuz, S. Skellenger, D.G. Bieger, E. Wise, D. Schmidt, and M. Edwards, 2003: The Presentation of Temperature Information in Television Broadcasts: What is Normal? *National Weather Digest*, **27:4,** 53 -58.

Neter, J., W. Wasserman, and G.A. Whitmore, 1988: *Applied Statistics, 3rd ed.*, Allyn and Bacon Press, Boston, MA.

Stephenson, D.B., K. Rupa Kumar, F.J. Doblas-Reyes, J.F., Royer, F. Chauvin, and S. Pezzulli, 1999: Extreme daily rainfall events and their impact on the ensemble forecasts of the Indian Monsoon. *Mon. Wea. Rev.*, **127,** 1954 – 1966.

Tukey, J. W., 1977: *Exploratory Data Analysis*, Addison-Wesley, Reading MA.

Ungar, S., 1999: Is strange weather in the air? A study of U.S. national network news coverage of extreme weather events. *Climatic Change*, **41**, 133-150.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 627 pp.

Footnote (Section 3)

1. University of Bergen (Norway)

   http://www.uib.no/people/ngbnk/kurs/notes/node31.html

Table 1. Calculated monthly average precipitation and standard deviations (inches) for the base period 1890 – 2007 using monthly precipitation data from the Columbia, Missouri Regional Airport.

Month	Average Precipitation	Standard Deviation
January	1.85	1.38
February	1.88	1.19
March	2.93	1.68
April	3.87	2.08
May	4.77	2.32
June	4.49	2.56
July	3.50	2.16
August	3.76	2.42
September	4.14	2.83
October	3.01	2.03
November	2.48	1.79
December	2.00	1.31

Table 2. Calculations of the monthly parameters for α and β from Equations 1 and 2 for the monthly precipitation for the base period 1890 – 2007 using monthly precipitation data from the Columbia, Missouri Regional Airport.

Month	α (shape parameter)	β (scale parameter)
January	0.163	11.34
February	0.128	14.43
March	0.166	33.52
April	0.073	53.24
May	0.065	73.27
June	0.116	38.53
July	0.118	29.75
August	0.156	24.05
September	0.148	27.96
October	0.124	24.19
November	0.155	16.02
December	0.110	18.19

Table 3. The calculated values for the maximum precipitation value (inches) for each quintile. The record value represents the top of the fifth quintile (100%).

Month	first quintile	second quintile	third quintile	Fourth quintile	record
January	0.82	1.22	1.91	2.48	6.87
February	0.90	1.40	1.97	2.50	6.80
March	1.51	2.41	3.04	3.72	10.09
April	2.16	2.86	4.21	5.07	11.69
May	2.89	4.05	4.94	6.21	13.34
June	2.35	3.30	5.02	6.71	14.86
July	1.70	2.69	3.72	4.97	12.14
August	1.45	2.80	4.08	5.83	10.19
September	1.75	2.99	4.33	5.79	13.34
October	1.37	2.17	2.86	4.38	13.44
November	1.00	1.64	2.49	3.81	10.42
December	0.99	1.46	1.94	2.72	7.82

Figure Captions

Figure 1. Binned monthly precipitation amounts (bars) for each January from 1890 – 2007. The first bin on the left is monthly precipitation amounts from 0 – 0.50 inches, and each bin is successively 0.50 inches greater. The dotted line represents a fitted gamma distribution, while the long dashed line represents a normal or Gaussian distribution.

Figure 2. Suggested templates for incorporating seasonal standard deviation information into weather graphics depicting daily temperature observations (adapted from Lupo et al. 2003 - a Tukey box plot (Tukey, 1977)).

Figure 3. A proposed sample template for a new precipitation amount graphic.

Figure 4. A second proposed sample template for a new precipitation amount graphic. This includes data from January 2005.

Figure 5. The generic template from Fig. 3 with data from January 2005. All precipitation amounts are shown in inches.

Figure 6. As in Fig. 5, except for March 2005. This figure shows a monthly progression for;

a) March 10, b) March 20, and c) the end of the month.

Figure 7. As in Fig. 5, except for August 2005.

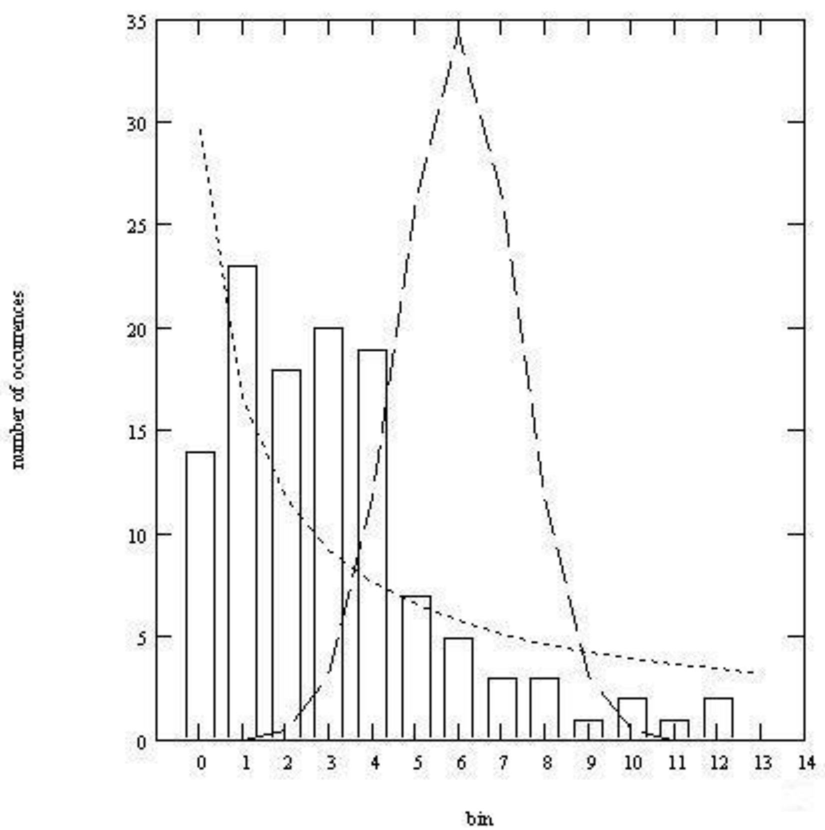


Figure 1. Binned monthly precipitation amounts (bars) for each January from 1890 – 2007. The first bin on the left is monthly precipitation amounts from 0 – 0.50 inches, and each bin is successively 0.50 inches greater. The dotted line represents a fitted gamma distribution, while the long dashed line represents a normal or Gaussian distribution.

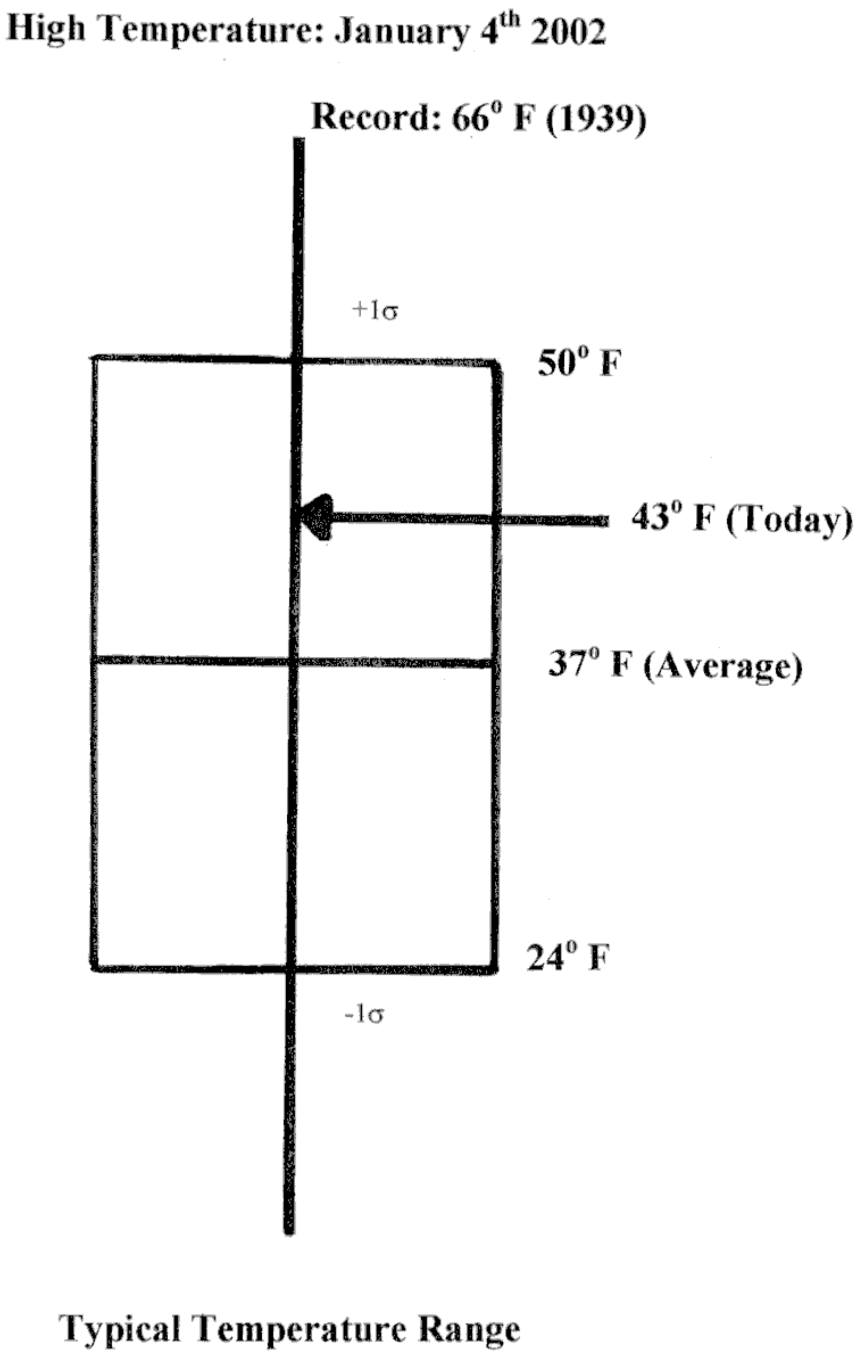


Figure 2. Suggested templates for incorporating seasonal standard deviation information into weather graphics depicting daily temperature observations (adapted from Lupo et al. 2003 - a Tukey box plot (Tukey, 1977)).

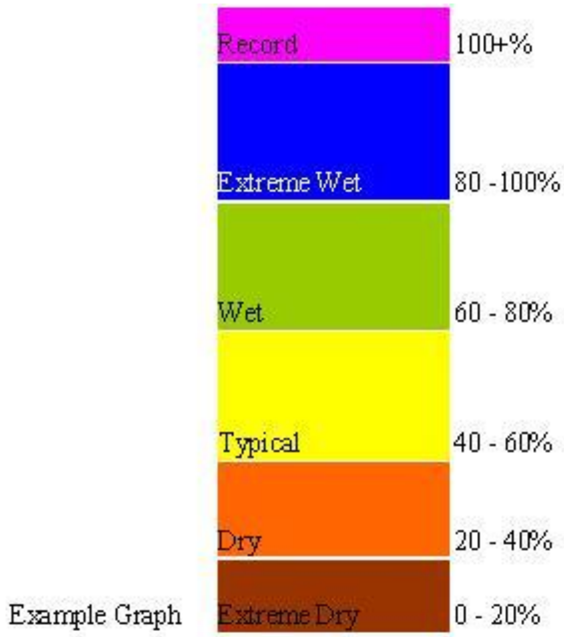


Figure 3. A proposed sample template for a new precipitation amount graphic.

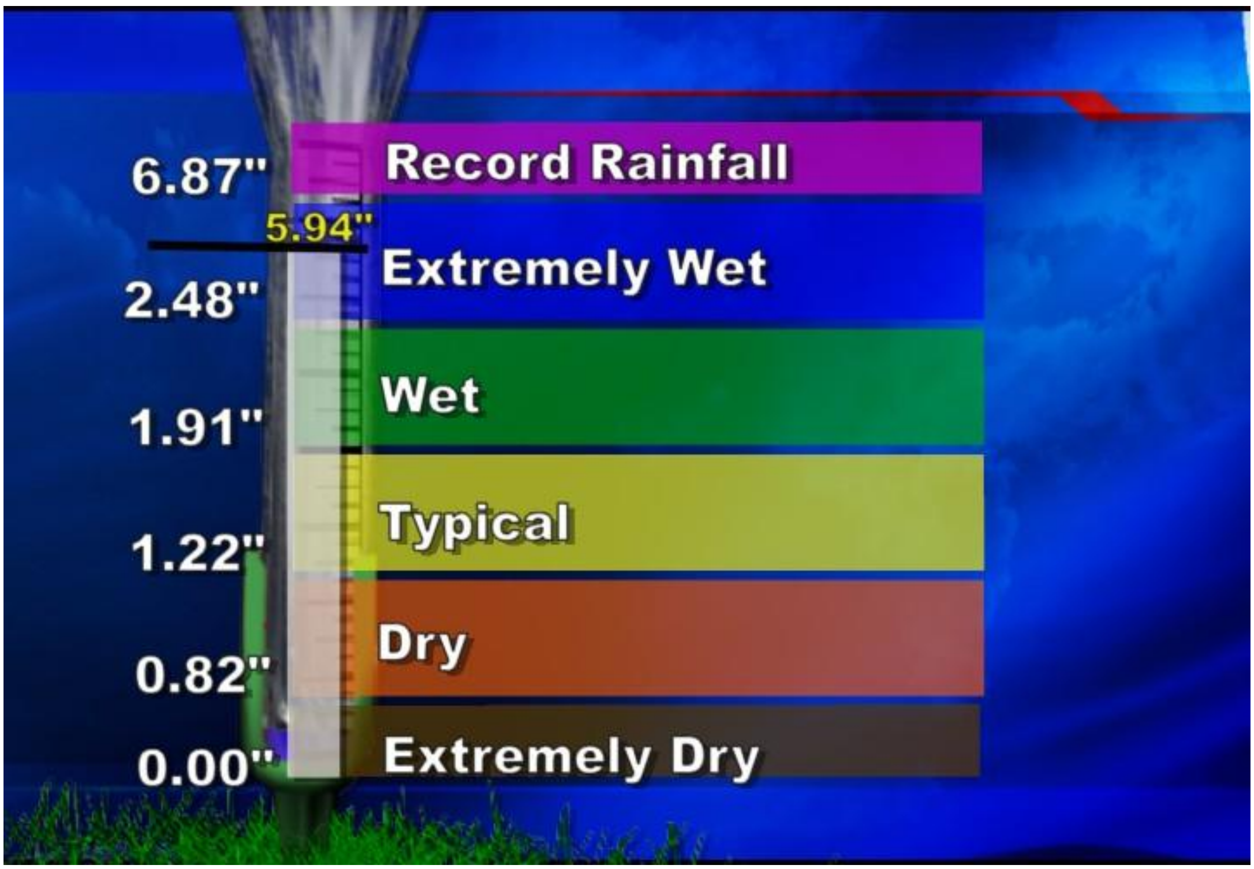


Figure 4. A second proposed sample template for a new precipitation amount graphic. This includes data from January 2005.

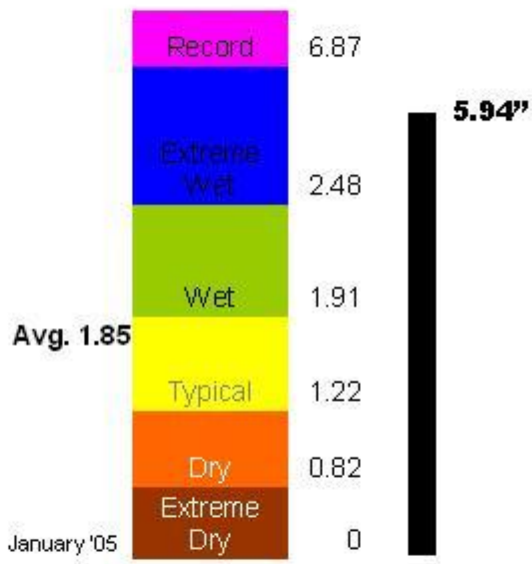
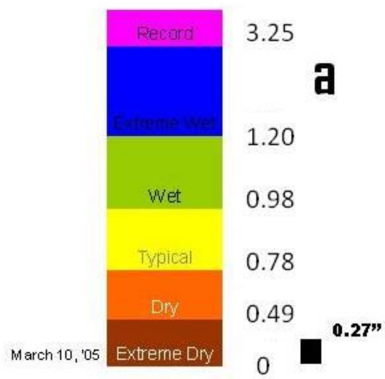
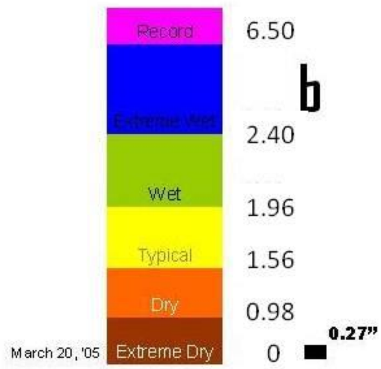


Figure 5. The generic template from Fig. 3 with data from January 2005. All precipitation amounts are shown in inches.


Record
3.25
a
Extreme Wet
1.20
Wet
0.98
Typical
0.78
Dry
0.49
0.27"
March 10, '05
Extreme Dry
0


Record
6.50
b
Extreme Wet
2.40
Wet
1.96
Typical
1.56
Dry
0.98
0.27”
March 20, '05
Extreme Dry
0

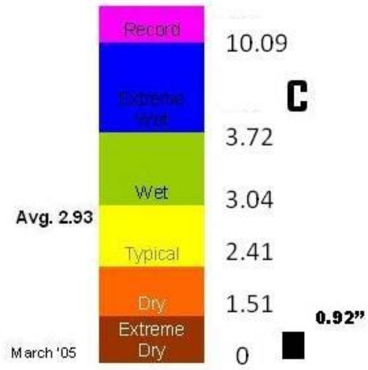
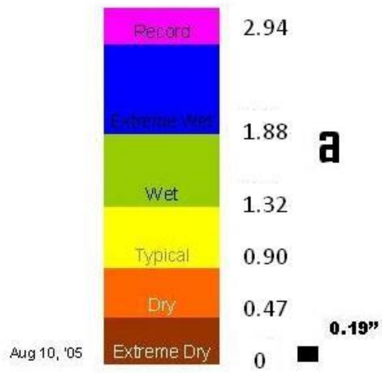


Figure 6. As in Fig. 5, except for March 2005. This figure shows a monthly progression for; a) March 10, b) March 20, and c) the end of the month.


Record
Extreme Wet
Wet
Typical
Dry
Extreme Dry
2.94
1.88
1.32
0.90
0.47
0
a
0.19"
Aug 10, '05

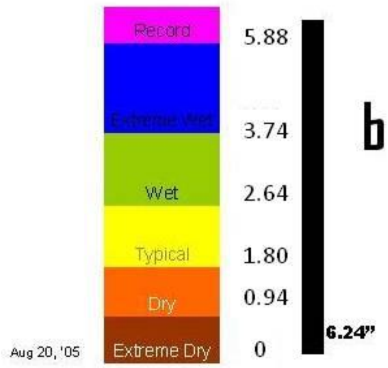


Figure 7. As in Fig. 5, except for August 2005.