5.4 ESTIMATING PARAMETERS FOR CHARACTERIZING TIMES BETWEEN STORMS

James V. Bonta*
Research Hydraulic Engineer
USDA - Agricultural Research Service
North Appalachian Experimental Watershed
Coshocton, Ohio

Greg Johnson
Applied Climatologist
USDA - Natural Resources Conservation Service
National Water and Climate Center
Portland, Oregon

[1]1. INTRODUCTION

1.1 *Background*

Hydrological modeling for engineering design often requires short-time increment precipitation data for use in newer watershed models. However, such data are seldom available, and computer simulation of short-time increment data is an alternative. One approach to simulation to meet this need is by simulating storms rather than using 24-hr totals, which are often disaggragated by inflexible design storms. Instead, actual storms start and end at any time of day and last from a few minutes to several days. Consequently, simulation of short-time increment precipitation “data” on a *storm* basis is needed. The present investigation is an exploratory study into simple parameterization methods for purposes of identifying and simulating the occurrence of storms.

1.2 *Storm-Occurrence Characterization and Simulation*

Simulating a time series of storms includes synthesizing a continuous record of the occurrence of *storms,* storm durations, storm depths, and within-storm intensities. This paper specifically addresses partial requirements for modeling and parameterization for a storm-generator model developed by Bonta (1997 and 2001a). The approach is statistical and storm physics is not considered. Statistical characteristics for a given location and time of year are maintained in the model.

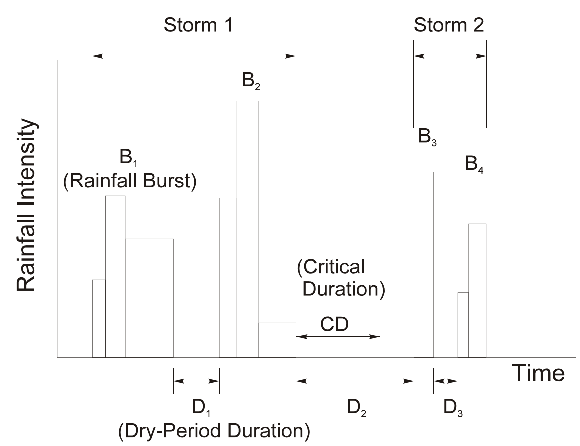

Figure 1. Definition for “critical duration” (CD) for identifying storms in a precipitation record.

Two characterization parameters are fundamental for storm-occurrence simulation - “critical duration” (CD) and average time between storms (TBS). CD is the minimum dry period between rainfall bursts that separates a record of precipitation into statistically independent storms (Fig. 1). Dry-period durations less than CD are incorporated into individual storms, and dry periods greater than CD separate a historic or simulated record into *storms* of varying durations. TBS is the average dry-period duration computed from the time-between-storm data resulting from storm identification using CD. CD and TBS parameters must be characterized on at most a monthly basis, and at specific locations, to account for observed temporal variations on seasonal or shorter time scales, and spatial variations (Bonta, 2001b).

Modeling of storm occurrence (the beginning of a storm - month, day, year, hour, minute) involves sampling from the exponential frequency distribution of times between storms,

$$F(TBS_i) = 1 - e^{(-TBSi / TBS)} \qquad (TBS_i >= CD) \qquad (1)$$

where $F(TBS_i)$ is the cumulative distribution function (fraction greater than), TBS is the average time between storms, and TBS_i is an individual value of TBS. Subsequent steps in simulation include sampling of the frequency distribution of storm duration to advance the simulation of storms in time, sampling of distributions of storm depths, and finally simulation of within-storm intensities. More detail on these subsequent steps can be found in Bonta (1997 and 2001a).

1.3 *Objectives*

The overall objective of the present study is to investigate potential practical methods for estimating CD and TBS that characterize the time between storms for a given location and month. Specific objectives are to investigate: 1) the adequacy of regressions between CD and average monthly precipitation (P_{mo}), TBS and P_{mo}, and CD and TBS; and 2) the direct mapping of TBS. More detail one estimation of CD and TBS discussed in this paper are found in Bonta (2002).

*Corresponding author address: James V. Bonta, USDA-ARS, NAEW, Box 488, Coshocton, OH 43812; email: bonta@coshocton.ars.usda.gov.

2. PROCEDURE

2.1 *Determination of CD - Identification of Storms*

Precipitation records are composed of bursts of precipitation (e.g., B_1, B_2, B_3, and B_4, Fig.1) and dry times between bursts (e.g., D_1, D_2, and D_3, Fig.1). The goal in the procedure is to determine the "critical duration" (CD), such that dry times (D_i) greater than CD separate storms from one another, and dry times less than CD are included in "storms". CD determined in this manner captures spatial and seasonal variability of dry periods that are useful for accurate simulation of storms.

Storms were identified in the present study by using the exponential method as described above (Restrepo and Eagleson, 1982). Dry times between storms (TBS) are found in recording rain-gauge data and subjected to a trial and adjustment process to find the CD at which durations longer than CD form an exponential distribution (eqn 1). The exponential distribution has the property that the mean and standard deviation are equal (the coefficient of variation, CV, is unity). CV is iteratively computed after removing lesser TBS (CV >1) until CV is less than unity. This process is repeated until CV becomes less than unity. At this TBS, a new TBS is interpolated between the current (TBS at CV<1) and previous (TBS at CV>1) values, and the interpolated value becomes the CD for the data set. The TBS that are greater than CD form the exponential distribution that is used in stochastic storm simulation, and which is characterized by the average TBS (eqn 1). Data were grouped by month, and the process was repeated for each month and rain gauge selected for analysis.

2.2 *Data Used*

Precipitation data from National Weather Service (NWS) 60-min recording rain gauges in the plains area of Eastern Colorado, Southern Wyoming, and Western Nebraska and Kansas were selected (Bonta, 2002). May, June, July, August, and September were selected for use in the present exploratory study to coincide with the same months used in a study of CD by Bonta (2001b). A total of 34 gauges were used, covering approximately 225,000 km^2 over parts of the four-state area (170 values each of CD and TBS).

2.3 *Regression Equations*

Exploratory plotting of the data led to evaluating rectangular- and log-log-grid equations between CD and TBS and P_{mo}, and CD and TBS. The linear form of each equation was fitted to the data through standard regression methods. For CD vs P_{mo}, the rectangular-grid and log-log linear equations were:

$$CD = a_R + b_R (P_{mo}) \qquad (1a)$$

$$\log_{10}(CD) = A_L + b_L \log_{10}(P_{mo}) \qquad (1b)$$

where subscript R associates parameters with rectangular coordinates and L associates parameters with log-log relationships, intercepts are a_R and $A_L = \log_{10}(a_L)$, and slopes are b_R and b_L.

For TBS vs P_{mo}, the rectangular-grid and log-log linear equations were:

$$TBS = c_R + d_R (P_{mo}) \qquad (2a)$$

$$\log_{10}(TBS) = C_L + d_L \log_{10}(P_{mo}) \qquad (2b)$$

where intercepts are c_R and $C_L = \log_{10}(c_L)$, and slopes are d_R and d_L

For CD vs TBS, the rectangular-grid and log-log linear equations were:

$$CD = e_R + f_R (TBS) \qquad (3a)$$

$$\log_{10}(CD) = E_L + f_L \log_{10}(TBS) \qquad (3b)$$

where intercepts are e_R and $E_L = \log_{10}(e_L)$, and slopes are f_R and f_L.

Eqns. 1b, 2b, and 3b are logarithmic forms of a power equation. The regressions based on the rectangular-grid are referred to as the "rectangular" regressions. Units for TBS and CD are minutes and for P_{mo} are mm in all equations above.

To illustrate the degree of fit of the regressions (best-fitted curves to one of the poorest-fitting curves), examples of regressions for the "station analyses" were plotted for three regressions: the best-fitted regression (100% of r^2 values [all 34 regressions] were less than this regression = largest r^2), median (50% of r^2 values were less than this regression {17 regressions]), and one of the poorest fits (10% of r^2 values were less than this regression - 90% were greater than this r^2 [30 regressions had r^2 greater than this]}.

Computed and measured TBS and CD values by different methods were compared by computing the median of the absolute value of deviations between calculated and observed values and results plotted.

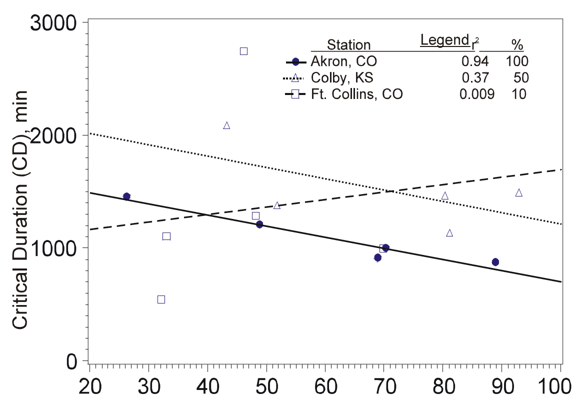

Figure 2 Sample good, medium, and poor regressions between CD and P_{mo}.

3. RESULTS AND DISCUSSION

3.1 *Estimating Critical Duration (CD vs P_{mo})*

Rectangular regression was as good as log-log regression so only rectangular-grid regressions were used. The analysis yielded generally poor r^2, with values of 0.94. 0.37, and 0.009 corresponding to cumulative r^2 percentages of 100%, 50%, and 10%, respectively. Data for Ft. Collins, CO (Fig. 2) had one outlier that skewed the regression line, compared with the trend of other points at this gauge. Data for Akron, CO showed the best regression. The generally poor correlations may be due to the observed sensitivity of the CD-estimation method to large values of TBS for a given month (Bonta, 2001b). This might occur because of poor data and/or because persistent dry periods caused by large-scale atmospheric forcings (e.g., SOI) that would tend to mix frequency distributions of “dry” and “normal” periods together, skewing both TBS and CD data towards longer values.

3.2 *Estimating Average Time Between Storms (TBS vs P_{mo})*

Log-log relationships consistently fitted the TBS vs P_{mo} data better than a rectangular equation form. For example, an r^2 of 0.60 or greater occurred for 94% of the log-log relationships, compared with 85% of the rectangular relationships. All three example log-log graphs (Fig. 3) appear to fit the data well. The worst sample regression (Garden City Station, KS - r^2=0.60) visually appears to be a good relation. The median example regression had a r^2 value of 0.81 (Phillips, WY), and the best regression had an r^2 of 0.97 (Akron, CO). Regression slopes (eqn. 2b) were always negative and varied from -2.60 to -0.80 (median=-1.09), and intercepts ranged from 3.84 to 8.63 (median= 5.74). The negative slopes imply that as average monthly precipitation increases, average time between storms decreases. This follows from the observation that *storm* depths and durations are positively correlated (Bonta, 1998). Therefore, if storm precipitation amounts decrease, then storm durations decrease, and dry times between storms increase. The strong correlations between TBS and P_{mo} suggest that PRISM P_{mo} maps (Daly et al., 2000, NRCS 1998a and b) can be used to estimate TBS for ungauged areas between stations if regression parameters can be determined, such as by mapping (Bonta, 2002). This is because of the generally good regressions found between P_{mo} and TBS. Furthermore, there is the possibility of parameter mapping (eqn. 2b) using the PRISM model. Mapping the slope and intercept parameters yields two maps, compared with five monthly maps required for the months of May through Sept. If a similar analysis was expanded to the entire year, two possible maps would be needed in place of 12 (one for each month), a possible 83% reduction in the size of a data base needed for estimating TBS.

It was found that slopes and intercepts in eqn. 2b were correlated. When intercept (C_L) is graphed against slope (d_L), a strong negative correlation is apparent (Fig. 4) with an r^2 of 0.99,

$$C_L = \log_{10}(c_L) = 3.703 - 1.862\, d_L. \qquad (4)$$

This equation has the potential to relate slope and intercept map surfaces over the study area. Eqn. 4 is encouraging for simplifying the estimation of TBS for storm-generator purposes for any specific location. The correlated parameters imply that only one map is necessary, reducing a data base needed for estimating TBS by 50% (potentially by 11/12 compared with direct monthly TBS mapping for an entire year). By substituting eqn. 4 for intercept into eqn. 2c and simplifying, the equation for TBS (min) for a particular month becomes:

$$TBS = 5047\,(0.0137\, P_{mo})^{dL} \qquad (5)$$

and a map of the spatial variation of the slope parameter (d_L) is required. By solving eqn. 4 for slope and substituting into eqn. 2b, the equation for TBS (min)

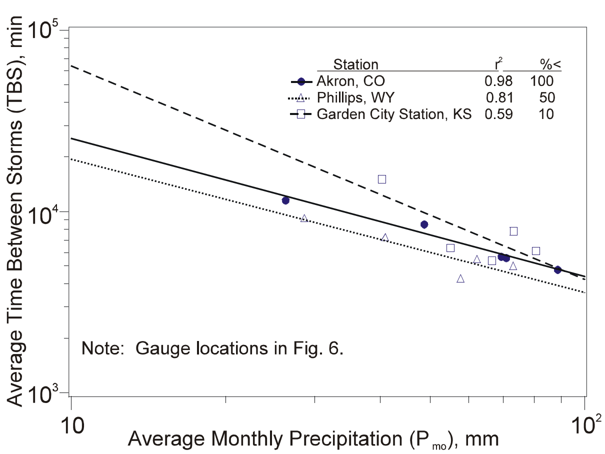

Figure 3. Three representative relationships between TBS and P_{mo}.

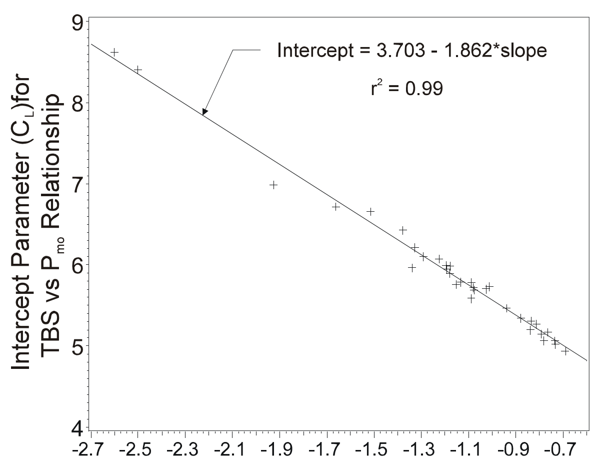

Figure 4. Relationship between intercept and slope parameters in eqn. 2b.

becomes,

$$TBS = c_L \, P_{mo}^{\,(1.989 - 0.537\,CL)} \qquad (6)$$

and a map of the spatial variation of the intercept parameters in eqn 6 (c_L and C_L) is required. The spatial extent, mathematical bounds, and applicability of eqns. 5 and 6 are unknown and require further study. Mapping the parameters, such as potentially with the PRISM model (as opposed to general contouring), may improve the utility of the equations particularly in areas with significant spatial complexity of climate such as in mountains.

It was found that eqn. 2b yielded better estimates of TBS compared to TBS estimated from eqns. 5 and 6 (median deviation=768 min - Bonta, 2002). TBS estimates resulting from eqn 6 that required the intercept performed slightly better than eqn 5 that required the slope (median deviation=1089 min compared with 1158 min). Fig. 5 shows the scatter between measured and computed monthly TBS using eqn 2b.

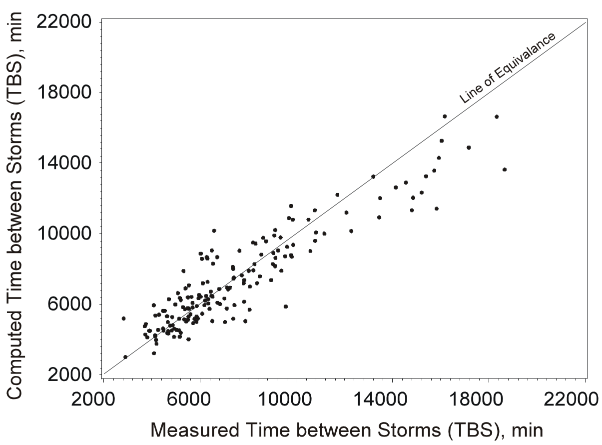

Figure 5 Comparison between measured and computed TBS (eqn 2b).

It can be concluded that eqn. 2b links the average seasonal climatology of the study area (in particular average monthly precipitation) to a parameter required for storm modeling (TBS). Mapping using the PRISM model to map parameters may reduce the error in computing TBS for ungauged areas and may make eqn 2b more feasible; however, this requires further study. A closer investigation of the data by identifying periods of dry-period persistence in the data (e.g., SOI) may reduce variability in all relationships because frequency distributions of TBS for persistently dry periods would not be mixed with those of "normal" periods, which currently skews the CD and TBS toward large values. Such data separation may lead to better correlations. Furthermore, a sensitivity study is required to evaluate how much error in estimating TBS is tolerable for specific objectives.

3.3 *Estimating Critical Duration (CD vs TBS)*

Log-log and rectangular-grid regressions yielded approximately equal, acceptable frequency distributions of r^2. Consequently, only the rectangular-grid equation is used in further development (eqn. 3a). The example graphs (Fig. 6) show that the best regression was found at Big Springs, NE (r^2=0.98), the regression having 50% greater r^2 values was found at Castle Rock, CO (r^2=0.78). While these are representative of some of the better regressions, some regressions had r^2 values as low as 0.06. For these few poorer regression, problems with data or the effect of mixing exponential distributions for different types of weather may be a factor affecting the regressions. For example, persistent dry periods due to SOI may be mixed with wetter periods, skewing the CD and TBS values. Separating data into these classes may improve correlations investigated in this paper. The strong correlation between TBS and the climate variable P_{mo}, and the correlation between CD and TBS, suggest that CD may be a predictable climate-related variable.

Slope parameters (eqn 3a) over the 4-state area ranged from -0.035 to 0.291 with a median of 0.131. Intercept parameters ranged from -944 to 1991 with a median of 260. The *slope* parameter ranges from less-than-zero to greater-than-zero values, suggesting that regressions with zero slope (independence between CD and TBS) are included in the data set. Also, the *intercepts* span zero, suggesting that CD depends only on slope for some gauges. Maps of the slope and intercept surfaces could potentially be used to compute estimates of CD at any point in the 4-state area (Bonta, 2002).

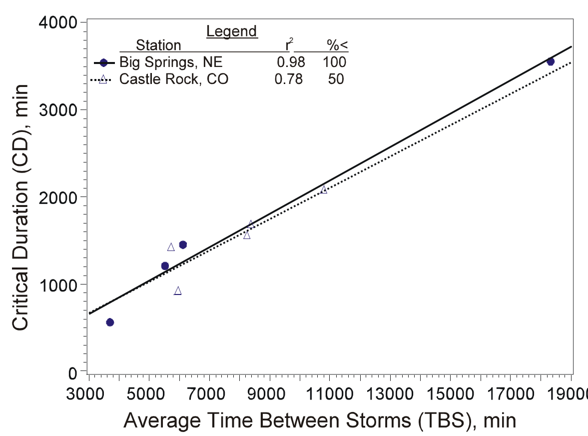

Figure 6 Regressions of critical duration vs. average times between storms for three example gauges.

It was found that there was an inverse relationship between e_R and f_R as found above for corresponding parameters for TBS vs P_{mo} (strong regression with an r^2 of 0.86).

$$e_R = 1294 - 7103\, f_R. \qquad (7)$$

Substituting eqn. 7 into eqn. 3a and simplifying yields,

$$CD = 1294 - f_R\,(\,TBS - 1703\,) \qquad (8)$$

Similarly, solving eqn. 7 for f_R and substituting into eqn. 3a yields,

$$CD = e_R + TBS\ (\ 0.182 - e_R\ /\ 7103) \qquad (9)$$

Eqn. 8 requires a map of slope for estimation, and eqn. 9 requires a map of intercept over the study area (Bonta, 2002). These equations reduce five maps required for CD estimation (e.g., Bonta, 2001b) to one map of either slope or intercept, in addition to a slope or intercept map for the TBS vs P_{mo} relationship, and a map of P_{mo} (e.g., from PRISM) for a maximum of three maps.

The individual regressions (eqn 3a) are noticeably better than the other regressions using correlated paramters using eqns. 8 and 9. Median deviations from eqn. 3a was 173 min, and from eqns. 8 and 9 were 891 and 485 min. Fig 7 shows relatively good agreement between measured and computed CD using equation 3a for individual gauges (all months and gauges for 170 values). As mentioned previously, measured CD is sensitive to large TBS values (e.g., as encountered with persistent dry periods), and correlations may be improved by avoiding the potential mixing of frequency distributions of two or more distinct set of TBS data (e.g., dry, wet, and “normal” periods). The association of CD with TBS suggests that PRISM mapping of CD or its regression parameters with TBS may have potential.

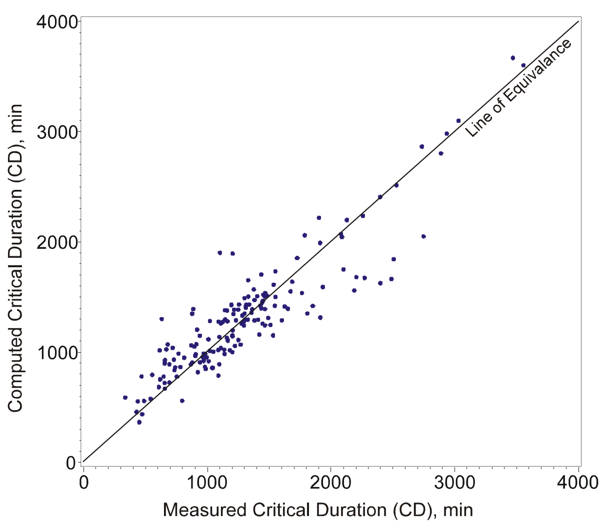

Figure 7. Comparison between measured and computed CD using individual regressions (eqn. 3a)

4. CONCLUSIONS

Methods were investigated to estimate critical duration (CD - the minimum dry time between bursts of precipitation that identify storms) and the average time between storms (TBS - the parameter of the exponential distribution of times between storms). The study area included 34 rain gauges covering approximately 225,000 km^2 of a four-state area including parts of Colorado, Wyoming, Kansas, and Nebraska, and for five months from May through September. The following conclusions can be made:

- Individual precipitation-gauge station analysis provided relatively good estimates of TBS and CD using regressions of TBS vs P_{mo} and CD vs TBS.
- Correlated parameters reduce the number of maps needed for estimation of CD and TBS, and consequently of the size of a data base that could be developed from the results in the present study. Slope and intercept parameters were strongly correlated with r^2=0.99 for TBS vs P_{mo} and r^2=0.87 for CD vs TBS.
- Reduction of regression error in CD and TBS estimation may be possible by separating data sets based on “wet”, “dry”, and “normal” dry periods, taking into account forcings such as SOI that cause persistence in the data, and by mapping regression parameters for individual stations for both CD and TBS using the PRISM model.

The results of this study are useful for guiding practical parameterization methods for CD and TBS for a stochastic storm-generator model. The results are also useful for other hydrological investigations such as floods and droughts. However, more research is needed to improve the accuracy of the methods for estimating CD and TBS. This includes investigation of climate variables other than, or in combination with, monthly average precipitation.

5. REFERENCES

Bonta, J.V., 1997: Proposed use of Huff Curves for hyetograph characterization. pp. 111-124. In: C.W. Richardson et. al., (ed.) *Proceedings of the Workshop on Climate and Weather Research. Denver, Colorado.* July 17-19, 1995. USDA-Agricultural Research Service, 1996-03, 223 pp.

Bonta, J.V., 1998: Generation of storm occurrence, depth, and duration. Presented at ASAE Meeting, July 12-16, 1998, Orlando, FL. Paper #982023.

Bonta, J.V., 2001a: Development of a New Storm Generator Model and Associated Precipitation Studies, *Proceedings 81st Annual Meeting of the Amer. Meteorological Soc.*, 14-19 Jan. 2001, Paper # J2.12.

Bonta, J.V., 2001b: Characterizing and Estimating Spatial and Temporal Variability of Times Between Storms. *Trans. ASAE*, **44**(6),1593-1601.

Bonta, J.V., 2002: Estimation of Parameters Characterizing Frequency Distributions of Times Between Storms. Submitted to *J. of Appl. Meteor.*

Daly, G.H. Taylor, W.P. Gibson, T.W. Parzybok, G.L. Johnson, P.A. Pasteris, 2000: High-quality spatial climate data sets for the United States and Beyond. *Trans. ASAE,* **43**(6),1957-1962.

NRCS, 1998a: PRISM climate mapping project, Precipitation, compact disk for the Central Region.

NRCS, 1998b: PRISM climate mapping project, Precipitation, compact disk for the West Region.

Restrepo, P.J. and Eagleson, P.S., 1982: Identification of independent rainstorms. *J. of Hydrology,* **55**,303-319.