STOCHASTIC SIMULATION OF STORM OCCURRENCE, DEPTH, DURATION, AND WITHIN-STORM INTENSITIES

J. V. Bonta

ABSTRACT. *Newer watershed models require detailed continuous temporal precipitation to drive modeled hydrologic processes. The variability of precipitation inputs to models is a major source of variability in watershed flows that is often evaluated in the context of risks. However, available short−term−increment rainfall data are not adequate. A storm−generator model (StormGen) that synthesizes storms directly (several storms per day to several days per storm) was developed and tested using precipitation data from Coshocton, Ohio. The concepts for modeling the four elements of actual storms (storm occurrence, storm duration and depth, and within−storm intensities) and storm−model characterization and parameterization using an empirical and statistical approach are presented. Supporting studies have been conducted to help with practical parameterization in ungauged areas. Times between storms (TBS) are represented and simulated well by exponential distributions. Storm durations are characterized by empirical distributions of durations in a month. Storm depths require conditional simulation given storm durations. Within−storm intensities utilize probabilistic information contained in Huff curves. Initial performance evaluation of three of the elements of the model for a 200−year simulation shows that the model works well. TBS was modeled best (monthly average deviation of −1.3%, with values ranging from −3.2% to 3.0%). Monthly average storm duration deviations ranged from −6.2% to +1.6% with an average of −1.3%. Monthly average storm depth deviations ranged from −17.1% to +0.1% with an average of −9.2%, although actual magnitudes ranged only from −2.2 to 0 mm. Average deviations between simulated and measured average monthly precipitation was +1.7%, ranging from −8.6% to +12.6%. Corresponding depth differences ranged from −10.0 to 10.8 mm, with an average of +0.5 mm. Total simulated precipitation for the entire period deviated from measured precipitation by +0.6%, corresponding to +0.5 mm. The Poisson assumption for statistical independence of storms in the model is validated using Coshocton data, with the ratio of average storm duration to average TBS of only 0.107. The results of this study suggest that the StormGen model is promising and that modeling concepts and characterization deserve further investigation.*

Keywords. *Disaggregation, Drought, Storm depth, Storm duration, Storm generator, Storm simulation, Storm synthesis, Time between storms.*

Watershed models are required for engineering design for runoff and erosion control, water−quality evaluations, and global−change investigations. These models are becoming more sophisticated and require detailed continuous temporal and spatial inputs of precipitation to drive the modeled hydrologic processes. The variability of precipitation and weather inputs to models is a major source of variability in outputs from watershed models. These outputs are often evaluated in the context of risks, frequencies of occurrence, and durations of exceedance of flow and water−quality constituents. Particularly needed are long records of data having short time increments (of the order of minutes) with fine depth resolution for advanced infiltration modeling, routing, and water−quality algorithms. However, such records that include measured precipitation extremes are generally unavailable, or they do not have good spatial coverage.

Article was submitted for review in February 2004; approved for publication by the Soil & Water Division of ASAE in June 2004.

The author is **James V. Bonta, ASAE Member Engineer,** Research Hydraulic Engineer, USDA−ARS North Appalachian Experimental Watershed, Box 488, Coshocton, Ohio 43812; phone: 740−545−6349; e−mail: bonta@coshocton.ars.usda.gov.

Precipitation and weather data for natural−resource models often use records of data collected over a 24 h period. These data can be measured from midnight to midnight, from 0800 to 0800, or between other observation times. When using these data, the actual observations times are ignored and data are often assumed to have occurred from 2400 to 2400, even though they were not measured during this time. "Daily" data with varying observation times lead to temporally and spatially noncomparable data sets for analyses that often are ignored.

Weather generators, e.g., GEM (Hanson et al., 1989, 1994, 2002) and CLIGEN (Nicks et al., 1995), can supply temporally distributed stochastic estimates of precipitation, but the smallest time resolution is 24 h. This is because these models are parameterized using the 24 h amounts described above. Stochastically generating storm hyetographs given a 24 h total generated by these models is complicated, because a single storm can span a 24 h period, several storms can comprise a 24 h total, or both situations can occur. Stochastically disaggregated 24 h amounts depend on quantification of serial dependence from one day to the next, which is difficult to accomplish. CLIGEN incorporates a 24 h disaggregation algorithm that simulates storm duration, time to peak, and peak intensity (Yu, 2000).

Typically, engineers resort to "design storms" to disaggregate daily−total precipitation to estimate within−storm intensities because of the unavailability of short−time increment precipitation data. However, design storms are fixed patterns of precipitation distributed in time, and do not represent the wide random variability of storms found in nature. Design storms require subjective assumptions regarding the equivalence of precipitation and flow frequency distributions, and they require assignment of the duration of precipitation, season of year, and antecedent moisture condition. Design storms are inappropriate for thoroughly analyzing watershed model outputs in terms of flow durations, load durations, and extremes. Consequently, stochastic synthesis of a continuous record of storms is an attractive alternative, and would eliminate the 24 h and design−storm constraints. Use of a storm model would allow generation of a time series of precipitation for locations for which data do not exist, and is the subject of this article.

Many stochastic models of the precipitation process have been developed through the years (Bras and Rodriguez−Iturbe, 1976; Corotis, 1976; Cox and Isham, 1998; Crovelli, 1971; Grace and Eagleson, 1966; Knisel and Snyder, 1975; Marien and Vandewiele, 1986; Morris, 1980; Morrison and Smith, 2001; Nguyen and Rousselle, 1981; Pattison, 1965; Rao and Chenchayya, 1974; Raudkivi and Lawgun, 1974; Todorovic and Woolhiser, 1974; Valencia and Schaake, 1973; Woolhiser and Osborn, 1985). A review of the literature reveals that stochastic rainfall generation models can be categorized in many ways including: degree of incorporation of season/time of year, whether the model is concerned with extreme precipitation or all precipitation, degree of autoregressive characterization, incorporation of space and time or just time at a point, sequences conditioned on whether a day/period was wet or dry, differing degrees of incorporation of storm occurrence, storm depth and duration, and within−storm intensities, many incorporate some form of the Poisson−process assumption, whether the model is characterized deterministically or probabilistically, fineness of short−time−increment rainfall, and number of parameters. Of particular interest are the varying approaches to models of within−storm intensities by Woolhiser and Osborn (1985), Knisel and Snyder (1975), Over and Gupta (1996), and Grace and Eagleson (1966). These approaches include cascades, probabilistic characterization of the progression of intensities, and an urn model. No models incorporate snowfall, and none have extensive supporting studies for general practical characterization in areas without data.

A stochastic storm generator (StormGen) has been developed at the USDA Agricultural Research Service (ARS) North Appalachian Experimental Watershed (NAEW) facility at Coshocton, Ohio. It stochastically simulates the four basic elements of storms at a point: storm occurrence (dry time since the end of the previous storm to the beginning of the next storm), storm duration, storm depth, and short−time increment (of the order of minutes) within−storm precipitation intensities. The model incorporates improved features of some models found in the literature and new concepts for storm modeling. It has been developed simultaneously with supporting studies of characterization to make it practical for ungauged areas, and to help guide the modeling concepts. Additionally, the effect of time of year is also incorporated by using a monthly time step for characterization and simulation. This article is part of a larger project to develop a practical spatial−temporal storm model.

OBJECTIVES AND SCOPE

The objectives of this study are to: (1) present the modeling concepts for development of a stochastic storm generator that removes the 24 h precipitation−total constraint and simulates storms, (2) present concepts for characterizing and parameterizing the model, and (3) conduct a preliminary evaluation of model performance. The version of StormGen described is version 2004.01. A more comprehensive application and testing of StormGen for individual storm elements in other climates is subject of other articles.

PROCEDURE

GENERAL APPROACH

The approach is statistical and empirical, and does not include storm physics. Four tasks comprise the present study: (1) identifying storms in a precipitation record, (2) characterizing and developing frequency distributions of storm duration and depth data, (3) developing a conceptual and mathematical model, and (4) conducting preliminary tests of the model performance. The storm model and characterization methods were developed simultaneously to maximize the practicality of the model for ungauged areas. Many supporting studies have been conducted to aid in practical parameterization (Bonta, 2001, 2003, 2004). Guiding principles for simplified parameterization include minimization of the number parameters and selection of parameters that could be used for global−change studies.

The underlying assumption in the model is that storm events follow a time−varying Poisson process, with exponentially distributed interarrival times. The validity of this assumption will be discussed later. While this process varies in time, monthly characterization and simulation assumes stationarity within a month. A monthly time step minimizes data characterization needed for the model, yet allows seasonal characterization of precipitation. Monthly frequency distributions and their parameters comprise the majority of the information required to run the model.

The model has been developed so that optimization and calibration of the model are not necessary. However, empirical data are used to develop parameters. Current studies involve developing relationships to estimate these parameters in the absence of data and evaluating the robustness of the characterization methods.

A previous simulation study with the model investigated interpolation methods to smoothly vary simulation between months for the underlying distributions of storm occurrence, durations, and depths (Bonta, 1998). Interpolation would avoid possible sudden changes in statistical characteristics in modeled output because of changes in monthly frequency distributions due to local climate. A combination of no month−to−month storm−duration and TBS interpolations, and simple semilog interpolation between storm−depth distributions for a given month yielded optimal modeling results.

In the present study, each storm element, the methods of characterizing, parameterizing, and interpolating each element, and the simulation concepts are presented. Only preliminary model performance evaluation of storm occurrence and storm depths and durations are studied, and evaluation of the within−storm simulation is not included. Evaluation of the simulation performance of within−storm intensities is the subject of another article. Comprehensive model performance evaluation of the four elements of the model in different climates will also be reported in other articles. Evaluation factors in the present study include comparison of measured and simulated frequency distributions, monthly totals, and period totals of precipitation.

Data Used

Data from rain gauge RG100 at the NAEW facility are used to illustrate the proposed concepts of the storm generator. Data from this gauge are tabulated at changes in precipitation intensity ("breakpoint" data) for the 61−year period of record from 1937 through 1997. The depth resolution of the Coshocton data is 0.01 in (0.25 mm), and the time resolution is 1 min. Average annual precipitation for RG100 at Coshocton, Ohio, is 950 mm. Only the months of April through October (referred to as the "period") are used to minimize the possibility of including snowfall, which must be treated separately. Average precipitation for the period is 602 mm.

Characterization and Simulation of the Four Storm Elements

Storm Occurrence

Storm Occurrence Characterization

Prior to characterizing storm occurrence and the other three storm elements, storms must be identified and separated within precipitation data to form the underlying database for further characterization of other storm elements. Several methods exist to identify storms (Bonta and Rao, 1988a); however, often a minimum dry period is used. In a precipitation record, bursts of precipitation are separated by periods of no precipitation (fig. 1a), or "dry−period durations" (D_i). Intuitively, at the extremes for a given location, dry−period durations of the order of minutes between bursts of precipitation (B_i) would belong in the same storm as the bursts. However, bursts of precipitation separated by dry−period durations of the order of days would not belong in the same storm. Consequently, there is a "minimum dry−period duration" (MDPD) that separates bursts of precipitation. For example, figure 1a shows that continuous precipitation bursts B_1 and B_2 are separated by dry−period duration D_1, bursts B_2 and B_3 are separated by D_2, and B_3 and B_4 are separated by D_3. A dry period greater than MDPD separates groups of bursts of precipitation and short dry periods from one another, and the D_i greater than MDPD are termed "times between storms" (TBS). Dry periods less than MDPD are included in "storms." In figure 1a, $D_3 < D_1 < \text{MDPD} < D_2$, and two storms are apparent. Often a constant, arbitrary MDPD value is used to separate storms (e.g., 6 h by Huff, 1967). However, Bonta (2001) showed that MDPD can depend on the time of year, and that median monthly MDPD ranged from 16.5 to 26.9 h over the plains area of Colorado and adjoining states in the U.S., encompassing an area of 225,000 km^2.

Restrepo and Eagleson (1982) developed the exponential method of computing the MDPD to identify storms in a precipitation record that is used in the storm model. Their iterative method assumes that the MDPD is found when D_i greater than MDPD form an exponential distribution:

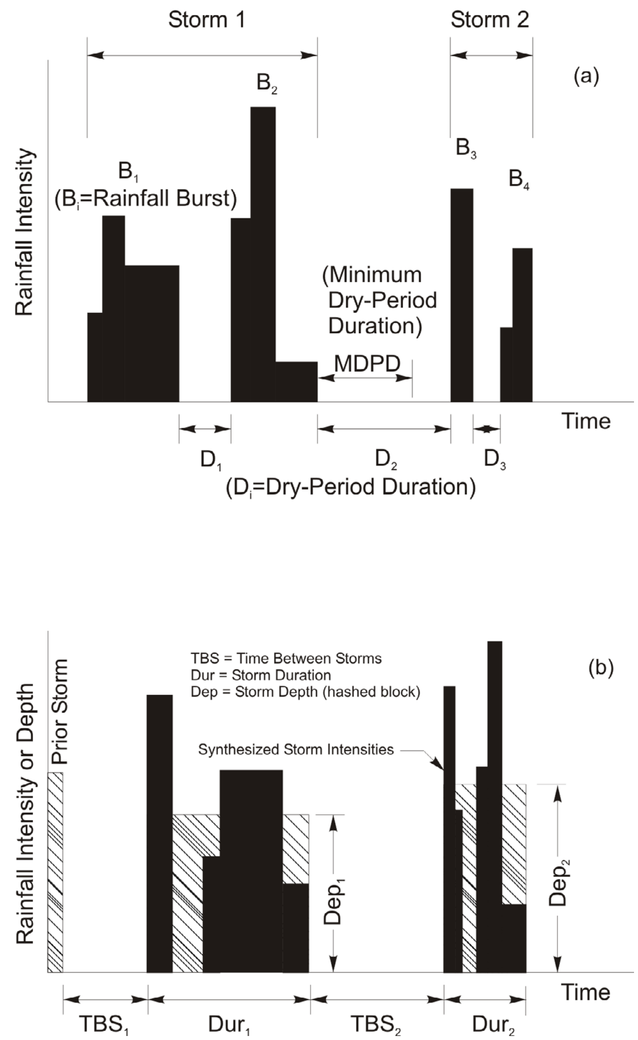

Figure 1. Schematic of (a) approach used to identify and separate storms and (b) simulation of storm occurrence, duration, depth, and within−storm intensities by StormGen.

$$F(\text{TBS}_i) = 1 - e^{(-\text{TBS}_i / \text{ATBS})} \quad (\text{TBS}_i \geq \text{MDPD}) \qquad (1)$$

where

$F(\text{TBS}_i)$ = cumulative distribution function (fraction greater than)
TBS_i = individual value of TBS (time units)
ATBS = average time between storms (time units).

The exponential frequency distribution has the property that the mean and standard deviation are equal, i.e., the coefficient of variation (CV) is unity, and consequently it is a one−parameter model (the mean, ATBS). The iterative process begins with computing CV for all dry periods for a given month collapsed across years. If CV is greater than unity, then the shortest duration is eliminated from the data set and CV is recomputed with the remaining data. The process is repeated until CV < = 1. The interpolated TBS duration at CV = 1 becomes the MDPD. Seasonal variation and changes in climate throughout a year are characterized by developing monthly frequency distributions.

Monthly TBS are associated with the beginning month of the dry period, and can extend into one or more succeeding months. This allows dry periods lasting longer than one

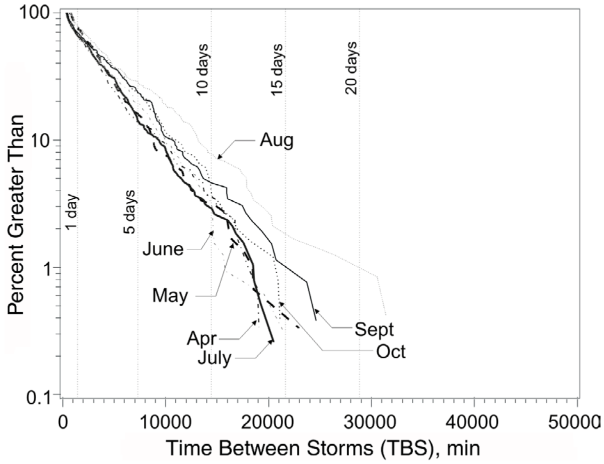

Figure 2. Empirical monthly exponential frequency distributions of TBS for RG100 at Coshocton, Ohio.

month, and allows a smooth transition in frequency distributions from month to month.

The significance of the exponential distribution of TBS is that interarrival times of "events" in a Poisson process follow this distribution, making the events (storms) statistically independent and simplifying simulation. However, storms have a duration associated with them, whereas the Poisson process assumes events to be instantaneous without duration. This assumption eliminates the need to condition storm occurrence on the occurrence of previous storms, simplifying storm modeling. Restrepo and Eagleson (1982) suggested that the Poisson−process analog is appropriate if the ratio of mean storm duration to the average TBS is much smaller than unity. They found that the exponential method generally gave storm durations ranging from 0.06 to 0.29 of ATBS (mean interarrival time), with a median of about 0.14 and with 70% of the values less than 0.19. Worldwide data were used and ranged from very dry (Saudi Arabia) to very wet (Colombia) climates. Examples of the good fit of the exponential method to TBS data are shown in Bonta (2001, 2003) in an area for which average annual precipitation is about 420 mm.

The exponential method is used to characterize and simulate storms in the StormGen model. Monthly characterization of the model requires 12 sets of 2 parameters (ATBS and MDPD on a monthly basis). Parameters ATBS and MDPD account for storm characteristics occurring during different times of the year, climates, and locations. StormGen can provide long−term simulations of a partial year (e.g., only the April through October period for each year of simulation), reducing parameter requirements and allowing flexibility for specific investigations.

For RG100, MDPD ranges from 330 min in October to 799 min in May (table 1). The corresponding exponential trends in the TBS data are presented in figure 2 (TBS data that fit an exponential distribution have a linear trend on a semilog grid). August has the longest average dry times between storms (ATBS = 5835 min), and April has the shortest dry times (ATBS = 3946 min). The Poisson assumption is validated for RG100 by observing that the ratio of the average duration to ATBS is much less than unity for all months (ranging from 0.074 to 0.176 in table 1, with an average of 0.107). The StormGen model has the capability to use empirical distributions, fitted exponential distributions, and parameter estimates from relations between ATBS and MDPD and average monthly precipitation (e.g., Bonta, 2003).

Storm Occurrence Simulation

Modeling of storm occurrence begins with sampling a TBS from the exponential frequency distribution of TBS for the current month since the end of the last storm (TBS_1 in fig. 1b). The "current" month at the beginning of simulation starts at the beginning of the month of interest specified in the input file of StormGen. A simulated TBS defines the beginning of a new storm in terms of the month, day, year, and 4−digit hour−minute of the day. The model can simulate varying time resolutions, including each minute, rounding to the nearest equal time increment (e.g., nearest 5 min, etc.), hourly, and daily. Simulation is advanced by sampling alternately between storm duration (Dur_i explained in the next section) and TBS distributions (fig. 1b; TBS_1, Dur_1, TBS_2, Dur_2, etc.).

Storm Duration

Storm Duration Characterization

Monthly storm−duration frequency distributions are formed using the database of storms identified with MDPD (e.g., fig. 3). Measured monthly storm durations for RG100 at Coshocton range from 299 min (July) to 753 min (May) (table 1). The importance of quantifying the effect of time of year (i.e., month) is apparent by observing the wide differences in horizontal positions of the distributions in figure 3.

Bonta and Rao (1992), Rao and Chenchayya (1974), and Grace and Eagleson (1966) found that the Weibull distribution fit storm−duration data well. StormGen has the capability to accept monthly Weibull and other distribution parame−

Table 1. Measured rainfall characteristics at Coshocton, Ohio, for RG100.

Month	Minimum Dry Period Duration, MDPD (min)	Average Time Between Storms, ATBS (min)	Ratio of Average Storm Duration to ATBS	Storm Duration (min)	Average Measured Storm Depth (mm)	Average Monthly Total Rainfall (Storms, mm)	Average Monthly Total Rainfall for NAEW (Rain gauge, mm)	Ratio of Storm Total to NAEW Total
April	660	3946	0.159	627	10.4	86.0	86.2	1.00
May	799	4286	0.176	753	12.3	102.1	97.0	1.05
June	489	4259	0.097	414	12.3	107.7	102.6	1.05
July	392	3967	0.075	299	12.0	115.8	108.7	1.07
August	638	5835	0.074	434	12.9	83.6	80.2	1.04
September	448	4809	0.078	373	10.7	77.1	70.6	1.09
October	330	4651	0.088	410	8.1	58.0	56.6	1.02
Average			0.107					1.05
Total						630.3	601.9	1.05

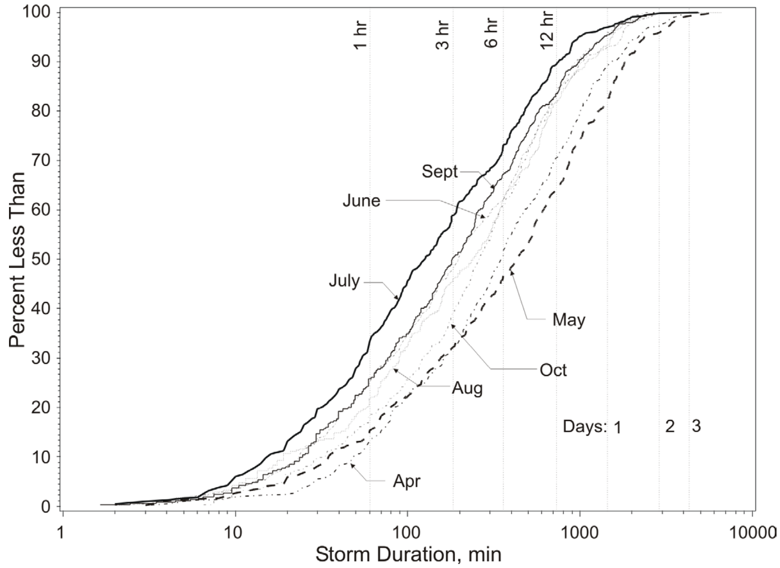

Figure 3. Monthly storm−duration frequency distributions for RG100 at Coshocton, Ohio.

ters, as well as to use empirical storm−duration distributions. The model also has the capability to simulate with varying time resolutions, including each minute, equal increment (e.g., every 5 min), hourly, and daily. Methods for improving characterization and for simplifying parameter estimation of these distributions in ungauged areas are currently being investigated.

Storm Duration Simulation

Storm duration (Dur_i) is simulated after synthesis of a TBS. The method proposed by Bonta (1997) is to sample the frequency distribution of storm durations for the current month (fig. 3). The month associated with storm duration is the month in which the beginning of the storm occurs, and storms can span months. Simulation time is advanced by storm duration since the time of the previous TBS, and is expressed in terms of the month, day, year, and 4−digit hour−min of the end of the storm. TBS and storm−duration distributions are alternately sampled to advance simulation time as described previously (fig. 1b; Dur_1, TBS_2, Dur_2, ..., TBS_n, Dur_n). Storms are statistically independent because of the Poisson−process assumption.

Storm Depth

Storm Depth Characterization

Monthly storm−depth frequency distributions are formed using the database of storms identified with MDPD. It is apparent that there is a statistically significant correlation between storm depths and durations, although there is wide scatter (points in fig. 4a; rank correlation coefficient = 0.57 at <0.0001 significance probability). Therefore, conditional distributions of storm depths must be developed for stochastic simulation that are dependent on storm duration. For example, it is highly unlikely there will be a large storm precipitation for a storm of short duration. There are at least three methods of characterizing conditional relations using data grouped by month: (1) regression of storm depths on durations, (2) determining regions of statistical independence between storm depth and duration, and (3) equal−duration−quantile characterization of storm duration and depths.

The first approach uses an equation (power equation in log−log linear form) to remove the positive correlation between storm depths and durations (slope = power = 0.655 in fig. 4a). The equation is forced through the point that describes the time and depth resolution of the data (1 min and 0.01 in. = 0.25 mm for Coshocton data; fig. 4a). Stochastic storm−depth simulation uses the regression equation and samples from the frequency distribution of residuals. However, the standard deviation of residuals varies with storm duration, with small values of standard deviation at small and large durations (fig. 4a). Furthermore, storm−depth distributions are truncated at the depth resolution of the data, leading to bounded distributions.

The regression approach was used by Grace and Eagleson (1966) and Rao and Chenchayya (1974), where a linear equation was used instead of a power equation. A reason for the difference in equation form may be the use of differing methods of developing an MDPD. The exponential method of identifying storms is used in the present study compared with the rank−correlation method used in the two cited references that yield shorter MDPD values (Bonta and Rao, 1988a).

A second approach is to identify regions of independent depth−duration points, and this method is used in the StormGen model. Frequency distributions of depth are found by iteratively finding adjacent independent regions of storm depths and durations on a plot of storm depth versus duration (fig. 4a). Starting at the smallest storm duration, the rank correlation coefficient is computed for a minimum set of depth−duration points and tested for statistical significance. If the correlation is significant, then the upper bound of the current region is set to the maximum duration of the set of points. If not, then the next larger depth−duration pair is added and a rank correlation is recomputed and tested for significance. This procedure is repeated until all points are included in independent regions. For example, figure 4a shows that ten independent depth−duration regions characterize the RG100 data for July. The ten corresponding empirical frequency distributions of storm depths between the duration bounds are developed by collapsing depth data within each duration region (fig. 4b). These empirical

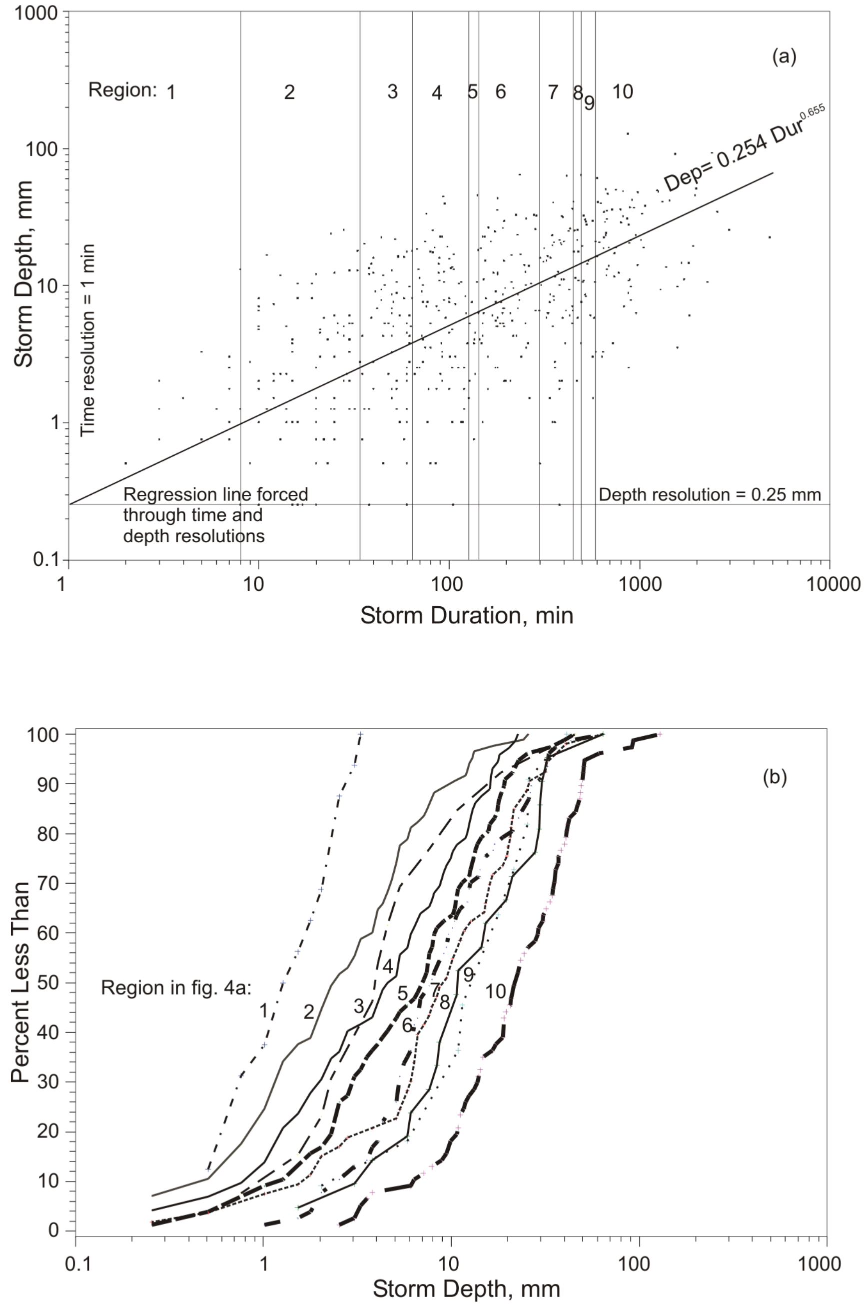

Figure 4. July storm precipitation for RG100 at Coshocton, Ohio: (a) correlation between storm depth and duration and regions of independence, and (b) conditional storm−depth distributions for duration regions above.

distributions are truncated at the depth resolution of the data (e.g., 0. 25 mm). This method of characterizing conditional storm depths does not yield unique independent duration regions because the procedure depends on the starting duration−depth pair.

The independent−region approach worked well for a study in which the magnitudes and frequencies of occurrence of peak runoff rates were determined (Bonta and Rao, 1992). Storms for a particular month, and corresponding empirical frequency distributions of depths, were developed for assigned storm−duration regions. The duration and depth distributions were stochastically sampled for watershed model inputs.

The third method varies the independent−region method by assigning storm−depth values to lie between fixed quantiles of storm duration. This removes the problem of nonuniqueness of independent regions. However, it also allows possible collapsing of data from different underlying depth distributions, resulting in statistically significant correlations within quantile regions. Practical parameterization would include relations between quantiles boundaries and other climatic data, and is an area for further research.

The Weibull distribution has been used successfully by Bonta and Rao (1992) to describe the frequency distributions of storm depths. StormGen can use empirical conditional depth distributions (fig. 4b), parameters for fitted Weibull distributions, and parameters for other distributions. The depth−duration input file to the current version of the model will accept either equal or unequal−quantile characterizations (methods 2 and 3).

Storm Depth Simulation

The approach used for simulating storm depths is to treat storm depth as dependent on storm duration using the independent−region method above (figs. 4a and 4b). For a given month, storm duration is sampled from the appropriate monthly distribution of duration (fig. 3). With this duration value, the appropriate duration region is determined (fig. 4a), and the corresponding depth distribution is sampled. For example, if storm duration of 10 min was simulated for July, then the depth distribution in region 2 would be sampled (fig. 4a and 4b). In figure 1b, precipitation depths of Dep_1 for the first storm and Dep_2 for second one would be stochastically simulated. StormGen allows the resolution of simulated depths to be assigned (e.g., 0.25 mm, 2.5 mm). The model can simulate TBS, duration, and depth without synthesizing the within−storm intensities described next. This option is used in the preliminary performance evaluation of the model.

WITHIN-STORM INTENSITIES

Storm Intensity Characterization

Within−storm intensities are stochastically modeled for each storm total duration synthesized above. There are an infinite number of variations of storm intensities possible in a storm, and fixed patterns such as design storms are not suitable.

The natural, random, within−storm variability of intensities can be summarized by using Huff curves (Huff, 1967). These curves were first developed by using heavy rain storm data from rain gauges located in Illinois and are isopleths of probability of storm intensities. Bonta and Rao (1992) developed a practical method for using more of the information contained in these curves for modeling peak flow rates and their frequencies of occurrence. Bonta (1997) first suggested the use of Huff curves for stochastic simulation of storm intensities.

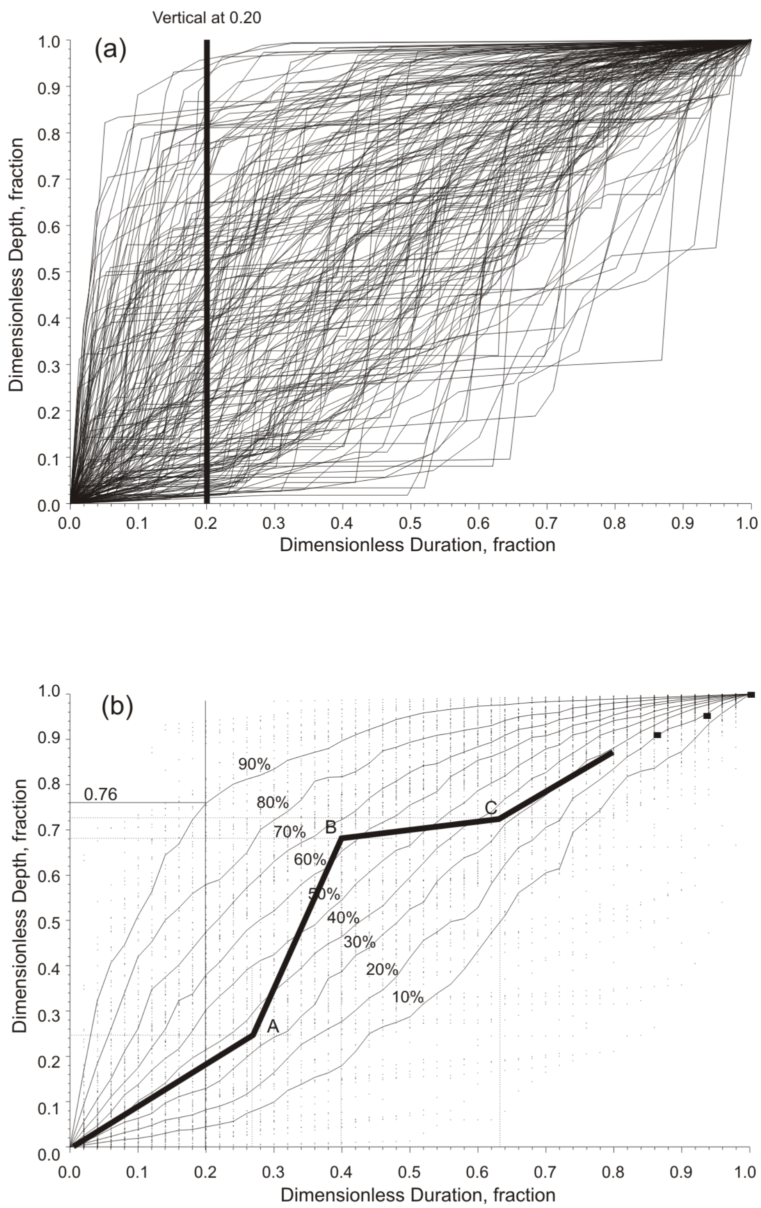

Figure 5. Huff curves for RG100 at Coshocton for May and June: (a) dimensionless mass curves for 182 storms, and (b) dimensionless mass−curve intersection points from figure 5a, isopleths of probability of dimensionless depths, and schematic of stochastic simulation of within−storm intensities.

Huff Curves

Bonta (2004) presents detail on the construction of Huff curves. From the storms identified in the precipitation record using the monthly MDPD values, storm mass curves are developed and nondimensionalized (fig. 5a). Order is made of the apparent disorder in figure 5a following a two−step process to add probability to the dimensionless mass curves. For a given vertical line representing a single dimensionless duration (e.g., dimensionless duration−axis value of 0.2 in fig. 5a), the intersection of each dimensionless mass curve with the vertical is interpolated ("interpolated mass−curve intersection"). Interpolated mass−curve intersections are computed for every vertical selected by an analyst (fig. 5b; the interpolated mass−curve intersections for dimensionless depths are found at verticals in increments of 0.02 along the dimensionless−duration axis). Deciles of the cumulative probability distribution of dimensionless depths at each vertical are found, and isopleths constructed. For example, 90% of dimensionless mass−curve intersections are less than 0.76 at 20% of elapsed time into storms (fig. 5b). These isopleths are Huff curves, and they summarize the probabilistic nature of storm intensity in terms of dimensionless accumulated depths for dimensionless elapsed storm times. A 90% probability curve (fig. 5b) may be interpreted as the precipitation events that occur in less than 90% of the storms for all durations. The probabilistic information contained in them can be used for stochastic generation of storm mass curves.

Factors affecting Huff curve development have been studied by Bonta and Rao (1987) and Bonta (2004). These studies have shown that Huff curves are robust with respect to MDPD and sampling interval of the data (i.e., 3 and 60 min data), both important features for practical application. Using available data, Bonta and Rao (1989) also showed the potential for regionalization of Huff curves. Bonta (2004) showed that the curves are stable across large distances (660 km) and cited other studies that suggest these curves can be regionalized over large areas. Bonta and Shahalam (1998, 2003) presented methods to compare Huff curves for developing stable sets of curves (i.e., minimum number of storms required for stability). Because of these supporting studies, the robustness of these curves, the simple probabilistic representation of precipitation intensities, and the potential for regionalization, Huff curves are proposed to stochastically synthesize within−storm intensities.

Storm Intensity Simulation

Dimensionless storm mass curves shown in figure 5b (Huff curves) from the data−derived dimensionless storm mass curves in figure 5a suggest a reverse procedure to stochastically generate storm mass curves. At a vertical selected near the beginning of a dimensionless storm (e.g., vertical at 0.20), the empirical cumulative distribution of dimensionless storm depths is sampled (e.g., point A = 0.25 in fig. 5b). At the next selected dimensionless storm duration (e.g., vertical at 0.40), the corresponding empirical frequency distribution of dimensionless depth is sampled (e.g., point B = 0.68). The vertical for any selected dimensionless duration can be doubly interpolated such as at 0.631 (fig. 5b). The first interpolation is between verticals for which the distributions of dimensionless mass−curve depth intersections have been developed (e.g., distributions of dimensionless depths between duration increments of 0.62 and 0.64; fig. 5b). The second interpolation is along the interpolated vertical for the random dimensionless depth (e.g., point C = 0.725 on the interpolated vertical at 0.631). The procedure is repeated until the dimensionless mass curve reaches the coordinates (1.0, 1.0). The intensity distributions within storms (e.g., storms described by Dur_1−Dep_1 and Dur_2−Dep_2 in fig. 1b) are synthesized by using this method.

Because a dimensionless depth must always remain the same or increase as a mass curve is formed, a simulated dimensionless depth must be greater than or equal to the previous depth. It can be equal because the exponential method of identifying storms includes dry periods less than MDPD (e.g., D_1 and D_3 in fig. 1a). While simulating using Huff curves incorporates serial correlation, initial testing suggests that Huff curves require additional information on serial correlation, and this is the subject of current research. Methodologies utilizing Huff curves that are being investigated include simultaneous simulation of time and depths for within−storm intensities, weighted averaging for adjacent depths, and rank correlation between adjacent depths that represent correlated intensities.

The result of the above procedure is a mass curve in dimensionless form. However, depth and time units are desired for practical applications. The procedure to assign units for a mass curve is to multiply each mass−curve point by the stochastically generated depth−duration pair as described above. The number of points comprising a mass curve is dependent on the time and depth resolution requested by the user. For example, storm−duration points cannot be modeled finer than a 5 min resolution, and storm depths cannot be less then 0.25 mm if these resolutions are initially specified. Bonta and Rao (1992) used this method of determining units for isopleths to estimate peak−flow rates and their return periods. Huff curves in StormGen can be empirical distributions for each vertical, fitted distributions for each vertical, or equations for individual curves (Bonta and Rao, 1988b, 1992).

Initial Testing of StormGen

Analysis

The StormGen model was evaluated by comparing measured and simulated frequency distributions for TBS, storm depth, and storm duration. Furthermore, precipitation totals for each month and for the period of simulation (200 years) were compared with corresponding measured values. These comparisons allow the determination of the proper functioning of the model, they can highlight areas for improvement in model characterization and simulation, and they can document the aggregated performance of all elements of the model (long−term total precipitation). Frequency distributions of simulated TBS, storm durations, and storm depths were compared with corresponding measured frequency distributions using the Kolmogorov−Smirnov (KS) test and significance probabilities, and by comparing mean monthly values, differences between measured and simulated values, and percent deviations. Measured monthly storm−precipitation totals were used because simulated storms can span months, and there is no midnight simulation at the end of the month when within−storm intensities are not simulated. Consequently, monthly totals are approximate because they can span a month. Table 1

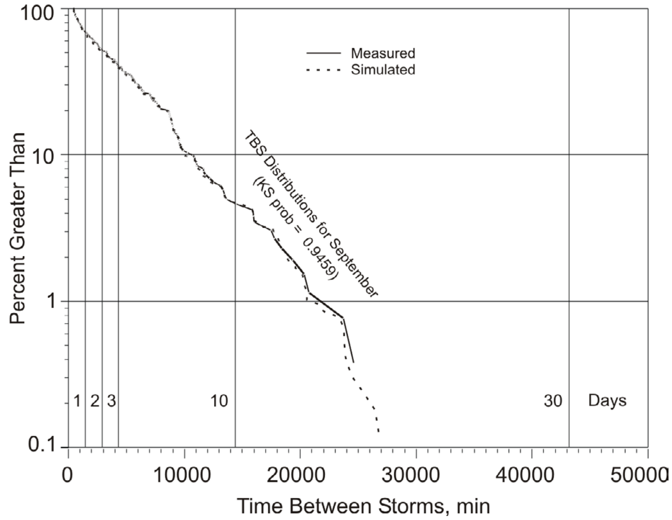

Figure 6. Comparisons of simulated with measured frequency distributions for RG100 at Coshocton, Ohio, for the worst comparison of time between storms in table 2.

shows that average monthly total precipitation at the NAEW computed using storm totals is about 1.05 times larger than average monthly totals computed using a recording rain gauge for which midnight values are available. Monthly and period total precipitation data are collapsed across the entire 200−year period of simulation. Graphs of the modeled and measured frequency distributions with the smallest KS significance probability ("worst" simulations) were superimposed.

Comparison of Frequency Distributions

Storm occurrence distributions are modeled exceptionally well (fig. 6), with large KS significance probabilities greater than 0.9459 (table 2; Sept.), not considering October. The simulated October distribution, while having a large probability of 0.4489, is not representative of the corresponding measured distribution because dry periods beyond the end of October are not recognized. This is because the model only retains information on dry periods that are followed by a simulated storm. The simulation period ends in October; consequently, no storms are simulated after TBS values extending beyond the end of October, truncating the TBS distribution at shorter values. Average monthly measured and simulated TBS values agree well (table 2). Percent monthly deviations ranged from −3.2% to +3.0%, excluding October, with an average of +0.4%. The good simulation of TBS is partly attributed to the characterization of dry times between storms that spans months and makes StormGen a useful tool for drought studies.

Storm duration is also modeled well (table 2), with KS probabilities greater than 0.5412 (April). April had the lowest KS significance probability; however, the average measured and simulated storm durations are nearly identical (627 min vs. 617 min, respectively; fig. 7). Percent monthly deviations ranged from −6.2% to +1.6%, with most months between −2.7% and +1.6% (average of −1.4%).

Average storm depths are least well simulated of the three storm elements modeled (table 3); however, the actual differences are small. KS probabilities for comparisons of frequency distributions ranged from 0.1262 in July to 0.9492 in June. Figure 7 (July) shows the typical trend found of oversimulation of depths for smaller storms and undersimulation for larger storms. Generally, an average monthly storm depth for simulated storms was less than for measured storms (table 3). The deviations between average measured and simulated monthly storm depth ranged from about 0 mm (April) to a −2.2 mm undersimulation (August). The corresponding percent deviations were 0.1 % and −17.1%. The larger percent deviations are due in part to the small actual deviation as a fraction of a relatively small number.

Table 2. Comparison between measured (Coshocton, Ohio) and simulated times between storms and storm durations.

	Time Between Storms (min)				Storm Duration (min)			
Month	Average Measured	Average Simulated	Percent Deviation (Sim−Avg)	KS Significance Probability	Average Measured	Average Simulated	Percent Deviation (Sim−Avg)	KS Significance Probability
April	3946	4021	1.9	1.0000	627	617	−1.6	0.5412
May	4286	4414	3.0	0.9880	753	706	−6.2	0.9182
June	4259	4190	−1.6	0.9970	414	419	1.3	1.0000
July	3967	3952	−0.4	1.0000	299	298	−0.3	0.7084
August	5835	5986	2.6	0.9998	434	422	−2.7	0.9995
September	4809	4655	−3.2	0.9459	373	366	−1.7	0.9978
October	4651	4111[a]	−11.6	0.4489	410	417	1.6	0.9869
Average			0.4[b]				−1.4	

[a] Nonrepresentative TBS, percent deviation, and KS probabilities because simulated dry periods extending beyond the end of October are not accounted for.
[b] Average does not include October.

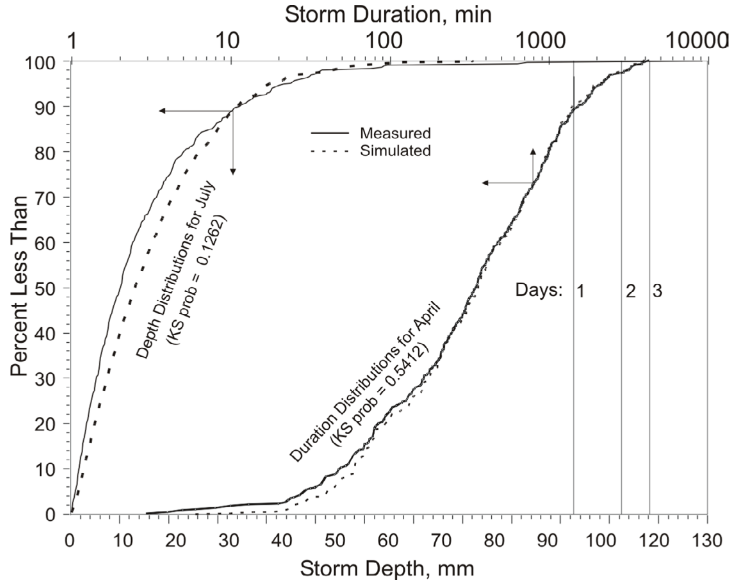

Figure 7. Comparisons of simulated with measured frequency distributions for RG100 at Coshocton, Ohio, for worst comparisons in tables 2 and 3 for storm durations and depths.

Table 3. Comparison of measured (Coshocton, Ohio) and simulated monthly storm depths.

Month	Average Measured (mm)	Average Simulated (mm)	Difference (Sim−Avg) (mm)	Percent Deviation (Sim−Avg)	KS Significance Probability
April	10.4	10.4	0.0[a]	0.1[a]	0.2783
May	12.3	11.4	−0.9	−7.3	0.8200
June	12.3	11.4	−0.9	−7.4	0.9492
July	12.0	10.2	−1.8	−15.0	0.1262
August	12.9	10.7	−2.2	−17.1	0.3004
September	10.7	9.8	−0.8	−7.8	0.8172
October	8.1	7.3	−0.8	−10.2	0.7759
Average			−1.1	−9.2	

[a] Round−off error explains the 0.1% deviation compared with the 0.0 mm deviation.

Table 4. Comparison of measured (Coshocton, Ohio) and simulated monthly and period total precipitation.

Month	Average Measured (mm)	Average Simulated (mm)	Difference (Sim−Avg) (mm)	Percent Deviation (Sim−Avg)
April	86.0	96.7	10.8	12.6
May	102.1	98.8	−3.3	−3.3
June	107.7	107.6	−0.1	−0.1
July	115.8	105.9	−10.0	−8.6
August	83.6	78.5	−5.1	−6.1
September	77.1	82.1	5.1	6.6
October	58.0	64.2	6.2	10.6
Average			0.5	1.7
Total Period Values (April−Oct.)	630.3	633.9	3.5	0.6

The larger disparity between measured and simulated storm depths compared with TBS and durations is attributed to three factors: (1) storm depths may be sensitive to storm−duration simulation (conditionally simulated as in fig. 4); (2) the independent−region method for characterizing storm depths may not be adequate for storm simulation, and other methods require investigation; and (3) the method of interpolation between adjacent independent depth distributions requires improvement. The last two factors are more likely to cause the deviations than the first factor because duration distributions are well simulated (table 2; fig. 7).

Monthly and Period Totals

Deviations between monthly total precipitation values were small, ranging from −2.2 to 0.0 mm (−17.1% to 0.1%, respectively; table 4). The average deviation for all months was −1.1 mm, corresponding to −9.2% deviation. Average period simulated total precipitation (April through October) was only +3.5 mm greater than measured storm totals. This corresponds to a 0.6% error in total precipitation for the 200−year simulation.

Discussion

The model results are generally favorable and promising in terms of simulated monthly TBS and storm depth and duration distributions, and in terms of monthly and period total precipitation depths. Storm depths were least well modeled, which is attributed principally to the method of interpolating between storm−depth distributions for a given month, and to the method of characterizing conditional storm depth distributions given storm duration. Larger average storm depth percent deviations have actual storm depth deviations of only −1.1 mm. Storm depths tend to be oversimulated at smaller depths and undersimulated at larger depths. The method of evaluation highlights the weaknesses in the depth element of the model.

With development and robustness of characterization methods documented in the literature, the StormGen model appears to be promising for practical application. However, further testing of each storm element in different climates and a comparison of simulated and measured results using independent rain data are required.

StormGen Model Inputs and Outputs (Version 2004.01)

The current version of the storm generator incorporates several features that allow expression of different forms of input and different outputs. Inputs include monthly MDPD estimates, frequency distributions of storm depths and durations, and Huff curves. The MDPD can be provided using assigned values and by using regression relationships between TBS and MDPD and monthly precipitation (Bonta, 2003). TBS and storm depth and duration frequency distributions can use empirical data directly or use parameters for a variety of fitted distributions. Output options are daily totals, 24 h totals at a user−specified time of day, or user−specified constant time intervals (e.g., 5 min, hourly, etc.). Furthermore, outputs can be only storm occurrences, durations, and depths (no within−storm intensities as in the present study). The general output is a month, day, year (starting at year 0), time of day to the minute, and accumulated precipitation. Special output formats are available for specific watershed models.

Research Needs

The current version of StormGen described in this article simulates precipitation at a point, and further research is needed to incorporate coherent spatially distributed storm precipitation. Further research is also needed to include persistence of droughts and wet periods, to simulate snowfall, to associate model parameters with outputs from global circulation models, and to improve the storm−depth characterization and simulation. The model can be modified to include these weather factors by developing conditional distributions (e.g., analyze observed TBS values within lagged classes of the Southern Oscillation Index that forecasts persistence in TBS). Comprehensive application and testing of StormGen for different simulated elements in other climates is the subject of other articles. Coupling the storm generator with a weather generator (e.g., GEM) would enable the storm generator to obtain information from the air temperature component to simulate snowfall. More research is needed to develop general characterization schemes for developing the distributions needed to model ungauged areas. Use of the storm model with watershed models is also a research need.

Conclusions

The concepts for modeling and characterizing point storm precipitation as incorporated into the StormGen storm−simulation model are presented (four elements of storm occurrence, storm durations and depths, and within−storm intensities). The model avoids the common assumption of daily or 24 h total precipitation and synthesizes storms directly (including several storms per day to several days per storm). The removal of the 24 h constraint eliminates the need to artificially synthesize within−storm intensities by using design storms or other patterns with limited flexibility and nonrepresentativeness. The storm generator synthesizes the month, day, year, hour, and minute of the beginning and end of individual storms (storm occurrence and duration), storm depth, and the within−storm intensities to the nearest time resolution (e.g., 1 min, 5 min, 24 h, etc.) and depth resolution specified (e.g., 0.25 mm, 2.5 mm, etc.). Characterization and simulation of the four storm elements is illustrated by using 61 years of rain gauge data from Coshocton, Ohio. Characterization studies and model development have occurred simultaneously to maximize the practicality of the model in ungauged areas. An initial performance evaluation of the storm generator was conducted for a period of 200 simulated years. The following conclusions can be made from this study:

- Dry times between storms (TBS) are characterized well by the exponential method of separating storms, yielding the minimum dry period duration (MDPD) and the exponential distribution of times between storms. MDPD is successfully used to identify storms to form a database for subsequent characterization.
- Storm duration is characterized well by monthly frequency distributions of storm duration formed by using MDPD to identify storms.
- Conditional simulation of storm depths on storm duration is required to account for the statistically significant correlation between these two storm elements.
- Determination of regions of independent storm durations and depths is adequate for characterizing and simulating storm depths, but improvement in characterizing the dependence of storm depth on storm duration is warranted.
- A power equation for storm depths (Dep) and storm duration (Dur) characterizes the correlation between these two variables for the range of small to large storm depths (Dep = $a\mathrm{Dur}^b$), unlike in other studies in which a linear equation was used (Dep = $c + d\mathrm{Dur}$).
- Monthly characterization of short−time increment precipitation data incorporates observed wide time−of−year variability of distributions into the storm simulation model.
- The Poisson−process assumption that characterizes statistically independent storm events is validated using Coshocton data: average storm duration was on average only 0.107 of average TBS (ranging from 0.074 to 0.176).
- Successful simulations of the frequency distributions of the three storm elements were obtained in the following order (starting with the best simulation): TBS, storm durations, and storm depths. Overall, the model performed well.
- Percent deviations between the simulated and measured distributions of the three storm elements are: For monthly TBS distributions: average of 0.4%, with a minimum and maximum of −3.2% and 3.0%, respectively (excluding October). For monthly duration distributions: average of −1.4%, with a minimum and maximum of −6.2% and 1.6%, respectively. For monthly depth distributions: average of −9.2%, with a minimum and maximum of −17.1% and +0.1%, respectively. Corresponding absolute monthly depth deviations ranged from −2.2 mm to 0 mm.
- Deviations between the simulated and measured total monthly and period precipitation are: For monthly totals: average of +1.7%, with a minimum and maximum of −8.6 and +12.6%, respectively. Corresponding depths range from −10.0 mm to 10.8 mm, with an average of +0.5 mm. For period totals: +0.6%, corresponding to +0.5 mm.

The storm−simulation model can be used for a variety of purposes (especially if coupled with a weather−generator model such as GEM) including risk analysis for studies of the spread of crop and animal pests and diseases, erosion and water−quality modeling, wind erosion, drought, groundwater recharge, climate change, soil−moisture depletion, evaporation and drainage between storms, leaf wetness, and natural resource applications. Furthermore, the storm generator can aid in other studies such as intensity−duration−frequency analysis, flood analysis forecasting, heating and air conditioning engineering, erosivity estimates, international agricultural trade, and engineering design. A major advantage of the storm generator is its ability to generate long records of storms in areas where there are few or no data. The results of this study are encouraging, and further research should help achieve the potential uses of the model.

References

Bonta, J. V. 1997. Proposed use of Huff curves for hyetograph characterization. In *Proc. Workshop on Climate and Weather Research*, 111−124. C. W. Richardson et al., eds. Denver, Colo.: USDA Agricultural Research Service.

Bonta, J. V. 1998. Generation of storm occurrence, depth, and duration. ASAE Paper No. 982023. St. Joseph, Mich.: ASAE.

Bonta, J. V. 2001. Characterizing and estimating spatial and temporal variability of times between storms. *Trans. ASAE* 44(6): 1593−1601.

Bonta, J. V. 2003. Estimation of parameters characterizing frequency distributions of times between storms. *Trans. ASAE* 46(2): 331−343.

Bonta, J. V. 2004. Development of Huff curves for design storms and short−time increment precipitation analysis. *Trans. ASAE* (accepted for publication).

Bonta, J. V., and A. R. Rao. 1987. Factors affecting development of Huff curves. *Trans. ASAE* 30(6): 1689−1693.

Bonta, J. V., and A. R. Rao. 1988a. Factors affecting the identification of independent storm events. *J. Hydrology* 98(3−4): 275?293.

Bonta, J. V., and A. R. Rao. 1988b. Fitting equations to families of dimensionless cumulative hyetographs. *Trans. ASAE* 31(3): 756?760.

Bonta, J. V., and A. R. Rao. 1989. Regionalization of storm hyetographs. *Water Resources Bulletin* 25(1): 211?217.

Bonta, J. V., and A. R. Rao. 1992. Estimating peak flows from small agricultural watersheds. *J. Irrigation and Drainage Eng.* 118(1): 122−137.

Bonta, J. V., and A. Shahalam. 1998. Investigation of techniques to compare Huff curves. ASAE Paper No. 982192. St. Joseph, Mich.: ASAE.

Bonta, J. V., and A. Shahalam. 2003. Cumulative storm rainfall distributions: Comparison of Huff curves. *J. Hydrology (New Zealand)* 42(1): 65−74.

Bras, R. I., and I. Rodriguez−Iturbe. 1976. Rainfall generation: A nonstationary time−varying multidimensional model. *Water Resources Research* 12(3): 450−456.

Corotis, R. B. 1976. Stochastic considerations in thunderstorm modeling. *J. Hydraul. Div. ASCE* 102(HY7): 865−879.

Cox, D. R., and V. S. Isham. 1998. Stochastic spatial−temporal models for rain. In *Stochastic Methods in Hydrology*, 1−24. O. E. Barnoff−Nielsen, V. K. Gupta, V. Perez−Abreu, and E. Waymire, eds. New York, N.Y.: World Scientific.

Crovelli, R. A. 1971. Stochastic models for precipitation. PhD diss. Fort Collins, Colo.: Colorado State University.

Grace, R. A., and P. S. Eagleson. 1966. The synthesis of short−time increment rainfall sequences. Report No. 91. Cambridge, Mass.: Massachusetts Institute of Technology, School of Civil Engineering, Hydrodynamics Laboratory.

Hanson, C. L., G. L. Johnson, and W. L. Frymire. 2002. The GEM (generation of weather elements for multiple applications) weather simulation model. In *Preprints, 13th AMS Conf. Appl. Climatology*, 117−121. Boston, Mass.: American Meteorological Society.

Hanson, C. L., H. B. Osborn, and D. A. Woolhiser. 1989. Daily precipitation simulation model for mountainous areas. *Trans. ASAE* 32(3): 865−873.

Hanson, C. L., K. A. Cumming, D. A. Woolhiser, and C. W. Richardson. 1994. Microcomputer program for daily weather simulation. Publ. No. ARS−114. Washington, D.C.: USDA−ARS.

Huff, F. A. 1967. Time distribution of rainfall in heavy storms. *Water Resources Research* 3(4): 1007−1019.

Knisel, W. G., Jr., and W. M. Snyder. 1975. Stochastic time distribution of storm rainfall. *Nordic Hydrology* 6: 242−262.

Marien, J. L., and G. L. Vandewiele. 1986. A point rainfall generator with internal storm structure. *Water Resources Research* 22(4): 475−482.

Morris, C. D. 1980. A stochastic precipitation simulator. In *International Symposium on Urban Storm Runoff*, 49−58. Lexington, Ky.: University of Kentucky.

Morrison, J. E., and J. A. Smith. 2001. Scaling properties of flood peaks. *Extremes* 4(1): 5−22.

Nicks, A. D., L. J. Lane, and G. A. Gander. 1995. Chapter 2: Weather generator. In *Hillslope Profile and Watershed Model Documentation*. D. C. Flanagan and M. A. Nearing, eds. NSERL Report No. 10. West Lafayette, Ind.: USDA−ARS National Soil Erosion Research Laboratory.

Nguyen, T.−T.−V., and J. Rousselle. 1981. A stochastic model for the time distribution of hourly rainfall depth. *Water Resources Research* 17(2): 399−409.

Over, T. M., and V. K. Gupta. 1996. A space−time theory of mesoscale rainfall using random cascades. *J. Geophysical Res.* 101(D21): 26319−26331.

Pattison, A. 1965. Synthesis of hourly rainfall data. *Water Resources Research* 1(4): 489−498.

Rao, R. A., and B. T. Chenchayya. 1974. Probabilistic analysis and simulation of the short time increment rainfall process. Technical Report No. 55. West Lafayette, Ind.: Purdue University, Water Resources Research Center.

Raudkivi, A. J., and N. Lawgun. 1974. Simulation of rainfall sequences. *J. Hydrol.* 22(3−4): 271−294.

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

Todorovic, P., and D. Woolhiser. 1974. Stochastic model of daily rainfall. In *Proc. Symposium on Statistical Hydrology*, 232−246. Miscellaneous Pub. No. 1275. Washington, D.C.: USDA−ARS.

Valencia, D. R., and J. C. Schaake, Jr. 1973. Disaggregation processes in stochastic hydrology. *Water Resources Research* 9(3): 580−585.

Woolhiser, D. A., and H. B. Osborn. 1985. A stochastic model of dimensionless thunderstorm rainfall. *Water Resources Research* 21(4): 511−522.

Yu, B. 2000. Improvement and evaluation of CLIGEN for storm generation. *Trans. ASAE* 43(2): 301−307.