1.0 OVERVIEW

This Technical Memorandum (TM) provides a brief summary of the algorithm that was created for development of watershed precipitation-frequency (PF) relationships and uncertainty bounds for the Mesoscale Storm with Embedded Convection (MEC) storm type (Figure 1). These are convective storms characterized by very high intensity precipitation including complex clusters of convective cells which are often supported by low to moderate intensity precipitation in areas surrounding the convective cells. This includes storms produced by Mesoscale Convective Complexes (MCC) and other convective storms commonly referred to as thunderstorms. These are storms that can produce flash floods on small to intermediate size watersheds generally less than about 1,000-mi$^2$.

The 6-hour duration is typically selected as the key duration which is representative of the time period when the majority of precipitation occurs from these mesoscale storm types. The algorithm utilizes stochastic generation methods for simulation of multiple storms per sample year from which the annual maximum watershed precipitation is produced.

This algorithm employs findings from regional point PF analyses, precipitation data from historical MEC storms observed on the watershed and large MEC storms observed in the region that could have occurred on/near the watershed of interest. The algorithm has numerous components which are discussed in the following sections. The analyses and findings for the 467-mi$^2$ watershed above
Watauga Dam on the Watauga River (Figures 1 and 3) are used as a case study to provide context for description of the stochastic storm generation algorithm.

2.0 FRAMEWORK FOR STOCHASTIC STORM GENERATION FOR A WATERSHED

The concept behind the Move-The-Earth (MTE) algorithm is to simulate the natural meteorological processes where many MEC events occur in a given year that produce precipitation at a given location in the TVA study area. The location where heavy precipitation occurs from an MEC event is random relative to the location of the precipitation stations in a network of stations in a given area. The precipitation annual maxima at a given station can be produced by a wide range of situations based on the storm magnitude, the nearness of the storm/storm center to the station and the spatial pattern of precipitation. This would include situations such as a small MEC event occurring nearby, or a larger MEC event occurring more distant from the station.

This spatially random storm behavior is simulated using precipitation fields (raster fields) obtained from spatial analyses of historical MEC storms (MetStorm17) where the precipitation fields are used for describing the spatial distribution of precipitation within and near the watershed of interest. Figure 2a depicts a precipitation field with a complex pattern of convective storm activity with several separate storm centers. The rectangular area encompassing a specific storm center (Figure 2a) has been given the name storm center zone. These separate storm zones are spatially distinct and there are likely to be timing differences for the 6-hour maxima precipitation between the storm center zones as well. Figure 2b depicts an example of a specific storm center zone that can be used for stochastic generation of precipitation on a user-specified watershed where only one storm center is displayed.

The position of the precipitation field relative to a watershed is randomly selected in a manner consistent with the spatial behavior of historical storms. The procedure is called “moving-the-earth” because the watershed raster field is randomly placed relative to the precipitation raster field rather than vice-versa as storms occur naturally. This is done for computational convenience because the watershed raster field is much smaller in size than the precipitation raster field. This approach also results in a much larger sample set of storms that can occur “over” the watershed by comparison to the data set of storms that actually occurred on the watershed of interest.

A virtual key station is positioned at the watershed centroid and is used for random placement of the watershed relative to a precipitation field. The watershed centroid is a logical choice for the spatial reference point because the subject of interest is areal-average precipitation for the watershed which is best represented by the centrally-located watershed centroid. Figure 3 depicts the network of four stations and a virtual key station at the watershed centroid used for the watershed above Watauga Dam on the Watauga River. The virtual key station is used for generating precipitation annual maxima at the watershed centroid and for scaling of the precipitation field as part of the stochastic storm generation process (Figure 4). Stochastically generated precipitation annual maxima at surrounding substations are later used for verification of the storm generating process. The areal-average watershed precipitation is computed from simple arithmetic accounting of the precipitation at each grid-cell in the watershed raster field.
The procedure for random placement of the storm relative to the watershed is to set limits on the location of the virtual key station relative to the storm center (maximum point precipitation). This is accomplished by setting a lower bound on the ratio of the precipitation at the virtual key station relative to the maximum point precipitation at the storm center. This procedure is discussed further in Section 5.2.

Figure 2a – Precipitation Field for MEC Storm of August 31, 1982 for the Tennessee Valley Study Area Showing Three Separate Storm Center Zones and Spatial Precipitation Patterns for the 6-Hour Duration

Figure 2b – Storm Center Zone for Portion of 6-Hour Precipitation Field for MEC Storm of August 31, 1982
2.1 Data Sources
A variety of precipitation data are required for use in conducting a watershed precipitation analyses for the MEC storm type which include:

1. Precipitation annual maxima data series for each station in the study area domain for the MEC storm type for the key duration (6-hours)
2. A network of precipitation stations location within and/or very near the watershed of interest that are sufficient in number and spacing to be representative of the spatial distribution of convective precipitation
3. Concurrent precipitation for the key duration for each station in the station network for each historical storm that occurred on the watershed or occurred within the study area
4. Radar reflectivity data for noteworthy MEC storms that occurred in the study domain after radar time series data became available in the mid-1990s
5. Spatial and temporal storm analyses (MetStorm17) for a large sample set of MEC storms observed within the study domain

2.2 Supporting Analyses – Prior Studies
Findings from several point and spatial precipitation analyses for the MEC storm type are needed to support the methods used for stochastic storm generation. These analyses are conducted to support a variety of applications and are typically conducted prior to analyses needed specifically for the move-the-earth algorithm. These analyses include:
1. Findings from regional point precipitation-frequency analyses (Hosking and Wallis\textsuperscript{11}, Schaefer\textsuperscript{23,25,26,27}, MGS et al\textsuperscript{18}, L-RAP\textsuperscript{13}) which includes spatial mapping of the at-site means, regional L-Cv, regional L-Skewness and identification of the regional probability distribution

2. Point PF relationship for 6-hour precipitation annual maxima series (AMS) data applicable to the virtual key station and point PF relationships for all other stations in the station network

3. Point PF relationship for 6-hour precipitation applicable to the virtual key station based on a peaks-over-threshold (POT) analysis for data from storms that occur many times in a given year

2.3 Supporting Analyses – Studies Specific to Move-the-Earth Algorithm
The following analyses are conducted specifically to support the move-the-earth algorithm.

1. Findings of spatial storm analyses (MetStorm\textsuperscript{17}) for 6-hour areal-average watershed precipitation resulting from MEC storms that occurred on the watershed

2. Probability-plot of historical 6-hour watershed-average precipitation annual maxima based on findings of spatial analyses of historical storms

3. Precipitation fields (raster field) obtained from spatial storm analyses (MetStorm\textsuperscript{17}) of major storms that have occurred in the climatic region where the watershed is located

4. Areal mask of watershed and subbasins (watershed raster fields) for use in moving-the-earth placement of the watershed of interest relative to the precipitation fields

5. Findings from spatial storm analyses with and without radar data to provide a bias correction for spatial analyses of storms that pre-dated radar coverage. The bias correction is needed because storm centers typically do not occur directly over stations in the station network

2.4 Stochastic Storm Generation Methods
Several probabilistic concepts and Monte Carlo sampling methods are used in the development of a watershed PF relationship and uncertainty bounds using the MTE stochastic storm generation approach. These include:

1. Application of the Total Probability Theorem using 6-hour precipitation at the virtual key station as the independent variable (Benjamin and Cornell\textsuperscript{1}, Kuczera and Nathan et al\textsuperscript{19})

2. Stratified sampling of 6-hour precipitation annual maxima series data (AMS) from the virtual key station

3. Generation of multiple storms per year with 6-hour precipitation less than the annual maxima at the virtual key station using standard Monte Carlo methods and the point PF relationship from the POT analysis

4. Latin-hypercube sampling (McKay et al\textsuperscript{16}) of L-moment parameters for the virtual key station to account for epistemic uncertainties in the at-site mean, regional L-Cv, regional L-Skewness and the regional probability distribution

5. Resampling (Efron\textsuperscript{3}) to produce a subset of possible spatial precipitation patterns for storm center zones selected from the full sample of storm center zones
6. Characterization of uncertainties in alternative sampling limits for placement of the virtual key station (watershed) relative to the location of the storm center

Figure 4 – Flowchart for Development of a Mean 6-Hour Watershed Precipitation-Frequency Relationship and Uncertainty Bounds for Mesoscale Storms with Embedded Convection (MEC)
2.5 Stochastic Storm Generation Flowchart
The sequence of tasks for stochastic storm generation are summarized in the Flowchart depicted in Figure 4. The inner (blue) loop is used for computing a watershed PF relationship for one plausible scenario of inputs and probabilistic model parameters and typically involves simulation of 6,000 sample years of storms with a range of 1 to 20 storms per sample year (see Section 7 for details). The outer loop is used for computing alternative plausible watershed PF relationships (simulation scenarios) that account for epistemic uncertainties through different scenarios of plausible inputs and probabilistic model parameters. The outer red loop typically involves about 200 scenarios to capture the effects of uncertainties which provides for computation of a mean 6-hour watershed PF relationship and 90% uncertainty bounds. This approach typically results in stochastic generation of about 12 million separate storms with 10 storms on-average generated each sample year.

3.0 POINT PRECIPITATION-FREQUENCY RELATIONSHIPS FOR THE VIRTUAL KEY STATION
As discussed previously, a virtual key station is located at the centroid of the watershed of interest to provide a representative indicator of mesoscale precipitation that likely extends over most, or all, of the watershed. The areal extent of the storm over the watershed depends upon the relative size of the watershed and spatial extent of the MEC storm. Figure 5 depicts the 6-hour point PF relationship applicable to the virtual key station for the Watauga watershed based on regional PF analyses conducted for the Tennessee Valley Study Area (MGS et al18). Figure 6 depicts the at-site Peaks-Over-Threshold (POT) solution for the virtual key station. The POT relationship was developed using the Langbein14 equation for conversion of findings from annual maxima series to POT data series. The 0.30-inch threshold for the POT analyses is associated with a storm magnitude that is exceeded about 10 times per year.

The point PF relationship in Figure 5 is used for generation of precipitation annual maxima at the virtual key station (2nd box in Flowchart, Figure 4). The POT point PF relationship in Figure 6 is used for generation of addition storms per sample year at the virtual key station which are smaller in magnitude than the precipitation annual maxima generated from the PF relationship shown in Figure 5.
4.0 SPATIAL ANALYSES OF HISTORICAL WATERSHED-AVERAGE PRECIPITATION
Spatial storm analyses (MetStorm17) are conducted for historical storms that occurred on the watershed to develop an annual maxima series dataset for 6-hour precipitation for the MEC storm type. The findings of this analysis are used to calibrate the outputs from the stochastic storm generation algorithm to reasonably replicate the PF behavior of historical storms. Figure 7 depicts a probability-plot of 6-hour areal-average watershed precipitation annual maxima for the watershed above Watauga Dam as compared to the areal-average point PF relationship for the watershed. The areal-average point precipitation PF relationship is obtained using the areal-average of the gridded at-site means, areal-average of the gridded regional values of L-Cv and L-Skewness, and the regional probability distribution. The differences between the areal-average point and watershed PF curves can be viewed in the context of a conventional areal reduction factor (ARF).
5.0 SPATIAL DISTRIBUTION OF PRECIPITATION

The natural variability in the spatial distribution of precipitation from convective storms is preserved through a resampling approach which uses historical precipitation fields (Figure 2) for random placement of the watershed of interest near the storm center. For the case of the Watauga watershed, spatial storm analyses were conducted using MetStorm17 software which produced 118 storm center zones based on MEC storm events that occurred on 27 separate storm dates. The majority of spatial analyses were specifically conducted for MEC storms where radar data were available; older (pre-radar) storms were then analyzed for larger MEC storms where a maximum point rainfall of 4.0-inches or greater had occurred in 6-hours.

Figure 8a depicts the rank order of maximum 6-hour point precipitation at the storm center for each of the 118 storm center zones. This dataset of 118 storm center zones is termed the Master Storm Center Dataset. Bootstrap resampling methods (Efron3) are used to randomly assemble a subset of storm center zones for use in stochastic storm generation which allows consideration of the representativeness of the sample set and supports the uncertainty analysis (Section 11.2). Figure 8b depicts an example of a dataset of 80 randomly selected storm center zones for one of the 200 scenarios that are used in the uncertainty analysis for developing watershed PF relationships.
5.1 Stratified Sampling of Spatial Precipitation Patterns for Storm Center Zones

Spatial storm patterns for convective precipitation are quite complex and there is high diversity amongst the spatial patterns. It is reasonable to expect that there may be some non-linearity of spatial patterns with storm magnitude, particularly for extreme storm magnitudes. This potential for non-linearity is addressed by using a stratified sampling approach in stochastic storm generation. Specifically, spatial patterns with smaller maximum point precipitation are used for stochastic storm generation for smaller values of precipitation at the virtual key station and spatial patterns with the largest maximum point precipitation are used for stochastic storm generation for the largest values of precipitation at the virtual key station. This is accomplished by selecting a subset of spatial patterns from the full sample set of spatial patterns for a given magnitude of point precipitation at the virtual key station. For the case of the Holston River system, this can be envisioned as a sliding window of 25 spatial patterns from a sample set of 80 spatial patterns (Figure 8b), where the individual 25 storm center zones selected varies with the magnitude of point precipitation at the virtual key station.

5.2 Constraints on Random Placement of Spatial Precipitation Patterns for Watersheds

The nearness of the random placement of the watershed (virtual key station at watershed centroid) relative to the spatial precipitation pattern and associated storm center (grid-cell of maximum point precipitation) must be done in a manner consistent with the behavior of historical storms and avoid over-scaling per Section 5.3. The basic concept is clear - a number of convective storms occur each year at random locations in the vicinity of any specified location. One of those storms has some combination of maximum point precipitation, spatial distribution of precipitation and nearness to the specified location to produce the 6-hour annual maximum at the specific location.

The critical criterion is how “near” must a convective storm occur to a given precipitation station to produce an annual maximum at that station? The answer is not straightforward as the answer relates to the magnitude of the storm center, the magnitude of precipitation gradients away from the storm center, or multiple storm centers, the shape of the spatial precipitation pattern, and the number of convective storm events in a given year.

Information on historical storm behavior is needed to set constraints on the random placement of the search area where the virtual key station can be placed relative to the storm spatial patterns. An analysis of 30 convective storms was conducted for the watershed annual maxima at each of the Watauga and South Holston watersheds. A measure of “nearness” was computed as the ratio of the maximum 6-hour precipitation at a station near the watershed centroid divided by the maximum 6-hour point precipitation for the storm center. This ratio is termed the \textit{Ratio to Storm Center (RSC)} and provides an objective measure of the nearness of the virtual key station to the storm center. In this context, the spatial pattern of 6-hour precipitation is treated as a dimensionless shape pattern where the maximum 6-hour point precipitation (storm center) is the normalizing value. Figure 9a depicts the storm center zone for the MEC storm of July 19, 1996. The dimensionless shape pattern for the storm and the Ratio to Storm Center concept are graphically depicted in Figure 9b along with the random placement of the virtual key station and Watauga watershed (see also Figure 3 for the geometry of the Watauga watershed and associated station network).

\[ \text{RSC} = \frac{\text{PPT}_{\text{key}}}{\text{PPT}_{\text{StormCenter}}} \]  

(1)
where: $PPT_{\text{key}}$ is the 6-hour precipitation at the virtual key station; and $P_{\text{StormCenter}}$ is the maximum 6-hour precipitation in the spatial storm pattern corresponding to the location of the storm center.
Probability-plots of RSC for the Watauga and South Holston watersheds are depicted in Figures 10 and 11 where the largest annual maxima at the stations are seen to be associated with the highest ratios to storm center. Conversely, the smallest station annual maxima are associated with the lowest ratios to storm center. These results are supportive of the heuristic logic presented earlier that the largest annual maxima at a station are typically produced when the storm center occurs near that station.

![Figure 10 - Probability-Plot of Ratio of Maximum 6-Hour Point Precipitation at Butler TN to 6-Hour Maximum Point Precipitation in the Vicinity of the Watauga Dam Watershed](image)

![Figure 11 - Probability-Plot of Ratio of Maximum 6-Hour Point Precipitation at Abingdon 7WSW VA to 6-Hour Maximum Point Precipitation in the Vicinity of the South Holston Dam Watershed](image)

These findings provide a practical approach for setting constraints on the search area for random placement of the watershed (virtual key station) near the storm center. The approach is to set a minimum threshold for the ratio to storm center ($RSC_{min}$) based on the magnitude of the target precipitation at the virtual key station during stochastic storm generation. This approach includes a
variable threshold which results in the storm center being nearer the virtual key station for extreme storms and allows the storm center to be more distant from the virtual key station for common storms.

5.3 Scaling of Spatial Patterns to Larger Magnitude Storms
Linear scaling of spatial precipitation patterns is used as part of the algorithm for stochastic storm generation. Linear scaling is possible because the spatial variability of convective precipitation is dominated by the convection processes with limited topographic (orographic) enhancement due to the relatively short period of high-intensity precipitation in MEC storms. This is indicated by the similarity in magnitude of the 6-hour at-site mean for the MEC storm type in the general vicinity of any given location where the watersheds affected by MEC storms are typically less than 1,000-mi² (Figure 12).

Storm scaling is accomplished following the random placement of the watershed of interest within the precipitation field for a specific storm center zone. This action results in a specific grid-cell in the precipitation field overlying the grid-cell where the virtual key station is located. The Storm Scaling Factor (SSF) is the ratio of the desired 6-hour precipitation at the virtual key station to the actual 6-hour grid-cell precipitation in the precipitation field (Equation 2). The SSF is then used to scale the precipitation for each grid-cell in the precipitation field to produce a precipitation field with the desired precipitation at the virtual key station. Precipitation at each of the substations in the station network are determined by querying the precipitation field at the grid-cell location for each of the substations.
$$SSF = \frac{P_{key}}{P_{Gckey}}$$  \hspace{1cm} (2)

where: $SSF$ is the storm scaling factor; $P_{key}$ is the desired 6-hour precipitation at the virtual key station; $P_{Gckey}$ is the 6-hour precipitation in the precipitation field at the grid-cell location of the virtual key station.

### 5.4 Adjustment to Account for Storm Center Not Being Recorded by a Precipitation Station

The findings of radar analysis of convective storms in the TVA study area (Martin and Caldwell\textsuperscript{15}) have demonstrated that the maximum 6-hour point precipitation is rarely recorded at one of the precipitation stations in a network. This results in an underestimation of the areal-average watershed precipitation which varies with the size of the watershed of interest. The areal-average watershed precipitation should be adjusted upward for precipitation fields for historical storms where radar data are not available. Equation 3 is limited to watershed areas less than 1,000-mi\textsuperscript{2}. Table 2 lists bias correction for a range of watershed areas.

\[ Bias_{nonradar} = 1.1213 + 0.0104 \times \log_{10}(A_w) - 0.0439 \times (\log_{10}(A_w))^2 + 0.0094 \times (\log_{10}(A_w))^3 \]  \hspace{1cm} (3)

where: $Bias_{nonradar}$ is a bias correction for areal-average 6-hour watershed precipitation where spatial storm analyses are conducted without radar data; $Bias_{nonradar} = 1$ for storm spatial analyses conducted using radar data; and $A_w$ is the watershed area in square miles.

It should be noted that the terms *Storm Area* (Table 1) and watershed area are used interchangeably. This is a practical necessity because the radar spatial analyses are conducted for historical storms without reference to any given watershed. However, the correction factors are applied to specific watersheds with a wide variety of shapes.

### Table 1 – Bias Corrections to Adjust Estimates of Areal-Average Precipitation for Convective Storm Events for Spatial Storm Analyses Conducted without Radar Data

<table>
<thead>
<tr>
<th>Storm Area (mi\textsuperscript{2})</th>
<th>Radar Bias Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.12</td>
</tr>
<tr>
<td>2</td>
<td>1.12</td>
</tr>
<tr>
<td>5</td>
<td>1.11</td>
</tr>
<tr>
<td>10</td>
<td>1.10</td>
</tr>
<tr>
<td>20</td>
<td>1.08</td>
</tr>
<tr>
<td>50</td>
<td>1.06</td>
</tr>
<tr>
<td>100</td>
<td>1.04</td>
</tr>
<tr>
<td>200</td>
<td>1.03</td>
</tr>
<tr>
<td>500</td>
<td>1.02</td>
</tr>
<tr>
<td>1000</td>
<td>1.01</td>
</tr>
<tr>
<td>2000</td>
<td>1.00</td>
</tr>
</tbody>
</table>
5.5 Computation of Areal-Average Watershed Precipitation

Areal-average watershed precipitation is computed by simple arithmetic accounting for the precipitation of each grid-cell that is within the watershed of interest. Adjustments to the scaled precipitation field can be made for calibration to the historical watershed PF annual maxima (Figure 7) and to replicate PF relationships at nearby stations as part of the process of computing areal-average watershed precipitation. This is accomplished by additive ($\theta_a$) and multiplicative ($\theta_m$) adjustment/calibration factors as shown in Equation 4. Experience to date indicates the $\theta_a$ typically has a value near zero and $\theta_m$ has a value near unity.

$$ P_{ij} = \theta_a + \theta_m \times SSF \times P_{x,y} $$

where: $P_{ij}$ is the precipitation in the scaled and calibrated precipitation field for a given storm center zone at grid-cell $i,j$ where an indexing shift of $x=i+\Delta x$ and $y=j+\Delta y$ aligns the two raster fields; $\theta_a$ and $\theta_m$ are additive and multiplicative adjustments for calibration of the stochastically generated watershed PF relationship to the historical watershed PF relationship and to PF relationships at nearby substations; $SSF$ is the storm scaling factor; and $P_{x,y}$ is the precipitation in the historical precipitation field at grid-cell $x,y$ for a given storm center zone.

The areal-average watershed precipitation ($P_w$) is computed as:

$$ P_w = Bias_{nonradar} \times \frac{\sum \sum (P_{ij} \times AGC_{ij})}{A_w} $$

where: $P_w$ is the areal-average watershed precipitation; $Bias_{nonradar}$ is the bias correction for use/non-use of radar data; $SSF$ is the storm scaling factor; $P_{ij}$ is the precipitation in the scaled and calibrated precipitation field for a given storm center zone at grid-cell $i,j$; $AGC_{ij}$ is the area of grid-cell $i,j$ that is within the watershed of interest; and $A_w$ is the watershed area.

6.0 MULTIPLE STORMS PER YEAR

The stochastic storm generation algorithm simulates many mesoscale storms per sample year to emulate the natural processes. The truncated Poisson distribution is used to select the number of storms to be generated in each sample year. The conventional Poisson distribution has a lower bound of zero and the truncated Poisson distribution has a lower bound of one, which assures at least one storm per year during the simulations. An analysis of historical MEC storms in the Holston River system yielded a mean value of 10 storms per year. Figure 13 depicts the truncated Poisson distribution with a mean value of 10 storms per year which shows a range of from 1 storm to about 20 storms per year.

The annual maxima for watershed-average precipitation and precipitation annual maxima at each of the substations in the station network are selected from the stochastically generated point precipitation from the multiple storms each sample year.
7.0 TOTAL PROBABILITY THEOREM

The Total Probability Theorem (Equation 6) is used to compute annual exceedance probabilities for specific magnitudes of watershed-average precipitation (Benjamin and Cornell\textsuperscript{1}). The law of total probability states that given \( n \) mutually exclusive events \( A_1 \) through \( A_n \) whose probabilities sum to unity, then:

\[
P(B) = P(B|A_1) P(A_1) + P(B|A_2) P(A_2) + \ldots + P(B|A_n) P(A_n)
\]

(6)

where: \( B \) is the event of interest and \( P(B|A_i) \) is the conditional probability of \( B \) given \( A_i \).

In this application, the 6-hour point precipitation at the virtual key station is represented by the variable \( A \) and the 6-hour watershed-average precipitation is represented by the variable \( B \). The Total Probability procedure used here is to divide the range of precipitation magnitudes at the virtual key station into intervals (bins) using the Extreme Value Type 1 variate (x-axis) for creation of the bins (Figure 14a). Note the bins have equal intervals based on the Extreme Value Type 1 variate, but are not equal intervals in terms of precipitation nor incremental probabilities.

The law of total probability for a continuous probability distribution for watershed-average precipitation (Kuczera and Nathan et al\textsuperscript{19}) can be expressed as:

\[
p(Q > q) = \sum_{i=1}^{m} p(Q > q|R_i) p(R_i)
\]

(7)

where: \( p(Q > q) \) is the probability of watershed-average precipitation \( (Q) \) being greater than a specified magnitude \( (q) \) and summed over all precipitation bins \( (i \text{ to } m) \); \( p(R_i) \) is the probability of precipitation \( (R) \) being within a specific precipitation bin \( i \); and \( p(Q > q|R_i) \) is the conditional probability of \( Q > q \) given the precipitation magnitude being within precipitation bin \( i \), which can be estimated as the number of exceedances of \( q \) divided by the number of simulations within a precipitation bin.
The Cumulative Distribution Function (CDF) for watershed-average precipitation can be determined by computing exceedance probabilities \( p(Q > q) \) for a series of \( q \) values (Figure 14b) over the range of magnitudes of watershed-average precipitation \( Q \) using Equation 7.

The Total Probability Theorem is also used to compute the point PF relationship for each of the substations. In this application, the precipitation annual maxima for each substation is determined by querying the scaled precipitation field (Equation 5) for the multiple storms that are generated each sample year. The point PF values at a given substation are represented by variable \( B \) in Equation 7 and by \( Q,q \) in Equation 7.

Figure 14a – Division of 6-Hour Precipitation at Key Station into Sampling Intervals (Bins)

Figure 14b – Cumulative Distribution Function for 6-Hour Watershed-Average Precipitation Defined for Selected Values of 6-Hour Watershed-Average Precipitation
8.0 CALIBRATION AND VALIDATION

Procedures are available for calibration and validation of the watershed-average PF relationship produced by the stochastic storm generation procedure (Flowchart Figure 4).

8.1 Calibration of Watershed-Average Precipitation-Frequency Relationship

The watershed-average PF relationship produced by the stochastic storm generation procedure can be calibrated to the historical watershed-average PF relationship by additive ($\theta_a$) and multiplicative ($\theta_m$) adjustment/calibration factors as shown in Equation 4. Experience indicates that only minor adjustments are usually needed to merge the stochastically generated watershed PF relationship with the historical watershed PF relationship. Figure 15 depicts an example of this type of calibration for the Watauga watershed where $\theta_a$ and $\theta_m$ were zero and unity, respectively (no calibration adjustments required).

Minor adjustments can also be made to the stochastically generated values of point precipitation at the substations. The base value and gradient for the storm centering ratio ($RSC_{\text{min}}$) can be adjusted to improve the fit between the stochastically generated point PF relationship at the substations and the point PF relationship obtained from the regional frequency analysis (MGS et al\textsuperscript{18}). These adjustments also result in changes to the areal-average watershed PF relationship.

Figures 16a through 16e depict comparisons of stochastically generated point precipitation at substations as compared to the point PF relationship obtained from the regional point PF analysis. The locations of these substations are referenced in Figure 3 for the Watauga watershed and the distance from the virtual key station at the watershed centroid is noted in each of the figures. It is important to note that this high level of replication was produced using the regional point PF relationship at the virtual key station and scaling procedures along with the spatial precipitation patterns observed in historical storms.
Figure 16a – Comparison of Stochastically Generated Point PF Relationship at Reese NC as Compared to the Point PF Relationship Obtained from the Regional Point PF Analysis

Figure 16b – Comparison of Stochastically Generated Point PF Relationship at Butler TN as Compared to the Point PF Relationship Obtained from the Regional Point PF Analysis

Figure 16c – Comparison of Stochastically Generated Point PF Relationship at Banner Elk NC as Compared to the Point PF Relationship Obtained from the Regional Point PF Analysis
8.2 Validation of Watershed Precipitation-Frequency Relationship

A suitable solution is obtained when the stochastically generated watershed PF relationship merges well with the historical watershed PF relationship (Figure 15) and the substation point PF relationships obtained from the stochastically generated storms reasonably match the point PF relationships obtained from regional PF analysis. Suitability of the watershed PF relationship is validated by several measures including:

- Point PF relationship at the virtual key station is preserved (Figure 5)
- Spatial structure of convective storms is preserved by the large sample set of historical spatial precipitation patterns (118 spatial precipitation patterns for Watauga watershed)
- Any non-linearities in the spatial precipitation patterns associated with storm magnitude that exist are reduced by using a stratified sample of historical storm patterns where the magnitude of the storm center and the general magnitude of spatial precipitation is comparable to the target value of precipitation at the virtual key station (section 5.0)
• Stochastically generated watershed PF relationship is calibrated to historical watershed-average annual maxima as represented by the historical watershed PF relationship (Figure 15)

• Stochastic generation of multiple storms per year results in point PF relationships at substations in the station network that reasonably replicate the point PF relationships determined from the regional PF analysis (Figures 16a through 16e)

10.0 UNCERTAINTY CONSIDERATIONS

The ability to characterize uncertainties is a major component of the algorithm for stochastic storm generation of a watershed PF relationship. Aleatoric and epistemic uncertainties are inherent to all aspects of the natural processes of Mesoscale Storms with Embedded Convection (MEC) precipitation and the algorithm created to simulate the natural processes. The aleatoric uncertainties are addressed in the simulation procedures in the inner (blue) loop of the Flowchart shown in Figure 4 and epistemic uncertainties are addressed in the outer (red) loop of the Flowchart.

Execution of the procedures depicted in the Figure 4 Flowchart allows for development of a mean-frequency curve (best-estimate) and uncertainty bounds for the 6-hour watershed PF relationship (Figure 19). The following sections provide a brief summary describing considerations and procedures for characterizing epistemic uncertainties for each of the components of the algorithm in conducting the uncertainty analysis.

10.1 Uncertainties in Precipitation-Frequency Relationship at Virtual Key Station

The 6-hour point PF relationship at the virtual key station is a primary driver for stochastic storm generation. There are four epistemic uncertainties that are addressed as part of the algorithm that include: the at-site mean; regional L-Cv and L-Skewness and the regional probability distribution applicable to the site of the virtual key station.

Virtual Key Station At-Site Mean – Epistemic uncertainty in the at-site mean at the virtual key station is modeled by a Normally distributed random variable with mean zero and a standard deviation based on the relative standard error of spatially mapped at-site means. This typically represents about a 4-5% standard error relative to the at-site mean. This information is obtained from the relative standard error computed in spatial mapping of at-site means for the project study area.

Regional L-Cv and Regional L-Skewness – Sampling properties of L-Moment ratios are near Normally distributed (Hosking and Wallis11). Epistemic uncertainties in the regional values of L-Cv and L-Skewness are based on the relative standard error of regional solutions for L-Cv and L-Skewness that are used in spatial mapping (MGS et al18). A Normally distributed random variable with mean zero and a standard deviation equal to the relative standard error are used in stochastic simulations. Relative standard errors are commonly on the order of 2% and 6% for regional L-Cv and L-Skewness, respectively, when very large regional datasets of precipitation annual maxima and storm typing are employed.
Regional Probability Distribution – The choice of the 4-parameter Kappa distribution (Hosking\textsuperscript{10}, Hosking and Wallis\textsuperscript{11}) provides the capability to emulate alternative probability distributions near the chosen 3-parameter regional probability distribution. Specifically, the second shape parameter of the Kappa distribution ($h$, Equation 10) can be treated as a fixed value which yields a 3-parameter distribution. In this context, the Generalized Logistic (GL), Generalized Extreme Value (GEV), Gaucho and Generalized Pareto (GP) are seen as special cases of the 4-parameter Kappa distribution where the second shape parameter ($h$) has values of -1, 0, +0.5 and +1, respectively. An L-Moment ratio diagram is depicted in Figure 17 which shows the regional L-Moment Ratio pairings for 18 homogeneous regions for the MEC storm type for the TVA study area and the regional probability distribution being very near the GEV.

The quantile function for the 4-parameter Kappa distribution is:

$$q(F) = \xi + \frac{\alpha}{\kappa}\left\{1 - \left(\frac{1 - F^h}{h}\right)^{\kappa}\right\}$$

where: $\xi$, $\alpha$, $\kappa$, and $h$ are location, scale and two shape parameters respectively.

In practice, the 4-parameter Kappa distribution is selected as the regional probability distribution and the second shape parameter is determined in increments of 0.05, which effectively yields a 3-parameter distribution. There is mathematical reasoning (Gumbel\textsuperscript{6}) to support the case that the GEV is approached asymptotically based on the mathematical form of the probability distribution that describes the precipitation magnitudes generated by a specific meteorological process and the number of storms generated in a given year from which the annual maxima are selected. This consideration supports the choice of a regional probability distribution that is near, but not yet asymptotic to, the GEV.

![Figure 17 – L-Moment Ratio Diagram Showing L-Skewness and L-Kurtosis Pairings for 18 Homogeneous Subregions for the Mesoscale Storm with Embedded Convection (MEC) Storm Type](image-url)
Epistemic uncertainty about the chosen 3-parameter regional probability distribution is modeled by preserving the functional relationship between L-Moment ratios for L-Skewness and L-Kurtosis for the fixed value of the second shape parameter while allowing for uncertainty (variability) in both L-Skewness and L-Kurtosis. This can be viewed as having a L-Moment ratio curve near and parallel to one of the 3-parameter distributions shown in Figure 17, such as the GEV. Uncertainty about L-Skewness is described in the prior section and uncertainty in L-Kurtosis is characterized by a Normally distributed random variable. An Equivalent Independent Record Length (EIRL) measure (MGS et al\textsuperscript{18}) is used in estimating the standard deviation for L-Kurtosis where epistemic uncertainties in L-Kurtosis are similar in magnitude to that for L-Skewness.

10.2 Uncertainties in the Spatial Structure of Precipitation
As discussed previously, the spatial patterns of convective precipitation can be very chaotic. Localized high intensities, steep gradients, and abrupt edges in the precipitation field are common with convective storm activity. These situations pose a major challenge for modeling the spatial storm structure using a spatial correlation structure and statistical methods. These conditions are the reason that a resampling approach using historical precipitation fields was chosen for modeling the spatial structure of convective precipitation.

Epistemic uncertainty in the spatial structure of convective precipitation is modeled using a type of jackknife resampling procedures. This is accomplished by using a randomly selected subset of historical precipitation fields from the full dataset of historical precipitation fields for each of the simulation scenarios (See Section 5.0). For the case of the Watauga watershed, a subset of 80 precipitation fields is selected from the full dataset of 118 precipitation fields for each simulation scenario.

10.3 Uncertainties in the Selection of the Ratio to Storm Center
Uncertainties are present in the magnitude of the Ratio to Storm Center (RSC) to be used in random placement of the virtual key station and associated watershed on the storm center zones of the historical precipitation fields. The first step can be considered as calibration to identify the values of RSC\textsubscript{min} that best replicates the historical watershed PF relationship (Section 5.2, Figure 7) and the substation PF relationships (Figures 15a-15e). Epistemic uncertainty in the RSC is modeled by allowing a range of RSC parameter values to be considered in the uncertainty analysis.

10.4 Uncertainties Due to Climate Change
Uncertainties due to climate change are not considered at this stage of the analysis. There may be changes in frequency and/or magnitude of MEC precipitation due to climate change. Climate change effects can be considered when stochastic flood modeling is conducted. This approach will allow greater flexibility in assessing effects from both the meteorological and hydrological perspective.

10.5 Uncertainties Associated with Spatial Storm Analyses Conducted by MetStorm
Stochastic storm generation is based on the findings of spatial storm analyses that are conducted using MetStorm\textsuperscript{17}. The MetStorm software is comprised of a collection of algorithms based on findings of prior studies, meteorological experience and judgment. The meteorological community has imperfect understanding of the meteorological processes associated with MEC precipitation and the resultant spatial and temporal characteristics. In addition, limited spatial and temporal data are
typically available for conducting spatial storm analyses. These factors combine to produce
epistemic uncertainties in the resultant estimates of areal-average watershed precipitation.
While these uncertainties are judged to be small relative to that for regional L-Skewness and the
regional probability distribution discussed earlier, they are contributors to the total uncertainty
(Figure 18). These epistemic uncertainties are not included in the uncertainty analysis at present.
These issues will be examined over time and may be included at some point in the future.

10.6 Relative Contributions of Components to the Total Uncertainty
A separate set of stochastic storm simulations are conducted to identify the relative contribution of
each component to the total uncertainty. This analysis is conducted by first computing the total
uncertainty variance where Latin hypercube sampling allows all components to vary through their
respective ranges of uncertainty. Subsequent analyses are then conducted wherein the value of one
component of the stochastic storm generation algorithm is fixed at its mean value and the resultant
uncertainty variance is compared to the simulations where uncertainties are considered for all other
components. Figure 18 shows a stacked histogram where the relative contribution of uncertainty for
each component for the Watauga watershed is depicted. The square root of the uncertainty variance
on the ordinate of Figure 18 corresponds to the standard deviation for the width of uncertainty
bounds shown in Figure 19.

The precipitation fields and Ratio to Storm Center (RSC) are relatively small contributors to the total
uncertainty beyond 10^{-4} AEP. In particular, the uncertainty in L-Skewness and the identification of
the regional probability distribution are the primary contributors to uncertainty for extreme
precipitation. These two components have the greatest influence on the shape of the upper tail of
the watershed PF relationship.

Further reduction in epistemic uncertainty for L-Skewness and the regional probability distribution is
possible in the future. The use of storm typing and large regional studies have the potential to provide
an increased understanding in the range of L-Skewness values and identification of the regional
probability distribution and associated probabilistic behavior for convective storm processes.

Figure 20 depicts the 6-hour point PF relationship for areal-average values of the at-site mean,
regional L-Cv and regional L-Skewness for the Watauga watershed. The point PF relationship has a
common level of annual exceedance probability at all locations throughout the watershed for a
given value of 6-hour precipitation. The difference between the point PF relationship and the areal-
average watershed PF relationship (Figure 20) can be viewed in the context of an Areal Reduction
Factor (ARF). It should be noted the ARFs are not fixed values but vary with annual exceedance
probability. Smaller areal reduction occurs for common storm events and larger areal reduction
occurs for extreme storm events. ARF relationships from detailed moving-the-earth analyses can be
used in developing scaling procedures for producing watershed PF relationships for use on other
watersheds.
Figure 18 – Stacked Histogram Showing Relative Contribution of Various Sources of Uncertainty to the Total Uncertainty in the Watershed Precipitation-Frequency Relationship for the Watauga Watershed TN

Figure 19 – Watershed 6-Hour Precipitation-Frequency Relationship and 90% Uncertainty Bounds for the Watauga Watershed for the Mesoscale Storm with Embedded Convection (MEC) Storm Type
12.0 REFERENCES AND SELECTED ARTICLES


