Central Region Technical Attachment
Number 23-03
December 2023

MRMS and FLASH Thresholds for Assessing Flash Flood Potential in Realtime for the Western Great Lakes Part 2: Land Cover and Soil Moisture Considerations

W. Scott Lincoln, NWS Chicago

Corresponding Author:
W. Scott Lincoln
Senior Service Hydrologist
NWS Chicago, Romeoville, IL
scott.lincoln@noaa.gov

1.0 Introduction

Recent improvements to National Weather Service flash flood warnings allow the ability to indicate the expected damage threat using three severity levels (base, considerable, and catastrophic). Thus, the most extreme flash flooding situations can be separated from the lower-end situations that impact a far smaller number of people. In Lincoln & Marquardt (2023), a technique is proposed for real-time flash flood warning decisions called the 4-Panel Technique. This technique uses experimental products from the Multi-Radar Multi-Sensor (MRMS; Zhang et al. 2016) system and the Flood Locations and Simulated Hydrographs (FLASH; Vergara et al. 2019) project to assist with estimating flash flood severity. With the 4-Panel Technique, four products, 1-hr Radar-Only QPE, Max QPE Average Recurrence Interval (ARI), Max QPE-to-GFFG Ratio, and Unit Streamflow, are used together to evaluate potential flash flooding severity. Lincoln & Marquardt (2018) collected and analyzed 190 cases, about half of which involved flash flooding of varying severities and half of which involved no reported flooding, and then calibrated the 4-Panel Technique to provide the best combination of probability of detection (POD) and false alarm ratio (FAR) using the critical success index (CSI). One limitation of this effort was that events were not analyzed based upon land cover, land surface slope, or soil moisture conditions, each of which may impact product thresholds relevant to flash flooding. In this study, we illustrate how optimal thresholds for MRMS and FLASH products may change based upon antecedent conditions and the specific locations where heavy rainfall occurs.

2.0 Methodology

2.1 Collecting Flash Flood Cases and Assigning Severity

As this current study directly builds upon Lincoln & Marquardt (2023), a more-detailed summary of how flash flood cases were collected can be found there. A list of flash flood events occurring in the NWS Chicago and NWS Milwaukee Hydrologic Service Areas (HSAs) was created covering summer 2016 through fall 2022. Impacts from each event were reviewed and a subjective assessment of flash flood severity was then assigned to each flash flood case. Each severity level was assigned a numeric value ranging from 0.0 (no flooding) to 4.0 (flash flood - catastrophic). See Table 1 for a list of each severity level and the warning level equivalent.

Peak values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE GFFG Ratio, and Unit Streamflow, within several miles of the reports of flooding, were collected and averaged over a four-pixel area. To reduce potential selection bias associated with selecting only cases where flooding was known to occur, null (no flooding) cases were added by searching MRMS/FLASH data during 2016-2022 for instances where elevated values occurred in the products.

Table 1. The numeric values used for classifying flash flood severity. From Lincoln & Marquardt (2023).

IBW Level	**Severity (Numeric)**
No Flooding	0.0
Advisory	1.0
Flash Flood Warning Base	2.0
Flash Flood Warning Considerable	3.0
Flash Flood Warning Catastrophic	4.0

2.2 The 4-Panel Technique

The 4-Panel Technique presented in Lincoln & Marquardt (2023) utilizes 1-hr Radar-Only QPE, Max QPE ARI, Max QPE GFFG Ratio, and Unit Streamflow with custom color tables aligned to the expected magnitude of flash flooding. When three panels (products) show the same flash flood impact level, a particular hazard product is considered “recommended.” The 4-Panel Technique is just part of the warning decision process, which also includes analyzing the spatial footprint of elevated MRMS/FLASH values, biases in MRMS data, trends in MRMS/FLASH values, meteorological considerations, and observed reports of flooding. Lincoln & Marquardt (2023) proposed recommended thresholds for each of the products for use in the western Great

Lakes region (Table 2). One strength of this technique is that thresholds for the MRMS/FLASH products can be changed to match regional or local studies, with the color tables shifted as needed. It is possible that thresholds may vary based upon infrastructure design standards and an area's vulnerability to flood disasters.

One limitation of the current implementation of the 4-Panel Technique, however, is that it does not take into account potential variability in rainfall-runoff response due to land surface slope (flat terrain versus hilly terrain), land cover (rural areas versus urban areas), or soil moisture (dry versus wet). The previously proposed thresholds thus include a wide range of conditions which may cause variability in the best threshold to use on a case-by-case basis. By collecting this extra information about each flash flood case, the cases can be subdivided to look for correlations and trends.

Table 2. Recommended threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow when used as part of the 4-Panel Technique. Note that due to the very small sample size for "flash flood catastrophic" (severity level 4) cases, borderline events (severity level 3.5) were also included; values should still be used with caution. From Lincoln & Marquardt (2023).

	Advisory	**Flash Flood Base**	**Flash Flood Considerable**	**Flash Flood Catastrophic**
1-hr Radar-Only QPE	1.5	2.0	2.5	2.7
Max QPE ARI	1	5	125	175
Max QPE GFFG Ratio	125	140	325	375
Unit Streamflow	200	230	850	1100
POD	0.91	0.83	0.58	1.0
FAR	0.35	0.42	0.30	0.60
CSI	0.61	0.52	0.47	0.40
Number of Cases Exceeding Severity	87 / 190	59 / 190	12 / 190	4 / 190

2.3 Matching Events to Slope, Land Cover, and Soil Moisture

Each flash flood case from Lincoln & Marquardt (2023) was tied to a land surface slope, land cover, and soil moisture. For slope, the extent of elevated MRMS/FLASH values was compared to generalized slope values, as illustrated by Figure 1, and classified as either “flat,” “hilly,” or “foothills.” To determine land surface slope categories, the 30-meter (1 arc second) elevation data from the USGS (https://apps.nationalmap.gov/downloader/) were first resampled to 1 km, and used to calculate percent slope. Cells with slope values <5% were classified as “flat,” cells with slope values >20% were classified as “foothills,” and the remaining cells were classified as “hilly.” Based upon the topography of the study area, the vast majority of flash flood cases were classified as “flat” slope unless they occurred in southwest Wisconsin, where “hilly” classifications were possible.

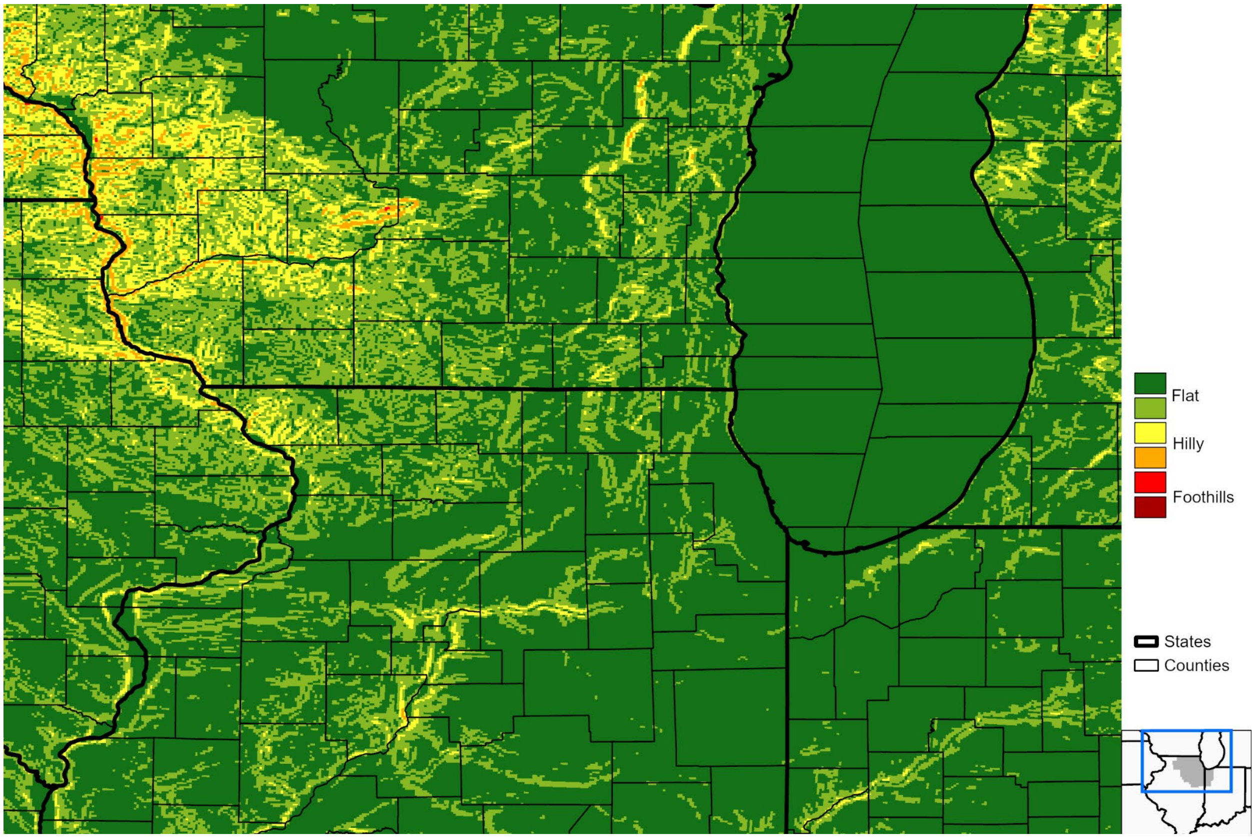

Figure 1. Land surface slope, or topography, used to classify each flash flood case.

For land cover, the extent of elevated MRMS/FLASH values were compared to generalized land cover classifications, as illustrated by Figure 2, and classified as either “rural,” “suburban,” or “urban.” To determine land cover categories, the 30-meter (1 arc second) impervious data from the USGS (https://www.usgs.gov/centers/eros/science/national-land-cover-database/) were first resampled to 500 m, and then separated into categories. Cells with imperious values <20% were classified as “rural,” cells with impervious values >45% were classified as “urban,” and the remaining cells were classified as “suburban.”

Based upon the topography of the study area, flash flood events generally were only classified as “suburban” or “urban” if they occurred in the Chicago, Rockford, Milwaukee, or Madison metropolitan areas.

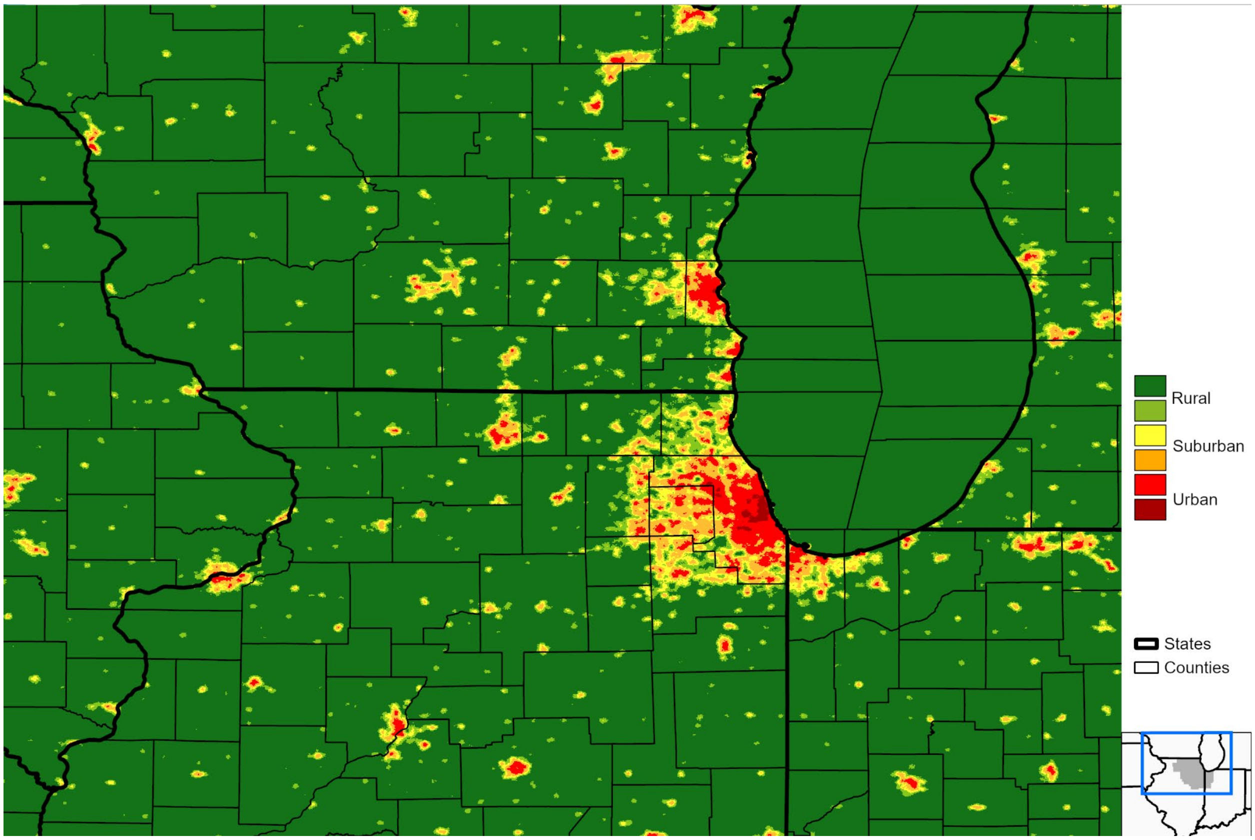

Figure 2. Land cover used to classify each flash flood case.

For soil moisture, the extent of elevated MRMS/FLASH values were compared to three different representations of soil moisture, due to limitations in available soil moisture data, each with pros and cons. The first measure of soil moisture used was the 7-day streamflow percentile from the U.S. Geological Survey (USGS; https://waterwatch.usgs.gov/index.php?id=pa07d). Although streamflow is not a direct measure of soil moisture, elevated streamflow levels are often tied to elevated soil moisture and groundwater levels and USGS data are readily available for the last several decades. Using USGS streamflow as a proxy for soil moisture also has a precedent in similar research; it was used as a soil moisture proxy in Lincoln & Thomason (2018). The 7-day streamflow percentile classifications, as illustrated by Figure 3a, are classified as “much below average,” “below average,” “average,” “above average,” or “much above average”; the value assigned to a given flash flood case was the predominant classification across the impacted area.

The second measure of soil moisture used was the modeled soil moisture percentile from the Climate Prediction Center (CPC; https://www.cpc.ncep.noaa.gov/soilmst/descrip.htm). The CPC soil moisture product is a direct attempt to model soil moisture (rather than a proxy, such as with USGS) and is readily available over the last several decades, but it represents a large depth of soil near the surface (potentially several feet). Soil moisture values most important to flash flood processes are generally limited to the upper several inches. The soil moisture percentiles, as illustrated by Figure 3b, range from 0 to 100; the value assigned to a given flash flood case was the average value across the impacted counties.

The third measure of soil moisture used was the modeled soil moisture from the Coupled Routing and Excess Storage (CREST) method, part of FLASH. This product is meant to represent a smaller portion of the soil column near the surface and is also the soil moisture product used by other hydrologic products from FLASH, but limited data availability precluded its use in 36 of the 190 cases (events occurring prior to March 2019). The soil saturation values from CREST, illustrated by Figure 3c, range from 0 to 100; the value assigned to a given flash flood case was the average value across the impacted area.

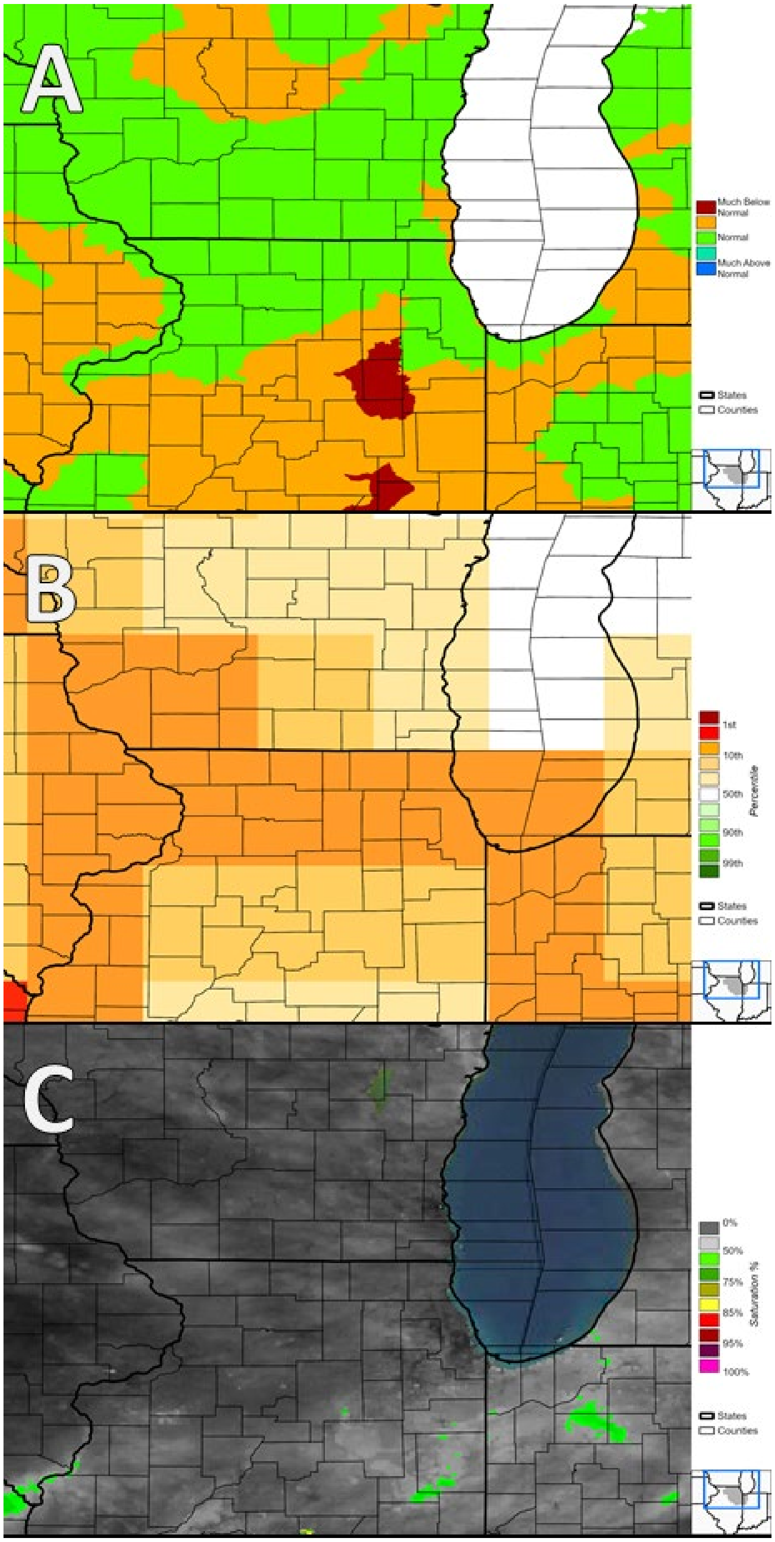

Figure 3. Example soil moisture data used to classify each flash flood case. Soil moisture data came from three sources, including USGS streamflow percentile (top), CPC-modeled soil moisture percentile (middle), and CREST-modeled soil moisture saturation (bottom).

2.4 Subdividing Cases by Slope, Land Cover, and Soil Moisture

To determine if there is a potential relationship between flash flood severity and land surface slope, land cover, and soil moisture, the individual flash flood cases were subdivided into groups. Flash flood cases were subdivided into two groups each such that the groups consisted of an approximately equal number of cases (Table 3). For land surface slope, flash flood events with a “flat” classification numbered 183, and classifications “hilly” and “foothills” were grouped together, numbering seven (Figure 4). For land cover, events with a “rural” classification numbered 96, and classifications “suburban” and “urban” were grouped together, numbering 94 (Figure 5). For soil moisture, as measured by USGS streamflow as a proxy, classifications “much above average” and “above average” were grouped together, numbering 129, and classifications “average,” “below average,” and “much below average” were grouped together, numbering 61 (Figure 6). For soil moisture modeled by CPC, values were split at the 95th percentile, with values <95 numbering 87 and values >=95 numbering 99 (Figure 7). For soil moisture modeled by CREST, values were split at 20% saturation, with values <20 numbering 80 and values >=20 numbering 72 (Figure 8).

Table 3. Summary of values and terms used for subdividing flash flood cases into two (2) groups, and the number of flash flood cases in each group.

	Group1	**Group2**
Land Surface Slope values	“Flat” (slope <5%)	“Hilly” or “Foothills” (slope >=5%)
Land Surface Slope number of cases	183	7
Land Cover values	“Rural” (impervious <20%)	“Suburban” or “Urban” (impervious >=20%)
Land Cover number of cases	96	94
USGS Streamflow values	“Average” or “Below Average” or “Much Below Average”	“Much Above Average” or “Above Average”
USGS Streamflow number of cases	61	129
CPC Soil Moisture values	<95th percentile	>=95th percentile
CPC Soil Moisture number of cases	87	99
CREST Soil Moisture values	<20% saturation	>=20% saturation
CREST Soil Moisture number of cases	80	72

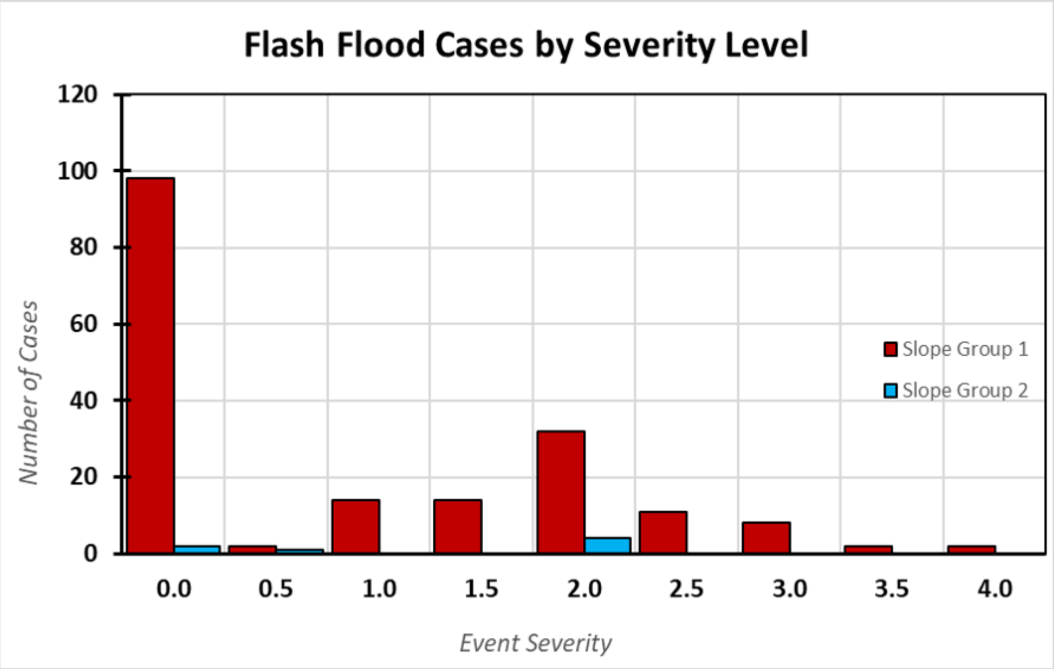

Figure 4. Numeric severity values for the collected flash flood cases, subdivided into two groups for land surface slope.

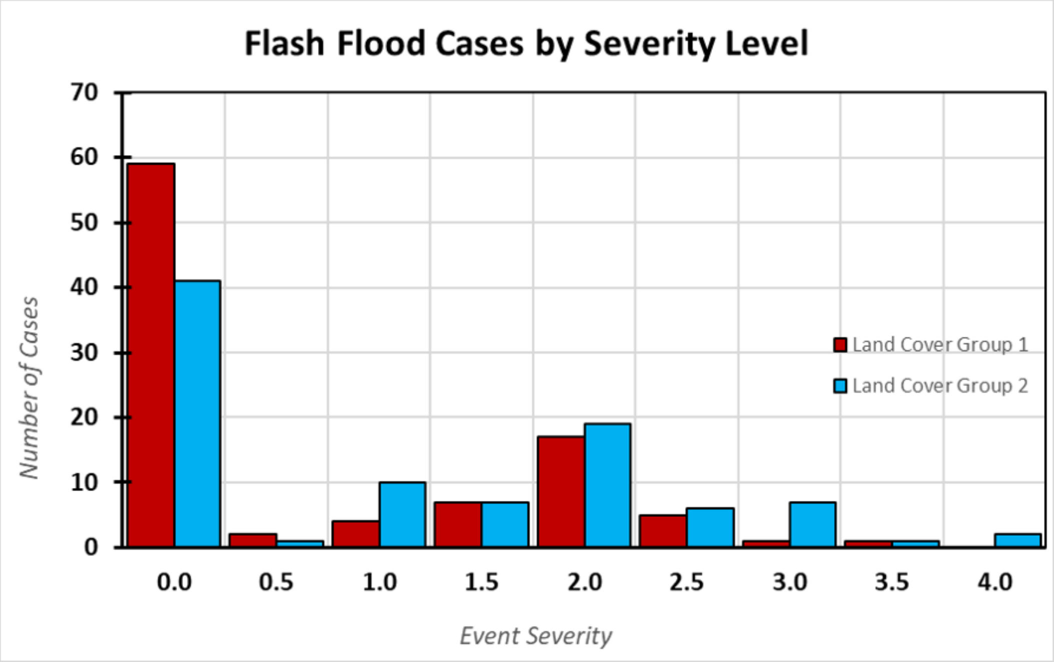

Figure 5. Numeric severity values for the collected flash flood cases, subdivided into two groups for land cover.

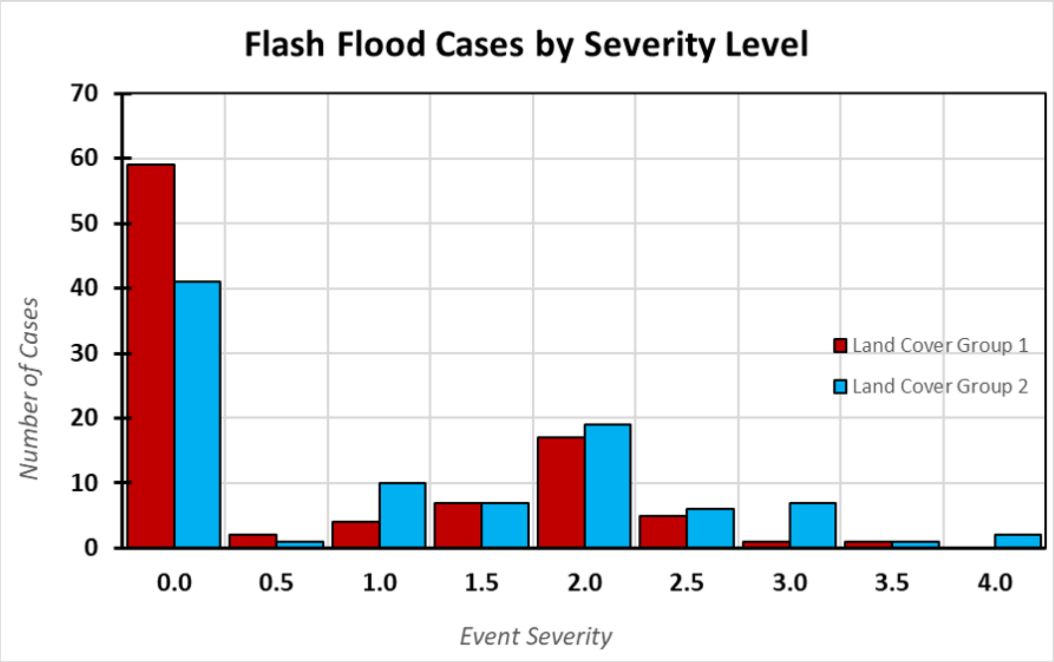

Figure 6. Numeric severity values for the collected flash flood cases, subdivided into two groups for USGS streamflow percentile.

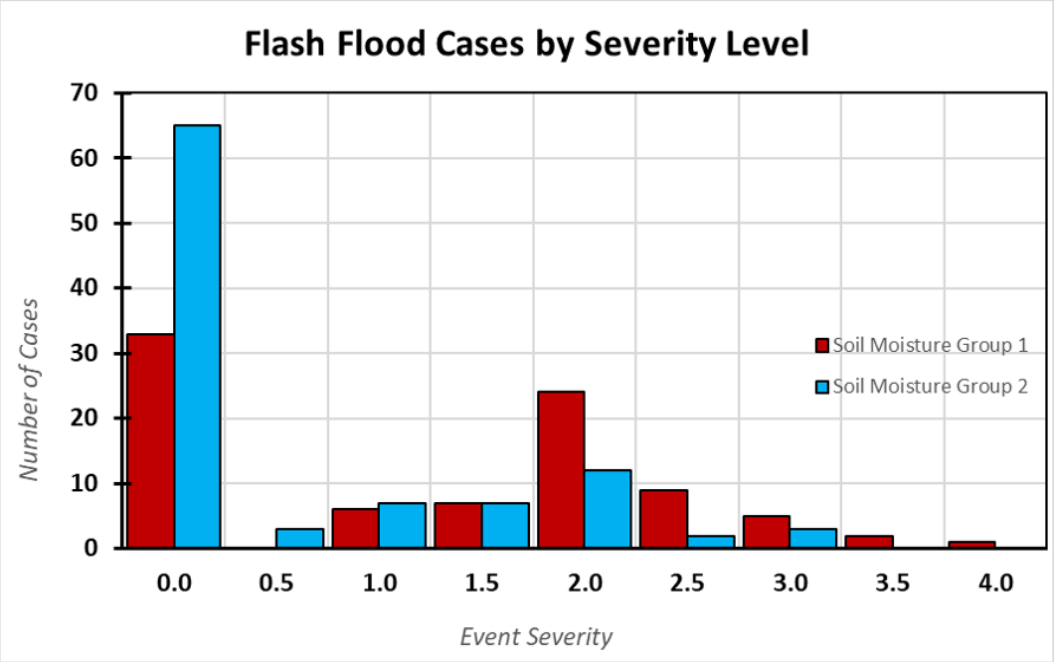

Figure 7. Numeric severity values for the collected flash flood cases, subdivided into two groups for soil moisture modeled by CPC.

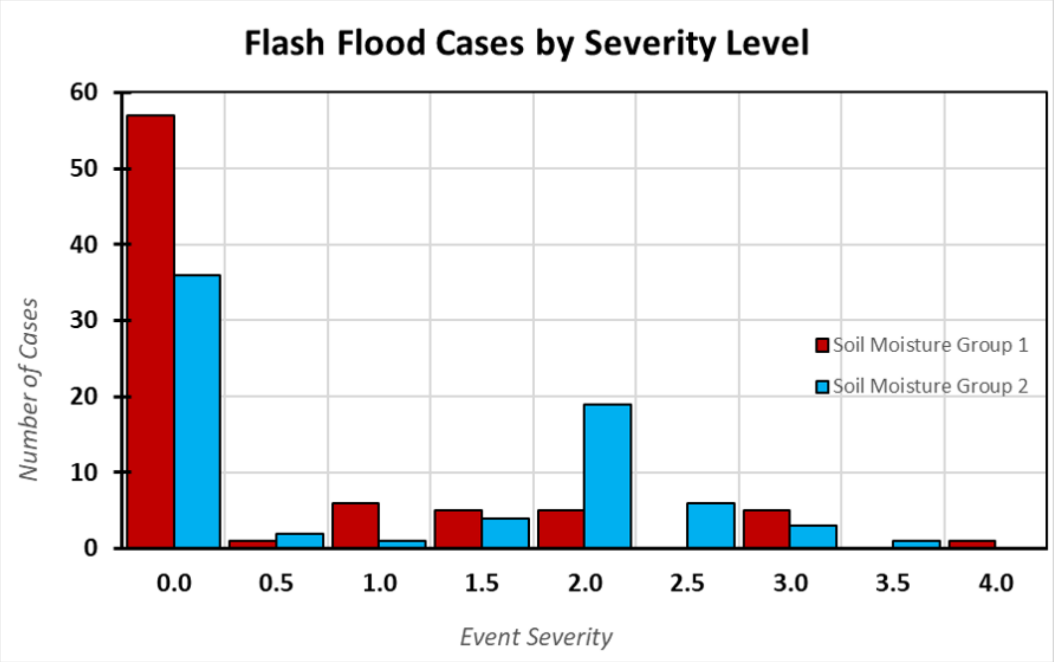

Figure 8. Numeric severity values for the collected flash flood cases, subdivided into two groups for CREST-modeled soil moisture.

2.5 Calibration

Using the calibration script developed for Lincoln & Marquardt (2023), the flash flood cases, subdivided as indicated by Table 3, were calibrated for best CSI. The calibrated thresholds for the subdivided groups were then compared to the calibrated thresholds for the full sample of flash flood cases to evaluate whether or MRMS/FLASH thresholds should be adjusted based upon land cover and soil moisture. The calibration script was run with the same minimum, maximum, and interval values used by Lincoln & Marquardt (2023).

3.0 Results and Discussion

3.1 Differences in Flash Flood Severity Based Upon Slope, Land Cover, and Soil Moisture

Values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE GFFG Ratio, and Unit Streamflow, were separated into six bins. For each of the bins, the subdivided flash flood cases were compared to determine if separate thresholds may be helpful for that particular MRMS/FLASH product. For land surface slope, the number of cases that could be broken out into group 2 is very small (only 7), which prevents a useful comparison of flash flood severity values. While the most appropriate MRMS/FLASH thresholds may still vary based upon land surface slope, from the flash flood cases available to this study a difference could not be discerned. For land cover, group 2 (“suburban” and “urban”) had higher flash flood severity values for almost all bins for each product (Figure 9). The biggest differences were noted for 1-hr Radar-Only QPE, Max QPE ARI, and Max QPE GFFG Ratio.

For soil moisture represented by USGS streamflow, differences were noted between group 1 and group 2, but they were inconsistent (Figure 10). A higher severity was noted for most values of 1-hr Radar-Only QPE with group 2, but for the other MRMS/FLASH products, severity values were unchanged or even decreased. A decrease in flood severity for higher streamflow values (a proxy for increased soil moisture) was an unexpected result, and it could not be easily explained by the available data. It is also worth noting that this finding was different from what was found by Lincoln & Thomason (2018), which used USGS streamflow as a soil moisture proxy and concluded that higher streamflow levels suggested lower QPE ARI values for warning thresholds.

For soil moisture modeled by CPC, group 2 (>=95th percentile) had lower severity levels for most bins for each product (Figure 11). A decrease in flood severity for higher soil moisture values was an unexpected result. A few factors might be contributing to this unexpected result, including the very high cutoff value (very moist, 95th percentile soil moisture) necessary to reach an approximately equal amount of cases in each group. CPC soil moisture is also based upon a deeper soil column than what is correlated to flash flood processes, and the period of time from which cases were selected included large stretches of elevated soil moisture over this soil depth.

For soil moisture modeled by CREST, group 2 (>=20% saturation) had higher flash flood severity values for almost all bins of each product (Figure 12). This result was more consistent with expectations and differed from the other soil moisture data used. Although CREST soil moisture data was not available for all flash flood cases, the fact that it models soil moisture at depths more closely related to flash flood processes suggested that it should be the primary soil moisture product used for further study.

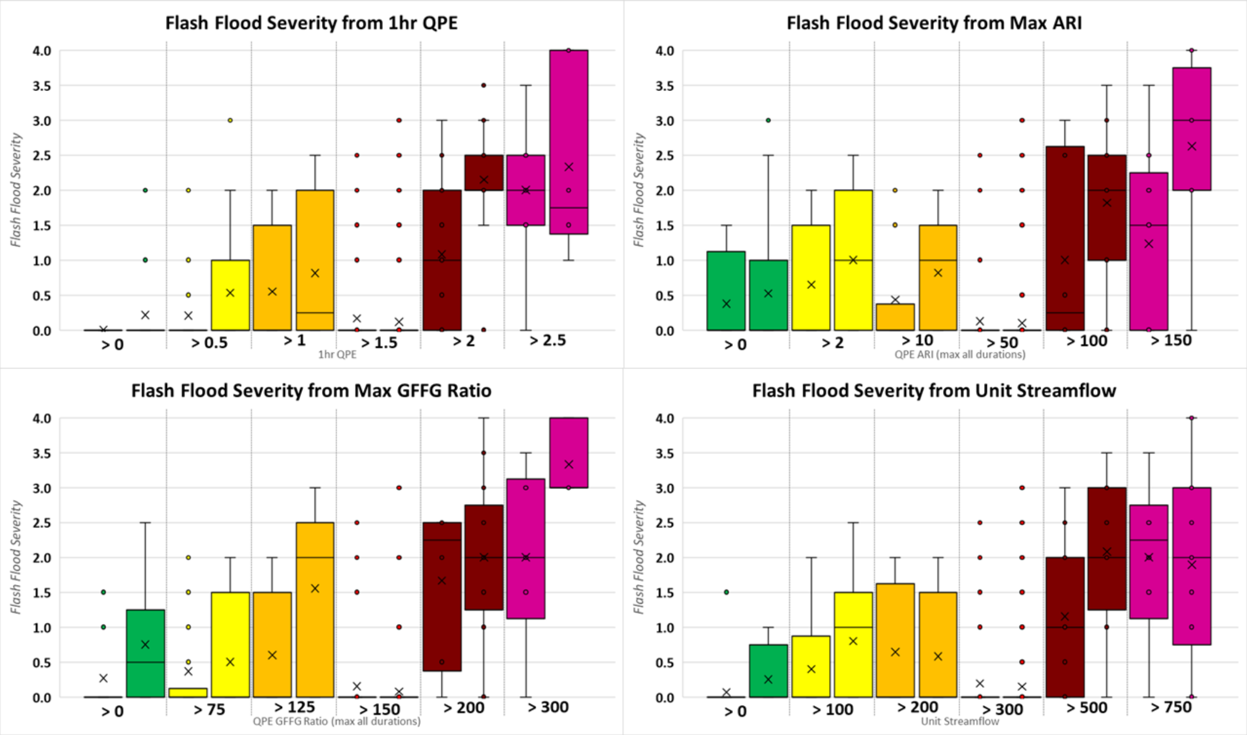

Figure 9. Comparison of flash flood severity values based upon differences in land cover. Group 1, cases occurring in generally rural areas, is shown by the left box for a given bin, and group 2, cases generally occurring in suburban or urban areas, is shown by the right box for a given bin. Flash flood severity values are shown for 1-hour QPE (top left), Max QPE ARI (top right), Max QPE-to-GFFG Ratio (bottom left), and Unit Streamflow (bottom right).

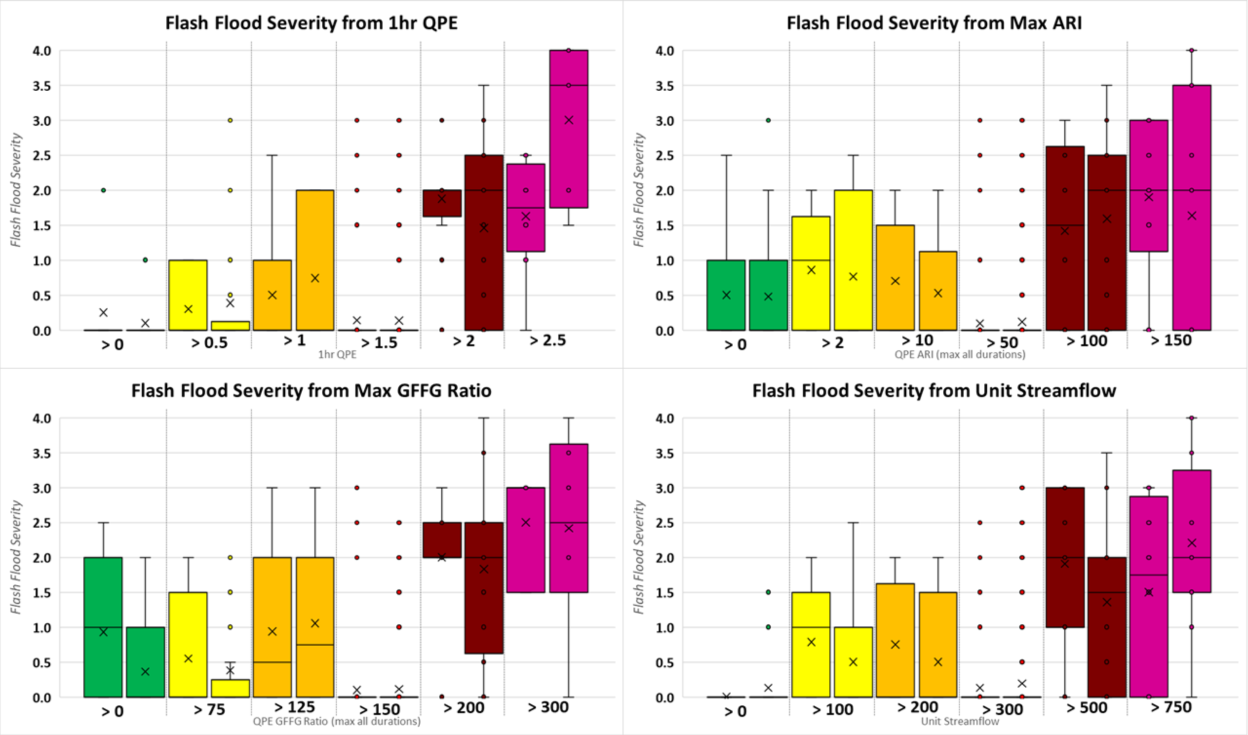

Figure 10. Comparison of flash flood severity values based upon differences in USGS streamflow. Group 1, cases occurring with below average or average streamflow, is shown by the left box for a given bin, and group 2, cases occurring with above average streamflow, is shown by the right box for a given bin. Flash flood severity values are shown for 1-hour QPE (top left), Max QPE ARI (top right), Max QPE-to-GFFG Ratio (bottom left), and Unit Streamflow (bottom right).

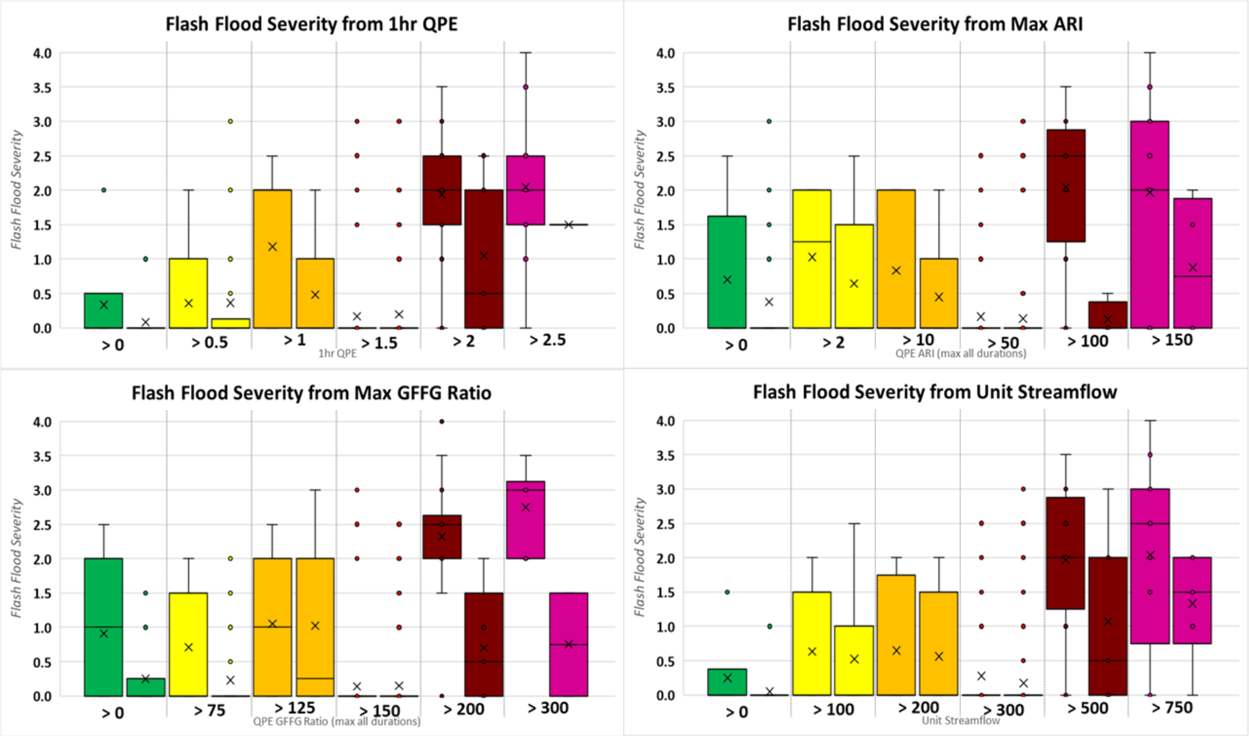

Figure 11. Comparison of flash flood severity values based upon differences in CPC-modeled soil moisture. Group 1, cases occurring with below average or average streamflow, is shown by the left box for a given bin, and group 2, cases occurring with above average streamflow, is shown by the right box for a given bin. Flash flood severity values are shown for 1-hour QPE (top left), Max QPE ARI (top right), Max QPE-to-GFFG Ratio (bottom left), and Unit Streamflow (bottom right).

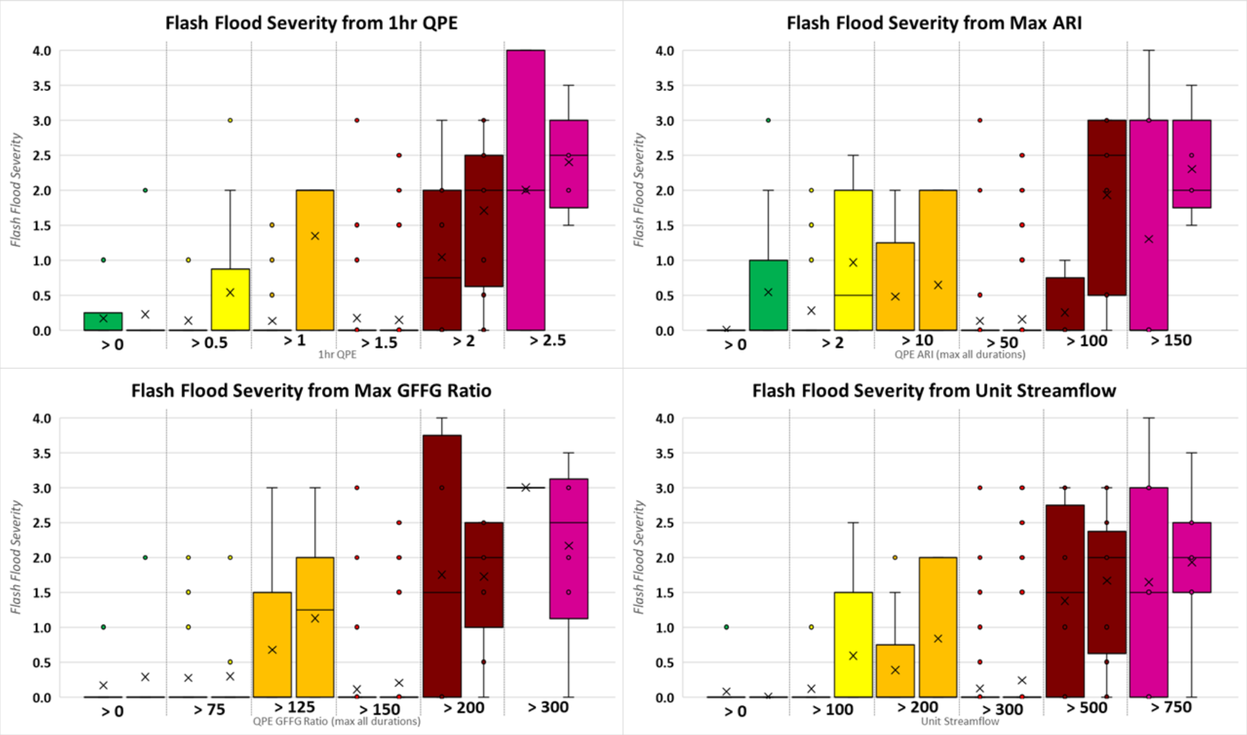

Figure 12. Comparison of flash flood severity values based upon differences in CREST-modeled soil moisture. Group 1, cases occurring with soil saturation <20%, is shown by the left box for a given bin, and group 2, cases occurring with soil saturation >=20%, is shown by the right box for a given bin. Flash flood severity values are shown for 1-hour QPE (top left), Max QPE ARI (top right), Max QPE-to-GFFG Ratio (bottom left), and Unit Streamflow (bottom right).

After evaluating land cover’s effect on flash flood severity, it was determined that areas classified as “urban” or “suburban” generally increased flash flood severity for the same MRMS/FLASH value. The exception to this trend was the Unit Streamflow product, where only a few bins showed a notable difference in severity. For the other products, the difference in flash flood severity ranged from about 0.5 to about 1.0 across the evaluated bins. Based upon this difference, the difference between impacts for the same MRMS/FLASH values in urban and suburban areas compared to rural areas could be up to an entire category of warning (Flash Flood Warning Base impacts in more urban areas and Flood Advisory impacts in rural areas).

Evaluating the effect of soil moisture on flash flood severity was more difficult. Using USGS Streamflow as a proxy for soil moisture did not yield the expected results, except for with the 1-hr Radar-Only QPE product. Many bins of MRMS/FLASH values showed no change or even decreases in flash flood severity with increased streamflow percentiles (and assumed increased soil moisture values). Because flash flood processes are most closely related to near surface soil moisture, USGS streamflow percentile, which is generally related to total column soil moisture and groundwater levels, may not be a good proxy for soil moisture in this application. The period of record from which flash flood cases were selected also included extended periods of time where soil moisture and river levels were above average. This caused the majority of cases - even with null cases added - to have elevated total column soil moisture and streamflow levels. Using CPC-modeled soil moisture also did not provide the expected results, in a similar way to USGS streamflow. Many bins of MRMS/FLASH values showed a noticeable decrease in flash flood severity. The very high cutoff for group 1 and group 2 values (95th percentile soil moisture) was evaluated next to determine if it could be a contributing factor. It was assumed that group 1, covering values from the 0th percentile to the 94th percentile, covered such a large range of situations compared to group 2, covering the 95th to 100th percentile, that no difference in flash flood severity could be easily discerned. Changing the cutoff for group 1/group 2 to lower percentile values did not provide significantly different results, with group 2 not showing a consistent increase in severity values for the same MRMS/FLASH value. It should be noted, however, that lowering the group1/group2 cutoff value drastically changed the ratio of cases in group 1 versus group 2, reducing the meaningfulness of the results. It is possible that high CPC soil moisture values skewed to the higher end of the distribution among available flash flood cases are affecting the comparison, similar to the issues with USGS streamflow. Another potential factor is related to the Max GFFG Ratio and Unit Streamflow products themselves. Each of these products is theoretically supposed to be adjusted based upon soil moisture. This adjustment, if representative of actual soil moisture conditions, might be expected to cause no difference between group 1 and group 2 for these products. This would not explain the results for 1-hr Radar-Only QPE or Max QPE ARI, however. Based upon the concerns noted with USGS streamflow and CPC soil moisture, they were excluded from further testing.

Using CREST-modeled soil moisture yielded results much closer to what was expected. It was determined that areas with an average soil moisture saturation of 20% or greater had consistently higher flash flood severity values for the same MRMS/FLASH value than soil moisture saturation of 0-19%. This result is consistent with the fact that the CREST soil moisture product is the best available representation of soil moisture that is most closely related to flash flood processes. The difference in flash flood severity ranged from about 0.5 to about 1.5 across the evaluated bins. Based upon this difference, MRMS/FLASH values that correlate to Flood Advisory-level impacts in areas with <20% soil saturation could potentially correlate with Flash Flood Warning Base-level impacts in areas with >20% soil saturation. The observed difference could be partially affected by the low cutoff for group 1 and group 2 values (20% soil saturation), although very high values (70% or greater) are rarely seen, even during periods of rainfall. Near-surface soil moisture rapidly begins to decline after rainfall.

3.2 POD/FAR/POFD/CSI for the 4-Panel Technique Using Subdivided MRMS/FLASH Values and Existing Thresholds

Tables 4, 5, and 6 indicate the POD, FAR, and CSI when using the 4-Panel Technique for the flash flood cases collected by Lincoln & Marquardt (2023). The lesser-used probability of false detection (POFD) values were also provided for reference. These statistics are based upon the recommended thresholds (Table 2) and are calculated based upon all cases, only cases in group 1, and only cases in group 2, for land cover and soil moisture, as indicated by section 2.4. While there is no consistent trend between the various groups and the full group of cases, values for POD, FAR, POFD, and CSI generally differed between each. Statistics for the Flash Flood Warning Catastrophic (4) severity level were not provided, as the sample size was so small that statistics based upon further subdivision would not be meaningful.

Differences that most closely match expectations occur with flash flood cases broken up by land cover and soil moisture (CREST). For those subdivisions of cases, FAR was often lower with group 2 (urban or high soil moisture) than group 1 (rural or low soil moisture). This behavior was not consistent with all severity levels, and became harder to discern at the Flash Flood Warning Considerable (3) and Flash Flood Warning Catastrophic (4) severity levels. This may be due to the much smaller sample sizes (12 and 2, respectively) at those severity levels. Regardless of the cause, this may indicate difficulty finding separate MRMS/FLASH thresholds for different land cover and soil moisture.

Table 4. POD, FAR, and CSI values for Advisory (1)-level flooding using the 4-Panel Technique, with values broken up by land cover and soil moisture. The first value is the value for all cases, the second value is for group 1, and the third value is for group 2. For an explanation of the different groups of flash flood cases, see section 2.4.

	Land Cover			**Soil Moisture (USGS)**			**Soil Moisture (CPC)**			**Soil Moisture (CREST)**		
POD	0.91	0.91	0.90	0.91	0.88	0.92	0.91	0.89	0.94	0.91	0.86	0.94
FAR	0.35	0.45	0.25	0.35	0.25	0.40	0.35	0.21	0.48	0.35	0.57	0.30
POFD	0.41	0.43	0.38	0.41	0.37	0.42	0.41	0.39	0.40	0.41	0.43	0.37
CSI	0.61	0.52	0.69	0.61	0.68	0.58	0.61	0.72	0.50	0.61	0.40	0.67

Table 5. POD, FAR, and CSI values for Flash Flood Warning Base (2)-level flooding using the 4-Panel Technique, with values broken up by land cover and soil moisture. The first value is the value for all cases, the second value is for group 1, and the third value is for group 2. For an explanation of the different groups of flash flood cases, see section 2.4.

	Land Cover			**Soil Moisture (USGS)**			**Soil Moisture (CPC)**			**Soil Moisture (CREST)**		
POD	0.83	0.83	0.83	0.83	0.91	0.78	0.83	0.85	0.76	0.83	1.00	0.83
FAR	0.42	0.46	0.40	0.42	0.34	0.47	0.42	0.30	0.58	0.42	0.63	0.31
POFD	0.27	0.24	0.32	0.27	0.29	0.27	0.27	0.33	0.22	0.27	0.28	0.26
CSI	0.52	0.49	0.54	0.52	0.62	0.46	0.52	0.62	0.37	0.52	0.37	0.60

Table 6. POD, FAR, and CSI values for Flash Flood Warning Considerable (3)-level flooding using the 4-Panel Technique, with values broken up by land cover and soil moisture. The first value is the value for all cases, the second value is for group 1, and the third value is for group 2. For an explanation of the different groups of flash flood cases, see section 2.4.

	Land Cover			**Soil Moisture (USGS)**			**Soil Moisture (CPC)**			**Soil Moisture (CREST)**		
POD	0.58	1.00	0.50	0.58	0.60	0.57	0.58	0.75	0.00	0.58	0.50	0.75
FAR	0.30	0.33	0.29	0.30	0.40	0.20	0.30	0.25	1.00	0.30	0.00	0.40
POFD	0.02	0.01	0.02	0.02	0.04	0.01	0.02	0.03	0.01	0.02	0.00	0.03
CSI	0.47	0.67	0.42	0.47	0.43	0.50	0.47	0.60	0.00	0.47	0.50	0.50

3.3 Calibrated Thresholds Using Subdivided MRMS/FLASH Values

The calibration script from Lincoln & Marquardt (2023) was run to determine the best combination of MRMS/FLASH values for use with the 4-Panel Technique, as determined by highest CSI value, but for the land cover and soil moisture groups instead of all cases. Based upon the noted differences in flash flood severity when cases were subdivided by land cover and the three different estimations of soil moisture, in combination with the POD, FAR, POFD, and CSI statistics, calibration was only performed for land cover and soil moisture modeled by CREST.

For Flood Advisory (severity level 1), no significant difference in calibrated MRMS/FLASH thresholds was noted for cases subdivided by land cover. For cases subdivided by CREST-modeled soil moisture, there was a notable difference in the 1-hr Radar-Only QPE and Unit Streamflow values that yielded the best CSI. For Flash Flood Warning Base (severity 2), there were notable differences between 1-hr Radar-Only QPE, Max QPE GFFG Ratio, and Unit Streamflow values that yielded the best CSI, when subdivided by both land cover and CREST-modeled soil moisture. For a comparison of statistics for the different subdivided groups of flash flood cases, see Tables 7, 8, and 9.

Table 7. Calibrated threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow that yield the best CSI for Flood Advisory (severity level 1). For comparison purposes, the calibrated and recommended values for all flash flood cases presented by Lincoln & Marquardt (2023) are indicated.

	All Cases		Land Cover		Soil Moisture (CREST)	
	Calibrated	**Recommended**	**Group 1**	**Group 2**	**Group 1**	**Group 2**
1-hr Radar-Only QPE	1.5	1.5	1.4-1.5	1.4-1.5	1.7-1.9	1.1-1.5
Max QPE ARI	1	1	1	1	1	1
Max QPE GFFG Ratio	140	125	140-150	140	150	140
Unit Streamflow	200	200	180-200	100-200	180-220	260-300
POD	0.90	0.91	0.91	0.88	0.77	0.91
FAR	0.30	0.35	0.38	0.22	0.41	0.23
POFD	0.32	0.41	0.33	0.31	0.21	0.24
CSI	0.65	0.61	0.58	0.71	0.50	0.72
Number of Cases Exceeding Severity	87 / 190		35 / 190	52 / 190	22 / 152	34 / 152

Table 8. Calibrated threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow that yield the best CSI for Flash Flood Warning Base (severity level 2). For comparison purposes, the calibrated and recommended values for all flash flood cases presented by Lincoln & Marquardt (2023) are indicated.

	All Cases		**Land Cover**		**Soil Moisture (CREST)**	
	Calibrated	**Recommended**	**Group 1**	**Group 2**	**Group 1**	**Group 2**
1-hr Radar-Only QPE	2.0-2.2	2.0	2.3-2.7	1.9-2.0	1.7-1.9	2.1-2.5
Max QPE ARI	1-2	5	1-6	1	1-6	1
Max QPE GFFG Ratio	140-145	140	110-140	140-150	190	140
Unit Streamflow	230-290	230	360-400	400-460	400	220
POD	0.81	0.83	0.83	0.83	0.91	0.86
FAR	0.34	0.42	0.35	0.25	0.29	0.24
POFD	0.18	0.27	0.15	0.17	0.06	0.19
CSI	0.57	0.52	0.57	0.64	0.67	0.68
Number of Cases Exceeding Severity	59 / 190		24 / 190	35 / 190	11 / 152	29 / 152

Table 9. Calibrated threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow that yield the best CSI for Flash Flood Warning Considerable (severity level 3). For comparison purposes, the calibrated and recommended values for all flash flood cases presented by Lincoln & Marquardt (2023) are indicated. Thresholds and statistics based upon a sample size of fewer than five are likely not useful and are not displayed.

	All Cases	**Land Cover**	**Soil Moisture (CREST)**			
	Calibrated	**Recommended**	**Group 1**	**Group 2**	**Group 1**	**Group 2**
1-hr Radar-Only QPE	2.5	2.5		1.7-1.9	1.7-1.9	
Max QPE ARI	150	125		80-100	35-100	
Max QPE GFFG Ratio	350	325		260-450	210-450	
Unit Streamflow	1100	850		500-540	500-540	
POD	0.58	0.58		0.80	0.83	
FAR	0.12	0.30		0.33	0.17	
POFD	0.01	0.02		0.05	0.01	
CSI	0.54	0.47		0.57	0.71	
Number of Cases Exceeding Severity	12 / 190		2 / 190	10 / 190	6 / 152	4 / 152

3.4 Discussion

The noted difference in flash flood severity with the two subdivided groups of cases generally supports the possibility of using different MRMS/FLASH thresholds depending on land cover and soil moisture conditions. Calibration of these subdivided cases provided different MRMS/FLASH thresholds than those found by Lincoln & Marquardt (2023), which calibrated all cases together, and further supports the possibility of utilizing varying thresholds. It should be noted, however, that the different MRMS/FLASH thresholds determined through calibration did not always follow the expected trends from land cover and soil moisture. Even if calibration more strongly supported the concept of using a different set of thresholds depending on a given situation or rainfall location, implementing this would present a challenge. While it may initially make sense to utilize varying thresholds, the possible benefit of doing so must be weighed against the difficulty of implementing this for operational warnings, and must also be thought of from the perspective of the warning forecaster.

One benefit to the existing implementation of the 4-Panel Technique is the simplicity of use for the warning forecaster. One set of thresholds for each MRMS/FLASH product is used to create one set of colors, each of which represents a flash flood severity level. Using different thresholds based upon, for example, urban or rural land cover could potentially require not just two different procedures but also two different sets of colors, with a forecaster needing to jump between different windows or panes depending on the location of rainfall. This problem expands when you add differences in soil moisture, which then could potentially imply the need for four different implementations of the 4-Panel Technique (rural-wet, rural-dry, urban-wet, urban-dry). This adds not just an IT management burden but also the potential for more work on the part of the warning forecaster and potential confusion between "versions" of the 4-Panel Technique. If the typical difference in flash flood severity levels between, for example, urban and rural land cover were approximately 0.5 or approximately 1.0, a possible solution to mitigate this challenge would emerge. Rather than implementing multiple color tables and multiple 4-Panel Techniques for each situation, warning forecasters could utilize awareness that the suggested flash flood severity (color) could be "off" by a factor of 0.5 or 1.0 when heavy rainfall moves into urban areas. This, after some practice and training, would allow for some quick mental adjustment to the 4-Panel Technique's recommendation when the land cover and soil moisture conditions deviated significantly from the "mean" (calibrated) condition.

Another item of note relates to the Max QPE-to-GFFG Ratio and Unit Streamflow products and their potential difference in flash flood severity with different land cover and soil moisture. In theory, GFFG (and thus, QPE-to-GFFG Ratio) and Unit Streamflow should account for differences in land cover and soil moisture, which would reduce or eliminate any potential difference in flash flood severity between events occurring with different conditions. In practice, it may not be this simple. GFFG's adjustment for land cover has two parts. First, the creation of a threshold runoff value is based upon simple modeling of a 5-year ARI design storm over a particular land cover type (Lincoln 2017). In urban areas, the modeled land cover would generate more runoff (similar to reality), leading to a higher threshold. This would prevent urban areas, which generate significantly more runoff from the same rainfall than rural areas, from hitting threshold runoff (and GFFG threshold) values much more easily/frequently. Second, some implementations of GFFG in the NWS provide an additional adjustment to GFFG during the final steps of creating the values in operations (separate from when threshold runoff is calculated during development) which artificially lowers GFFG values in urban areas, or even sets them to a fixed level based upon local experience. Unit Streamflow is similar in that it is based upon hydrologic modeling and should provide different values depending upon where rainfall occurs. Changes in soil moisture are considered in this modeling, based upon the CREST-modeled soil moisture reviewed in this study, and different land cover also affects the rainfall-runoff relationship. In practice, this would generally mean an increase in streamflow (and thus, the streamflow divided by upstream area, or unit streamflow) in urban areas or areas with higher soil moisture.

When infrastructure is designed, at least in recent times, hydrologic modeling is used with a selected rainfall design storm (2-year ARI, 5-year ARI, etc.) to determine the appropriate specifications. In practice, this means that for the same design storm, infrastructure would generally need to be built larger in urban areas compared to rural areas. Because GFFG takes into account land cover when modeling the threshold for initial development, one might expect the same GFFG ratio to imply the same flash flood severity regardless of land cover. Conversely, because Unit Streamflow is not based upon a ratio of a threshold, one might expect a higher value in urban areas for the same flash flood severity seen in rural areas. This behavior was not necessarily observed with the available flash flood cases. Flash flood severity often changed for the same Max QPE-to-GFFG Ratio bins, and flash flood severity often stayed the same for the same Unit Streamflow bins. The available data and analysis are not robust enough to make strong conclusions from differences between expected and observed behavior, especially for higher-severity flood events. There remain uncertainties in the subjective methodology used to determine flash flood severity, the spatial coverage of flash flood events differed significantly between events, and multiple events cover areas of varying land cover and soil moisture, making selection of an average value or “predominant” value potentially problematic.

Contrary to expectations, when looking at Flash Flood Warning Base (severity level 2), the calibrated threshold for Max QPE-to-GFFG Ratio increased for land cover group 2 (suburban and urban) compared to group 1 (rural), and the calibrated threshold for 1-hr Radar-Only QPE increased for soil moisture group 2 (wetter) compared to group 1 (drier). Calibrated thresholds also did not always change in a consistent way between Flood Advisory (severity level 1) and Flash Flood Warning Base (severity level 2). One difficulty with interpreting the results from calibration is that there are many combinations of different MRMS/FLASH values that yield similar POD, FAR, and CSI values, and there can be different combinations of FAR/POD values that yield very similar CSI values. Although calibration can assist with determining recommended thresholds, calibration may need to be used along with other methods.

The distribution of flash flood severity values for each subdivided group was also reviewed. For cases subdivided by CREST-modeled soil moisture, group 2 (wetter) had higher severity values at the 50th and 75th percentile levels, and a higher mean value. For cases subdivided by land cover, group 2 (suburban & urban) also had higher severity values at the 50th and 75th percentile levels, and a higher mean value. The difference in flash flood severity ranged from 0.0 to about 1.0, with a typical value near 0.5. While the exact difference in severity may vary based upon the range of values (bins) chosen, on the whole, severity was generally higher for cases with higher soil moisture and rain events occurring in more urbanized areas. While it is possible that some of this difference could be due to the available cases (random chance), it also generally supports the idea of different MRMS/FLASH values based upon land cover and soil moisture, especially for 1-hr Radar-Only QPE and Max QPE ARI.

3.5 Recommendations for Using the 4-Panel Technique with Varying Land Cover and Soil Moisture

Viewing all available information suggests that the best thresholds for MRMS/FLASH products vary, but typically by small amounts, based upon land cover and soil moisture. To prevent confusion on the part of the warning forecaster and to limit IT workload related to using varying sets of thresholds, the best recommendation would be for a warning forecaster to make adjustments in real time, by mentally adjusting the threshold for 1-hr Radar-Only QPE and Max QPE ARI downward by up to half a severity level (half a color) when the heavy rainfall is occurring in urban areas or areas with very high CREST-modeled soil moisture, and upward by up to half a severity level when the heavy rainfall is occurring in very rural areas or areas with very low CREST-modeled soil moisture. An example of how to apply this recommendation to a situation where CREST-modeled soil moisture values are very high across the region is illustrated by Figure 13. Although the best threshold for QPE-to-GFFG Ratio and Unit Streamflow may also vary, the conflicting information currently available, and the fact that these products should theoretically take land cover and soil moisture into account (at least somewhat), suggests that warning forecasters should typically keep these warning thresholds near the recommended values. Recommended thresholds for MRMS/FLASH products at the Flood Advisory (1), Flash Flood Warning Base (2), and Flash Flood Warning Considerable (3) severity levels for use with the 4-Panel Technique are provided by Tables 10, 11, and 12.

It is possible that the warning forecaster will encounter situations that do not exactly meet the criteria provided here. Note that the values in Table 10, 11, and 12 provide adjustments to thresholds based upon *both* land cover and soil moisture deviating from typical by a *set* amount. In certain situations, the warning forecaster will need to make further “on the fly” mental adjustments to MRMS/FLASH thresholds. Some possible examples include:

- A situation where only land cover or only soil moisture deviates from the “typical” condition. An example would be CREST soil moisture saturation near 20%, but the rainfall occurring in a rural area. In this case, the warning forecaster would consider making an adjustment to 1-hr Radar-Only QPE and Max QPE ARI, but the adjustment would be smaller than that indicated in the tables (smaller than half a severity level or half a color).
- A rare situation where the land cover and soil moisture significantly exceed the criteria indicated for making MRMS/FLASH threshold adjustments. For example, heavy rainfall occurring in a highly urbanized and flood-prone area with CREST soil moisture saturation >60%. In this case, the warning forecaster would consider making an adjustment to 1-hr Radar-Only QPE and Max QPE ARI that is slightly larger than indicated in the tables.

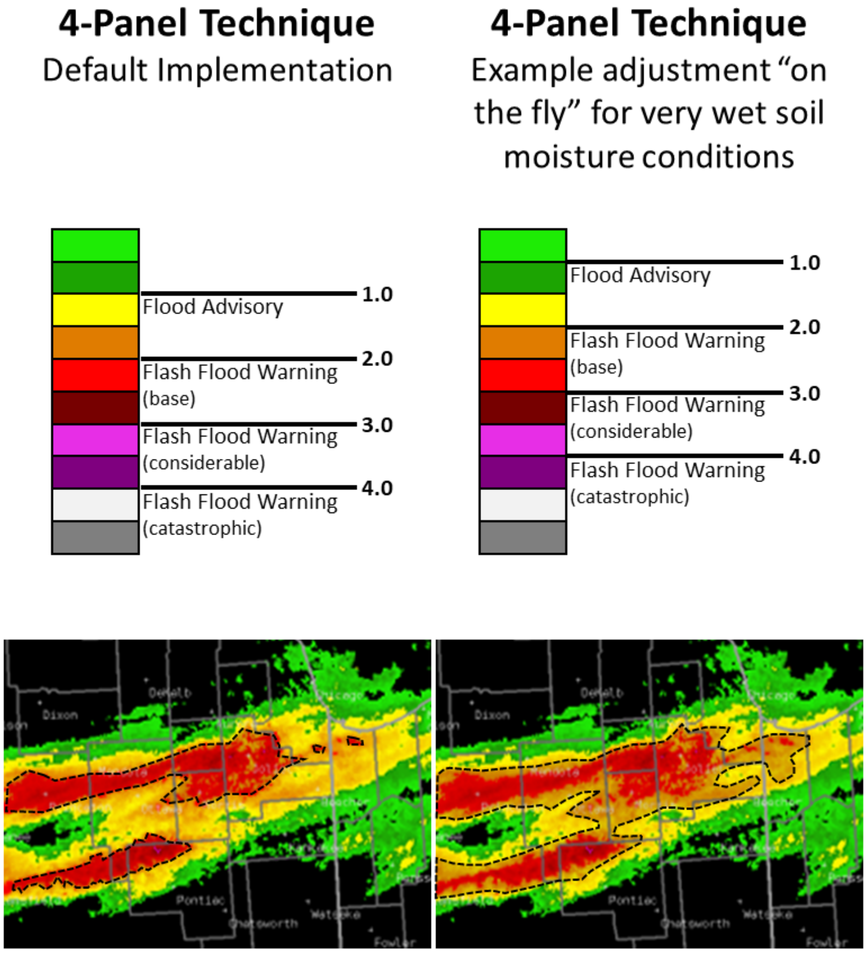

Figure 13. Example "on the fly" adjustment to usage of the 4-Panel Technique based upon soil moisture conditions. In this hypothetical example, very wet soil moisture conditions modeled by CREST exist across the region. In this illustration, the warning forecaster mentally adjusts the threshold for Flash Flood Warning Base (severity level 2) downward by up to 0.5 severity (half a color). For this example panel (representing either 1-hr Radar-Only QPE or QPE ARI), more areas are now indicated as potentially experiencing flash flooding of the given magnitude, and this potentially increases the number of panels at that severity level for a given location. As a reminder, three of four panels indicating the same severity level at the same location is considered a recommendation for a given flood hazard when using the 4-Panel Technique.

Table 10. Recommended Flood Advisory (severity level 1) threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow, based upon differences in land cover and soil moisture, when used as part of the 4-Panel Technique.

	Typical Thresholds	**Rural Land Cover and CREST Soil Moisture <15% Saturation**	**Urban Land Cover and CREST Soil Moisture >25% Saturation**
1-hr Radar-Only QPE	1.50	1.75	1.25
Max QPE ARI	1	2	1
Max QPE GFFG Ratio	125	125	125
Unit Streamflow	200	200	200

Table 11. Recommended Flash Flood Warning Base (severity level 2) threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow, based upon differences in land cover and soil moisture, when used as part of the 4-Panel Technique.

	Typical Thresholds	**Rural Land Cover and CREST Soil Moisture <15% Saturation**	**Urban Land Cover and CREST Soil Moisture >25% Saturation**
1-hr Radar-Only QPE	2.00	2.25	1.75
Max QPE ARI	5	25	2
Max QPE GFFG Ratio	140	140	140
Unit Streamflow	230	230	230

Table 12. Recommended Flash Flood Warning Considerable (severity level 3) threshold values for 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow, based upon differences in land cover and soil moisture, when used as part of the 4-Panel Technique.

	Typical Thresholds	**Rural Land Cover and CREST Soil Moisture <15% Saturation**	**Urban Land Cover and CREST Soil Moisture >25% Saturation**
1-hr Radar-Only QPE	2.50	2.50	2.25
Max QPE ARI	125	150	75
Max QPE GFFG Ratio	325	325	325
Unit Streamflow	850	850	850

4.0 Conclusions and Future Work

The MRMS and FLASH products 1-hr Radar-Only QPE, Max QPE ARI, Max QPE-to-GFFG Ratio, and Unit Streamflow have now been used for several years to provide useful insights into possible flash flooding. One method for using these products to inform flood-related warning decisions is the 4-Panel Technique, where three out of four products indicating the same severity level is considered a recommendation for a flood hazard, assuming no biases in MRMS QPE. In Lincoln & Marquardt (2023), 190 flash flood cases (including 100 null cases) were reviewed to provide recommendations for Flood Advisory, Flash Flood Warning Base, Flash Flood Warning Considerable, and Flash Flood Warning Catastrophic thresholds. One limitation with this effort was that the documented cases covered a wide range of land cover and soil moisture values, which could limit applicability on a case-by-case basis. The predominant land cover type - rural, suburban, or urban was collected for each of the flood cases, along with three different estimates for soil moisture, including USGS streamflow, CPC-modeled soil moisture, and CREST-modeled soil moisture. Land cover and CREST-modeled soil moisture were selected for further review and calibration. Potential methods for using varied MRMS/FLASH thresholds in operations were discussed, and a recommended method for taking into account land cover and soil moisture was presented. Although these recommendations may improve the usage of the 4-Panel Technique, warning forecasters must also monitor the spatial footprint of MRMS/FLASH values, trends in MRMS/FLASH values, rainfall observations, and meteorological conditions to determine the best time to issue a given flood hazard product. Warning forecasters may also have to make small mental adjustments to the indicated thresholds based upon the situation. Warning forecaster training can reduce confusion and unease with using the 4-Panel Technique, and can increase confidence in the types of adjustments that may need to be made in real time. It is recommended that warning forecast training include using the 4-Panel Technique to assess potential flash flood severity, including situations that deviate from the exact criteria presented in this study.

Future work on this topic could include the collection of more cases, specifically cases which include impacts at the Flash Flood Warning Considerable (3) and Flash Flood Warning Catastrophic (4) severity levels, which could improve recommended thresholds for higher-end events. It remains possible that the optimal thresholds for MRMS/FLASH products differ between regions of the country or even small areas within a given NWS forecast office warning area. Offices in other parts of the country should consider replicating the study of flash flood cases presented in both Lincoln & Marquardt (2023) and this study to ensure that results presented here are applicable.

5.0 Acknowledgements

The author would like to thank Kevin Donofrio, NWS Chicago Science and Operations Officer, for his guidance on both part one and part two of this research effort and for a helpful review of this research paper. The author would like to thank NWS Milwaukee Senior Service Hydrologist Sarah Marquardt for collaboration on the original effort to evaluate the performance of the 4-Panel Technique developed at NWS Chicago. The author would also like to thank Jeff Manion of NWS Central Region Headquarters for his helpful review of both manuscripts in this series.

5.0 References

Lincoln, W.S., 2017: 2017 Update to Gridded Flash Flood Guidance at the Lower Mississippi River Forecast Center. *NWS Technical Report.* https://www.weather.gov/media/lmrfc/tech/2017_LMRFC_GFFG_Update.pdf

Lincoln, W. S., and R. F. L. Thomason, 2018: A preliminary look at using rainfall average recurrence interval to characterize flash flood events for real-time warning forecasting. *J. Oper. Meteor.*, 6, 13–22.

Lincoln, W.S., and S. Marquardt, 2023: MRMS and FLASH Thresholds for Assessing Flash Flood Potential in Realtime for the Western Great Lakes. *NWS Central Region Technical Attachment,* 23-01.

Vergara, H.J., J.J. Gourley, and A. Arthur, 2019: Lessons learned in the application of the MRMS-FLASH system at scales ranging from the river reach to the Conterminous United States for Impact-based Flood Prediction. American Geophysical Union, Fall Meeting 2019, abstract #H43I-2108.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, 97, 621–637.