Geometrical Characterization of Precipitation Patterns

AMIR AGHAKOUCHAK, NASRIN NASROLLAHI, JINGJING LI, BISHER IMAM, AND SOROOSH SOROOSHIAN

Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California

(Manuscript received 1 April 2010, in final form 20 August 2010)

ABSTRACT

Satellite estimates and weather forecast models have made it possible to observe and predict precipitation over large spatial scales. Despite substantial progress in observing patterns of precipitation, characterization of spatial patterns is still a challenge. Quantitative assessment methods for spatial patterns are essential for future developments in prediction of the spatial extent and patterns of precipitation. In this study, precipitation patterns are characterized using three geometrical indices: (i) a connectivity index, (ii) a shape index, and (iii) a dispersiveness index. Using multiple examples, the application of the proposed indices is explored in pattern analysis of satellite precipitation images and validation of numerical atmospheric models with respect to geometrical properties. The results indicate that the presented indices can be reasonably employed for a relative comparison of different patterns (e.g., multiple fields against spatial observations) with respect to their connectivity, organization, and shape.

1. Introduction

Spatial patterns in precipitation fields are fundamental to hydrologic modeling and streamflow analysis. Many studies highlight the importance of precipitation space–time variability (Fiener and Auerswald 2009; Haile et al. 2009; Corradini and Singh 1985), which has been proven to affect the quality of runoff predictions (Goodrich et al. 1995; Schuurmans and Bierkens 2007). Historically, most studies focus on temporal patterns in precipitation data (see Grayson and Blöschl 2001b and references therein). By nature, distributed hydrological modeling requires some assumptions regarding the spatial heterogeneity of precipitation. In practical applications, the variability of precipitation is often described by the error in the magnitude of precipitation in space and time. In addition to the magnitude of precipitation, the general pattern and its location are also important (Foufoula-Georgiou and Vuruputur 2001; Sorooshian et al. 2010, manuscript submitted to *Bull. Amer. Meteor. Soc.*). One interesting issue regarding the spatial variability is the extent to which precipitation patterns affect hydrologic predictions. So far, patterns and structures in precipitation fields are the least researched mainly because of the lack of areal representation of precipitation.

In recent years, weather radar systems and satellites have provided detailed information on spatial patterns of precipitation that were not available from traditional point measurements. Even when high-resolution data were available, the characterization and numerical representation of patterns are still challenging issues (Grayson and Blöschl 2001a).

There are a myriad of applications for a method of numerical–geometrical representation of patterns. For example, a meteorologist may want to quantitatively compare the predicted rainfall fields obtained from a weather forecast model with the actual radar–satellite observations. Zepeda-Arce et al. (2000) proposed several statistical measures that can describe how well the variability and organization of the predicted fields matches that of the observed fields. The study confirms that such statistical measures are very informative and may lead to future advancements in atmospheric models. While pixel- and event-based statistical methods offer invaluable information, they do not provide any information on errors resulting from the displacement of spatial patterns (see Marzban and Sandgathe 2006 and Brown et al. 2004 for details). In addition, pixel-based validation approaches have the tendency to overpenalize

Corresponding author address: Amir AghaKouchak, Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, CA 92617.
E-mail: amir.a@uci.edu

DOI: 10.1175/2010JHM1298.1

© 2011 American Meteorological Society

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

displacement errors, which may result in a low score for a good displaced forecast. The increasingly popular object-based verification methods intend to address this problem by focusing on the ability of models to capture attributes such as the shape, volume, and distribution of precipitation objects. Ebert and McBride (2000) and Ebert and Gallus (2009) introduced the concept of contiguous rain area (CRA) to decompose the total error into volume, pattern, and displacement error. Brown et al. (2004) defined composite rain objects as adjacent rain objects that may appear as parts of the same rain system and developed a method to delineate such composite objects. Subsequently, they computed conventional verification measures for each matched pair (observed–estimated) of composite objects. The method was later revised and extended by Davis et al. (2009) and termed the Method for Object-Based Diagnostic Evaluation (MODE), which relies on the distance separating the objects' centroids to match objects in observed–estimated spatial fields, along with a temporal criterion to determine whether objects identified in two time steps belong to the same rain system as it evolves. Another class of object-oriented methods focuses on the entire verification domain instead of individual precipitation objects. Among others, Wernli et al. (2008) introduced a domain-based object-oriented approach that takes into account several components of precipitation fields in the verification domain but does not deal with individual pairs of matched single or composite objects. While object-oriented methods are vital for reliable validation analysis, they do not address geometrical features of precipitation objects.

Another application of the geometrical representation of patterns could be validation of satellite estimates against radar observations. The majority of satellite and radar validation studies are limited to statistical analysis [e.g., Villarini et al. (2009); Tian et al. (2009); AghaKouchak et al. (2011); Tabary et al. (2007)]. However, geometrical properties of precipitation may also have a physical significance in practical applications. Moreover, many studies address stochastic simulation of spatial fields (e.g., AghaKouchak et al. 2010a and references therein); however, it is essential to develop methods of testing simulated spatial patterns against spatial measurements (Grayson and Blöschl 2001b). So far, only a few studies are devoted to comparisons of simulated and measured patterns (e.g., Bronstert and Plate 1997; Welch et al. 1988a,b; Gilleland et al. 2009; Whelan and Anderson 1996; Davis et al. 1995; Blöschl and Kirnbauer 1991; Ahijevych et al. 2009).

Prior to an introduction to the methodology, it is essential to define patterns. In mathematics, a pattern is often defined as a process or any sequence of numbers that may be described by a mathematical function or stochastic process (Devlin 1994; Resnik 1999). Fractals (Mandelbrot 1983; Falconer 1987) are also a scale-invariant mathematical representation of patterns that are commonly used in hydrological applications (e.g., Lovejoy and Mandelbrot 1985; Schertzer and Lovejoy 1989). In this study, however, a geometric framework for the description of rainfall patterns is used that is solely based on the geometry of the rainfall pixels that are above a certain threshold. The presented approach utilizes the following simple geometrical concepts: 1) a minimum theoretical perimeter; 2) a convex hull; 3) an actual perimeter; 4) a number of isolated structures; and 5) an area of isolated clusters to characterize patterns. Using these geometrical properties, a connectivity index between 0 and 1 is defined, where 1 indicates a connected (organized) structure, whereas 0 implies an unorganized (random) structure. Furthermore, a shape index is introduced that identifies the departure of a pattern from a virtual pattern of the same area with the minimum theoretical perimeter. The latter can be used to differentiate between a strip-shaped and a circle-shaped precipitation pattern. An area index is defined to describe the degree of dispersiveness of the patterns. These indices are meant to be used for relative comparison of multiple patterns (e.g., radar–satellite images) with respect to their shape, dispersiveness, and connectivity (e.g., connected rainfall pixels in a satellite image) and not their values (e.g., rain rates). For example, one can compare the connectivity (organization) of multiple images with that of the reference image. In the subsequent sections, a number of examples are provided to illustrate the application.

The paper is organized into four sections. After the introduction, limitations of a well-known method of pattern characterization are discussed. Then, the geometrical indices are introduced and validated using stochastically generated random fields. The third section is devoted to the application of the proposed indices to pattern analysis of satellite precipitation fields and validation of weather prediction models. The fourth section summarizes the conclusions and remarks.

2. Characterization of patterns

In digital topology, one important property of an image is its connectivity, which shows how the structures are interconnected. One fundamental geometrical property of an image (here, precipitation field) is the number of isolated structures or clusters (Vogel 2002). Another basic property of an image is the number of redundant connections, which is defined as the number of holes within the pattern (Lin et al. 2007). The Euler

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

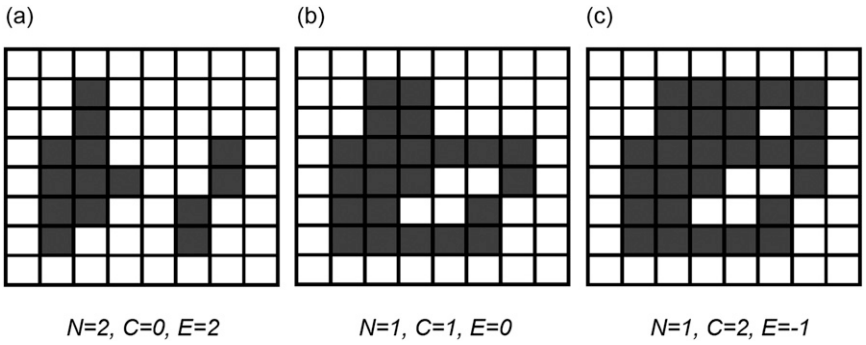


FIG. 1. Euler number for three patterns: (a) $E = 2$, (b) $E = 0$, and (c) $E = -1$.

number combines these two properties to produce a measure of connectivity (Vogel 2002):

$$E = N - C, \tag{1}$$

where E = the Euler number, N = the number of isolated structures (clusters), and C = the number of redundant connections.

The Euler number, alternatively known as the Euler characteristics, ranges from $-\infty$ to $+\infty$. A positive Euler number indicates a poorly connected pattern, whereas a negative E implies a more connected pattern. Figure 1 illustrates how the Euler number describes the connectivity of patterns. In Fig. 1a, two isolated structures (N) are distinguishable with no redundant connections, indicating the Euler number of 2 [see Eq. (1)]. In Fig. 1b, one isolated structure and one hole (C) can be identified ($E = 0$). One can see that the Euler number for Fig. 1b, which is obviously more connected, is less than Fig. 1a. In Fig. 1c, the Euler number is even smaller ($E = -1$), indicating a more connected pattern than the other two figures. Notice that this measure of connectivity is defined for values above or below a given threshold. For rainfall patterns, a hole (C) is a set of no-rain pixels (or rainfall below a certain threshold, e.g., 1 mm hr^{-1}) enclosed by a connected set of nonzero pixels (or rainfall above a certain threshold). The Euler number has been widely used in various fields including image processing, pattern recognition, and the characterization of a network of pores in porous media (e.g., Chen 1988; DeHoff 1987; Mecke and Wagner 1991; Bieri and Nef 1985; Vogel 2002).

One limitation of this measure of connectivity is that if there are a large number of holes or redundant connections in patterns, the Euler number may result in a misleading assessment of pattern connectivity. This issue is illustrated in Figs. 2a–c using three actual radar images obtained from the Memphis Next Generation Weather Radar (NEXRAD) station (KNQA site, 75 km × 75 km images, 3 May 2002). The figures show three precipitation patterns of rainfall rates above the threshold of 1 mm hr^{-1}. A visual comparison confirms that Fig. 2b is more connected than Fig. 2a, however, the Euler number values indicate vice versa. Furthermore, Figs. 2b and 2c, which have the same Euler number, are significantly different from a geometrical viewpoint. These examples indicate the Euler number may not be sufficient to characterize rainfall patterns.

a. Connectivity index

To overcome the above limitations, the connectivity index (C_{index}) is defined as

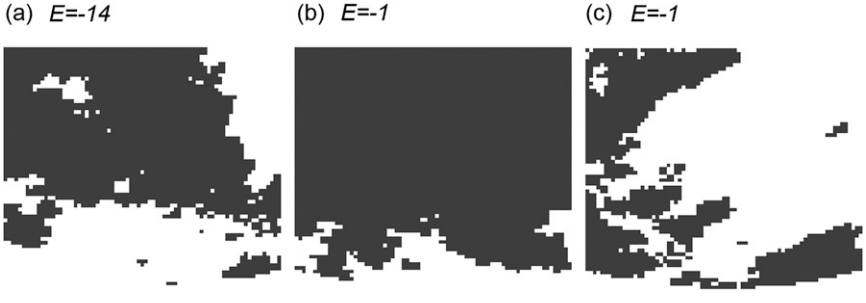


FIG. 2. The Euler number for three radar rainfall images: (a) $E = -14$, (b) $E = -1$, and (c) $E = -1$.

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

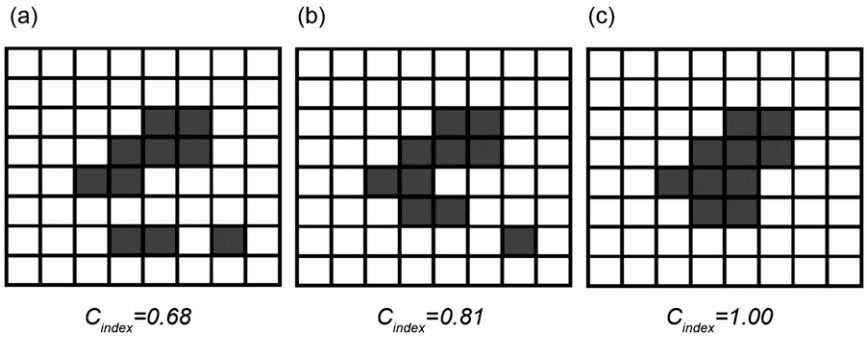


FIG. 3. The connectivity index (C_{index}) for three patterns.

$$C_{\text{index}} = 1 - \frac{\text{NC} - 1}{\sqrt{\text{NP}} + \text{NC}}, \qquad (2)$$

where, C_{index} = the connectivity index, NP = the number of (nonzero) pixels, and NC = the number of isolated clusters.

The connectivity index [Eq. (2)] is calculated based on the number of isolated clusters (NC), the total number of nonzero pixels (or pixels above a given threshold), and the dimension of the domain. Figure 3 provides C_{index} values for three patterns. As shown, the pattern in Fig. 3a has a less connected structure than those of Figs. 3b and 3c. Furthermore, Fig. 3c is the most connected pattern among the others. This is confirmed with the connectivity indices (C_{index}) of 0.68, 0.81, and 1.00 for Figs. 3a, 3b, and 3c, respectively. For precipitation fields presented in Fig. 2, the connectivity indices are computed as (i) 0.95, (ii) 0.99, and (iii), 0.89, respectively. The computed connectivity indices confirm that Fig. 2b is more connected than Figs. 2a and 2c. Unlike the Euler number that ranges from $-\infty$ to $+\infty$, the C_{index} is bounded between 0 and 1 as long as there is at least one nonzero pixel. It should be noted that this index is defined to compare multiple fields with respect to a reference image (e.g., simulated versus observed) and is not designed to compare patterns with significantly different spatial scales and/or number of pixels.

The performance of the C_{index} is evaluated using multiple stochastically generated patterns with a different range (r) values. The range of a spatial structure is the distance at which the random function (process) becomes uncorrelated. That is, 2 points further apart than the range of the function will be considered as statistically independent. Figs. 4a–d display 4 patterns generated with the following variogram ranges (r): 5, 15, 70, and 100 and unit sills (see Deutsch and Journel 1998 for details). As one expects, the patterns with larger ranges (correlation distance) exhibit higher levels of connectivity and organization (e.g., compare Figs. 4a and 4d). In Fig. 5, the solid lines show the number of isolated structures (Fig. 5a) and the connectivity indices (Fig. 5b) of the patterns shown in Fig. 4. One can see that as the number of connected structures decrease (see the solid line in Fig. 5a), the connectivity index increases (see the solid line in Fig. 5b). As expected, the connectivity index of the pattern shown in Fig. 4d is higher than the other three, whereas the pattern shown in Fig. 4a exhibits the least connectivity among the others (see Fig. 5b). This is consistent with the ranges used for the simulation of the fields (Fig. 4), which indicates the C_{index} can be used to reasonably compare multiple patterns in terms of connectivity and organization. The robustness of this index is evaluated by generating 100 random patterns with the same correlation ranges mentioned above (5, 15,

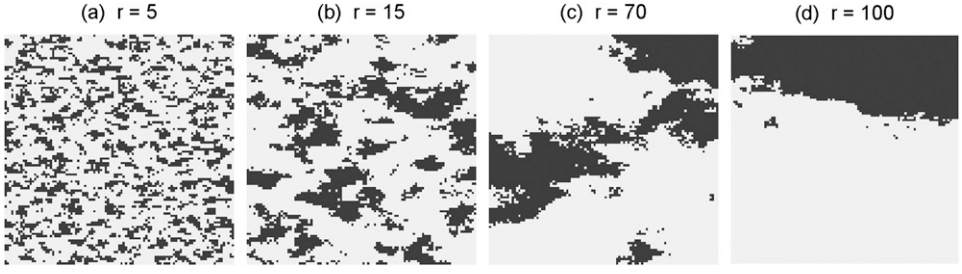


FIG. 4. Stochastically generated patterns with different correlation ranges: (a) $r = 5$; (b) $r = 15$; (c) $r = 70$; and (d) $r = 100$.

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

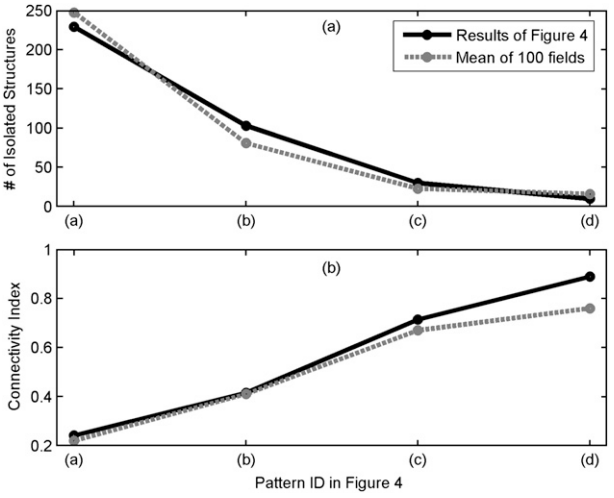


FIG. 5. (a) The number of isolated structures and (b) the connectivity indices of the patterns shown in Fig. 4.

70, and 100). The dashed lines in Figs. 5a and 5b represent the mean number of isolated structures and the mean connectivity indices of all 100 randomly generated patterns, respectively. One can see that the connectivity indices are stable and, as expected, the mean values of indices increase as the correlation ranges increase.

b. Shape index

Both the Euler number and C_{index} provide some information regarding the connectivity of patterns, and they do not offer any additional information on the geometry of clusters. For example, Figs. 6a–c display three patterns with the same Euler number and C_{index} ($E = 1$, $C_{\text{index}} = 1$). To characterize rainfall patterns, however, one may also need additional information on the geometry of patterns (e.g., to distinguish between strip-shaped and circular patterns). The following simple shape index (S_{index}) provides additional information on the geometrical shape of patterns:

$$S_{\text{index}} = \frac{P_{\min}}{P}, \tag{3}$$

where S_{index} = the shape index, $P_{\min}$ = the theoretical minimum perimeter, and P = the actual perimeter of the pattern.

For a given area, the theoretical minimum perimeter is achieved when the pattern is closest to a circle. This issue is illustrated in Fig. 7 using three 8-pixel patterns. As shown, Fig. 7c exhibits the least perimeter among the others. The theoretical minimum perimeter of an n-pixel pattern can be obtained as (for proofs and details, the interested reader is referred to Yackel et al. 1997; Donaldson 2000)

$$P_{\min} = \begin{cases} 4 \times \sqrt{n} & \text{if} \quad \lfloor \sqrt{n} \rfloor = \sqrt{n} \\ 2 \times (\lfloor 2 \times \sqrt{n} \rfloor + 1) & \text{if} \quad \lfloor \sqrt{n} \rfloor \neq \sqrt{n} \end{cases}, \tag{4}$$

where $\lfloor x \rfloor$ = floor function (the largest integer not greater than x).

Using Eq. (4), $P_{\min}$ for the patterns shown in Fig. 7 can be obtained as $12 \times [2 \times (5 + 1)]$. Consequently, the shape indices (S_{index}) for Figs. 7a–7c can then be derived as 0.56, 0.75, and 1.0, respectively. Using this index, some characteristics of precipitation patterns can be recognized. For example, in Figs. 6a–6c with the same Euler number and C_{index} of 1, S_{index} can be obtained as $12/32 = 0.38$, $12/22 = 0.54$, and $12/12 = 1.00$, respectively. As shown, the index (S_{index}) is significantly different for a strip-shaped (Fig. 6a) than a square-shaped (Fig. 6c) pattern.

In the above examples, one shape index is derived for each image. Some geometrical properties such as area and shape index can be estimated for all clusters within each image separately (if the image includes multiple clusters). The histogram of the areas/indices may provide additional geometrical information. For example, one can plot the histogram of areas of clusters (clusters area–frequency graph) to investigate what percentage of clusters exceed a certain area. For the 100 stochastically generated fields (r: 5, 15, 70, and 100) discussed above (see section 2a), Fig. 8 displays the number of clusters that fall in each deciles (10 quantiles) of area. Since the number of fields and thus the number of clusters are large, the natural log of the number of clusters is plotted

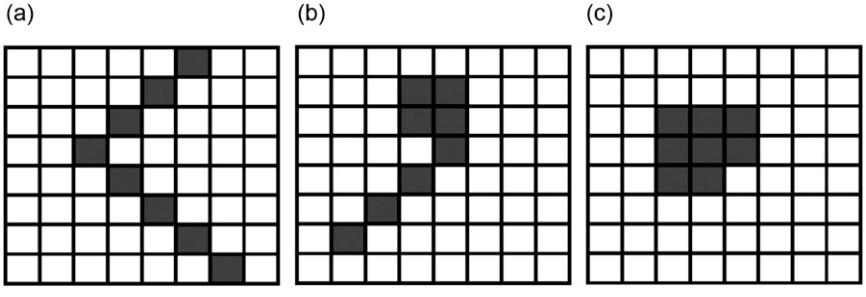


FIG. 6. Three patterns with similar Euler numbers and connectivity indices of 1.

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

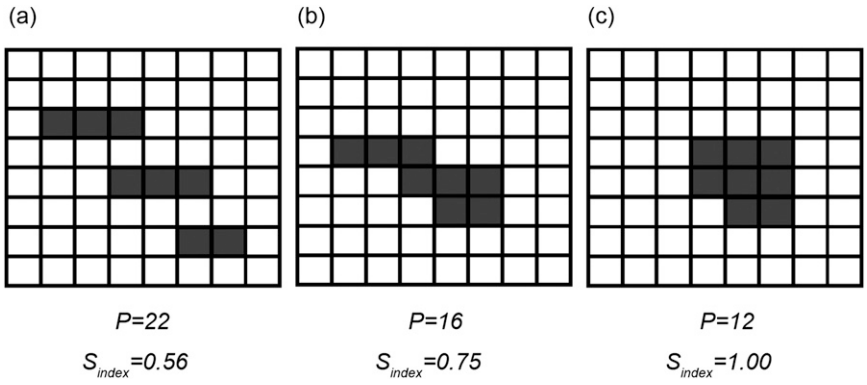


FIG. 7. The perimeter and S_{index} of 3 patterns that consist of 8 pixels with a $P_{\min} = 12$: (a) $P = 22$, (b) $P = 16$, and (c) $P = 12$.

versus the deciles of area. The figure indicates that more large clusters are present in the fields with larger r (dashed lines). On the other hand, the patterns of fields generated with $r = 5$ and $r = 15$ include a higher number of small clusters (solid lines in Figs. 8a and 8b) compared to the fields generated with $r = 70$ and $r = 100$ (dashed lines in Figs. 8a and 8b). In general, this graph may offer further information on the clusters within a given image.

c. Area index

Another geometrical indicator, termed as area index $\text{Area}_{\text{index}}$, is defined based on the concept of convex hull. The term convex hull refers to the boundary of the minimal convex set containing a finite set of points in space (see the solid dashed lines in Fig. 9). The $\text{Area}_{\text{index}}$ is described as the ratio of the actual area of pattern divided by the area of its convex hull:

$$A_{\text{index}} = \frac{A}{A_{\text{Convex}}}, \tag{5}$$

where A = the area of the pattern, and A_{Convex} = the area of the convex hull.

Notice that if the center of each pixel is used for estimation of the convex hull, the following equation should be used instead of Eq. (5):

$$A_{\text{index}} = \frac{A}{A_{\text{Convex}} + 0.5 \times P_{\text{Convex}}}, \tag{6}$$

where P_{Convex} is the perimeter of the convex hull. The second term in the denominator is added to make sure that the convex hull's area includes parts of the pixels that fall outside the boundary, when centers of pixels are used to derive the convex hull. In this equation, unit size is assumed for the pixels, and hence the area of the pattern is equivalent to the number of pixels. When comparing multiple patterns, those with indices closer to 1 are likely to be more structured. On the other hand, a low index implies higher dispersiveness of the pattern. Figures 9a–c display three example patterns and corresponding A_{index} values computed using Eq. (6). One can visually confirm that Figs. 9b and 9c with lower A_{index} values exhibit higher dispersiveness compared to Fig. 9a. Care should be taken when using this index because one single pixel away from the others can significantly change the area of the convex hull. To avoid this situation, one can ignore extraneous pixels (e.g., single pixels or a few pixels that result in a drastic change in the convex area).

3. Application

In this section, the methods discussed above are applied for validation of a weather prediction model and

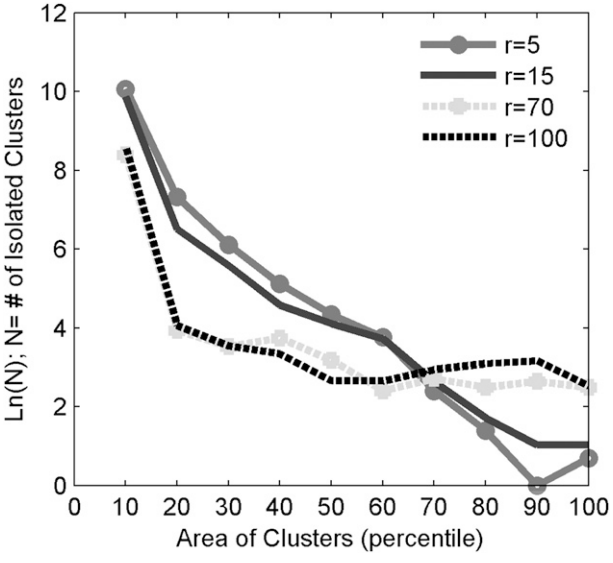


FIG. 8. Natural log of the number of clusters vs the area of clusters (in percentile) for 100 stochastically simulated patterns: (a) $r = 15$ (solid line) and $r = 70$ (dashed line), and (b) $r = 5$ (solid line) and $r = 100$ (dashed line).

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

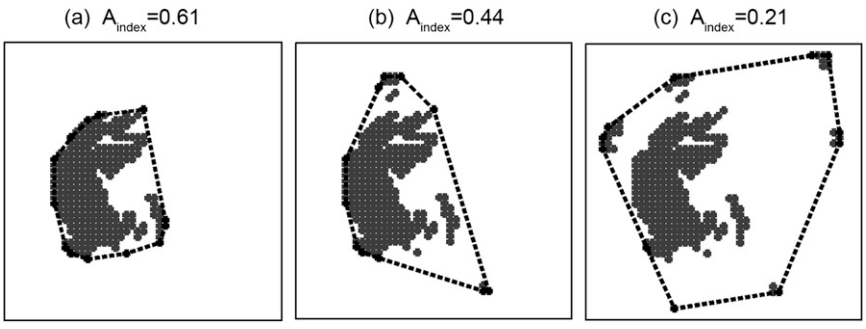


FIG. 9. The A_{index} for three example patterns: (a) $A_{\text{index}} = 0.61$, (b) $A_{\text{index}} = 0.44$, and (c) $A_{\text{index}} = 0.21$.

satellite precipitation estimates against radar data from a geometrical viewpoint. It should be noted that the following examples are provided solely to demonstrate the application of geometrical indices in describing spatial patterns. Statistical validation of weather forecast models and satellite data is not within the scope of this study.

a. Validation of satellite images with respect to precipitation patterns

In this example, different satellite-based precipitation patterns are compared to the stage II radar-based precipitation pattern. Figures 10a–10c show three satellite images, which occurred at 0900 UTC 24 September 2005 during Hurricane Rita, with the spatial and temporal resolutions of 0.25° × 0.25° and 3 h [Fig. 10a: Tropical Rainfall Measuring Mission (TRMM) 3B42, Huffman et al. (2007); Fig 10b: the Climate Prediction Center's morphing technique (CMORPH), Joyce et al. (2004); Fig. 10c: Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), Sorooshian et al. (2000)]. Hurricane Rita was one of the most intense tropical cyclones that made landfall on the U.S. Gulf Coast. Notice that in Figs. 10a–10c, only values above the 50th percentile threshold are considered to avoid small rainfall rates. Figure 10d displays the corresponding stage II image. The stage II data provide estimates of precipitation using a combination of radar and rain gauge measurements. The data is available on the Hydrologic Rainfall Analysis Project (HRAP) grid, with a spatial resolution of approximately 4 km. The stage II data are aggregated in space to match the spatial resolution of TRMM, CMORPH, and PERSIANN data (0.25°). Figures 10e–10h and 10i–10l present similar figures for precipitation values exceeding the 75th and 90th percentiles, respectively. The domain of all figures includes 94 × 47 pixels, each being 0.25° × 0.25°. Table 1 lists the C_{index}, S_{index}, and the area of the patterns, shown in Fig. 10. For the threshold of the 50th percentile, the C_{index} of the CMORPH pattern is closer to that of stage II. With respect to the shape of patterns (S_{index}), the TRMM and CMORPH seem to be closer to the stage II data (see columns 5–7 in Table 1). For a higher threshold of the 75th percentiles, the C_{index} values indicate that the connectivity of PERSIANN patterns is more similar to those of stage II. However, the S_{index} values imply that the TRMM is closer to stage II with regard to the pattern's shape. The C_{index} and S_{index} indices indicate that for rainfall values above the 90th percentile thresholds, the PERSIANN fields are more in agreement with the radar estimates. Comparing the areas of the satellite images, one can see that the PERSIANN and CMORPH fields are preferred over the TRMM patterns for all thresholds (see columns 8–10 in Table 1). The A_{index} values indicate that TRMM and CMORPH images are closer to observations with respect to the dispersiveness of patterns (see columns 11–13 in Table 1). It is worth mentioning that the TRMM data have 1.5-h shift compared to the other datasets. In the above application, where the idea is to compare 2 or more images (e.g., satellite and radar) with each other, the closest satellite image to the reference measurements is used. Obviously, this may have some drawbacks (e.g., unfair comparison) in validation analysis. It should be noted that the above example is provided to illustrate the application of geometrical indices in validation studies and quantitative comparison of multiple patterns. This example should not be considered a comprehensive validation of satellite precipitation data.

b. Validation of atmospheric models with respect to geometrical properties

The Weather Research and Forecasting Model (WRF), which is widely used for operational applications, features various parameterization schemes for modeling weather systems. In the following example, Hurricane Rita, which occurred in September 2005, is modeled using this model with different cumulus parameterization schemes:

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

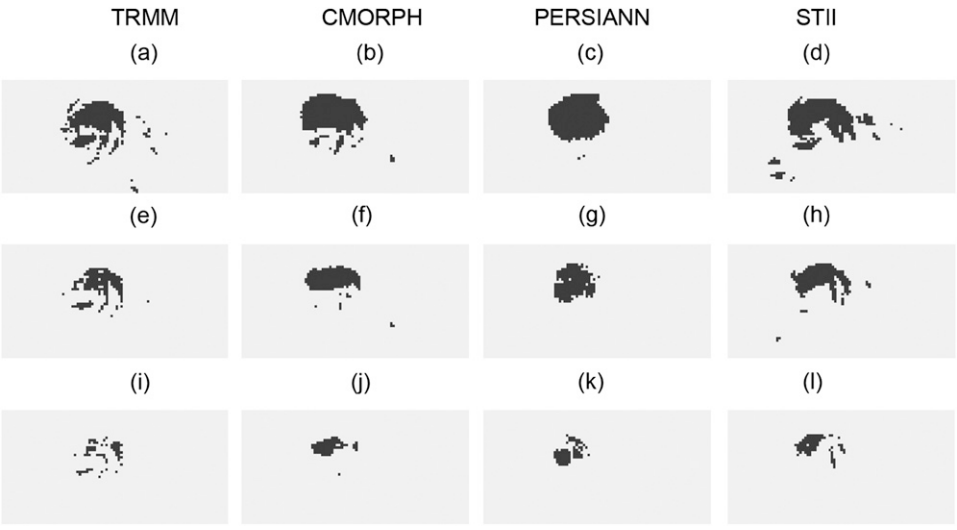


FIG. 10. The TRMM, CMORPH, PERSIANN, and stage II precipitation patterns for rainfall rates above (a)–(d) the 50th, (e)–(h) 75th, and (i)–(l) 90th percentiles.

no cumulus parameterization (NCP); Kain–Fritsch (KF; Kain 2004; Kain and Fritsch 1993; Kain and Fritsch 1990); Betts–Miller–Janjic (BMJ; Betts and Miller 1993); and Grell–Devenyi (GD; Grell and Devenyi 2002). In all simulations, the 6-hourly outputs from the Global Forecast System (GFS), developed by the National Centers for Environmental Prediction (NCEP), are used as the initial and boundary conditions of the model. A nested domain system is used with the grid resolutions of 4 km and 1 km, where the initial and boundary conditions for the high-resolution domain are obtained from simulations of the coarser domain. The simulation starts at 0600 UTC 20 September 2005 and continues until 0600 UTC 26 September 2005. The model settings are based on the Noah land surface model (Chen and Dudhia 2001), Longwave radiation (RRTM, Mlawer et al. 1997), shortwave radiation (Dudhia 1989), the Yonsei University (YSU) planetary boundary layer scheme (Hong and Dudhia 2003), and Lin et al. (1983) microphysics. Figures 11a–11d present the precipitation patterns (0900 UTC 24 September 2005) simulated using the NCP, KF, BMJ, and GD parameterization, respectively (threshold: 50th percentile). Figure 11e shows the corresponding stage II rainfall estimates. Figures 11f–11j and 11k–11o display similar figures for rainfall values above the 75th and 90th percentiles, respectively. The domain of all figures includes 300×180 pixels, each being $0.04° \times 0.04°$. One can see that the observed (Fig. 11e) and simulated (Figs. 11a–11d) patterns are quite different from a geometrical viewpoint. Table 2 summarizes the C_{index}, S_{index}, area, and A_{index} of the patterns, shown in Fig. 11, for rainfall values above the following: the 50th, 75th, and 90th percentile thresholds. As shown, with respect to the area of the storm, using the BMJ cumulus parameterization results in patterns closer to those of the stage II data (see columns 8–10 in Table 2). Comparing S_{index} value confirms that the shape indices of patterns simulated using BMJ and GD schemes are more in agreement with the radar observations (columns 6–7 in Table 2). For instance, the shape indices of patterns simulated with NCP, KF, BMJ, and GD parameterizations are computed as 0.39, 0.4, 0.21, and 0.23, respectively (above the 50th percentile threshold). The same index for radar observations is derived as 0.17, which is closer to those of BMJ and GD schemes. While the area indices of BMJ correspond better to those of radar patterns, the A_{index} values of the NCP and KF schemes deviate significantly from reference observations. For the 75th and 90th

TABLE 1. The C_{index}, S_{index}, and area of the patterns, shown in Fig. 10, for rainfall values above the 50th, 75th, and 90th percentiles.

	C_{index}			S_{index}			Area (km^2)			A_{index}		
Data	50%	75%	90%	50%	75%	90%	50%	75%	90%	50%	75%	90%
TRMM	0.73	0.87	0.75	0.23	0.30	0.24	126 665	63 140	25 795	0.32	0.35	0.24
CMORPH	0.86	0.88	0.96	0.39	0.51	0.53	143 605	71 610	28 490	0.53	0.37	0.38
PERSIANN	0.97	0.94	0.95	0.72	0.49	0.38	144 760	73 150	28 875	0.79	0.77	0.58
Stage II	0.88	0.92	0.91	0.34	0.38	0.37	147 070	73 535	29 260	0.47	0.46	0.39

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

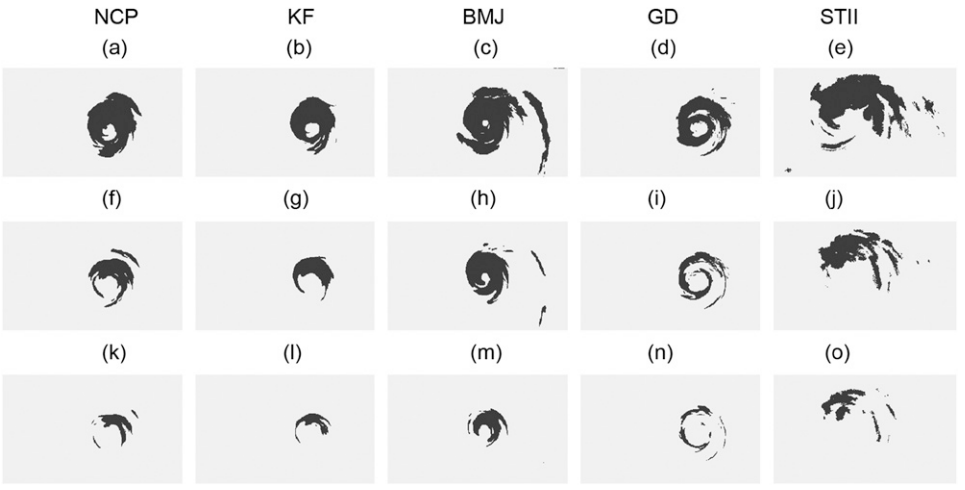


FIG. 11. Simulated Hurricane Rita precipitation patterns for rainfall values above (a)–(e) the 50th, (f)–(j) 75th, and (k)–(o) 90th percentiles. NCP, KF, BMJ, and GD are shown.

percentile thresholds, GD-derived patterns also exhibit a good agreement with radar area indices (see columns 12–13 in Table 2). With respect to the connectivity, although all parameterization schemes show higher C_{index} values than radar data, BMJ and GD are in better agreement with reference patterns. Overall, the C_{index}, S_{index}, and A_{index} values indicate that the BMJ and GD are superior to the NCP and KF (see columns 2–7 and 11–13 in Table 2). This conclusion is based on the geometrical properties of precipitation fields and cannot be generalized. The objective here is not to conduct an extensive validation study but to demonstrate the concept of using geometrical properties for validation purposes. In a thorough validation analysis, other statistical measures and physical details are to be considered that are beyond the scope of this work.

4. Summary and conclusions

In one of the earliest studies on the patterns in hydrology, Grayson and Blöschl (2001b) predicted that, in the future, testing spatial patterns would receive more attention and would provide a quantum advance in the reliability of hydrologic predictions. Recent technological developments in the field of remote sensing have made it possible to observe precipitation (and many other weather variables) over large spatial scales. These data allow researchers to investigate issues that were not possible from point measurements. Despite substantial progress in observing patterns of weather variables, characterization of spatial patterns is still a challenge. While a great deal of effort is put to statistical validation of satellite estimates, uncertainties associated with spatial patterns are not well researched. New models and spatial analysis tools are to be developed for such validation studies.

In addition to the remotely sensed patterns, many numerical weather forecast models also predict precipitation patterns for the near future. However, the accuracy of the predicted patterns from weather forecast models is still an open question. Future improvements in modeling techniques largely depend on comprehensive validation studies. The importance of this issue has motivated research initiatives such as the Spatial Forecast Verification Methods Intercomparison Project, which aims to develop new validation and verification methods for spatial forecasts (see Gilleland et al. 2009 and Ahijevych et al. 2009 for an extensive review and qualitative comparisons of model verification techniques). Without developing quantitative assessment methods for spatial

TABLE 2. C_{index}, S_{index}, and area of the patterns, shown in Fig. 11, for rainfall values above the 50th, 75th, and 90th percentiles.

	C_{index}			S_{index}			Area (km^2)			A_{index}		
Data	50%	75%	90%	50%	75%	90%	50%	75%	90%	50%	75%	90%
NCP	1.00	0.99	0.99	0.39	0.25	0.30	93 104	46 528	18 608	0.83	0.53	0.33
KF	0.99	1.00	1.00	0.40	0.43	0.31	66 656	33 344	13 328	0.75	0.58	0.40
BMJ	0.95	0.96	0.99	0.21	0.27	0.32	136 224	68 128	27 200	0.40	0.33	0.32
GD	0.96	0.97	0.95	0.23	0.15	0.16	74 096	37 056	14 816	0.60	0.39	0.18
Stage II	0.85	0.89	0.93	0.17	0.16	0.19	148 144	73 536	29 248	0.35	0.36	0.24

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

patterns, future improvements in prediction of the spatial extent and patterns of precipitation may be limited.

In this study, connectivity, as a spatial property of patterns, is characterized using a geometrical index (C_{index}). The connectivity index is validated using stochastically generated patterns with different levels of connectivity (organization). A shape index (S_{index}) is introduced that describes patterns quantitatively based on their departure from a virtual pattern of the same area with the minimum theoretical perimeter. Furthermore, an index is introduced to describe the dispersiveness of patterns (A_{index}).

The application of the proposed indices is demonstrated in the validation of various satellite precipitation estimates and a weather prediction model from a geometrical point of view. The results showed that the indices can be used to compare multiple patterns with respect to a reference pattern. In fact, all indices are designed to evaluate the organization, structure, and dispersiveness of one or more patterns (e.g., multiple satellite data) with respect to a reference (e.g., radar image). The indices are not meant to compare fields in different spatial scales (e.g., fields with a significantly different number of pixels and dimensions). For example, a precipitation field with a spatial resolution of 1° should not be compared with the same image with a spatial resolution of 0.25° because the number of pixels in the two fields are significantly different, which may result in unreliable measures of geometrical properties. One possible solution to this limitation is to aggregate the data into the same spatial resolution before comparing geometrical patterns. Additionally, other factors such as domain size may be important for the interpretation of the results. For example, if the domain size is too big it may not be possible to reasonably capture the geometrical features. For geometrical analysis, the optimal size of the domain depends on the object(s) being investigated. If the size of the domain is too big, one can divide the domain to 4, 8, or 16 panels and investigate each panel separately.

In the examples provided here, stage II estimates are used as the reference dataset. Intuitively, unreliable surface precipitation data may strongly affect validation and verification analysis regardless of the choice of the metric (e.g., statistical, geometrical, and object-oriented metrics). It is well known that radar estimates are subject to various types of errors including beam blockage, anomalous propagation, and beam overshooting, among others (Krajewski and Smith 2002). Based on the available data, gauge-adjusted radar estimates are the best area approximation of precipitation (AghaKouchak et al. 2010b). In general, the consequences of radar uncertainties in validation analysis are rather general and are not limited to this study.

In addition to the aforementioned indices, different ways of numerically representing patterns (e.g., cluster area–frequency plot) are discussed. The so-called cluster area–frequency graph can be used when a large number of simulated patterns are compared to spatial observations. This tool can also be useful for the relative comparison of multiple sets of stochastically generated fields with certain statistical properties. This will allow one to find out, for example, which set (of fields) is spatially more connected and/or which set is more random.

Along with other means of statistical validation techniques, the proposed indices can be used for validation of satellite precipitation data. Furthermore, the indices can be employed to evaluate the performance of numerical weather prediction models with respect to geometrical properties of predicted precipitation patterns. Studying patterns of precipitation may also lead to a better understanding of the spatial variability of precipitation. There are various representations of space–time rainfall organization (see Zepeda-Arce et al. 2000 and references therein); however, most models are based on some statistics of precipitation data and hence do not consider the geometrical details of rainfall patterns. In fact, preserving the statistical properties of a process does not necessarily lead to the representation of entire geometrical patterns (Grayson and Blöschl 2001a). One can argue that the inability to capture the geometrical properties of precipitation may limit the quality of precipitation forecasts–estimation. However, detailed studies are required to evaluate the significance of spatial patterns in the uncertainty of precipitation estimates–forecasts.

As mentioned earlier, traditional statistical verification measures summarize pixel-to-pixel differences (e.g., mean square error, probability of detection, and false alarm ratio) between observation and forecasts–estimates but do not address errors related to spatial patterns, displacement, and location of the storm event (Brown et al. 2004; Marzban and Sandgathe 2006). In addition, traditional pixel-based verification measures tend to overpenalize displacement errors, which results in many good forecasts–estimates being overlooked. One advantage of the introduced geometrical indices is that they do not overpenalize good forecasts–estimates because of displacement errors. On the hind side, however, geometrical indices cannot detect displacement and storm location errors. Therefore, one may opt to employ the geometrical indices along with other object-oriented verification methods such as CRA (Ebert and McBride 2000; Ebert and Gallus 2009) and MODE (Davis et al. 2009), among others.

In addition to the validation of weather prediction models and satellite estimates, assessing errors in pattern and location attributes are important for distributed hydrologic modeling [e.g., the effects of connectivity (organization) of precipitation patterns on hydrologic

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

predictions]. That is, how an organized precipitation pattern is different than an unorganized (random) pattern with respect to the output hydrograph of a distributed model. Efforts are underway by the authors to demonstrate whether the geometrical properties of rainfall fields have physical significance in practical applications (e.g., hydrologic modeling). Similar research questions are also considered when reconstructing small-scale precipitation variability from regional-scale averages (statistical downscaling). For reliable rainfall downscaling, it is important to know how geometrical properties of precipitation change when rainfall fields are downscaled to finer scales (e.g., from 0.25° to 0.04°). Future research in this area may answer to what extent the geometrical features of precipitation change with scale, and more importantly, whether the geometrical properties of precipitation change with scale systematically, such that they can be formulated. This information can significantly advance statistical downscaling techniques.

Given the available spatial data, we believe that in the near future more effort will be concentrated to improve spatial analysis techniques by including geometrical properties and other means of pattern representation. We need to emphasize that geometrical properties of precipitation are not limited to those introduced in this paper, and we also do not claim that the presented indices are sufficient to distinguish all geometrical features. Furthermore, the authors acknowledge that the indices may not necessarily distinguish differences between geometrical properties of multiple patterns due to a variety of reasons (e.g., significant differences in numbers of pixels, domain size, and dispersiveness of pixels among others). Defining reliable and meaningful geometrical representations merits a great deal of research. This study was an attempt to contribute to understanding the geometrical properties of precipitation. It is worth pointing out that the presented geometrical indices, similar to other validation metrics, provide the tools to detect potential deficiencies in estimates–simulations with respect to observations. Interpretation of their meanings, however, depends largely on the problem at hand. It is hoped that future advancements in this area lead to better prediction–representation of precipitation from a geometrical point of view.

Acknowledgments. The authors thank the editor and reviewers for their constructive comments and suggestions, which led to improvements in the revised version of the manuscript. The financial support for the authors of this study is made available from NOAA/NESDIS/NCDC (Prime Award NA09NES4400006, NCSU CICS Sub-Award 2009-1380-01) and two NASA Grants NNX09AO67G and NNX08AU78H.

REFERENCES

AghaKouchak, A., A. Bárdossy, and E. Habib, 2010a: Conditional simulation of remotely sensed rainfall data using a non-Gaussian *v*-transformed copula. *Adv. Water Resour.,* **33,** 624–634.

——, ——, and ——, 2010b: Copula-based uncertainty modeling: Application to multi-sensor precipitation estimates. *Hydrol. Processes,* **24,** 2111–2124.

——, A. Behrangi, S. Sorooshian, K. Hsu, and E. Amitai, 2011: Evaluation of satellite-retrieved extreme precipitation rates across the central United States. *J. Geophys. Res.,* **116,** D02115, doi:10.1029/2010JD014741.

Ahijevych, D., E. Gilleland, B. Brown, B. Casati, and E. Ebert, 2009: Application of spatial verification methods to idealized and NWP gridded precipitation forecasts. *Wea. Forecasting,* **24,** 1485–1497.

Betts, A., and M. Miller, 1993: The Betts–Miller scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 24, Amer. Meteor. Soc., 107–121.

Bieri, H., and W. Nef, 1985: A sweep-plane algorithm for computing the Euler characteristic of polyhedra represented in boolean form. *Computing,* **34,** 287–302.

Blöschl, G., and R. Kirnbauer, 1991: Point snowmelt models with different degrees of complexity internal processes. *J. Hydrol.,* **129,** 127–147.

Bronstert, A., and E. Plate, 1997: Modeling of runoff generation and soil moisture dynamics for hillslopes and micro-catchments. *J. Hydrol.,* **198,** 177–195.

Brown, B., and Coauthors, 2004: New verification approaches for convective weather forecasts. Preprints, *11th Conf. Aviation, Range, and Aerospace Meteorology,* Hyannis, MA, Amer. Meteor. Soc., 9.4. [Available online at http://ams.confex.com/ams/11aram22sls/techprogram/paper_82068.htm.]

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585.

Chen, M.-H., 1988: A fast algorithm to calculate the Euler number for binary images. *Pattern Recognit. Lett.,* **8,** 295–297.

Corradini, C., and V. Singh, 1985: Effect of spatial variability of effective rainfall on direct runoff by geomorphoiogic approach. *J. Hydrol.,* **81,** 27–43.

Davis, C. A., B. Brown, R. Bullock, and J. Gotway, 2009: The method for object-based diagnostic evaluation (mode) applied to WRF forecasts from the 2005 SPC spring program. *Wea. Forecasting,* **24,** 1327–1342.

Davis, R., J. McKenzie, and R. Jordan, 1995: Distributed snow process modelling: An image processing approach. *Hydrol. Processes,* **9,** 865–875.

DeHoff, R., 1987: Use of the disector to estimate the Euler characteristic of three-dimensional microstructures. *Acta Stereo.,* **6,** 133–140.

Deutsch, V., and A. Journel, 1998: *GSLIB—Geostatistical Software Library and User's Guide*. Oxford University Press, 384 pp.

Devlin, K., 1994: *Mathematics: The Science of Patterns*. Scientific American Library, 224 pp.

Donaldson, W., 2000: Grid-graph partitioning. Ph.D. dissertation, Computer Science, University of Wisconsin, 140 pp.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

Ebert, E., and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.,* **239,** 179–202.

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020

——, and W. A. Gallus Jr., 2009: Toward better understanding of the contiguous rain area (CRA) method for spatial forecast verification. *Wea. Forecasting,* **24,** 1401–1415.

Falconer, K., 1987: *The Geometry of Fractal Sets*. Cambridge University Press, 180 pp.

Fiener, P., and K. Auerswald, 2009: Spatial variability of rainfall on a sub-kilometer scale. *Earth Surf. Processes Landforms,* **34,** 848–859.

Foufoula-Georgiou, E., and V. Vuruputur, 2001: Patterns and organization in precipitation. *Spatial Patterns in Catchment Hydrology: Observations and Modeling,* B. Grayson and G. Blöschl, Eds., Cambridge University Press, 82–104.

Gilleland, E., D. Ahijevych, B. Brown, B. Casati, and E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting,* **24,** 1416–1430.

Goodrich, D., J. Faures, D. Woolhiser, L. Lane, and S. Sorooshian, 1995: Measurement and analysis of small-scale convective storm rainfall variability. *J. Hydrol.,* **173,** 283–308.

Grayson, R., and G. Blöschl, Eds., 2001a: *Spatial Patterns in Catchment Hydrology: Observations and Modeling*. Cambridge University Press, 404 pp.

——, and ——, 2001b: Spatial processes, organisation and patterns. *Spatial Patterns in Catchment Hydrology: Observations and Modeling,* R. Grayson and G. Blöschl, Eds., Cambridge University Press, Cambridge, 3–16.

Grell, G., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.,* **29,** 1693–1696.

Haile, A., T. Rientjes, A. Gieske, and M. Gebremichael, 2009: Rainfall variability over mountainous and adjacent lake areas: The case of Lake Tana basin at the source of the Blue Nile River. *J. Appl. Meteor. Climatol.,* **48,** 1696–1717.

Hong, S.-Y., and J. Dudhia, 2003: Testing of a new nonlocal boundary layer vertical diffusion scheme in numerical weather prediction applications. Preprints, *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction,* Seattle, WA, Amer. Meteor. Soc., 17.3. [Available online at http://ams.confex.com/ams/84Annual/techprogram/paper_72744.htm.]

Huffman, G., R. Adler, D. Bolvin, G. Gu, E. Nelkin, K. Bowman, E. Stocker, and D. Wolff, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.,* **8,** 3855.

Joyce, R., J. Janowiak, P. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.,* **5,** 487–503.

Kain, J., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.,* **43,** 170–181.

——, and J. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.,* **47,** 2784–2802.

——, and ——, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models. Meteor. Monogr.,* No. 24, Amer. Meteor. Soc., 165–170.

Krajewski, W., and J. Smith, 2002: Radar hydrology: Rainfall estimation. *J. Hydrol.,* **25,** 1387–1394.

Lin, X., J. Ji, and Y. Gu, 2007: The Euler number study of image and its application. *Proc. Second IEEE Conf. on Industrial Electronics and Applications,* Harbin, China, Institute of Electrical and Electronics Engineers, 910–912.

Lin, Y.-L., R. Rarley, and H. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Appl. Meteor.,* **22,** 1065–1092.

Lovejoy, S., and B. Mandelbrot, 1985: Fractal properties of rain, and a fractal model. *Tellus,* **37A,** 209–232.

Mandelbrot, B., 1983: *The Fractal Geometry of Nature*. Freeman, 468 pp.

Marzban, C., and S. Sandgathe, 2006: Cluster analysis for verification of precipitation fields. *Wea. Forecasting,* **21,** 824–838.

Mecke, K., and H. Wagner, 1991: Euler characteristic and related measures for random geometric sets. *J. Stat. Phys.,* **64,** 843–850.

Mlawer, E., S. Taubman, P. Brown, and M. Iacono, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-*k* model for the longwave. *J. Geophys. Res.,* **102,** 16 663–16 682.

Resnik, M., 1999: *Mathematics as a Science of Patterns*. Clarendon Press, 304 pp.

Schertzer, D., and S. Lovejoy, 1989: Generalized scale-invariance and multiplicative processes in the atmosphere. *Pure Appl. Geophys.,* **130,** 57–81.

Schuurmans, J., and M. Bierkens, 2007: Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. *Hydrol. Earth Syst. Sci.,* **11,** 677–693.

Sorooshian, S., K. Hsu, X. Gao, H. Gupta, B. Imam, and D. Braithwaite, 2000: Evolution of the PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.,* **81,** 2035–2046.

Tabary, P., J. Desplats, K. Do Khac, F. Eideliman, C. Gueguen, and J.-C. Heinrich, 2007: The new French operational radar rainfall product. Part II: Validation. *Wea. Forecasting,* **22,** 409–427.

Tian, Y., and Coauthors, 2009: Component analysis of errors in satellite-based precipitation estimates. *J. Geophys. Res.,* **114,** D24101, doi:10.1029/2009JD011949.

Villarini, G., W. Krajewski, and J. Smith, 2009: New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25° 3-hourly estimates over Oklahoma. *J. Geophys. Res.,* **114,** D12106, doi:10.1029/2008JD011475.

Vogel, H., 2002: Topological characterization of porous media. *Morphology and Condensed Matter: Physics and Geometry of Spatially Complex Systems,* Vol. 600, *Lecture Notes in Physics,* Springer, 75–92.

Welch, R., K. Kuo, B. Wielichi, S. Sengupta, and L. Parker, 1988a: Marine stratocumulus cloud fields off the coast of southern California observed using Landsat imagery. Part I: Structural characteristics. *J. Appl. Meteor.,* **27,** 341–362.

——, S. Sengupta, and D. Chen, 1988b: Cloud field classification based upon high spatial resolution textural features. 1. Gray level co-occurrence matrix approach. *J. Geophys. Res.,* **93,** 12 663–12 681.

Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.,* **136,** 4470–4487.

Whelan, M., and J. Anderson, 1996: Modelling spatial patterns of throughfall and interception loss in a Norway spruce (*picea abies*) plantation at the plot scale. *J. Hydrol.,* **186,** 335–354.

Yackel, J., R. Meyer, and I. Christou, 1997: Minimum-perimeter domain assignment. *Math. Program.,* **78,** 283–303.

Zepeda-Arce, J., E. Foufoula-Georgiou, and K. Droegemeier, 2000: Space–time rainfall organization and its role in validating quantitative precipitation forecast. *J. Geophys. Res.,* **105,** 10 129–10 146.

Downloaded from http://journals.ametsoc.org/jhm/article-pdf/12/2/274/4107296/2010jhm1298_1.pdf by guest on 21 August 2020