A Region-Tree-Based Approach for the Verification of Precipitation Forecasts

JINYI HOU AND PING WANG

School of Electrical and Information Engineering, Tianjin University, Tianjin, China

(Manuscript received 13 September 2018, in final form 18 December 2018)

ABSTRACT

This paper describes a region-tree-based precipitation forecast verification approach, which is then used to assess the performance of a numerical model forecast at various scales. In this verification approach, the forecast and observation fields are first represented using region-tree structures. The verification approach consists of three procedures: calculation of spatial structural properties of the forecast and observation trees, matching the forecast tree with the observation tree, and deforming the forecast tree to resemble the observation tree. Based on these procedures, the tree-based verification approach aims at determining the spatial structural differences, displacement errors, and intensity errors between the forecast and the observation at different scales. The behavior of the proposed verification approach is investigated by applications to a standard verification dataset from the spatial forecast verification intercomparison project.

1. Introduction

Verification of numerical weather model forecasts is useful for authorities to rank and select forecasting systems as well as to quantify the quality and uncertainty of forecasts. Forecast verification is also critical for model developers to identify the shortcomings and system errors of models, and thus to improve their models. Numerical models, with increasing resolution, are better able to reproduce features observed at finer scales. This leads to more challenges related to developing forecast verification approaches. The traditional gridpoint-based verification method compares the forecasts and observations for each grid by calculating intensity errors and categorical scores, such as probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). A limitation of the gridpoint-based verification method is that it does not consider the structural information of precipitation fields, and also cannot give the displacement errors of the forecasts.

Recently developed spatial forecast verification approaches focus on comparing spatial structures between gridded forecasts and observations. Gilleland et al. (2009) grouped these approaches into four categories: scale separation (e.g., Harris et al. 2001; Casati et al. 2004; Yano and Jakubiak 2016), neighborhood (e.g., Germann and Zawadzki 2004; Theis et al. 2005; Rezacova et al. 2007;

Roberts and Lean 2008), object based (e.g., Ebert and McBride 2000; Davis et al. 2006a,b; Micheas et al. 2007; Wernli et al. 2008; Gallus 2010; Lack et al. 2010; Duff et al. 2012; Li et al. 2015; Fox et al. 2016; Giannakaki and Martius 2016), and field deformation (e.g., Keil and Craig 2007, 2009; Ziegeler et al. 2012; Han and Szunyogh 2016). The first two categories calculate verification statistics on spatial filtered fields to obtain information about the scale. Specifically, scale separation approaches use bandpass spatial filters (e.g., Fourier, wavelets), whereas the neighborhood approaches use smoothing filters. The last two categories describe how much spatial movement needs to match the forecast to the observation. The difference between these two categories is that the object-based approaches process objects of interest, whereas the field deformation approaches analyze the entire field. There are also some verification approaches that do not belong to any one of these four categories. For example, Lakshmanan and Kain (2010) approximated the observed and forecasted fields by Gaussian mixture models (GMMs), and analyzed the parameters of the Gaussian components to infer forecast errors. Marzban and Sandgathe (2006, 2008) proposed a cluster analysis-based approach that has the ability to perform object-oriented verification on different spatial scales.

This paper proposes a spatial verification approach based on a region-tree analysis method. The basis of the region-tree analysis method is to represent a grid-based field using a region-tree structure, which is a set of

Corresponding author: Jinyi Hou, houjinyi@tju.edu.cn

DOI: 10.1175/MWR-D-18-0325.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

regions that are segmented from the field and are connected according to their overlapping relationships. The region-tree structure (also termed binary partition tree) is commonly used in image processing communities for image segmentation (Salembier and Garrido 2000; Lu et al. 2007; Alonso-Gonzalez et al. 2012), object detection (Vilaplana et al. 2008; Liu et al. 2011; Li et al. 2016), etc. Recently, Hou and Wang (2017) used the region-tree structure to represent radar reflectivity images, and adopted a tree-matching method to solve the problem of storm tracking. In the present study, the region-tree structure was used to solve the problem of precipitation forecast verification. When compared with previous work of Hou and Wang (2017), the notation of this work is how to use the region-tree structure to analyze the spatial structure of a precipitation field.

When both forecast and observation fields are represented using the region-tree structure, it is natural to verify the numerical model forecast by comparing the structural properties of these two region trees. It is further natural to verify the forecast based on the matching results of these two region trees. Additionally, it is possible to deform the region tree of the forecast to match with the region tree of observation, and then compute the intensity error between the modified forecast and observation. The above contents constitute the theme of this paper: a spatial verification approach based on region-tree analysis of forecast and observation fields.

Analogous to the scale separation and neighborhood verification approaches, the proposed region-tree-based approach also has the ability to verify a forecast at various scales. This is achieved by assigning a scale property to each region of the tree structure, and then applying the proposed verification approach to subsets of regions that are at certain scale ranges. When compared with the scale separation and neighborhood verification approaches, the region-tree-based multiscale verification approach does not filter precipitation fields initially.

This paper is organized as follows. Section 2 initially introduces the region-tree analysis method. Based on the analysis method in section 2, section 3 provides a forecast verification approach followed by section 4, which validates this verification approach. Finally, section 5 provides a summary.

2. Region-tree analysis method

The region-tree-based field analysis method includes four parts: tree structure representation of the field, reconstruction of the field from the region-tree structure, region-tree-based scale analysis, and region-tree-based structure analysis. Details of this analysis method are illustrated in this section with the use of some artificial

precipitation fields, which were generated using a GMM (Lakshmanan and Kain 2010).

a. Region-tree representation of a field

The process of building a region-tree structure from a grid-based field has been described in the paper of Hou and Wang (2017). Here, the process is recalled using an example shown in Fig. 1. Figure 1a shows an artificial precipitation field P generated from a GMM with three Gaussian components. Five steps are used to build its region-tree structure:

- Step 1. Convert the field P to binary images $\{P_i\}$ using a group of thresholds $\{t_i\}_{i=1,2,\dots,n}$ (see Fig. 1b). Specifically, this step first converts P to P_1 using $t_1 = 0 \text{ mm h}^{-1}$, then updates the threshold by $t_i = t_{i-1} + d_i$, $i = 2, 3, \dots, n$ and converts P to P_i using t_i , where d_i is the threshold interval. This step stops when t_i is larger than the maximum intensity of the field P . To obtain P_i , pixels in P with an intensity greater than or equal to t_i are replaced with the value of 1 and other pixels are replaced with the value of 0;
- Step 2. Identity all regions from P_i using the region growth method (Gonzalez and Woods 2007), where each region is a group of connected pixels that have a value of 1 (see Fig. 1c). The pixel information within each region such as pixel coordinates and intensities are recorded for all regions.
- Step 3. Recognize the overlapping relationships between the regions identified in step 2. Specifically, suppose r_i and r_{i+1} are two regions obtained from P_i and P_{i+1} ($i = 1, \dots, n-1$), respectively. If the coordinates of r_{i+1} are a subset of the coordinates of r_i , then r_i and r_{i+1} are overlapped. In Fig. 1c overlapped regions are connected using lines.
- Step 4. Calculate properties of the regions identified in step 2, such as the area, centroid coordinate, scale, and structural properties. Calculation of the structural properties will be illustrated in a later section. All region properties are recorded.
- Step 5. Store the obtained region-tree structure using a data structure, which is composed of sets of nodes and links. Each node stores the information of a region, including pixel information and regional properties, and each link stores the connection between regions. An instance of the data structure is shown in Fig. 1d. When referring to a region-tree structure, terms listed in Table 1 are often needed.

In Fig. 1, the threshold interval d_i is set to 10 mm h^{-1} , and the field P is converted to six binary images (see Fig. 1b). In general, adopting a smaller d_i will yield a more complete region-tree structure, but this requires more computation time. When using the region tree for

FIG. 1. Illustration of building a region-tree structure for a precipitation field. (a) An artificial field, (b) segmented binary images, (c) region-tree structure, and (d) the data structure of the region tree.

precipitation forecast verification, d_t could be set to 1 mm h^{-1} , which is small enough. Also, note that the first threshold adopted in step 1 must be set to 0 mm h^{-1} . In this way, the entire field will be identified as a region in step 2. This region is the root of the constructed region-tree structure in step 3. With the usage of this root region, later tree analysis processes will be simplified.

b. Field reconstruction from a region-tree structure

The above region-tree representation process could be viewed as a decomposition process, which decomposes a two-dimensional (2D) precipitation field into a group of structured 2D regions. Based on these regions, it is possible to reconstruct the precipitation field by simply

overlapping these regions together. The reconstruction process is an important part of the region-tree analysis method and is also essential for the later tree-based verification approach.

Specifically, the reconstruction process starts by building a blank field P' using the root of the region-tree structure (see the rectangular region in Fig. 1c). The blank field P' has the same size as the original field P , but all pixels in P' have a value of zero. Then, the reconstruction process iteratively goes through the region-tree structure from the root to leaves in a depth-first way (starts at the root node and explores as far as possible along each branch before backtracking), and overlays each region r_i on P' ; that is, finds the pixels in P' that are located in region r_i , and

TABLE 1. Terms related to a tree structure.

Term	Description
Child	Any successor of a node
Parent	Predecessor of a node
Path	A sequence of nodes and edges connecting from parent to child
Length of path	Number of nodes in a path
Subtree	A node together with its descendants
Root	The node that has no parent
Leaf	The node that has no child
Depth of node	Length of the path connecting a node with the root

replaces these pixels in P' with those pixels in r_i . The reconstruction process stops when all regions have been overlaid on the field P' .

In addition, the reconstruction process could be conducted using a subset of regions that meet certain conditions. For instance, one could first remove small regions from the region-tree structure, and then reconstruct a field P' using the remaining regions. In this way, noise in the original field P could be removed.

c. Region-tree-based scale analysis

The region-tree-based scale analysis method describes the spatial structure of a precipitation field using a scale spectrum of the region-tree structure, which is a distribution of region scales. To obtain the scale spectrum, the scale property for each region in the tree structure must first be calculated.

In previous works (e.g., Marzban and Sandgathe 2006, 2008), the scale of a region has been estimated using its area, and by assuming that larger regions have larger scales. However, this assumption is not suitable in some cases. Considering the two regions R_1 and R_2 shown in Fig. 2, they have equal areas but have different scales. Here R_2 has a larger scale because it has a larger spatial span than R_1 . In this work, the scale of a region was estimated using the diameter of its minimum circumscribed circle, which is a measure of the region's spatial span. The instances of the region's minimum circumscribed circle are shown in Fig. 2 using dashed circles. Estimated region scales for R_1 and R_2 are circle diameters d_1 and d_2 , respectively.

Once scales for all the regions in the tree structure have been calculated, the scale spectrum can be obtained by partitioning these region-scale values into a group of scale ranges and calculating the count of regions in each scale range. The selection of scale ranges is very important during the application of precipitation forecast verification. In this work, the scale ranges were set to <2 km (micro), 2–20 km (meso- γ), 20–200 km (meso- β), 200–2000 km (meso- α), and >2000 km (macro) (Orlanski 1975). Nevertheless,

FIG. 2. Illustration of region scale applied in this work. (a),(b) Regions with equal areas but different scales.

the threshold interval d_i also determines the shape of the obtained scale spectrum, because adopting a smaller d_i would generate more regions, and therefore produces a smoother scale distribution. However, when the threshold interval d_i is small enough (e.g., 1 mm h^{-1}), the shape of the distribution will not change significantly. An example of the scale spectrum is shown in Fig. 3g, which displays the scale spectrum using a bar graph. The width along with the abscissa and height of each bar represent the scale range and the number of regions in the scale range, respectively.

A comparison between the scale and Fourier spectra of a precipitation field is shown in Fig. 3. Given an artificial field in Fig. 3a, its Fourier spectrum, region-tree structure, and scale spectrum are shown in Figs. 3b, 3f, and 3g, respectively. Here, the Fourier spectrum was obtained through a two-dimensional discrete Fourier transform. Figure 3 also shows the decomposition results of the field based on the Fourier and scale spectra. In the decomposition process, the spectrum (Fourier or scale) is divided into three parts, and each part is used to reconstruct a field independently. As shown in Fig. 3, the filter-based scale analysis method (e.g., Fourier analysis) always blurs the spatial structure of the reconstructed field, whereas the region-based scale analysis method could keep the reconstructed spatial structure consistent with the original field. Thus, the region-based scale analysis method is intuitive as well as easy to understand and use.

d. Region-tree-based structure analysis

The region-tree-based structure analysis method is similar to the scale analysis method. It describes the spatial structure of a precipitation field using the structural spectrum of the region-tree structure of the field. The structural spectrum is a group of structural property distributions, in which the region-scale property is the independent variable. To obtain the structural spectrum

FIG. 3. Comparison between the region-based scale spectrum and the Fourier spectrum. (a) An artificial precipitation field, (b) the Fourier spectrum of (a), (c) the low-frequency part of (a), (d) the medium-frequency part of (a), (e) the high-frequency part of (a), (f) the region-tree structure of (a), (g) the region-based scale spectrum of (a), (h) the large-scale part of (a), (i) the medium-scale part of (a), and (j) the small-scale part of (a).

of a region-tree structure, structural properties of the regions in the tree structure are first calculated. Then, regions in the tree are divided into bins of different scale ranges. Finally, the regions' structural property distributions are calculated on each bin. These structural property distributions compose the structural spectrum of the region-tree structure. The key point of the region-tree-based structure analysis method is knowing how to calculate the structural properties of a region in the tree structure.

For each region in the tree structure, its properties can be categorized into two classes: geometrical and structural properties. The geometrical properties of a region are properties that can be calculated without considering the intensities within the region. In other words, the geometrical properties of a region are calculated only using the region's binary image. Typical geometrical properties of a region include the area, scale, centroid position, etc. In contrast, the structural properties of a region are computed by considering the variety of intensity or internal structure within the region. In this work, we focus on the region's structural properties, because geometrical properties have been well studied by previous object-based verification approaches.

When a precipitation field is represented using a region-tree structure, the internal structure of a region in the field could be described via the spatial structural properties of the region's subtree. An example of the region's subtree is illustrated in Fig. 4. Figures 4a and 4b show a precipitation field and its corresponding region-tree structure,

respectively. The region marked using a dotted contour in Fig. 4a corresponds to the marked subtree in Fig. 4b. Clearly, each region in the precipitation field corresponds to a subtree structure (a leaf's subtree is itself and the root's subtree is the entire tree structure). In this work, we describe the internal structure of each region using four spatial structural properties of its corresponding subtree structure. These four properties are the size reduction rate (SRR), the dispersion of leaves (DL), the mean absolute difference of leaf depths (MADLD), and overlap ratio (OR).

To calculate the structural properties of a subtree, paths of the subtree should be identified initially. An example of a subtree's paths is shown in Fig. 4c, where each path is a list of regions linking from the root of the subtree to a leaf. Based on the identified paths, definitions of the four structural properties of a subtree are given as follows.

1) SRR:

The SRR of a subtree is defined as the average of the paths' SRRs. The SRR of a path is calculated as $SRR = (s_r - s_l)/n_p$, where s_r and s_l are the areas of the path's root and leaf, respectively; n_p is the length of the path.

2) DL:

For a subtree, the dispersion of its leaves is calculated as $1/N \sum_{i=1}^N |x_i - \bar{x}|$, where x_i is the centroid position of the i th of N leaves, $\bar{x}$ is the average centroid position of these leaves, and $||$ indicates Euclidean distance.

FIG. 4. Illustration of the region's subtree structure. (a) A precipitation field, (b) region-tree structure, and (c) paths of the subtree.

3) MADLD:

The MADLD of a subtree is defined as $1/N \sum_{i=1}^N |l_i - \bar{l}|$, where l_i is the depth of the i th of N leaves, and $\bar{l}$ is the average depth of these leaves.

4) OR:

For a subtree structure, paths from the leaves to the root have many nodes in common. For example, in Fig. 4c region r_1 is shared by three paths and r_2 is shared by two paths. The measure OR responds to this property. It is defined as $OR = n_c/n_t$, where n_c is

the number of regions that are shared by different paths, and n_t is the number of regions in the subtree.

The above four properties are further illustrated below using examples. Figure 5 shows some artificial precipitation fields, their region-tree structures, and their structural properties. The artificial fields are composed of some local elliptical fields; each elliptical field was delineated by three parameters: centroid position, intensity, and scale. When a precipitation field is represented

FIG. 5. Illustration of the region's structural properties. (a) Size reduction rate (SRR), (b) the dispersion of leaves (DL), (c) the mean absolute difference of leaf depths (MADLD), and (d) overlap ratio (OR).

FIG. 6. Typical weather systems and their structural properties. (a) Stratiform precipitation, (b) convective system, (c) mixed precipitation, (d) convective system, (e) clustered convective system, (f) squall line, (g) growing convective system, and (h) mature convective system.

using the region-tree structure, each local elliptical field corresponds to a path of the tree. For a precipitation field, its SRR is determined by the scales and intensities of the local elliptical fields within it. In general, large precipitation fields with less intense local elliptical fields often have a large SRR (Fig. 5a). Additionally, as shown in Fig. 5, DL measures the dispersion of the centroids of the local elliptical fields (Fig. 5b), MADLD measures the dispersion of the intensities of the local elliptical fields (Fig. 5c), and OR measures the overlap degree of the local elliptical fields (Fig. 5d).

When using the above four structural properties in forecast verification, they could be used to recognize patterns of precipitation fields. The proposed structural property SRR could be used to distinguish between stratiform precipitations and convective systems. Stratiform precipitations (Fig. 6a) usually have a relatively larger SRR than convective systems (Fig. 6b), because stratiform precipitations have larger areas and lower intensities; MADLD could be used to recognize mixed precipitations. Because different kinds of precipitations usually have different intensities, when they are mixed together, the local intensities within a mixed precipitation present a wide distribution, which leads to a relatively larger MADLD (Figs. 6c and 6d); DL could be used to distinguish between squall lines and clustered convective systems. In a clustered convective system, cells are often gathered around the centroid of the system. Thus, the DL of a clustered convective system is relatively

smaller than that of a squall line (Figs. 6e and 6f). Meanwhile, OR could be used to recognize the evolution stage of a convective system. During the growth of the convective system, cells are usually gradually enlarged and merged. Thus, the OR of a growing convective system is smaller than that of a mature convective system (Figs. 6g and 6h).

Overall, steps of the region-tree-based structure analysis of a precipitation field can be summarized as follows: 1) build the region-tree structure of the precipitation field; 2) identify the subtree structure for each region; 3) calculate the four structural properties for each region according to its subtree structure; 4) divide the regions in the tree structure into bins according to their scales; and 5) obtain the structural property distributions at each bin. The obtained structural spectrum of the field is a group of property distributions at different scale ranges (e.g., Fig. 12). It should be noted that structural properties designed based on the region-tree structure are not limited to the above four properties. One could incorporate new structural properties into the structural spectrum to analyze the spatial structure of a precipitation field.

3. Tree-based verification methodology

Figure 7 describes the framework of this region-tree-based verification approach. The inputs of this framework are forecasted and observed precipitation fields. They are first represented by region-tree structures,

FIG. 7. Framework of the region-tree-based verification approach.

which are termed as forecast tree and observation tree, respectively. After the construction of region trees, the verification approach includes three procedures: comparing the spatial structures of the forecast and observation trees, matching the forecast tree with the observation tree and then counting the number of matched regions, and deforming the forecast tree to resemble the observation tree and then calculating the intensity error between the modified forecast and the observation. Figure 7 also shows the relationships between these three procedures. The matching procedure requires that the properties of the regions have been previously obtained, while the morphing procedure requires that the displacement vectors of the regions in the forecast tree have been obtained by the matching procedure.

Through the above three procedures, the verification approach verifies three aspects of a numerical model: the spatial structural similarity between the forecast and observation trees, the displacement errors and categorical scores (e.g., region-based CSI) of matched regions, and the intensity errors between the modified forecast and observation. Details of these procedures are described in the following sections.

a. Tree structural similarity-based verification

When precipitation fields of a forecast and its corresponding observation are represented using the region-tree structure, the performance of the forecast is first evaluated by comparing the spatial structure of the forecast tree with that of the observation tree. It is assumed that a better numerical model should produce a more similar region-tree structure.

Using the region-tree analysis method (see section 2), the spatial structure of a region tree is described via its

scale and structural spectra. Thus, we could compare the spatial structures of the forecast and observation trees on two aspects: 1) calculating the relative difference between their scale spectra and 2) calculating the degree of similarity between their structural spectra.

1) Scale spectrum difference:

The scale spectrum of a region-tree structure is a distribution of region scales, which can be denoted as $P(s)$, where s is scale range, $s \in S = \{\text{micro, meso-}\gamma, \text{meso-}\beta, \text{meso-}\alpha, \text{macro}\}$, and $P(s)$ indicates how many regions have a scale in the scale range s . Let the scale spectra of the forecast and observation trees be $P_F(s)$ and $P_O(s)$, respectively. When comparing $P_F(s)$ and $P_O(s)$, the relative difference $e(s)$ between the $P_F(s)$ and $P_O(s)$ is calculated as

$$e(s) = \frac{P_F(s) - P_O(s)}{P_O(s)}. \quad (1)$$

The value of $e(s)$ varies from $-\infty$ to ∞ . A negative (positive) value of $e(s)$ indicates that the forecast produces fewer (more) spatial structures at the scale of s . Generally, a forecast performs better if its corresponding $e(s)$ is close to zero at all scale ranges.

2) Structural spectrum similarity:

The structural spectrum of a region-tree structure is a group of structural property distributions with respect to scale range. Using mathematical language, the structural spectrum is composed of four sets of structural property distributions: $\{P_s(\text{SRR})\}_{s \in S}$, $\{P_s(\text{DL})\}_{s \in S}$, $\{P_s(\text{MADLD})\}_{s \in S}$, and $\{P_s(\text{OR})\}_{s \in S}$. Totally, there are 20 structural property distributions in a structural spectrum.

FIG. 8. Illustration of the similarity between two distributions. (a) The intersection of the distributions and (b) the union of the distributions.

When calculating the structural spectrum similarity between a forecast tree and an observation tree, the structural spectrum of the forecast tree could be denoted as $\{P_s^F(\text{SRR})\}_{s \in S} \cup \{P_s^F(\text{DL})\}_{s \in S} \cup \{P_s^F(\text{MADLD})\}_{s \in S} \cup \{P_s^F(\text{OR})\}_{s \in S}$, and the structural spectrum of the observation tree could be denoted as $\{P_s^O(\text{SRR})\}_{s \in S} \cup \{P_s^O(\text{DL})\}_{s \in S} \cup \{P_s^O(\text{MADLD})\}_{s \in S} \cup \{P_s^O(\text{OR})\}_{s \in S}$. The similarity between each pair of structural property distributions is first calculated as follows.

Suppose $P_s^F(\text{val})$ and $P_s^O(\text{val})$ are two corresponding region property distributions of the forecast and observation, respectively. The similarity $\text{sim}[P_s^F(\text{val}), P_s^O(\text{val})]$ between $P_s^F(\text{val})$ and $P_s^O(\text{val})$ is calculated as

$$\text{sim}[P_s^F(\text{val}), P_s^O(\text{val})] = \frac{\text{area}[P_s^F(\text{val}) \cap P_s^O(\text{val})]}{\text{area}[P_s^F(\text{val}) \cup P_s^O(\text{val})]}, \quad (2)$$

where function $\text{area}[P(\text{val})]$ denotes the area of distribution $P(\text{val})$, and operators $\cap$ and $\cup$ denote the intersection and union of two distributions, respectively. The similarity measure ranges from 0 to 1. A value of 1 denotes these two distributions are equal, while a value of 0 indicates that these two distributions do not overlap at all. Figure 8 illustrates how this similarity measure is calculated. Both P_1 and P_2 in Fig. 8 are two region property distributions. The areas of their intersection and union are shown in Figs. 8a and 8b, respectively.

Using Eq. (2), 20 similarity measures of structural property distributions between the forecast and observation trees are obtained: $\{\text{sim}[P_s^F(\text{SRR}), P_s^O(\text{SRR})]\}_{s \in S}$, $\{\text{sim}[P_s^F(\text{DL}), P_s^O(\text{DL})]\}_{s \in S}$, $\{\text{sim}[P_s^F(\text{MADLD}), P_s^O(\text{MADLD})]\}_{s \in S}$, and $\{\text{sim}[P_s^F(\text{OR}), P_s^O(\text{OR})]\}_{s \in S}$. These 20 similarity measures depict the similarity between the structural spectra of the forecast and observation trees. Each similarity measure indicates how similar the forecast and observation are regarding a structural property (e.g., SRR, DL, MADLD, or OR) and at certain scale range (e.g., micro, meso- γ , meso- β , meso- α , or macro). Take $\text{sim}[P_{\text{micro}}^F(\text{SRR}), P_{\text{micro}}^O(\text{SRR})]$ for instance; it

indicates the similarity between the forecasted and observed SRR distributions at the micro scale.

Also, one could obtain a comprehensive similarity measure by averaging similarity measures that are calculated using different structural properties. The comprehensive similarity measure $\overline{\text{sim}}(s)$ is a function of scale range s . It is calculated as

$$\begin{aligned} \overline{\text{sim}}(s) = & \alpha_1 \text{sim}[P_s^F(\text{SRR}), P_s^O(\text{SRR})] \\ & + \alpha_2 \text{sim}[P_s^F(\text{DL}), P_s^O(\text{DL})] \\ & + \alpha_3 \text{sim}[P_s^F(\text{MADLD}), P_s^O(\text{MADLD})] \\ & + \alpha_4 \text{sim}[P_s^F(\text{OR}), P_s^O(\text{OR})], \end{aligned} \quad (3)$$

where $\{\alpha_i\}_{i=1:4}$ are weight coefficients, and they satisfy $\sum_{i=1}^4 \alpha_i = 1$. In this work, α_i is set to 1/4. Here $\overline{\text{sim}}(s)$ measures the spatial structural similarity between the forecast and observation at the scale range s .

b. Tree-matching-based verification

The verification approach employs a tree-matching algorithm to match regions between the forecast and observation trees. Based on the matching results, the displacement errors of the regions in the forecast tree can be determined. In addition, the categorical scores of the matching results, such as region-based CSI, can be calculated.

The tree-matching algorithm employed here combines the object-matching method with the region-tree structure to match regions in the tree structures. Many object-matching methods have been proposed for object-based forecast verification. For example, the method for object-based diagnostic evaluation (MODE) matches two objects r_1 and r_2 if $d(r_1, r_2) < \sqrt{A_1} + \sqrt{A_2}$, where $d(r_1, r_2)$ is centroid distance between r_1 and r_2 ; A_1 and A_2 are areas of r_1 and r_2 , respectively. Additionally, objects of both the forecast and the observation could be matched using a combinatorial optimization method. Suppose that there are n_1 forecasted objects and n_2 observed objects; the combinatorial optimization method determines the most

FIG. 9. Comparison of matching results between the tree-based matching method and the method for object-based diagnostic evaluation (MODE). (a) A forecast field, (b) the corresponding observation of (a), (c) tree-based matching results, and (d) matching results of MODE.

likely correspondences between n_1 forecasted objects and n_2 observed objects by minimizing an objective function:

$$Q = \sum_{1 \leq i \leq n_1, 1 \leq j \leq n_2} C_{ij}, \quad (4)$$

where C_{ij} is the cost of matching forecasted object i with observed object j . The term is defined as $C_{ij} = d_p + d_a$, where d_p is Euclidean distance between the centroid points of the two objects and d_a is the square root of the area difference between the two objects. In practice, the minimal objective function Q could be obtained using the Hungarian algorithm (Jonker and Volgenant 1986).

In this work, the combinatorial optimization method is combined with the region-tree structure to form a region-tree matching algorithm, which iteratively matches regions between the forecast and observation trees from the roots to the leaves. Steps of the tree matching algorithm are given as follows:

- 1) At the first iteration, roots of the forecast and observation trees are matched directly, and then the children of these two roots are matched using the combinatorial optimization method.
- 2) In the k th ($k > 1$) iteration, for each pair of matched regions that have a depth of k , find their children and then match their children using the combinatorial optimization method. Let $k = k + 1$.
- 3) Step 2 continues until all regions within the forecast and observation trees have been processed.

Figure 9 shows an example to illustrate the tree-matching method. Figures 9a and 9b are precipitation fields of a forecast and an observation, respectively. Their corresponding region-tree structures are given in Figs. 9c and 9d. In the matching process, the first iteration matches two roots directly and then matches the children of these two roots using the combinatorial optimization method, resulting in region pairs

FIG. 10. Illustration of the region-tree morphing-based verification method. (a) A forecasted field and an observed field, (b) matching results of the forecast tree and the observed tree, (c) the morphing process of the forecasted tree, and (d) the morphing process of the forecasted tree under a given scale threshold.

(R_1, R_7) and (R_2, R_8) . In the second iteration, find the matched regions that have a depth of two [region pairs (R_1, R_7) and (R_2, R_8)]. For each region pair, find their children and match the children using the combinatorial optimization method. The second iteration results in four region pairs. Finally, when all regions have been processed, the tree-matching process ends. Matching results are shown in Fig. 9c using dotted lines.

As a comparison, Fig. 9d also shows the matching results of intensive regions in Figs. 9a and 9b using the MODE method. These regions are mismatched because of the displacement between the forecast and the observation. Such mismatches could be avoided in the tree matching process because the tree matching algorithm can use the spatial structure of the region-tree to improve the matching effect. Specifically, the tree matching algorithm first matches large regions, then matches the smaller regions within the large regions, thus avoiding mismatches of small regions that belong to different large regions [e.g., mismatch of (R_5, R_{10}) in Fig. 9d].

The result of the tree matching algorithm is a set of region pairs between the forecast and observation trees, which is denoted as $\{(r_{fi}, r_{oi})\}_{i=1:M}$, where r_{fi} and r_{oi} are the i th of M matched forecasted and observed regions, respectively. Based on the matching result, a forecast is verified on two aspects. First, the displacement error $\mathbf{e}_f$ of the forecast is calculated as

$$\mathbf{e}_f = \frac{1}{M} \sum_{i=1}^M (\mathbf{x}_{oi} - \mathbf{x}_{fi}), \quad (5)$$

where $\mathbf{x}_{oi}$ and $\mathbf{x}_{fi}$ are centroid coordinates of the regions r_{oi} and r_{fi} , respectively. A forecast that produces small

displacement error will have a better performance. Second, the forecast is verified using the region-based CSI. If a region in the forecast tree is matched with any region in the observation tree, it is defined as a “hit”; otherwise, it is defined as a “false alarm.” If a region in the observation tree is not matched with any other region in the forecast tree, it is defined as a “miss.” The region-based CSI is defined as $\text{CSI} = X/(X + Y + Z)$, where X and Y are numbers of hits and false alarms, respectively, and Z is the number of misses. The region-based CSI ranges from 0 to 1. The larger the region-based CSI is, the better the forecast performs. In addition, the displacement error and the region-based CSI of a forecast could be calculated as a function of scale. This is achieved by dividing region pairs and unmatched regions into bins according to their scales and then calculating the displacement error and the region-based CSI in each bin. Here, the scale of a region pair is defined as the larger scale of the region in the pair.

c. Tree morphing-based verification

The tree morphing-based verification method first modifies the forecast tree to form a region tree that resembles the observation tree, then reconstructs a modified precipitation forecast field using the modified forecast tree (see section 2b), and finally computes the intensity error between the modified forecast field and the observation field by averaging the point to point differences. The aim of this tree morphing-based verification method is to discover how a numerical model performs if its forecast displacement error is corrected.

An example (Fig. 10) is used to illustrate this tree morphing-based verification method. Figure 10a shows a

forecasted field (left) and its corresponding observed field (right). Figure 10b shows the forecast and observation trees, the tree-matching results (see dotted lines in Fig. 10b), and the displacement errors of the regions in the forecast tree (see arrows in Fig. 10b). The tree morphing-based verification method is carried out in Fig. 10c, which contains three steps:

- Step 1. A tree morphing process is applied to eliminate displacement errors of the forecasted regions. For each forecasted region r in the forecast tree, suppose that its displacement error is (d_x, d_y) and its internal pixels are $\{f(x_i, y_i)\}_{i=1, m_r}$; then, the tree morphing process revises the coordinates (x_i, y_i) of each pixel f_i to $(x_i + d_x, y_i + d_y)$. In this way, region r is translated and its displacement error is illuminated. When all regions in the forecast tree are translated, the obtained region-tree is called a modified forecast tree (see the middle tree in Fig. 10c).

In Step 1, if a region r in the forecast tree has a displacement error d_r , but its descendants do not have displacement errors (the descendants are not matched with any observed region), then d_r needs to be assigned to the descendants of r before the tree morphing process begins. Thus, the descendants would translate along with region r . For instance, in Fig. 10b region r_4 has a displacement error, but its child, region r_6 , does not have a displacement error. Before the tree morphing process, the displacement error of r_4 is assigned to r_6 (see r_6 in Fig. 10c).

- Step 2. Using the reconstruction process described in section 2b, a modified forecast field is reconstructed from the modified forecast tree (see the right field in Fig. 10c). Note that this process could merge two or more forecasted regions into a new region in the modified field (e.g., forecasted regions r_5 and r_6 are merged together after the modification). This is because the tree morphing process could translate a forecasted region to a place which has been taken by another forecasted region and the reconstruction process would overlap these two regions to form a merged region.
- Step 3. When the modified forecast is reconstructed, the mean absolute error (MAE) between the modified forecast and the observation is calculated as $\text{MAE} = 1/N \sum_{i=1}^N |O_i - F_i|$, where O_i and F_i are i th of N points in these two fields. The tree morphing-based verification approach uses the MAE to measure how a forecast performs when it has been modified. A better numerical forecast should produce a smaller MAE value.

The tree morphing-based verification approach can also verify a forecast as a function of scale. To this end, a

scale threshold c is added to tree morphing process. Specifically, in the forecast tree, if a region is larger than c , it is translated to eliminate its displacement error; if a region is less than c and its parent is larger than c , the region and its descendants are translated according to the displacement error of its parent. In this way, only regions in the forecast tree that are larger than the given scale threshold are translated to eliminate their displacement errors. By varying the scale threshold c , a series of modified forecast fields could be obtained. Then, by calculating the MAE between each modified forecast and the observation, the $\text{MAE}(c)$ could be obtained, which is a function of scale c .

Figure 10d shows an instance to illustrate how a scale threshold c would affect the tree morphing process. Suppose that in the forecasted tree (the left tree in Fig. 10d) only regions r_1 and r_2 are larger than c , in the tree morphing process regions r_1 and r_2 are first translated to eliminate their displacement errors. Then, small regions $r_3 \sim r_6$ are translated according to the displacement error of r_2 ; that is, the tree morphing process adds the displacement error of r_2 to all pixel coordinates in regions $r_3 \sim r_6$. Finally, the modified forecast under the scale threshold c is shown on the right of Fig. 10d. When compared with the original forecast in Fig. 10a, we can see that in Fig. 10d the spatial structure within region r_2 has not changed in the modified forecast field. This means that when we modify the large-scale parts of the forecast (larger than the scale threshold), local small-scale parts of the forecast could be preserved.

4. Results and discussion

a. Data

The performance of the proposed verification approach is validated by applying to nine cases of forecasts that were produced by three models from the Spatial Verification Methods Intercomparison Project (ICP; Ahijevych et al. 2009). The first model forecasts were produced at the National Centers for Environmental Prediction (NCEP) using a Weather Research and Forecasting (WRF) Model whose core was a Nonhydrostatic Mesoscale Model (Janjić et al. 2005) with 4.5-km grid spacing and 35 vertical levels; the second model forecasts were produced at the National Center for Atmospheric Research (NCAR) using the Advanced Research version of the WRF (ARW; Skamarock et al. 2005) core with 4-km grid spacing and 35 vertical levels; the last model forecasts were produced at the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma (also using the ARW core) with 2-km grid spacing and 51 vertical levels. For the verification of the forecasts, the ICP dataset includes

Stage II precipitation analysis (Lin and Mitchell 2011) of the NCEP. The observations from 1 June 2005 are compared with 24-h forecasts of 1-h rainfall accumulation carried out on 31 May 2005. All forecasts and observations were remapped onto the same (~ 4 km) grid. It should be noted that the following tests are provided solely to demonstrate the behaviors of the proposed verification approach.

b. Tree structural similarity comparison

In the first test, we applied the region-tree analysis method to both forecast and observation cases to explore their spatial structural features. Next we assessed the performances of the forecasts by comparing their scale and structural spectra with those of the observations. When applying the region-tree structure to represent precipitation fields in these cases, the threshold interval d_t was set to 1 mm h^{-1} , which resulted in 14 317, 10 250, 20 282, and 8655 regions for the precipitations of CAPS, NCAR, NCEP, and observation, respectively. Properties of these regions, such as area, scale, coordinate, and structural properties, were calculated as described in section 2. Based on these properties, the scale and structural spectra of CAPS, NCAR, NCEP, and observation were obtained. Since the precipitation fields have a horizontal resolution of about 4 km, these regions are mainly mesoscale regions.

Scale spectra of the forecasts and observations are first compared in Fig. 11. Figure 11a shows the scale spectra of CAPS, NCAR, NCEP, and observation via bar graphs. In general, the three forecasts and the observation present similar scale distributions of region numbers. The region numbers reach the maximum values at the meso- γ scale and then decrease rapidly as the scale increases. Figure 11a also shows that the differences of region numbers between the forecasts and the observation are obvious at each scale range. These differences are further assessed in Fig. 11b.

Figure 11b shows the relative differences of region numbers with respect to scale for three forecasts. Obviously, all forecasts produce positive relative differences at all scale ranges, suggesting the overforecasts for the three models. Among them, NCAR presents the best performance (it produces the smallest relative differences at all scale ranges), CAPS follows, and NCEP performs most poorly. From the perspective of relative number difference, the ranking of the three models is consistent with that given by experts (Keil and Craig 2009). Figure 11b also shows that these three models exhibit totally different performances as the scale range varies. For example, NCAR performs worse as the scale increases; NCEP performs better as the scale increases; and CAPS produces similar performances at all scale ranges.

FIG. 11. Comparison of scale spectra between three numerical models and the corresponding observations. Distributions of (a) region number and (b) relative number difference.

Structural spectra of the forecasts and the observation are shown in Fig. 12 via boxplots of the predefined four structural properties of forecast and observation regions at different scale ranges. Each boxplot presents the first and third quartiles, and the median. Red crosses in Fig. 12 denote outliers, which are greater than $q_3 + 1.5(q_3 - q_1)$ or less than $q_1 - 1.5(q_3 - q_1)$. Here q_1 and q_3 are the 25th and 75th percentiles of the sample data, respectively. An intuitive comparison of the structural spectra suggests that the three forecasts and the observation possess similar structural spectra, which indicates that these numerical models can reproduce the spatial structures of the observations.

To quantify the similarities of the structural spectra between the forecasts and the observation, we calculated the distribution similarity for each structural property at each scale range and then averaged the four similarity measures to obtain overall similarity measures at different scale ranges (Fig. 13). From Fig. 13, some

FIG. 12. Comparison of structural spectra between three numerical models and the corresponding observations. Boxplots of (a) size reduction rate (SRR), (b) the dispersion of leaves (DL), (c) the mean absolute difference of leaf depths (MADLD), and (d) overlap ratio (OR).

observations are found. First, it is hard to rank the three models, because the performances of the forecasts are related to scales. NCAR, NCEP, and CAPS perform best (the similarity measure is the largest) at meso- γ , meso- β , and meso- α scales, respectively. Second, the differences of the similarity measures between three models at the meso- α scale are smaller compared with at the meso- γ and meso- β scales. This probably occurs because the larger spatial structures of an observation are easier to predict. Many previous works also suggest that these three models present similar performances at large scales (e.g., Lack et al. 2010). Finally, when scale varies, CAPS presents smaller fluctuation than NCAR and NCEP.

c. Matching results of perturbed cases

In the second test, the behavior of the tree-matching-based verification method is demonstrated using a set of perturbed precipitation forecasts. The verification field is a 24-h forecast of 1-h accumulated precipitation field with a horizontal resolution of about 4 km. Perturbed cases are obtained by shifting the entire verification field to the right and southward by different amounts (see

Table 2). Note that the field of case 7 is multiplied by 1.5 and the field of case 8 has 1.27 mm subtracted from it.

Precipitation fields in cases 1–8 were first represented using the region-tree structure. Then, using the

FIG. 13. The structural similarity measurements between the forecast and observation trees vs scale ranges.

TABLE 2. Perturbed cases and their applied errors.

Case ID	Case name	Applied error
1	Pert000	No error/observation
2	Pert001	3 points right, 5 points down
3	Pert002	6 points right, 10 points down
4	Pert003	12 points right, 20 points down
5	Pert004	24 points right, 40 points down
6	Pert005	48 points right, 80 points down
7	Pert006	12 points right, 20 points down, $\times 1.5$
8	Pert007	12 points right, 20 points down, -1.27 mm

region-tree-matching method described in section 3b, the region trees of cases 1–8 were matched with the region tree of case 1, respectively. Since the true displacement errors of the regions in cases 1–8 are known, it is possible to determine which region pair in the tree matching results is correct. A matched region pair is correct if its displacement error is equal to the ground truth. When all correct matches of the regions are determined, the true region-based CSI (denoted by CSI_T) can be defined as $CSI_T = n_x / (n_x + n_y + n_z)$, where n_x is the number of correct matches, n_y is the number of perturbed regions that are not correctly matched with verification regions or not associated with any verification region, and n_z is the number of verification regions that are not associated with any perturbed region. The region-based CSI defined in section 3b (denoted by CSI_E) is essentially an estimation of CSI_T . For a good tree-matching method, its CSI_E should be close to CSI_T .

Both CSI_E and CSI_T for eight perturbed cases were analyzed in this test (Fig. 14a). In addition to the region-based categorical scores (CSI_E and CSI_T), based on the true displacement errors of the perturbed fields, we are also able to calculate the relative displacement error (RDE) of the matching results (see Fig. 14b). The RDE is defined as

$$RDE = \frac{1}{N} \sum_{i=1}^N (d_i - \bar{d}), \quad (6)$$

where d_i is the length of the displacement error of the i th of the N region pairs, and $\bar{d}$ is the length of the displacement error of the perturbed field. An effective tree-matching algorithm should produce a very small RDE.

In addition, we combined the MODE matching method with the region-tree structure to form a new region-tree-matching method (hereafter TMODE), and then used the TMODE to match the region trees of the perturbed cases 1–8 with the region tree of case 1. The TMODE method is similar to the proposed region-tree-matching method. The difference is that, in the iterative matching step (see section 3b), the children of matched regions were

FIG. 14. (a) Region-based CSIs and (b) relative displacement errors (RDEs) for the tree-based matching method and the method that combines the region-tree structure with the method for object-based diagnostic evaluation (TMODE).

matched using the MODE matching method, instead of the combinatorial optimization method. Similarly, CSI_{ES} , CSI_{TS} , and RDEs of eight perturbed cases for the TMODE method were calculated (Fig. 14).

From Fig. 14, the following observations are obtained:

- First, for the tree-matching method, its CSI_{ES} are larger than its CSI_{TS} for all cases. However, the differences between the CSI_{ES} and the CSI_{TS} are not significant. Especially for cases 1 and 2, the CSI_{ES} are nearly equal to the CSI_{TS} ; for cases 3–8, where the displacement errors are larger than 23 grids (about 93 km), the differences between the CSI_{ES} and the CSI_{TS} are less than 0.2. Figure 14a suggests that the CSI_{TS} could be roughly deduced from the CSI_{ES} when using the tree-matching-based verification method.
- Second, the effects of the tree-matching method highly depend on the displacement error between the perturbed and verification fields. As the displacement error

of a perturbed field increases, both CSI_E and CSI_T for the tree-matching method decrease, while RDE increases. However, it can be found that the tree-matching method is valid ($CSI_E \geq 0.5$ and $RDE \leq 8$ grids, or about 32 km) when the displacement error between the perturbed and verification fields is less than or equal to 47 grids, or about 188 km (cases 1–5).

- Third, both CSI_E and CSI_T for the tree-matching method are sensitive to the intensity of the perturbed field (see CSI_{ES} and CSI_{TS} for cases 4, 7, and 8). This occurs because, in the region-tree-matching process, only regions with the same intensities can be associated. When the intensity of the perturbed field is changed, intensities of the regions in the corresponding tree structure are also changed, and correct matches no longer exist. However, Fig. 14b shows that RDEs of the tree-matching algorithm for cases 4, 7, and 8 are very small (≤ 4 grids, about 16 km), suggesting that when the intensity of a perturbed field changes, the real displacement error can still be deduced from the matched region pairs by averaging their displacement errors. The effect of the intensity disruption is that the variance of the estimated displacement error is enlarged (see the variances of cases 4, 7, and 8).
- Fourth, the tree-matching method performs better than TMODE in this test. The TMODE-matching method produces smaller CSI_{ES} and CSI_{TS} , as well as larger RDEs than the tree-matching method. Additionally, significant differences are observed between the CSI_{ES} and CSI_{TS} for the TMODE-matching method, suggesting that the TMODE-matching method produced many incorrect matches. One possible reason for this poor performance is that the MODE was designed to associate large precipitation objects. However, when using the TMODE-matching method to match regions in the tree structures, we did not remove small regions initially, which led to many misses and incorrect matches.

Matching results that are shown in Fig. 14 do not take into account region scales; however, these matching results are highly affected by regions' scale property. We also explored how the region scale will affect the tree matching results. Once the region trees of the perturbed cases were matched with the verification case, the matched region pairs and unmatched regions on eight cases were divided into bins according to their scale properties. Then, CSI_{ES} and CSI_{TS} were calculated on each bin (Fig. 15). It can be seen that the performance of the tree-matching method increases with scale; the largest CSI_E and CSI_T occur in the largest-scale range. Figure 15 also shows that the differences between the CSI_{ES} and CSI_{TS} do not vary significantly.

FIG. 15. Estimated and real region-based CSIs vs scale ranges for the tree-matching method.

d. Intensity errors of the modified forecasts

The previous test shows that the region-tree-based verification method is able to estimate the displacement error of a forecast. Therefore, if this displacement error is corrected, the forecast would be further verified through its intensity error. In the third test, we explored how a model forecast performs when its spatial structure was modified. Furthermore, the modification process was carried out at different scales.

In this test, we first used the region-tree structure to represent nine cases of three forecasts and the corresponding observations. Then, we matched the forecast trees with the observation trees to obtain the displacement errors of the regions of the forecast trees. Next, we modified the forecast trees at different scales and used the modified region trees to reconstruct modified forecasts. Finally, the modified forecasts were compared with the observations to calculate their MAEs. In this way, we obtained MAE curves that vary over the scale.

The proposed forecast-morphing process is first illustrated using an example. Morphing results of a forecast at two different scales are shown in Fig. 16. The forecast field and its corresponding observation are shown in Figs. 16a and 16b, respectively. In Fig. 16c, regions with a scale larger than 240 km are translated to eliminate their displacement errors; if a region is larger than 240 km and all regions within it are less than 240 km, these small regions are translated according to the displacement error of the large region. Figure 16d is generally the same as Fig. 16c, but in Fig. 16d regions that have a scale larger than 160 km are modified independently. Note that the internal structure of the largest precipitation field in Fig. 16d has been revised. Figure 16 also reveals

FIG. 16. Morphing results of a forecast at two scale thresholds. (a) A forecast field, (b) an observation field, and the modified forecasts of (a) at scales of (c) 240 and (d) 160 km.

that the proposed tree-based morphing approach is different from previous pixel-based morphing approaches. It operates on regions instead of pixels. As a result, the tree-based morphing forecast is not identical to the observation, so that the MAE value of the modified forecast is not zero. Therefore, this region-tree morphing-based verification approach is only used to explore how a forecast performs in terms of MAE when it is modified at different scales.

Figure 17 shows the MAE values of modified forecasts with respect to the region-scale thresholds. Five scale thresholds were adopted. A threshold indicates that, in a forecast tree, displacement errors of the regions that have a scale less than the threshold were not modified. For example, the points at 0 km denote MAEs of forecasts where all regions were modified, while the points at 2000 km correspond to MAEs of forecasts where all regions were not modified. From Fig. 17, the following observations are obtained. 1) The MAEs of the modified forecasts become greater as the scale threshold increases. This indicates that the proposed tree-based morphing process could effectively modify the spatial

structures of the forecasts, making the forecasts more similar to the observation. 2) Over all scales, NCEP performs most poorly (it produces the largest MAEs for all scales). CAPS produces the smallest MAEs at

FIG. 17. Mean absolute errors (MAEs) of modified forecasts with respect to scales.

large scales, suggesting that CAPS performs better in forecasting larger spatial structure of the precipitation field; when compared with CAPS, NCAR can provide a better forecast of small-scale structures (it produces the smallest MAEs at small scales). 3) The ranking of these models changes as the scale threshold varies. When all regions of the three forecasts are modified, the ranking of the modified forecasts is NCAR > CAPS > NECP (see points at 0 km), which is consistent with the ranking given by other experts (Keil and Craig 2009). This indicates that the morphing process is effective.

5. Summary

This paper proposes a region-tree-based numerical model forecast verification approach. The basis of this approach is that both forecast and observation precipitation fields could be described using a region-tree structure, which is composed of a group of regions that are connected according to their overlapping relationships. The verification approach includes three procedures: comparing the tree structures of the forecast and the observation, matching the forecast tree with the observation tree, and deforming the forecast tree to resemble the observation tree. Through these procedures, the proposed verification approach evaluates a forecast on three aspects: spatial structural similarity with the observation, region-based categorical scores and displacement errors, and modified intensity errors with the observation. The proposed verification approach could be used to verify a forecast at arbitrary scale range by applying it to subtrees of the forecast tree that meet the scale condition.

When compared with previous spatial verification approaches, the approach in this work has the following features: 1) this verification approach applies a region-tree analysis method, which is a novel scale analysis tool that is intuitive and is easy to understand and use; 2) the proposed verification approach could quantitatively measure the spatial structural similarity between the forecast and the observation by comparing their structural spectra—a group of distributions of tree structural properties; 3) the verification approach estimates the displacement error of the forecast through matching of regions in the forecast and observation trees, and improves the matching effect by combining the region-tree structure with the conventional object matching method; and 4) the proposed verification approach provides a way to combine the filter and displacement verification approaches. Specifically, in the forecast tree, regions that pass through a scale filter are modified to eliminate their displacement errors. Thus, the tree-based verification could answer the question of

how a numerical model performs in forecasting local phenomena if the displacement errors of its forecasts are corrected at larger scales.

The proposed region-tree-based verification approach was applied to a standard dataset to investigate its behaviors. The dataset includes nine cases of three numerical model forecasts and corresponding observations. The results demonstrate that the region-tree-based verification approach could evaluate and compare the performances of numerical models at different scales. The performances of a model include the ability of reproducing the spatial structure of the observation, region-based categorical scores and displacement errors, and intensity errors after modification.

Acknowledgments. Funding for this work was provided by China Special Fund for Meteorological Research in the Public Interest (GYHY201406004) and the Natural Science Foundation of Tianjin, China (14JCYBJC21800).

REFERENCES

- Ahijevych, D., E. Gilleland, B. G. Brown, and E. E. Ebert, 2009: Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts. *Wea. Forecasting*, **24**, 1485–1497, <https://doi.org/10.1175/2009WAF2222298.1>.
- Alonso-Gonzalez, A., C. Lopez-Martinez, and P. Salembier, 2012: Filtering and segmentation of polarimetric SAR data based on binary partition trees. *IEEE Trans. Geosci. Remote Sens.*, **50**, 593–605, <https://doi.org/10.1109/TGRS.2011.2160647>.
- Casati, B., G. Ross, and D. B. Stephenson, 2004: A new intensity scale approach for the verification of spatial precipitation forecasts. *Meteor. Appl.*, **11**, 141–154, <https://doi.org/10.1017/S1350482704001239>.
- Davis, C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, <https://doi.org/10.1175/MWR3145.1>.
- , —, and —, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785–1795, <https://doi.org/10.1175/MWR3146.1>.
- Duff, T. J., D. M. Chong, P. Taylor, and K. G. Tolhurst, 2012: Procrustes based metrics for spatial validation and calibration of two-dimensional perimeter spread models: A case study considering fire. *Agric. For. Meteorol.*, **160**, 110–117, <https://doi.org/10.1016/j.agrformet.2012.03.002>.
- Ebert, E., and J. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, **239**, 179–202, [https://doi.org/10.1016/S0022-1694\(00\)00343-7](https://doi.org/10.1016/S0022-1694(00)00343-7).
- Fox, N. I., A. C. Micheas, and Y. Peng, 2016: Applications of Bayesian Procrustes shape analysis to ensemble radar reflectivity nowcast verification. *Atmos. Res.*, **176–177**, 75–86, <https://doi.org/10.1016/j.atmosres.2016.02.001>.
- Gallus, W. A., 2010: Application of object-based verification techniques to ensemble precipitation forecasts. *Wea. Forecasting*, **25**, 144–158, <https://doi.org/10.1175/2009WAF2222274.1>.
- Germann, U., and I. Zawadzki, 2004: Scale dependence of the predictability of precipitation from continental radar

- images. Part II: Probability forecasts. *J. Appl. Meteor.*, **43**, 74–89, [https://doi.org/10.1175/1520-0450\(2004\)043<0074:SDOTPO>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0074:SDOTPO>2.0.CO;2).
- Giannakaki, P., and O. Martius, 2016: An object-based forecast verification tool for synoptic-scale Rossby waveguides. *Wea. Forecasting*, **31**, 937–946, <https://doi.org/10.1175/WAF-D-15-0147.1>.
- Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, <https://doi.org/10.1175/2009WAF2222269.1>.
- Gonzalez, R. C., and R. E. Woods, 2007: *Digital Image Processing*. 3rd ed. Pearson, 976 pp.
- Han, F., and I. Szunyogh, 2016: A morphing-based technique for the verification of precipitation forecasts. *Mon. Wea. Rev.*, **144**, 295–313, <https://doi.org/10.1175/MWR-D-15-0172.1>.
- Harris, D., E. Foufoula-Georgiou, K. K. Droegemeier, and J. J. Levit, 2001: Multiscale statistical properties of a high-resolution precipitation forecast. *J. Hydrometeorol.*, **2**, 406–418, [https://doi.org/10.1175/1525-7541\(2001\)002<0406:MSPOAH>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0406:MSPOAH>2.0.CO;2).
- Hou, J., and P. Wang, 2017: Storm tracking via tree structure representation of radar data. *J. Atmos. Oceanic Technol.*, **34**, 729–747, <https://doi.org/10.1175/JTECH-D-15-0119.1>.
- Janjić, Z., T. Black, M. Pyle, E. Rogers, H. Y. Chuang, and G. DiMego, 2005: High resolution applications of the WRF NMM. *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., 16A.4, http://ams.confex.com/ams/WAFNWP34BC/techprogram/paper_93724.htm.
- Jonker, R., and T. Volgenant, 1986: Improving the Hungarian assignment algorithm. *Oper. Res. Lett.*, **5**, 171–175, [https://doi.org/10.1016/0167-6377\(86\)90073-8](https://doi.org/10.1016/0167-6377(86)90073-8).
- Keil, C., and G. C. Craig, 2007: A displacement-based error measure applied in a regional ensemble forecasting system. *Mon. Wea. Rev.*, **135**, 3248–3259, <https://doi.org/10.1175/MWR3457.1>.
- , and —, 2009: A displacement and amplitude score employing an optical flow technique. *Wea. Forecasting*, **24**, 1297–1308, <https://doi.org/10.1175/2009WAF2222247.1>.
- Lack, S. A., G. L. Limpert, and N. I. Fox, 2010: An object-oriented multiscale verification scheme. *Wea. Forecasting*, **25**, 79–92, <https://doi.org/10.1175/2009WAF2222245.1>.
- Lakshmanan, V., and J. S. Kain, 2010: A Gaussian mixture model approach to forecast verification. *Wea. Forecasting*, **25**, 908–920, <https://doi.org/10.1175/2010WAF2222355.1>.
- Li, J., K. Hsu, A. AghaKouchak, and S. Sorooshian, 2015: An object-based approach for verification of precipitation estimation. *Int. J. Remote Sens.*, **36**, 513–529, <https://doi.org/10.1080/01431161.2014.999170>.
- Li, M., A. Stein, W. Bijker, and Q. Zhan, 2016: Region-based urban road extraction from VHR satellite images using binary partition tree. *Int. J. Appl. Earth Obs. Geoinfo.*, **44**, 217–225, <https://doi.org/10.1016/j.jag.2015.09.005>.
- Lin, Y., and K. E. Mitchell, 2011: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.
- Liu, Z., L. Shen, and Z. Zhang, 2011: Unsupervised image segmentation based on analysis of binary partition tree for salient object extraction. *Signal Processing*, **91**, 290–299, <https://doi.org/10.1016/j.sigpro.2010.07.006>.
- Lu, H., J. C. Woods, and M. Ghanbari, 2007: Binary partition tree for semantic object extraction and image segmentation. *IEEE Trans. Circ. Syst. Video Tech.*, **17**, 378–383, <https://doi.org/10.1109/TCSVT.2006.888943>.
- Marzban, C., and S. Sandgathe, 2006: Cluster analysis for verification of precipitation fields. *Wea. Forecasting*, **21**, 824–838, <https://doi.org/10.1175/WAF948.1>.
- , and —, 2008: Cluster analysis for object-oriented verification of fields: A variation. *Mon. Wea. Rev.*, **136**, 1013–1025, <https://doi.org/10.1175/2007MWR1994.1>.
- Micheas, A. C., N. I. Fox, S. A. Lack, and C. K. Wikle, 2007: Cell identification and verification of QPF ensembles using shape analysis techniques. *J. Hydrol.*, **343**, 105–116, <https://doi.org/10.1016/j.jhydrol.2007.05.036>.
- Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, **56**, 527–530.
- Rezacova, D., Z. Sokol, and P. Pesice, 2007: A radar-based verification of precipitation forecast for local convective storms. *Atmos. Res.*, **83**, 211–224, <https://doi.org/10.1016/j.atmosres.2005.08.011>.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Salembier, P., and L. Garrido, 2000: Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. Image Process.*, **9**, 561–576, <https://doi.org/10.1109/83.841934>.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp., <https://doi.org/10.5065/D6DZ069T>.
- Theis, S. E., A. Hense, and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. Appl.*, **12**, 257–268, <https://doi.org/10.1017/S1350482705001763>.
- Vilaplana, V., F. Marques, and P. Salembier, 2008: Binary partition trees for object detection. *IEEE Trans. Image Process.*, **17**, 2201–2216, <https://doi.org/10.1109/TIP.2008.2002841>.
- 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, <https://doi.org/10.1175/2008MWR2415.1>.
- Yano, J., and B. Jakubiak, 2016: Wavelet-based verification of the quantitative precipitation forecast. *Dyn. Atmos. Oceans*, **74**, 14–29, <https://doi.org/10.1016/j.dynatmoce.2016.02.001>.
- Ziegeler, S. B., J. D. Dykes, and J. F. Shriver, 2012: Spatial error metrics for oceanographic model verification. *J. Atmos. Oceanic Technol.*, **29**, 260–266, <https://doi.org/10.1175/JTECH-D-11-00109.1>.