A CLUSTERING METHOD FOR RAIN-CELL DETECTION IN WEATHER NOWCASTING APPROACHES

Felipe Minotta-Zapata and Rafael A. Rodríguez-Solís

Electrical and Computer Engineering Department, University of Puerto Rico, Mayagüez, PR

ABSTRACT

This work focuses on the problem of forecasting the energy availability of renewable sources such as solar and wind in smart grids. To solve this problem, we propose to use weather radar information to make a short-time prediction of the weather conditions in the area where the renewable sources are located. For this purpose, an object tracking method will be used to make the predictions by using as cues the reflectivity and the velocity. This paper centers in an object modeling approach based on clustering, which will be used for the tracking algorithm.

***Index Terms*—** Clustering, weather nowcasting, smart grids

1. INTRODUCTION

Smart grids are an efficient manner to manage energy grid systems. In a smart grid, digital technology and instrumentation systems are incorporated to the power generation, transmission, distribution, and utilization components. The integration allows to gather the necessary information to efficiently manage the offer and demand of energy. This organization allows to incorporate and maximize the use of renewable energy sources in such a way that the overall operational cost is minimized [1]. One problem that arises with the use of renewable energy sources, such as solar and wind, is that they fluctuate during day, as they are affected by weather conditions. An energy availability prediction from these renewable sources based on the weather conditions would allow the operators to plan ahead energy dispatch from conventional sources or consumers to adjust their demand to the renewable sources availability. Therefore, weather information becomes of critical importance in this planning. Energy markets at distribution level, usually sample the power requirements every 15 minutes. Consequently, we propose to develop a system that will be able to predict the weather conditions 15 minutes ahead in the areas where the renewable energy sources are located.

Different weather nowcasting approaches has been proposed in recent years. An algorithm to determine the strike probability of thunderstorms in a chosen area is proposed in [2]. They used constant thunderstorm speed and direction with constant standard deviation, which was useful for short-lifetime storms. Ruzansky et al. presented a method to predict weather conditions 10 minutes ahead using least squares estimation formulated in the Fourier domain [3]. The proposed equation used the sequence of radar reflectivity fields, velocity fields, growth, and decay of intensity. Ruzansky and Chandrasekar in another work proposed to use specific differential phase instead of reflectivity to estimate rainfall [4]. Shukla et al. developed a source apportionment (SA) technique for tracking and nowcasting mesoscale convective systems (MCS) using satellite image sequences [5]. The convective areas were detected using neighborhood search criteria to select contiguous pixels. Thong et al. proposed a hybrid method combining interpolative picture fuzzy rule technique and particle swarm optimization for weather nowcasting [6]. The algorithm showed an improvement in prediction against other methods, but it requires more time to compute. Rossi et al. presented an algorithm for tracking convective storms using Kalman filtering [7]. They used a reflectivity threshold to identify convective cells in every time step for then apply a cluster algorithm to detect the centroid position of the storms. Bechini and Chandrasekar developed a weather prediction algorithm that attempts to exploit all the available dual-polarization and Doppler information provided by a single radar [8].

In this paper, we present the pre-processing steps and the object detection algorithm to make a further short-time prediction. All this based on the reflectivity and velocity information provided by the UPRM TropiNet weather radar network. The final objective of this project, which is currently in development, is to use an object tracking approach to make the correspondent predictions.

The rest of this paper is organized as follows. Section 2 describes the radar network. Section 3 explains the preprocessing made to the data supplied by the radars. Section 4 details the implemented method to merge the images provided by the radar network. Section 5 describes the proposed clustering algorithm for the high reflectivity zones. Section 6 shows the initial results obtained in this early stage. Finally, Section 7 provides concluding remarks.

This material is based upon work supported by the National Science Foundation under Grant No. ACI-1541106.

2. TROPINET RADAR NETWORK

TropiNet is a radar network located in the western side of Puerto Rico and it comprises three radars as seen in Fig. 2. Each node is a RXM-25 unit which is a polarimetric and Doppler weather radar, operating at X-band [9], with a temporal resolution of one minute with a scan range of 40.161 km, approximately. The full specifications for each radar node is documented in [10]. In order to take maximum advantage of the radar network configuration, the geographic zone in which the algorithms will make predictions is the zone where the three radar coverage overlap.

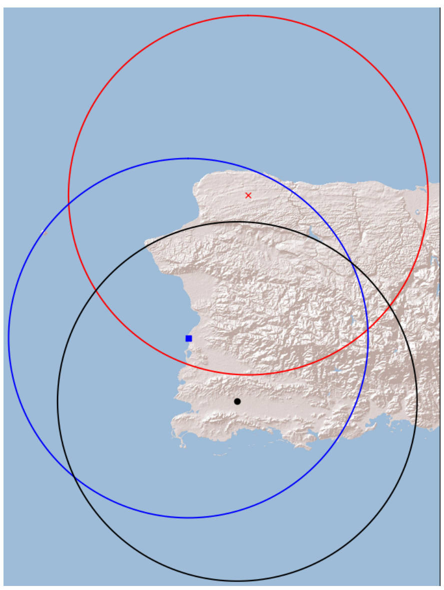

Fig. 1. Coverage of Isabela (red cross), Cabo Rojo (blue square), and Lajas (black circle) radars

3. RADAR INFORMATION PRE-PROCESSING

The radars provide images in polar form arranged in matrices, where each row of the matrix correspond to different radii. To simplify the image processing, a conversion to cartesian coordinates is carried out. This process presents two different challenges: data mapping and interpolation. The data mapping refers to the place where every sensed point will be located in the image after the conversion. The interpolation refers to process of covering the areas where the radar did not sense, but it can be inferred from the adjacent points of these areas, this process is carried out by a median filter. The process is identical for the reflectivity and the velocity matrices.

4. IMAGE RADAR FUSION

Since the purpose is to exploit the spatial radar network configuration, a process to merge the images of the three radars is performed. The first step is to determine the pixel distance between the centers of the three images. Vincenty's formula is a widely used method to calculate the distance between two points in the earth's surface, since it is accurate within 0.5mm. To establish the pixel distance between the three radars, it is necessary to first calculate the x- and y- physical distance displacements. The pixel distance can be then calculated by using rule of three.

4.1. Image Reflectivity Fusion

There are several methods to accomplish the fusion process of reflectivity images. For example, the fusion process could be made by taking the maximum value, the minimum value, or by calculating the mean between the different images. In this work, a weighted mean is used to carry out the fusion process. The weights of every image is established by giving the weight according to the closeness to the physical position of each radar. Consequently, the final value of each point in the merged image is obtained by

$$M_{ij}^{R} = \frac{\omega_{ij}^{1} R_{ij}^{1} + \omega_{ij}^{2} R_{ij}^{2} + \omega_{ij}^{3} R_{ij}^{3}}{\omega_{ij}^{1} + \omega_{ij}^{2} + \omega_{ij}^{3}}, \tag{1}$$

where M_{ij}^{R} is the ij-th pixel of the merged reflectivity image M^{R}, ω_{ij}^{k} is the weight of the ij-th pixel in the k radar, and R_{ij}^{k} is the reflectivity value of the k radar.

4.2. Velocity Data Fusion

A Doppler radar provides a negative number of velocity if the sensed point is moving towards the radar and a positive number if the sensed point is moving away the radar. However, the real direction of the sensed point may be different. In the overlapping zone, the velocity information provided by each radar helps to better estimate the direction of each sensed point.

In order to accomplish the fusion process, for each radar velocity matrix, it is necessary to create a set of unitary vectors pointing to the physical location of the radar. These unitary vectors can be represented as a pair of matrices containing the x- and y- component of each vector. Then, the set of velocity vectors is obtained by making a Hadamard multiplication between the velocity information provided by the radar and the components of the unitary vectors. The merged velocity for the x- and y- components M^{V_x} and M^{V_y} is obtained by summing their respective components.

5. RAIN-CELLS DETECTION ALGORITHM

Traditional tracking approaches are centered in new manners to perform the tracking of objects that can suffer occlusions, scale or appearance changes, rotations, etc. Moreover, the extracted features are mainly based in geometric characteristics, e.g., centroid, height, width, etc. However, applications with high uncertainty, like cloud tracking in weather prediction, cannot be easily addressed using these approaches since the clouds or rain cells change shape constantly, can completely disappear, scatter or combined with others.

To attack this problem, we propose to use a clustering method to model the cells as unshaped and unsized objects. We define zones of high reflectivity as the ones that are larger than a threshold. Those salient points per frame is the dataset in which the clustering method will be applied. To develop the clustering algorithm, it is necessary to take into account the next characteristics: there has to be an unsupervised clustering algorithm since there is no categories or labels and the number of cluster per frame is also unknown. Our clustering algorithm is based on divisive hierarchical clustering, which is an unsupervised machine learning algorithm. The following are the steps for developing the hierarchical clustering: 1) binarize each frame by using,

$$B_{ij}^{R} = \begin{cases} 1, & \text{if } M_{ij}^{R} \leq T_Z \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where B^R is the binarized image and T_Z is the reflectivity threshold, 2) retrieve the location of every salient point in the binarization, 3) form a single cluster with all points, 4) start by joining points that its distance is equal or less than a threshold distance, a new cluster is formed if the distance from the formed cluster to the point is greater than the threshold distance. To define if a point belongs to a cluster, we define a threshold distance that is given by the k-neighbor adjacent pixels from all pixels in the cluster. The process is repeated until all the points are assigned to a cluster.

From the clusters formed, we will extract the following features that will be used for the tracking algorithm. 1) Centroids: the x- and y-average of every point in the cluster. 2) Inter-cluster distances: distance between the nearest neighbors (single linkage), it will be used to predict a possible merge of clusters. 3) Intra-cluster distances: measure of how sparse is the cluster. 4) Average velocity: a wind vector that shows the probable speed and direction of the cluster. 5) Maximum velocity: the maximum wind vector associated with each cluster, it will be used to predict a possible cluster break. 6) Cluster size and 7) number of clusters.

6. INITIAL RESULTS

The tests of the processing algorithms shown in this article were made over data collected in 2016 [11]. Fig. 6 depicts an event of rainfall in the overlapping zone of the three radar nodes. The black bounding boxes enclose the zones of high reflectivity, which we consider for this example, the zones that have over 35 dBz [7]. The number of adjacent neighbors used for the clustering algorithm was 99, which can be seen as a high number, but in experiments we have found that lower adjacent neighbors lead to redundant prediction over the same weather system. In the tracking system, we plan to use a similar number of adjacent neighbors as in this example in order to avoid over processing.

Since the radar network can provide a measure every minute, our prediction system must be able to make a forecast within this time range. Therefore, we have also evaluated the elapsed time for the image pre-processing, fusion, and clustering. We have performed 100 experiments in order to extract the average time of every step and the entire process. The results are shown in Table 1. For performance analysis purposes, we have included the time reported by The Python ARM Radar Toolkit Py-ART [12] for a fusion of two radar images. The reason for the large time processing difference between our algorithm and the Py-ART method could be due to the more complex pre-processing interpolation used by Py-ART, which could lead to better results, than the used in our algorithm. Our time for clustering presents a large standard deviation because is dependent on the number of rain cells present in every sample.

Table 1. Time performance comparison

	Our Algorithm	Py-ART
Pre-processing	$\bar{X} = 7.02s$ $s = 0.13s$	16.63s
Image Fusion	$\bar{X} = 0.3s$ $s = 0.009s$	
Clustering	$\bar{X} = 8.34s$ $s = 1.18s$	–

7. SUMMARY AND FUTURE WORK

This paper presented the initial radar image processing for a future weather nowcasting system which it could be use as a service of renewable source energy availability for smart grids. The pre-processing stage comprises a conversion to cartesian coordinates and a smoothing process made by two median filters. The radar image fusion stage uses a weighted mean scheme for the reflectivity cue and a vector component summation for the velocity cue. The rain-cell detection is carried out by a clustering algorithm that uses a pixel distance threshold for determining the pixel association. The outputs from this clustering process, which are going to be used as input in the tracking algorithm, are the cluster centroids, the inter- and intra-cluster distance, the average and maximum velocity of every cluster, the cluster size, and the number of clusters per sampled image. Our current work is set on the cluster matching between consecutive sampled images as well as the formulation of a dynamic model which it can be used to the development of a state estimation framework.

8. REFERENCES

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid the new and improved power grid: A survey," *IEEE*

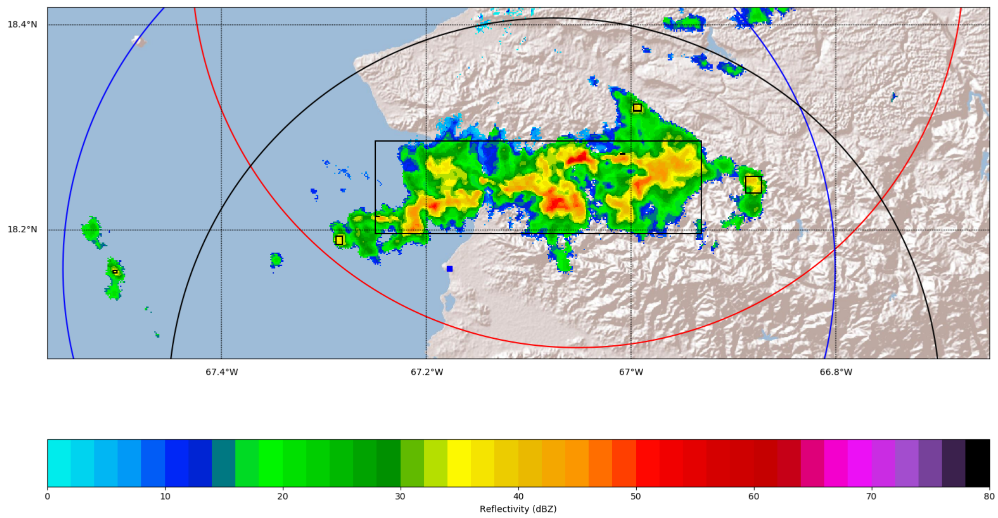

Fig. 2. Reflectivity image of the overlapping zone of the three radar nodes. The colormap used was "pyart_NWSRef" [12]

Communications Surveys Tutorials, vol. 14, no. 4, pp. 944–980, Fourth 2012.

[2] Sandy Dance, Elizabeth Ebert, and David Scurrah, "Thunderstorm strike probability nowcasting," *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 1, pp. 79–93, 2010.

[3] Evan Ruzanski, V. Chandrasekar, and Yanting Wang, "The casa nowcasting system," *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 5, pp. 640–655, 2011.

[4] Evan Ruzanski and V. Chandrasekar, "Nowcasting rainfall fields derived from specific differential phase," *Journal of Applied Meteorology and Climatology*, vol. 51, no. 11, pp. 1950–1959, 2012.

[5] B. P. Shukla and P. K. Pal, "A source apportionment approach to study the evolution of convective cells: An application to the nowcasting of convective weather systems," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 1, pp. 242–247, Feb 2012.

[6] P. H. Thong, L. H. Son, and H. Fujita, "Interpolative picture fuzzy rules: A novel forecast method for weather nowcasting," in *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, July 2016, pp. 86–93.

[7] Pekka J. Rossi, V. Chandrasekar, Vesa Hasu, and Dmitri Moisseev, "Kalman filteringbased probabilistic nowcasting of object-oriented tracked convective storms," *Journal of Atmospheric and Oceanic Technology*, vol. 32, no. 3, pp. 461–477, 2015.

[8] Renzo Bechini and V. Chandrasekar, "An enhanced optical flow technique for radar nowcasting of precipitation and winds," *Journal of Atmospheric and Oceanic Technology*, vol. 34, no. 12, pp. 2637–2658, 2017.

[9] M. B. Gálvez, J. Colom, V. Chandrasekar, F. Junyent, S. Cruz-Pol, and R. Rodriguez, "Salient features of the radar nodes in the puerto rico tropical weather testbed," in *2009 IEEE International Geoscience and Remote Sensing Symposium*, July 2009, vol. 3, pp. III–841–III–844.

[10] M. B. Gálvez, J. G. Colom, V. Chandrasekar, F. Junyent, S. Cruz-Pol, R. A. Rodrguez Solis, L. Len, J. J. Rosario-Coln, B. De Jess, J. A. Ortiz, and K. M. Mora Navarro, "First observations of the initial radar node in the puerto rico tropinet x-band polarimetric doppler weather testbed," in *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, July 2013, pp. 2337–2340.

[11] N. Hosannah, J. Gonzlez, R. Rodriguez-Solis, H. Parsiani, F. Moshary, L. Aponte, R. Armstrong, E. Harmsen, P. Ramamurthy, M. Angeles, L. Len, N. Ramrez, D. Niyogi, and B. Bornstein, "The convection, aerosol, and synoptic-effects in the tropics (cast) experiment: Building an understanding of multiscale impacts on caribbean weather via field campaigns," *Bulletin of the American Meteorological Society*, vol. 98, no. 8, pp. 1593–1600, 2017.

[12] J.J. Helmus and S.M. Collis, "The python arm radar toolkit (py-art), a library for working with weather radar data in the python programming language," *Journal of Open Research Software*, vol. 4, no. 1, pp. e25, 2016.