Convective cells algorithm for storm data tracking

Piotr Szuster
Department of Computer Science
Cracow University of Technology
ul. Warszawska 24, 31-155 Kraków, Poland
Email: piotr.szuster@pk.edu.pl

Joanna Kołodziej
Research and Academic Computer Network
ul. Kolska 12, 01-045 Warsaw, Poland
Email: joanna.kolodziej@nask.pl

KEYWORDS

radar, tracking, weather, thunderstorm, data fusion

ABSTRACT

Atmospheric conditions, such as thunderstorms, are significant factors that influence human activity. Harsh weather may severely impact both daily life and professional activities. Severe thunderstorms are a considerable hazard – they can generate heavy rainfall, high winds, large hail and tornadoes. Tracking of thunderstorms is necessary to gain situational awareness - knowledge of present and future storm-related threats and their corresponding significances. Thunderstorms are weather phenomena associated with cumulonimbus clouds. Those clouds are formed in deep, moist convection and are composed of liquid and solid water particles. Weather radars can detect those particles. Cumulonimbus-related particle concentration areas are represented in weather radar data as convective cells, making that measurement technique useful for storm-tracking applications. This paper proposes a new algorithm for storm data tracking in the data fusion process. The algorithm has been tested with real data from the POLRAD weather radar network and upper-air observations. The efficiency of the proposed algorithm has been justified in the empirical analysis. The algorithm projections can help generate weather warnings due to accurate forecasts of storm movement.

I. INTRODUCTION

A thunderstorm is a phenomenon related to a cumulonimbus cloud. Cumulonimbus is a dense, vertically developed cloud of water droplets and ice crystals. In most cases, they also produce rain, snow and hail precipitation. All of these solid and liquid forms of water particles can be detected by weather radar. As Cumulonimbuses are very dense, they are represented in weather radar data as high-concentration areas - convective cells. Convective cells are very distinguishable. Thus, they can be easily extracted as blobs. Current thunderstorms' positions can be estimated from blobs set and then used for target tracking applications [13].

Thunderstorms are one of the major weather-related hazards in Europe[12]. During the stormy days, proper, timely and complete assessments of storm positions, threats they pose, and threats–related significance are defined as *Situational Awareness (S-A)* conditions. Object tracking is locating and monitoring specific objects and their behaviour in sequential images. This paradigm is used successfully in biology and medicine [**cell**].

This paper uses this paradigm to define a new *Convective Cells (CC) tracking algorithm* for storm tracking. Our CC algorithm is a class of data fusion methods designed to solve storm tracking problems under S-A conditions[5].

The data for such monitoring is gathered from the observation (in our case, upper-air observations) and several instruments, such as weather radars. Usually, a weather radar covers a vast area. Thus, multiple convective cells can be detected, and the Multi-Target Tracking (MTT) technique can be applied to identify each cell and timely associate its position [10].

At least two timely separated measures are required to assess convective cell movement and different positions, namely *movement model* and *initial motion vectors*. Initial motion vectors can be generated in the sounding process, i.e. upper-air observations performed by radiosondes. Radiosondes measure pressure, height, temperature, dew point, wind direction and velocity. Such measurement is performed by a radiosonde connected by a string to a weather balloon and then released into the air. The instrument gathers data characterizing the vertical profile of the Earth's atmosphere during its ascent.

The rest of the paper is organized as follows. Sec. II defines the thunderstorm tracking problem and briefly surveys the possible methods described for solving such a problem. We define the connective cells tracking algorithm in Sec. III. The algorithm has been empirically evaluated in Sec. IV. We concluded the paper in Sec. V.

II. THUNDERSTORM TRACKING PROBLEM

Thunderstorms are organized disturbances in the Earth's atmosphere produced by cumulonimbus clouds that are always accompanied by lightning and thunder[9]. Cumulonimbus clouds are composed of both solid and liquid water particles. Moreover, they are very dense and vertically developed. Cumulonimbuses are formed during deep, moist convection in unstable atmospheres[3]. They can cause hazards such as excessive precipitation, large hail, severe wind gusts and

Communications of the ECMS, Volume 37, Issue 1,
Proceedings, ©ECMS Enrico Vicario, Romeo Bandinelli,
Virginia Fani, Michele Mastroianni (Editors) 2023
ISBN: 978-3-937436-80-7/978-3-937436-79-1 (CD) ISSN 2522-2414

occasionally tornadoes. To avoid the negative ravages of such hazards, there is a need to gain proper situational awareness [11].

Situational awareness (S–A) is "the perception of entities in the environment, comprehension of their meaning, and projection of their status in the near future". In this paper, we define S–A as the state of properly understanding thunderstorms, their relationships, their present and future threats and their significance. *Numerical Weather Prediction (NWP)* models can predict thunderstorms and related hazards. However, NWP models encounter significant difficulties in predicting both the correct location and expected intensity of thunderstorms as they depend on small-scale factors (mesoscale or smaller). Theoretically, NWP models of the denser grids and smaller scales could be used to acquire higher-quality results. Still, they are vulnerable to micro-physics schemes, grid resolution and initial conditions [4].

Another approach, storm nowcasting, could overcome these problems [14]. Nowcasting is a forecasting technique based on a very short period of up to 6 hours. This method utilizes the latest data measured by many sensors, and it is more suitable for small-scale event forecasting with reasonable accuracy. Cumulonimbus clouds, as they are very dense and vertically developed, are also composed of solid and liquid water particles. Thus, they can be represented in radar imagery as convective cells[1]. Storm nowcasting techniques often are based on treating a time series of historical radar images as the input and prediction of the radar image as the output because convection cells can be extracted from that data [6]. Extrapolation techniques commonly translate (meteorologically: advect) convection cells to predict their positions and properly associate after prediction. Generally, extrapolation techniques are connected with the storm tracking process - a nowcasting technique of storm's track assignment and prediction. A storm track is a time series of storm's positions. Storm tracking techniques can answer the question of where and when the storm will move in the future. Thus it can facilitate the gaining of proper situational awareness leading to avoidance of some of the negative consequences of the storm's presence.

A. RELATED WORK

In [8], authors proposed The Storm Cell Identification and Tracking Algorithm designed for WSR-88D Radar. The algorithm uses 3D-volumetric data. The default motion vector is defined either as user input or either as an average of motion vectors delivered by the previous scan. Latter is used as an initial motion vector if no average storm motion vectors are delivered from the previous scan. Authors suggested that mean wind between 0-6km above the ground layer should be used to deliver the initial SMV. Time association is the match with the smallest distance to the previous detection. The match distance has to have a value within a specific threshold. The updated storm motion vector is calculated as using a linear least squares fit that incorporates the current position and up to the 10 previous scans. The method has disadvantages: time association flexibility is limited by the threshold. The algorithm does not use automatic wind velocity and direction data gathering and thus requires user input to set accurate initial motion vectors. Moreover, the calculation of updated SMV's does not take into account acceleration.
TITAN algorithm [2] utilizes the Hungarian method for the solution of optimal assignment problems. TITAN is based on the following assumptions: the correct set of matches incorporates shorter paths rather than longer, the characters of the matched cells are similar, and the upper bound of distance is defined relative to the maximum expected velocity. The matching rule is based on overlapping consecutive entity areas. The authors also discussed the problem of merging and splitting the complex storm's systems. The idea of solving the merging/splitting problem is based on track extension/termination. TITAN does not accurately track fast-moving storm systems. Single threshold identification scheme limit tracking of small-scale thunderstorms. In [7], authors proposed another approach. THOR uses the gating function incorporating a dynamic search radius, which depends on storm speed and length. THOR is designed as an offline algorithm and is unsuitable for nowcasting.
TITAN and SCIT were designed for single radar applications. THOR utilize multi-sensor data. However offline nature of the algorithm makes it unsuitable for real-time processing. To overcome the limitations of the aforementioned solutions new approach is necessary.

III. CONNECTIVE CELLS TRACKING ALGORITHM

Weather radar data is available online, and remote repositories distribute it. Usually, data is available as images showing a horizontal cross-section of a three-dimensional radar field of view in a standard Cartesian coordinate system. The image usually has a timestamp to identify when the data was gathered. Each pixel of the radar image has a colour corresponding to the reflectivity value. An example radar scan bitmap is shown in Fig. 1.

Information about colour and corresponding reflectivity values is usually available as a radar scale. To perform the tracking process, the bitmap must be transformed into a reflectivity matrix using a scale, and all non-convective pixels' reflectivity values must be set to 0. Non-convective pixels can be identified by corresponding reflectivity lower than threshold $\tau < 44dBz$. For given radar image taken at time t_i of width w and height h, a *Reflectivity Matrix* $\mathbf{R_{t_i}}$ with values $r_{t_{i_{yx}}} \geq 44dbZ \quad x \in \{1, \ldots, w\}, y \in \{1, \ldots, h\}$ is generated as follows:

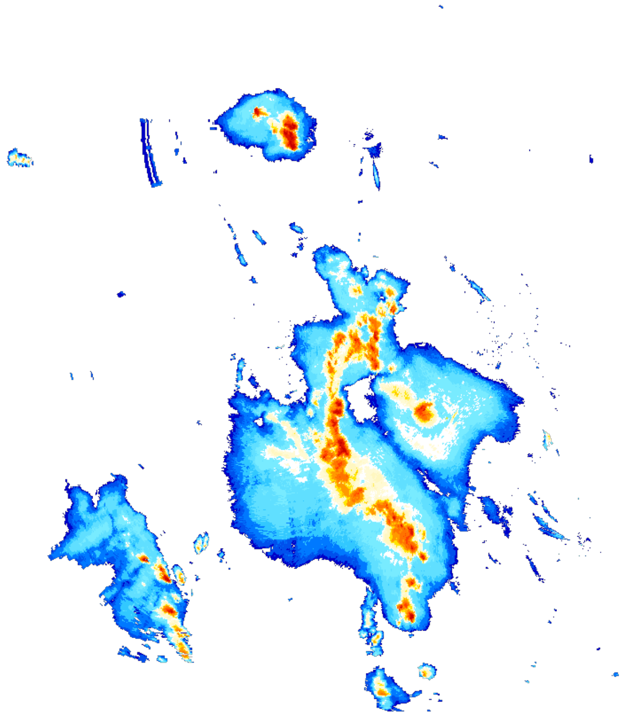

Fig. 1. Column max composite image taken at 18:00 UTC of July 14th, 2021

$$\mathbf{R_{t_i}} = \begin{bmatrix} r_{t_i,11} & r_{t_i,12} & \cdots & r_{t_i,1w} \\ r_{t_i,21} & r_{t_i,22} & \cdots & r_{t_i,2w} \\ r_{t_i,31} & r_{t_i,32} & \cdots & r_{t_i,3w} \\ \vdots & \vdots & \ddots & \vdots \\ r_{t_i,h1} & r_{t_i,h2} & \cdots & r_{t_i,hw} \end{bmatrix} \tag{1}$$

The matrix $\mathbf{R_{t_i}}$ defined by Eq.1 is then used for the definition of a set B_{t_i} of n_i blobs in the following way:

$$B_{t_i} = \{b_{t_i,1}, \ldots, b_{t_i,n_i}\} \tag{2}$$

Let us define *Blobs* in the set B_{t_i} as image regions with approximate constant properties, such as brightness, colour, contrast etc. In most cases, convective cells are depicted as well-defined, continuous areas of relatively high reflectivity ($\geq 44dBz$) surrounded by weak ones. Therefore, the blob $b_{t_i,k} \quad k = \{1, \ldots, n_i\}$, which is a set of m_k adjacent non-zero reflectivity values indexes pairs, can be formally defined in the following way:

$$\begin{split} b_{t_i,k} = \{(x_{t_i,k,1}, y_{t_i,k,1}), (x_{t_i,k,2}, y_{t_i,k,2}), \ldots, \\ (x_{t_i,k,m_k}, y_{t_i,k,m_k})\}, \end{split} \tag{3}$$

where: $k \in 1, \ldots, n_i$

We defined *Adjacency rule* for ensuring the continuity of the considered regions. We say that a pair $u = (x_u, y_u)$ is adjacent $\sim$ to pair $v = (x_v, y_v)$, if the following condition holds:

$$u \sim v \Leftrightarrow \begin{cases} x_u = x_v - 1 \wedge y_u = y_v \\ x_u = x_v + 1 \wedge y_u = y_v \\ x_u = x_v \wedge y_u = y_v - 1 \\ x_u = x_v \wedge y_u = y_v + 1 \end{cases} \tag{4}$$

All blobs with less than 7 pairs are removed from the set B_{t_i}. From u_{t_i} remaining blobs in the set set B_{t_i}, a new set of u_{t_i} *Measurements* $M_{t_i} = \{m_{t_i,1}, \ldots, m_{t_i,u_{t_i}}\}$ is created. A *Measurement* $m_{t_i,v} \quad v \in \{1, \ldots, u_{t_i}\}$ is defined as follows:

$$m_{t_i,v} = \{(\mu_{x_{t_i,v}}, \sigma_{x_{t_i,v}}), (\mu_{y_{t_i,v}}, \sigma_{y_{t_i,v}})\} \tag{5}$$

$$\mu_{x_{t_i,v}} = \frac{1}{m_v} \Sigma_{q=1}^{m_v} x_{t_i,v,q}, \tag{6}$$

$$\sigma_{x_{t_i,v}} = \sqrt{\frac{1}{m_v} \Sigma_{q=1}^{m_v} (x_{t_i,v,q} - \mu_{x_{t_i,v}})^2}, \tag{7}$$

$$\mu_{y_{t_i,v}} = \frac{1}{m_v} \Sigma_{q=1}^{m_v} y_{t_i,v,q}, \tag{8}$$

and

$$\sigma_{y_{t_i,v}} = \sqrt{\frac{1}{m_v} \Sigma_{q=1}^{m_v} (y_{t_i,v,q} - \mu_{y_{t_i,v}})^2}. \tag{9}$$

$E_{t_i} = \{e_1, \ldots, e_{q_{t_i}}\}$ is a multi-set of q_{t_i} entities tracked at time t_i. Each entity $e_p : p \in \{1, \ldots, q_{t_i}\}$ is a set of l_{t_i} previously associated measures and their timestamps (prior to time t_i):

$$e_p = ((m_{p,1}, t_{p,1}), \ldots, (m_{p,l}, t_{p,l_{t_i}})) \tag{10}$$

For measurement set M_{t_i}:

$$M_{t_i} = \{m_{t_i,1}, \ldots, m_{t_i,u}\} \tag{11}$$

If $E_{t_i} = \emptyset$:

$$E_{t_{i+1}} = \{((m_{t_i,1}, t_i)), \ldots, ((m_{t_i,u}, t_i))\} \tag{12}$$

As each entity, e_p is a set of measures and timestamps associated before time t_i predicted motion vector $\mathbf{v_{p,t_i}} = [\mathbf{v_{x_{p,t_i}}}, \mathbf{v_{y_{p,t_i}}}]$ for time t_i is delivered by motion model function:

$$\mathbf{v_{p,t_i}} = \mathbf{smv(e_p, t_i)} \tag{13}$$

The predicted entity e_p position is defined as:

$$\mathbf{l_{p,t_i}} = [\mu_{\mathbf{x_{t_{p,l_{t_i}}},p}}, \mu_{\mathbf{y_{t_{p,l_{t_i}}},p}}] + \mathbf{smv(e_p, t_i)} \tag{14}$$

where $[\mu_{x_{t_{p,l_{t_i}}},p}, \mu_{y_{t_{p,l_{t_i}}},p}]$ is last known entity e_p position prior to t_i. To simplify now we consider $\mathbf{l_{p,t_i}}$ from eq. 14 as:

$$\mathbf{l_{p,t_i}} = [\mu_{\mathbf{x_{p,t_i}}}, \mu_{\mathbf{y_{p,t_i}}}] \tag{15}$$

The predicted position has corresponding x and y deviations:

$$\sigma_{x_{p,t_i}} = \sigma_{x_{t_{p,l_{t_i}}},p} \tag{16}$$

$$\sigma_{y_{p,t_i}} = \sigma_{y_{t_{p,l_{t_i}}},p} \tag{17}$$

Thus predicted positions distributions can be described as normal distributions:

$$N_{x_{p,t_i}}(\mu_{x_{p,t_i}}, \sigma_{x_{p,t_i}}) \tag{18}$$

$$N_{y_{p,t_i}}(\mu_{y_{p,t_i}}, \sigma_{y_{p,t_i}}) \tag{19}$$

where $N(\mu, \sigma)$ denotes normal distribution with mean μ and standard deviation σ. For each of that distributions, the entity e_p at time t_i, a marginal probability density function is delivered:

$$g_{x_{p,t_i}}(x) = \frac{1}{\sigma_{x_{p,t_i}}\sqrt{(2\pi)}} \exp\left(\frac{-(\mathrm{x} - \mu_{\mathrm{x_{p,t_i}}})^2}{2\sigma_{\mathrm{x_{p,t_i}}}{}^2}\right) \tag{20}$$

$$g_{y_{p,t_i}}(y) = \frac{1}{\sigma_{y_{p,t_i}}\sqrt{(2\pi)}} \exp\left(\frac{-(\mathrm{y} - \mu_{\mathrm{y_{p,t_i}}})^2}{2\sigma_{\mathrm{y_{p,t_i}}}{}^2}\right) \tag{21}$$

The entity e_p joint probability density function at time t_i is defined:

$$g_{p,t_i}(x, y) = g_{x_{p,t_i}}(x) \cdot g_{y_{p,t_i}}(y) \tag{22}$$

To perform association for each entity e_p :

$$g_{p,t_i}(\mu_{x_{t_i,v}}, \mu_{y_{t_i,v}}) \tag{23}$$

has to be calculated for each measurement $m_{t_i v}$. An association set A_{t_i} of $j = u_{t_i} \times q_{t_i}$ i delivered:

$$A_{t_i} = \{a_{t_i,1}, \dots, a_{t_i,j}\} \tag{24}$$

Each association $a_{t_i,o} \quad i \in \{1, \dots, j\}$ is a triple:

$$a_{t_i,o} = (e_p, m_{t_i,v}, g_{p,t_i}(\mu_{x_{t_i,v}}, \mu_{y_{t_i,v}})) \tag{25}$$

All elements $a_{t_i,o}$ with $g_{p,t_i}(\mu_{x_{t_i,v}}, \mu_{y_{t_i,v}}) < 0.05$ are removed. Then A_{t_i} is sorted in descending order of $g_{p,t_i}(\mu_{x_{t_i,v}}, \mu_{y_{t_i,v}})$.

As the first step of calculating global marginal probability density, a set of all associations with the entity of the first association is created.

$$B_{t_i} = \{b_{t_i,1}, \dots, b_{t_i,w}\} : \tag{26}$$

For each chosen association $b_{t_i,q} : \quad q \in \{1, \dots, w\}$ of B_{t_i} a separate set of associations without both entity and measurement of chosen association $C_{t_i,q}$ is created and organized:

$$C_{t_i,q} = \{c_{t_i,q,1}, \dots, c_{t_i,q,1}\} \tag{27}$$

Then recursively, the first step of calculating global marginal probability density is repeated for each $C_{t_i,q}$ assumed as A_{t_i}. At point when $|C_{t_i,q}| = 1$ returned value is $g_{p,t_i}(\mu_{x_{t_i,v}}, \mu_{y_{t_i,v}})$ of remaining in $C_{t_i,q}$ association. At any shallower step of recursion $|C_{t_i,q}| > 1$ returned value is maximum of $g_{p,t_i}(\mu_{x_{t_i,v}}, \mu_{y_{t_i,v}})$ of entities of associations in current B_{t_i} multiplied by return from deeper step. The algorithm computes all possible joint probabilities association combinations. The global maximum corresponds to the final global association. All measurements without globally chosen associations are assigned to new entities. Measurements with the globally chosen association are given to their selected entities.

A. STORM MOTION VECTOR

As each entity $e_p : p \in \{1, \dots, q_{t_i}\}$ is a set of l_{t_i} previously associated measures and their timestamps (prior to time t_i):

$$e_p = ((m_{p,1}, t_{p,1}), \dots, (m_{p,l}, t_{p,l_{t_i}})) \tag{28}$$

For $k \in \{1, \dots, l_{t_i}\}$ k-th measurement of entity e_p can be denoted as:

$$m_{p,k} = \{(\mu_{x_{p,k}}, \sigma_{x_{p,k}}), (\mu_{y_{p,k}}, \sigma_{y_{p,k}})\} \tag{29}$$

Time at which measurement was taken respectively as $t_{p,k}$. As velocity can be understood as a distance travelled over time, instantaneous entity e_p velocity components at time t_{p_k} can be denoted as difference quotients:

$$\frac{\mu_{x_{p,k}} - \mu_{x_{p,k-1}}}{t_{p,k} - t_{p,k-1}} \tag{30}$$

$$\frac{\mu_{y_{p,k}} - \mu_{y_{p,k-1}}}{t_{p,k} - t_{p,k-1}} \tag{31}$$

$\forall k \in \{2, \dots, l_{t_i}\}$. We use the mean of velocity quotients to assess predicted storm motion $\mathbf{smv(e_p, t_i)}$ to preserve the stability of the track. Similarly, we calculate accelerations (as difference quotients of velocity) to include changes in the trajectory of convection cells. The mean velocity vector is multiplied by the time difference between the prediction time and the last measurement timestamp. Moreover, to maintain efficiency maximum history parameter is introduced to motion calculation to limit the number of required measurements.

For initial vectors (when $l_{t_i} = 1$), we used mean wind measures gathered by radiosondes during upper-air observations. Mean vectors are calculated for each upper-air station and then interpolated by inverse distance weighting for given $(\mu_{x_{p,k}}, \mu_{y_{p,k}})$.

B. LIGHTNING DETECTION

Thunderstorm-related convection cells are accompanied by lightning; thus, we use lightning detection data to distinguish thunderstorm cells. Lightning detection data are available as coordinates couples. They can be associated in the same way as measurements. When at least one lightning is associated with entity e_p, the entity e_p is marked as a thunderstorm.

IV. EXPERIMENTS

In the empirical evaluation of the proposed algorithm, we used real data gathered by the POLRAD weather radar network[1], real soundings available at the Wyoming University repository [2] and real detection data from the Blitzortung lightning detection network[3]. Radar data was in the form of a COLUMN MAX reflectivity product. Three memorable severe weather situations were selected to evaluate our algorithm:

- **24.06.2021** – Two tornadic supercells:

[1]`https://pl.wikipedia.org/wiki/POLRAD`
[2]`https://weather.uwyo.edu/upperair/sounding.html`
[3]`https://www.blitzortung.org/pl/live_lightning_maps.php`

1. in Hodonin, Czech Republic,
2. near Nowy Sacz, Poland,
3. record breaking hail of 13.5 cm in diameter recorded in Tomaszow Mazowiecki),

- **14.07.2021** – severe wind gusts generated by a supercell near Chrzanow, Poland),
- **20.08.2022** –flash flood near Czarny Dunajec, Poland.

42 representative entities were selected as the test set to cover a variety of kinematic characteristics of the atmosphere (wind direction and speed distribution). The longest track in the set, shown in Fig. 2, comprises 28 associations representing the 4,5 hours motion of long-lived severe supercell. The average length of the track was 6.7, corresponding to one hour lifetime. Duplicates of track lengths are caused by the analysis of multiple scenarios in which similar track lengths can generate different errors. This error variety is caused by storm shape and movement and strongly depends on the synoptic situation (wind dynamics, storms' rapid growth and decay caused by thermodynamic conditions). Direct visual comparison between predicted and actual movement based on examples may be difficult. Visual representation, in that case, is illegible.

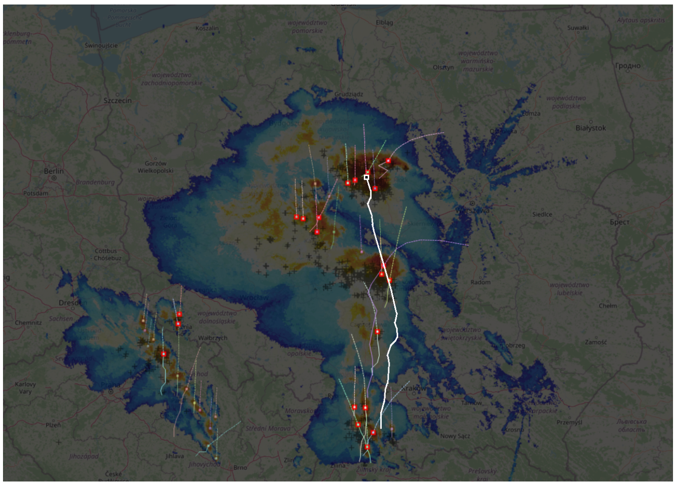

Fig. 2. Longest track (white) at 19:50 UTC recorded on July 14th, 2021. The track was associated with a supercell thunderstorm that produced severe wind gusts in Chrzanow, Poland

The mean error of the predicted storm position was equal to 2.58. The maximum error recorded was 7.31. The error was defined as the difference between the predicted storm position and the associated measurement at the next time point. The relation between track length and mean error is shown in Fig. 3.

Preliminary results show that mean error depends on track length. The longer track has more associations. Thus, blob merging/splitting occurrence influencing error is more likely. Fig. 3 also shows that a mean error can vary for relatively equal track length. Individual track analysis shows that error is more significant in situations with many convective cells in a small general area. The error also depends on the quality of initial motion vectors delivered from soundings. Additional research, including the involvement of upper air pro-

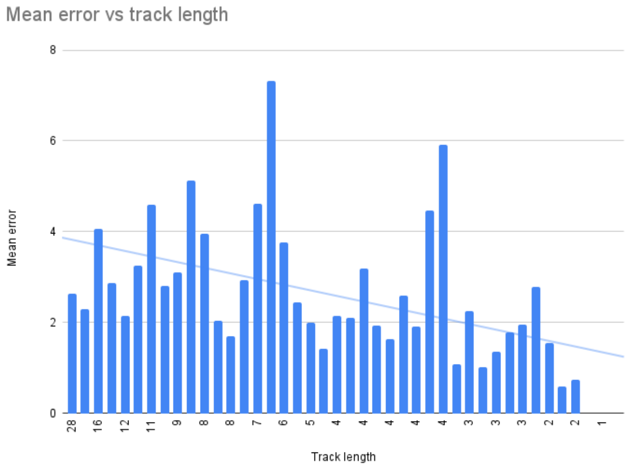

Fig. 3. Mean error versus track length with trend line

files from classical NWP models to produce initial motion vectors, is advisable to lower errors. The tracking process is very dependent on the quality of the blob extraction algorithm. More accurate blob extraction algorithms can be utilized to improve tracking results. Individual track analysis shows that High-quality results were delivered for isolated cells. Two fascinating examples are shown below. On 24.06.2021, a severe supercell produced a tornado near Breclav and Hodonin cities in the Czech Republic. The tornado intensity was estimated at F3/F4 on the Fujita scale, killing 6 people and injuring 200. The algorithm produced a very accurate projection of the storm movement 40 minutes before the tornado entered Hodonin. The prediction is shown in Fig. 4.

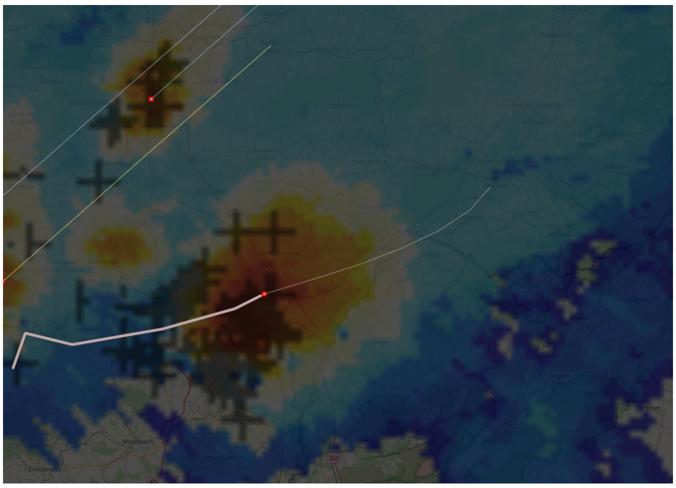

Fig. 4. Track of Hodonin supercell at 16:50 UTC. Continuous thick lines show past movement. Dashed lines depict predicted movement. The red point shows the current position. The cluster in the upper left corner is a new cell; thus, it has no past movement.

The second example is a supercell near Chrzanow, Poland, on July 14th, 2021. An isolated storm cell was born around 14:20 UTC on the Slovakian side of the

Slovakian-Polish border. At 14:50 UTC, the algorithm projected that the storm cell (now organized as a supercell) would impact Chrzanow, Poland. At 15:10 UTC local fire department was informed about massive destruction related to severe wind gusts. The supercell storm track is shown in Figure 5

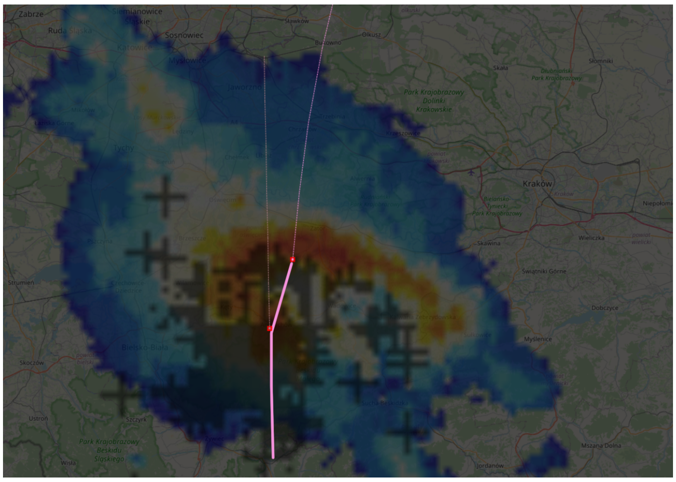

Fig. 5. Track of Chrzanow supercell at 14:50 UTC. Continuous thick lines show past movement. Dashed lines depict predicted movement. The red point shows the current position.

In the situations mentioned above, the algorithm accurately assessed storm motion. That assessment could issue a proper warning for people in the influenced area. Another conclusion is related to initial storm motion vector definitions from the literature. Tracking experiments during our research showed that the upper height limit for mean wind SMV should depend on thermodynamic equilibrium height as the 0-6km definition is not suitable for winter/early spring storms. In that case, storm tops are located at lower altitudes.

V. CONCLUSIONS

The main topic of this work was the definition and preliminary analysis of the effectiveness of the connective cells algorithm for storm tracking based on observation and measurement data using weather radiosondes. The algorithm belongs to the class of data fusion methods and conventions.

The empirical results of a simple experimental analysis indicate that the proposed approach is promising. The algorithm produced good results, especially for well-defined, severe thunderstorms. The algorithm projections can help deliver severe weather warnings due to accurate predictions of storm movement. Further research will incorporate classical NWP model results to provide accurate initial motion vectors and test additional blob extraction methods.

References

[1] Luca Baldini et al. "Ground-Based Weather Radar to Investigate Thunderstorms". In: Feb. 2018, pp. 113–135. DOI: 10.1007/978-3-319-72583-3_4 (cit. on p. 2).

[2] Michael Dixon and Gerry Wiener. "TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting—A Radar-based Methodology". In: *Journal of Atmospheric and Oceanic Technology* 10.6 (1993), pp. 785–797. DOI: https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atot/10/6/1520-0426_1993_010_0785_ttitaa_2_0_co_2.xml (cit. on p. 2).

[3] Brooks Harold Edward. *Ingredients-based forecasting*. Springer, 2007, pp. 133–140 (cit. on p. 1).

[4] Sabrina Guastavino et al. "Prediction of severe thunderstorm events with ensemble deep learning and radar data". In: (Sept. 2021). DOI: 10.21203/rs.3.rs-1928945/v1 (cit. on p. 2).

[5] David Hall and James Llinas. "Handbook of Multisensor Data Fusion". In: (June 2001). DOI: 10.1201/9781420038545 (cit. on p. 1).

[6] Alessandro Hering et al. "Nowcasting thunderstorms in the Alpine region using a radar based adaptive thresholding scheme". In: (2004) (cit. on p. 2).

[7] Adam L. Houston et al. "Thunderstorm Observation by Radar (ThOR): An Algorithm to Develop a Climatology of Thunderstorms". In: *Journal of Atmospheric and Oceanic Technology* 32.5 (2015), pp. 961–981. DOI: https://doi.org/10.1175/JTECH-D-14-00118.1. URL: https://journals.ametsoc.org/view/journals/atot/32/5/jtech-d-14-00118_1.xml (cit. on p. 2).

[8] J. T. Johnson et al. "The Storm Cell Identification and Tracking Algorithm: An Enhanced WSR-88D Algorithm". In: *Weather and Forecasting* 13.2 (1998), pp. 263–276. DOI: https://doi.org/10.1175/1520-0434(1998)013<0263:TSCIAT>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/wefo/13/2/1520-0434_1998_013_0263_tsciat_2_0_co_2.xml (cit. on p. 2).

[9] Philip Krider. *Thunderstorm*. Mar. 2023 (cit. on p. 1).

[10] Laura Leal-Taixé et al. "Tracking the Trackers: An Analysis of the State of the Art in Multiple Object Tracking". In: *CoRR* abs/1704.02781 (2017). URL: http://arxiv.org/abs/1704.02781 (cit. on p. 1).

[11] Arslan Munir, Alexander Aved, and Erik Blasch. "Situational Awareness: Techniques, Challenges, and Prospects". In: *AI* 3.1 (2022), pp. 55–77 (cit. on p. 2).

[12] "Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD". In: *Atmospheric Research* 93.1 (2009). 4th European Conference on Severe Storms, pp. 575–586. ISSN: 0169-8095. DOI: 10.1016/2008.10.020 (cit. on p. 1).

[13] Piotr Szuster. "Blob Extraction Algorithm in Detection of Convective Cells for Data Fusion". In:

Journal of Telecommunications and Information Technology 4 (Jan. 2020), pp. 65–73. DOI: `10.26636/jtit.2019.135319` (cit. on p. 1).

[14] James W. Wilson et al. "Nowcasting Thunderstorms: A Status Report". In: *Bulletin of the American Meteorological Society* 79.10 (1998), pp. 2079–2100. DOI: 10.1175/1520-0477(1998)079<2079:NTASR>2.0.CO;2 (cit. on p. 2).

BIOGRAFIES

PIOTR SZUSTER received his B.Sc. and M.Sc. degrees with distinctions in Computer Science at Cracow University of Technology, Poland, in 2015 and 2016, respectively. He is a Research and Teaching Assistant at the Cracow University of Technology and PhD student at AGH University of Science and Technology. His research interests include data fusion techniques, atmospheric thermodynamics and severe weather forecasting. Piotr Szuster is a member of the Skywarn Poland management board. More details are available at www.retsuz.pl

JOANNA KOŁODZIEJ is the Professor of Computer Science at the Research and Academic Computer Network (NASK) Institute in Warsaw. Prof. Kolodziej is the President of the Polish Chapter of the IEEE Computational Intelligence Society. She participated in several international and national projects and is a member of the EU Blockchain Partnership Society. Her research is focused on machine learning and Big Data, security aspects and energy awareness in resource management and data scheduling in clouds, cloud and edge computing, cybersecurity aspects in ICT systems and recently – blockchain technologies.