



# 1 **Analysis of the phase space of the downburst that occurred on 25** 2 **June 2021 in Sânnicolau Mare (Romania)**

3 Andi Xhelaj<sup>1</sup>, Massimiliano Burlando<sup>1</sup>

4 <sup>1</sup> *Department of Civil, Chemical and Environmental Engineering*  
5 *Polytechnic School, University of Genoa, Via Montallegro 1, 16145 Genoa, Italy*

6 *Correspondence to: Andi Xhelaj ([andi.xhelaj@edu.unige.it](mailto:andi.xhelaj@edu.unige.it))*

7 **Abstract.** Downburst winds, characterized by strong, localized downdrafts and subsequent horizontal straight-line  
8 winds, presents significant risk to civil structures. The transient nature and limited spatial extent present measurements  
9 challenges, necessitating analytical models for accurate understanding and predicting their action on structures. This study  
10 analyzes the Sânnicolau Mare downburst event in Romania, from June 25, 2021, using a bi-dimensional analytical model  
11 coupled with the Teaching Learning Optimization Algorithm (TLBO). The intent is to understand the distinct solutions  
12 generated by the optimization algorithm and assess their physical validity. Supporting this examination is a damage survey  
13 and wind speed data recorded during the downburst event. Employed techniques include agglomerative hierarchical  
14 clustering with the K-means algorithm (AHK-MC) and principal component analysis (PCA) to categorize and interpret  
15 the solutions. Three main clusters emerge, each displaying different storm characteristics. Comparing the simulated  
16 maximum velocity with hail damage trajectories indicates that the optimal solution offers the best overlap, affirming its  
17 effectiveness in reconstructing downburst wind fields. However, these findings are specific to the Sânnicolau Mare event,  
18 underlining the need for a similar examination of multiple downburst events for broader validity.

19 **KEYWORDS:** Downburst analytical model, Metaheuristic optimization algorithm, Multivariate data analysis, Downburst  
20 kinematic and geometric parameters, Damage survey.

## 21 **1 Introduction**

22 The wind climatology of Europe and several mid-latitude countries are primarily dominated by the presence of extra-  
23 tropical cyclones and thunderstorms. The understanding of the formation and evolution of extra-tropical cyclones dates  
24 back to the 1920s (Bjerknes and Solberg, 1922). The atmospheric boundary layer (ABL) winds generated during such  
25 systems are well recognized, and their influence on structures has been extensively studied and coded starting from the  
26 1960s (Davenport, 1961). These established models continue to be employed in contemporary engineering practice  
27 (Solari, 2019). Thunderstorm winds known as “downburst” consists of a strong and localized downdraft of air generated  
28 within a convective cell. These downdrafts after reaching the ground begins to spread horizontally, resulting in the  
29 formation of the downburst gust front, also known as the downburst outflow. The presence of strong turbulent wind within  
30 the downburst outflow poses significant risk to civil structures. Downburst may be generated by isolated thunderstorms,  
31 with length scales less than few kilometers. Additionally, they can be originated from more complex convective systems  
32 such as squall lines and bow echoes, in this case the spatial length scale which can potentially be affected by downbursts  
33 or downburst clusters is in the order of hundreds of kilometers (Fujita, 1978, Hjelmfelt, 2007). The size of the downburst  
34 outflow area of strong winds exhibits variability, leading to the classification of this phenomenon as either a microburst



35 or macroburst. A microburst is characterized by a strong outflow size that is less than 4 km, whereas a macroburst  
36 corresponds to an outflow size of intense wind greater than 4 km (Fujita, 1985). For the past four decades, the study of  
37 intense downburst wind and their impact on the built environment has constituted a prevailing subject of research in the  
38 field of Wind Engineering (Letchford, 2002). Since downburst event have high frequency of occurrence, they can be  
39 considered as one of the most severe meteorological phenomena. Thunderstorm, occurring at the mesoscale, exhibit  
40 nonstationary behaviour. Their origin is due to an instable convection condition in the atmosphere and the resulting  
41 horizontal wind profiles are significantly different from those usually observed in the ABL. From a statistical point of  
42 view, wind velocities, characterized by a mean return period greater than 10 or 20 years, are often due to these phenomena  
43 (Solari, 2014). The lack of a unified model for downburst outflows and their actions on structures, similar to Davenport's  
44 (1961) model for extra-tropical cyclones, is primarily due to significant uncertainties arising by the inherent complexity  
45 of downburst winds. Indeed, the transient nature and limited spatial extent of downbursts presents challenges in their  
46 measurements and restrict the availability of an adequate number of test cases. In 2020, Xhelaj et. al. presented an  
47 analytical model thought to simulate the bi-dimensional structure of downbursts. The model depends on 11 parameters  
48 that are estimated using a global metaheuristic optimization algorithm described in Xhelaj et. al. (2022). The integration  
49 between the analytical model and the optimization algorithm, as well as the estimation of the kinematic parameters of the  
50 downburst outflow, is based on the Teaching Learning Based Optimization (TLBO) algorithm. The TLBO algorithm  
51 operates with a population of solutions and emulates a teaching and learning activity through iterative process to attain  
52 the best solution within the population (Rao et al., 2011). Due to the stochastic nature of the TLBO algorithm when  
53 coupled with the analytical model, the procedure can produce different optimum (or best) solutions each time the  
54 algorithm is executed. This variability arises from the initial random population of solutions generated at the beginning  
55 of the algorithm and the intermediate transformations of the set of solutions carried out by the algorithm in order to  
56 converge towards the best solution. This study aims to examine the characteristics of the optimal solutions obtained  
57 through multiple runs of the optimization procedure. It seeks to investigate the variability of the best solutions when  
58 applying the optimization algorithm to reconstruct the wind field during an intense downburst event. The main objective  
59 is to assess the extent to which the solutions differ from each other and from the solution with the lowest objective function  
60 value. Additionally, the study explores whether these alternative solutions can be considered physically valid, particularly  
61 when additional data describing the downburst event is incorporated. The selected downburst event occurred in western  
62 Timis region of Romania on 25 June 2021 and was produced during the passage over the town of Sânnicolau Mare of an  
63 intense mesoscale convective system of bow echo type. This event was recorded by a bi-axial anemometer and  
64 temperature sensor, both placed on a telecommunication tower 50 m above the ground level. The telecommunication  
65 tower lies approximately 1 km south of Sânnicolau Mare. The downburst that occurred in Sânnicolau Mare was of  
66 significant magnitude, resulting in extensive hail damage of the facades of numerous buildings within the city. Subsequent  
67 to the occurrence of this intense event, a comprehensive damage survey was undertaken through a collaborative  
68 partnership between University of Genoa (Italy) and the University of Bucharest (Romania). The survey (Calotescu et.,  
69 al., 2022 and Calotescu et. al., 2023 (submitted)) pinpoints the GPS position of the buildings within the city that were  
70 predominantly impacted by the downburst. Moreover, a comprehensive map illustrating the hail damage of the building  
71 facades was generated. The map provides important information regarding the wind velocity experienced at urban scale,  
72 which has been used to validate the reconstruction/simulation of the downburst by the optimization procedure.

73 The analysis of the different optimal solutions (i.e., the data set) generated by the optimization algorithm was conducted  
74 through multivariate data analysis (MDA). This involved the joint application of cluster analysis and principal component

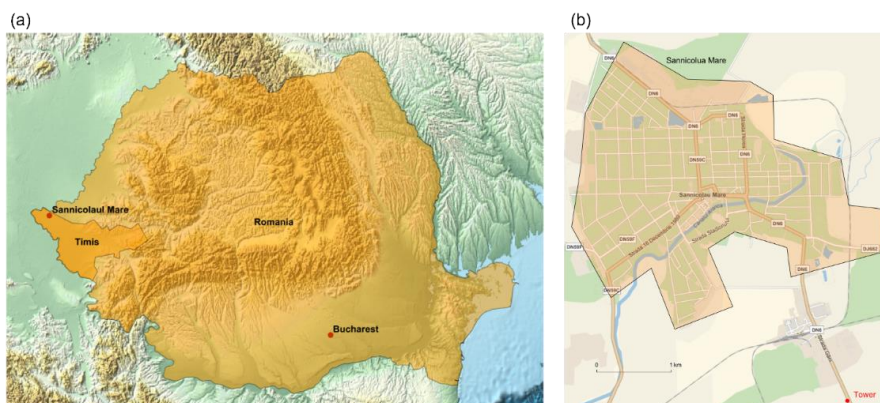


75 analysis to effectively examine and interpret the dataset. Cluster analysis (CA) is a data mining technique that groups  
76 similar solutions together, aiming to identify patterns in the data. It is commonly used in fields like meteorology and  
77 climatology to identify clusters of weather phenomena or geographical regions with similar weather patterns (Burlando  
78 et al., 2008; Burlando et al., 2009). Principal component analysis (PCA) is a mathematical technique used to decrease the  
79 dimensionality of a dataset while minimizing the loss of information within the data. This analysis is commonly used in  
80 meteorology and climatology to decrease the number of variables required for representing weather pattern or climate  
81 trends and to identify regions with similar weather patterns (Amato et. al., (2020); Jiang et. al., (2020)). Principal  
82 component analysis is utilized in this context to enhance the interpretation of the different optimal solutions.

83 The present work is structured in 7 Sections. Following the introduction, Section 2 provides a description of the  
84 monitoring system that acquired the full-scale measurement employed in this research. Section 3 provides a brief  
85 meteorological description of the downburst event in Sănnicolau Mare (Romania). Section 4 describes the data set  
86 employed for performing cluster analysis and principal component analysis. Section 5 describes the implementation of  
87 these analyses. Section 6 presents an in-depth account of the main results derived from the CA and PCA. In conclusion,  
88 Section 7 offers a summary of the principal findings derived from this research.

## 89 **2 Monitoring system and data acquisition**

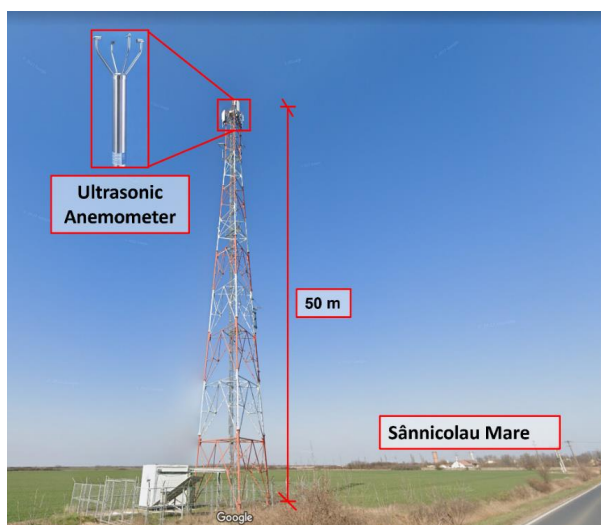
90 The complete set of measurements employed in this research were obtained through a monitoring system installed in  
91 Romania. Relevant information of this monitoring network can be accessed in the publications by Calotescu et al., (2021),  
92 Calotescu and Repetto, (2022) and Calotescu et. al., (2023) (submitted). The monitoring network received funding from  
93 the THUNDERR Project (Solari et al., 2020), which was conducted by the "Giovanni Solari – Wind Engineering and  
94 Structural Dynamics" Research (GS-Windyn) Group at the Department of Civil, Chemical, and Environmental  
95 Engineering (DICCA) of the University of Genoa. GS-Windyn, with a keen interest in monitoring poles and towers  
96 exposed to thunderstorm actions worldwide, secured funding for the acquisition of a full-scale structural monitoring  
97 network. This monitoring system was deployed on top of a 50 m lattice tower. The primary focus of this project revolves  
98 around three key objectives: first, the detection of thunderstorms; second, the analysis of wind parameters associated with  
99 these phenomena; and third, the experimental assessment of the structural response of telecommunication lattice towers  
100 to the forces generated by both synoptic and thunderstorm winds. Thunderstorms are local phenomena that occur in  
101 conditions of atmospheric instability, being characterized by the existence of vertical air currents that lead to the  
102 development of cumulonimbus clouds, the production of electric discharges, rain, and hail as well as strong downdrafts  
103 inducing damaging winds in proximity to the Earth's surface. The vertical profile of horizontal wind velocity in downburst  
104 winds showcases distinct characteristics when compared to the traditional velocity profile observed within the boundary  
105 layer. Notably, downburst winds exhibit a nose-like shape profile, with a pronounced maximum intensity near the ground.  
106 This specific profile presents a considerable risk, particularly for structures of low to medium height. The monitoring  
107 tower, named TM\_424, is property of the SC TELEKOM ROMANIA SRL and is located in the western part of Romania,  
108 Timis county, at approximately 1 km south of Sănnicolau Mare (Figure 1). The site is an open field, the terrain is flat with  
109 low grass vegetation.



110

111

Figure 1. (a) Location of the telecommunication tower TM\_424, situated 1 km south of Sannicolau Mare in Timis County, Romania. (b) Expanded view of the Sannicolau Mare town with the telecommunication tower TM\_424 represented by the red dot on the map. Maps generated using Mathematica (Wolfram Research, Inc., Version 13.3, 2023, <https://www.wolfram.com/mathematica>).



112

Figure 2. TM\_424 Telecommunication tower and sensors position at the top of the tower. On the horizon, approximately 1 km from the tower lies the small city of Sannicolau Mare. Image courtesy of © Google Street View, 2022 (<https://www.google.com/maps>).

113

114 Figure 2 shows the dimension of the tower. Among the various networks for the monitoring systems, the tower is  
115 equipped with a GILL WindObserver 70 ultrasonic anemometer at the top (Figure 2). The anemometer has a data

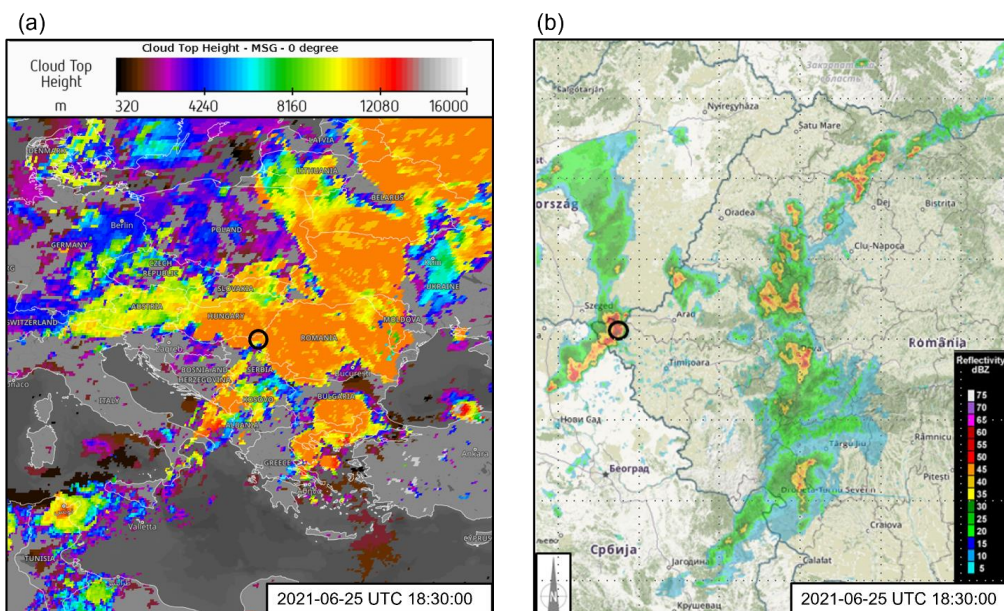


116 acquisition rate of 4 Hz, can measure the wind speed up to 70 m/s. In addition to the anemometer sensor, the tower is  
117 equipped with a temperature sensor installed near the location of the anemometer. The sensor was encased by a protective  
118 case. The working temperature range for this sensor is between -55 and 70 °C.

119

### 120 3 The Sânnicolau Mare (Romania) downburst event of 25 June 2021

121 In this section, a brief overview of the meteorological aspects pertaining to the downburst event in Sânnicolau Mare on  
122 25 June 2021 is provided. In the late afternoon of 25 June 2021, a severe downburst event affected the extreme western  
123 region of Romania. The downburst event took place in the Timis county (Figure 1a) between 18:00 and 19:00 UTC and  
124 struck the little town of Sânnicolau Mare (Figure 1b). At 17:30 UTC, a strong mesoscale convective system moving  
125 toward the east was approaching the town of Sânnicolau Mare. Figure 3a, acquired from Eumetsat, captures an image of  
126 a deep convective cell at 18:30 UTC. This weather phenomenon exhibits cloud tops ascending over 12 km above mean  
127 sea level, signifying the mature stage of the convection cycle. This mature storm cell was observed to have directly  
128 impacted the town under study. Figure 3b presents composite radar reflectivity data, indicating that this meteorological  
129 phenomenon can be classified as a mesoscale convective system known as bow echo. Radar reflectivity values at or above  
130 60 dBZ, as seen in this event, are typically indicative of severe weather conditions. Such conditions are often associated  
131 with the production of hailstones, with an average diameter of approximately 2.5 cm.



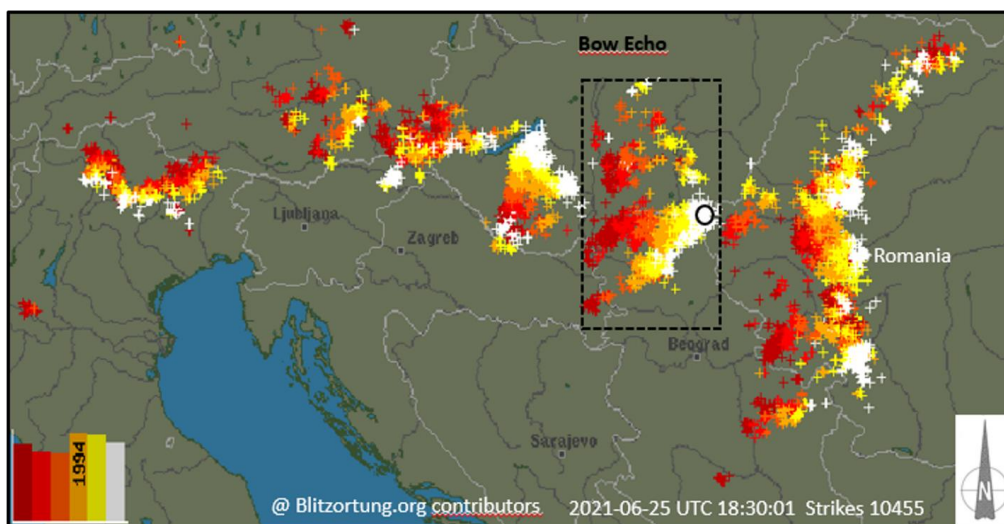
132 **Figure 3. (a) Distribution of cloud top heights derived from Meteosat Second Generation (MSG) valid for 25 June 2021 at 18:30**  
133 **UTC. Data and map obtained from ©EUMETSAT 2022 (<https://view.eumetsat.int>). (b) Composite radar reflectivity (dBZ) for**  
134 **June 25, 2021, at 18:30 UTC. The geographical location of Sânnicolau Mare and the apex of the bow echo are indicated by the**  
135 **black circle. Data and map obtained by ©2018 Administratia Nationala de Meteorologie (<https://www.meteoromania.ro>).**

136 The existence of a robust convective motion, indicative of the typical kinematic structure of a bow echo, is distinctly  
137 portrayed through the distribution of intensive lightning activity, as displayed in Figure 4a. As the figure illustrates, an  
138 approximate total of 10455 lightning strikes were recorded by the Blitzortung.org network across Eastern Europe between

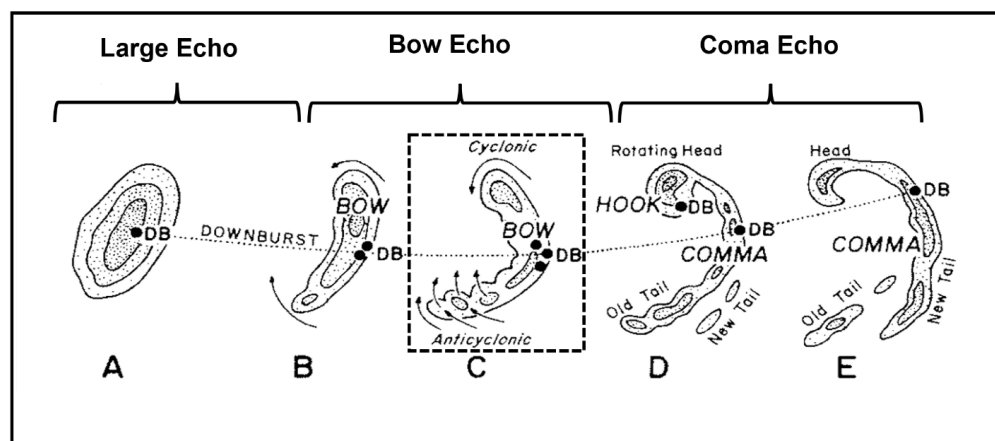


139 16:30 to 18:30 UTC. A significant concentration of these strikes correlates with the bow echo structure near the western  
140 Timis County in Romania. Bow echoes are a prevalent form of severe convective organization. These mesoscale  
141 convective systems can generate straight-lines surface winds that lead to extensive damage associated with downbursts.  
142 On occasion, they may also give rise to tornadoes.

(a)



(b)

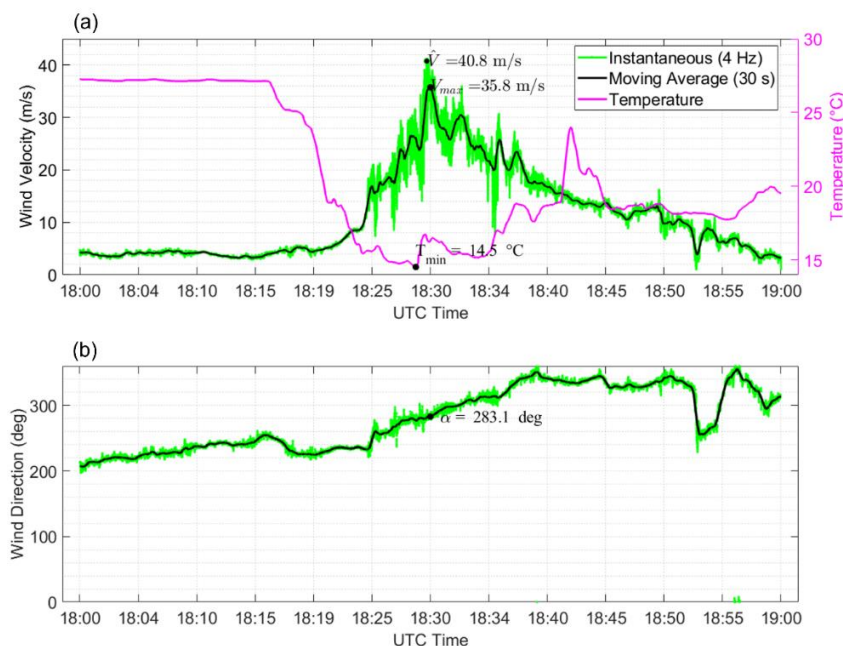


143 Figure 4. (a) Lightning strikes recorded between 16:30 to 18:30 UTC on June 25, 2021, sourced from the Blitzortung.org  
144 network archive for lightning and thunderstorms ([www.blitzortung.org](http://www.blitzortung.org)). The black circle marks the geographic location of  
145 Sânnicolau Mare, situated near the apex of the observed bow echo. (b) Typical radar echo morphology commonly observed in  
146 bow echoes, characterized by the generation of strong downbursts at the bow apex, denoted as DB. Adapted from Fujita (1978).

147 Figure 4b illustrates the characteristic kinematic structure of a bow echo as outlined by Fujita (1978). Typically, the  
148 system originates as a singular, prominent convective cell, either isolated or embedded within a broader squall line system  
149 (Phase A). As the surface winds strengthen, the parent cell undergoes transformation, evolving into a line segment of cells



150 with a bow-shaped configuration (Phase B). During the maximum intensity, the bow's center might develop a spearhead  
151 echo (Phase C), characterized by the occurrence of the most severe downburst winds at the apex of the spearhead. During  
152 the decay phase, the wind system frequently evolves into a comma-shaped echo (Phase E) (Weisman, 2001). The  
153 comparisons between Figures 3b, 4a, and 4b elucidate that the bow echo positioned above Sănnicolau Mare at 18:30 UTC  
154 is in its most intense stage (Phase C), as evidenced by the formation of the characteristic spearhead echo shape. The  
155 intense downburst event generated at the apex of the bow echo was recorded by the anemometer and temperature sensor  
156 situated 50 meters above the ground on the TM\_424 tower. The time histories of the moving average wind speed and  
157 direction (averaged over 30 seconds) (Solari et al., 2015; Burlando et al., 2017) for the recorded one-hour duration of the  
158 downburst event are given in Figure 5a and Figure 5b, respectively. At approximately 18:30 UTC the anemometer  
159 recorded an instantaneous maximum velocity (sampled at 4 Hz) of  $\hat{V} = 40.8$  m/s while the maximum moving average  
160 wind velocity was  $V_{\max} = 35.8$  m/s. This notable high velocity clearly evidences of the occurrence of an intense  
161 downburst. The time interval spanning from 18:20 to 18:45 UTC represents the primary indicator of the downburst's  
162 occurrence in the proximity of the telecommunication tower. This period is characterized by a sudden surge in wind speed,  
163 commonly referred intensification stage followed by a subsequent decrease in velocity after 18:30 UTC. Throughout the  
164 initial phase of intensification, the wind direction exhibited a clockwise rotation, ranging from  $235^\circ$  and extending to  
165 approximately  $360^\circ$ . Additionally, Figure 5a also includes 1-hour time series of the recorded temperature data. The  
166 temperature sensor is positioned at the same location of the anemometer. Before the passage of the downburst, the  
167 environmental temperature was on average  $27^\circ\text{C}$ , while at approximately 18:20 UTC the temperature dropped very  
168 sharply reaching the minimum value of  $14.5^\circ\text{C}$  at approximately 18:30 UTC. After the sharp drop the temperature started  
169 to rise and eventually returned to its pre-storm level (not shown).

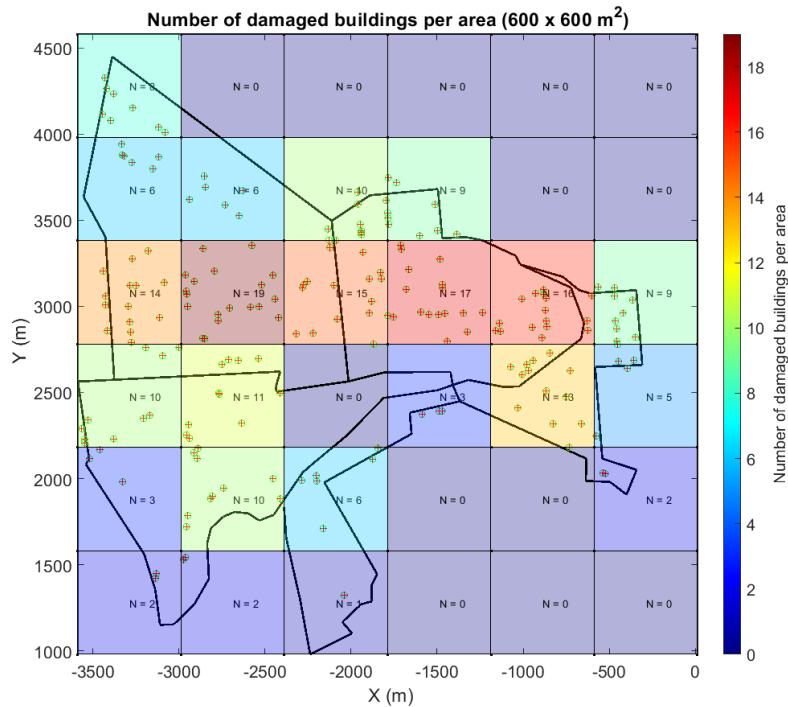


**Figure 5. Telecommunication tower monitoring network measurements from 18:00 to 19:00 UTC on June 25, 2021: (a) Time history of the instantaneous wind speed (green), moving average mean wind speed (black) and temperature record (magenta); (b) Instantaneous (green) and moving average mean wind direction (black).**

170

171 The downburst in Sânnicolau Mare caused widespread hail damage to the facades of numerous buildings. A collaborative  
172 damage survey was conducted by the University of Genoa (Italy) and the University of Bucharest (Romania) (Calotescu  
173 et al., 2022; Calotescu et al., 2023, submitted). The survey identified the affected buildings and produced a comprehensive  
174 map illustrating the hail damage. Figure 6 shows a schematic representation of the distribution of hail damage per area  
175 ( $600 \times 600 \text{ m}^2$ ) and the position of the buildings that suffers hail damage in the town of Sânnicolau Mare.





176

**Figure 6. Spatial distribution of damaged buildings and locations of hail-damaged structures within 600 x 600 m<sup>2</sup> area in the town of Sânnicolau Mare during the downburst event on June 25, 2021. The city boundaries of Sânnicolau Mare are delimited by the black line.**

#### 177 **4 Downburst reconstruction**

178 This section focuses on the modeling, optimization, and reconstruction of the Sânnicolau Mare downburst event. Section  
179 4.1 delves into the modeling and optimization approach used for downburst reconstruction. Section 4.2 introduces  
180 metaheuristic optimization and its application in the reconstruction of the specific downburst event under study. Finally,  
181 Section 4.3 outlines the multivariate data analysis employed to examine the solutions generated by the optimization  
182 algorithm.

#### 183 **4.1 Modeling and optimization approach for downburst reconstruction**

184 In this study, the authors utilize the computational model developed in a previous work by Xhelaj et al. (2020) for the  
185 reconstruction and simulation of the Sânnicolau Mare downburst event discussed in Section 3. The Xhelaj et al. (2020)  
186 model can simulate the spatiotemporal evolution of the bi-dimensional moving average (30 second window) wind speed  
187 and direction experienced during a typical downburst event at a specified height  $z$  above ground level (AGL). The Xhelaj  
188 et al., (2020) model is able to reconstruct/simulate the space-time evolution of the bi-dimensional moving average wind  
189 speed and direction produced during a generic downburst event at a height  $z$  above the ground level (AGL). The wind  
190 system simulated by the model represents the outflow structure of a translating downburst embedded in a synoptic scale  
191 wind, which is considered as constant across the simulation domain. The model comprises 11 variables that describe the  
192 kinematic structure of the downburst wind. Table 1 presents a short description of the 11 variables upon which the model



193 relies. As a result, the model allows for the reconstruction of the time-evolving moving average wind speed and direction  
 194 generated by the simulated downburst at every point within the simulation domain. By employing anemometric wind  
 195 speed and direction data collected during the Sănnicolau Mare downburst event, an optimization procedure can be  
 196 formulated to minimize the relative error (objective function  $F$ ), which quantifies the discrepancy between the observed  
 197 time series of the moving average wind speed and direction and the corresponding simulations generated by the model.  
 198 Since the Sănnicolau Mare downburst event was recorded by an anemometer positioned at a height of 50 meters AGL,  
 199 the analytical model will reconstruct the wind speed and direction at the corresponding height.

200 **Table 1. Variables of the Xhelaj et., al. (2020) analytical model.**

1	X-component touchdown location (at $t = 0$ ) ( $m$ )	$X_{C0}$
2	Y-component touchdown location (at $t = 0$ ) ( $m$ )	$Y_{C0}$
3	Downdraft radius ( $m$ )	$R$
4	Normalized radial distance from the center of the downburst where $V_{r,max}$ occurs (-)	$\rho = \frac{R_{max}}{R}$
5	Maximum radial velocity ( $m/s$ )	$V_{r,max}$
6	Duration of the intensification period ( $min$ )	$T_{max}$
7	Total duration of the downburst event ( $min$ )	$T_{end}$
8	Storm translational velocity ( $m/s$ )	$V_t$
9	Storm translational direction ( $deg$ )	$\alpha_t$
10	ABL wind speed below the cloud base ( $m/s$ )	$V_b$
11	ABL wind direction below the cloud base ( $deg$ )	$\alpha_b$

201

202 The reconstruction procedure gives rise to a mathematical optimization problem characterized by being single-objective,  
 203 nonlinear, and bound constrained, as discussed in Xhelaj et al. (2022). To tackle this optimization problem, the analytical  
 204 model is integrated with a global metaheuristic optimization algorithm. Specifically, the Teaching Learning Optimization  
 205 Algorithm (TLBO) proposed by Rao et al. (2011) is employed. The details pertaining to the integration of the analytical  
 206 model with the optimization algorithm, as well as the estimation of the kinematic variables associated with the downburst  
 207 event, are explained in detail in Xhelaj et al. (2022). The TLBO algorithm it is an iterative, stochastic, and population-  
 208 based algorithm comprising two distinct phases: the Teacher Phase and the Learner Phase. In the Teacher Phase, the best  
 209 solution in the population (the teacher) shares its knowledge (objective function) with the other solutions (the students)  
 210 to enhance their performance. In the Learner Phase, the students interact with each other to further improve their  
 211 performance. TLBO requires only two user-specified parameters: the maximum number of iterations  $T$  and the population  
 212 size  $N_p$ . When incorporating the objective function into a stochastic metaheuristic optimization algorithm, running the  
 213 algorithm independently multiple times is crucial to reach the optimal solution. This iterative approach allows for a deeper  
 214 exploration of the variable space, reducing the risk of getting trapped in local optima. However, it is important to note  
 215 that in the context of metaheuristic optimization, there is no guarantee of attaining a globally optimal solution. As a result,  
 216 the procedure can yield a range of solutions ordered based on the values assumed by the objective function, with some  
 217 being better than others. In this study, the TLBO algorithm is executed 1024 times independently, with each run producing  
 218 an optimal solution. Consequently, 1024 solutions are obtained. The reconstruction of the downburst event can be  
 219 accomplished by selecting the solution with the lowest objective function value, as it is considered the best representation



220 of the event based on the optimization process. This study aims to analyze and clarify the nature of all the solutions  
 221 generated by means of the TLBO algorithm for the downburst outflow reconstruction. This choice was made for a twofold  
 222 reason.

- 223 • The first reason is to determine the best possible solution among the 1024 totals, where best solution is the one  
 224 that minimizes the objective function  $F$ , and allows to reconstruct the Sânnicolau Mare downburst event.
- 225 • The second reason, which is the primary objective of this study, is to analyze these 1024 solutions using  
 226 multivariate data analysis (MDA). The method used in MDA are the Agglomerative hierarchical clustering  
 227 (AHC) coupled with the K-Means algorithm and principal component analysis (PCA).

228 The objective is to investigate the distinct characteristics of the different solutions provided by the TLBO algorithm,  
 229 enabling an understanding of their divergence from the optimal solution. If alternative solutions do exist, it signifies that  
 230 the algorithm's solution is not unique. As such, a more comprehensive definition of the objective function is necessary to  
 231 accurately discern between the optimal solution and its alternatives.

#### 232 4.2 Metaheuristic optimization and reconstruction of the Sânnicolau Mare downburst

233 In metaheuristic optimization, a commonly used guideline suggests setting the population size  $N_p$  as ten times the number  
 234 of variables to estimate  $D$  (Storn, 1996). In this study, where  $D$  corresponds to 11 variables, a population size of  $N_p = 110$   
 235 has been chosen. Additionally, considering the reported fast convergence rate of the TLBO algorithm (as mentioned in  
 236 Xhelaj et al., 2022), the maximum number of iterations  $T$  for this study has been set to  $T = 100$ . Table 2 displays the lower  
 237 and upper bounds of the optimization problem pertaining to the reconstruction of the Sânnicolau Mare downburst. These  
 238 parameter values have been determined based on a comprehensive literature review, available in Xhelaj et al. (2022).

239 **Table 2. Lower and upper bound of the decision variable parameters for the reconstruction of the Sânnicolau Mare**  
 240 **downburst. Table form Xhelaj et al. (2022).**

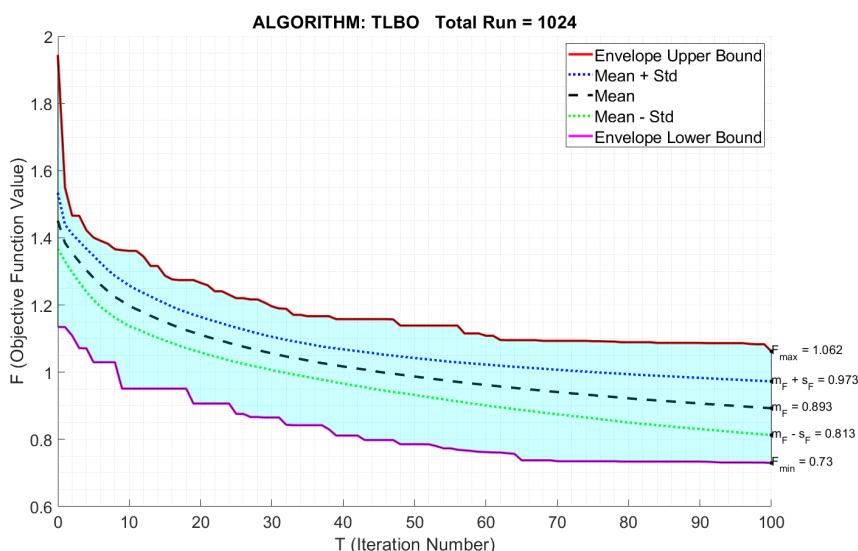
	Parameters/Variables	Lower Bound	Upper Bound
1	$X_{C0} (m)$	-10000	-10000
2	$Y_{C0} (m)$	-10000	-10000
3	$R (m)$	200	2000
4	$\rho = \frac{R_{max}}{R} (-)$	1.6	2.6
5	$V_{r,max} (m/s)$	0	40
6	$T_{max} (min)$	2	15
7	$T_{end} (min)$	15	60
8	$V_t (m/s)$	0	40
9	$\alpha_t (deg)$	0	359.9
10	$V_b (m/s)$	0	40
11	$\alpha_b (deg)$	0	359.9

241

242 The spatial domain of the downburst simulation covers an area of 20 x 20 km<sup>2</sup> while the grid resolution in both the X and  
 243 Y directions is set at 50 m. At the center of the domain is placed the probe that sense the time histories of the wind velocity  
 244 and direction due to the passage of the simulated downburst. Figure 7 illustrates the "performance chart" depicting the  
 245 convergence pattern of the objective functions during the reconstruction of the Sânnicolau Mare downburst using the



246 TLBO algorithm. The performance chart in Figure 7 illustrates the convergence pattern of the objective functions as  
 247 iterations progress. It shows the upper and lower envelopes that encapsulate all 1024 independent runs. The region within  
 248 the envelopes represents the objective function values' trend for all runs. At the end of the 100 iterations, the lower  
 249 envelope represents to the best objective function value obtained, while the upper envelope corresponds to the worst  
 250 objective function value obtained by the TLBO algorithm. The performance chart in Figure 7 includes additional visual  
 251 representations: a dashed line representing the mean convergence curve, and dotted lines representing the mean  
 252 plus/minus one standard deviation curves. These curves provide insights into the average behavior and deviation of the  
 253 objective function values across the 1024 runs. Based on the analysis of the performance charts, it can be observed that  
 254 the TLBO algorithm attains convergence after approximately 70 iterations. At the conclusion of 100 iterations, the best  
 255 and worst objective function values correspond to  $F_{min} = 0.730$  and  $F_{max} = 1.062$ , respectively. The mean and standard  
 256 deviation of the objective function values are determined as  $m_F = 0.893$  and  $s_F = 0.080$ , respectively.



257

258 **Figure 7. Performance chart for the reconstruction/simulation of the Sânnicolau Mare downburst using the TLBO algorithm.**

259

#### 260 4.3 Multivariate data analysis of solutions for the Sânnicolau Mare downburst reconstruction

261 The optimization algorithm provides in output a data table, where each row of the table is a solution of the optimization  
 262 problem. Therefore, the data table is composed of 1024 rows (solutions). The table has 12 columns, where 11 columns  
 263 represent the 11 variables/parameters of the analytical model, while the last column contains the values assumed by the  
 264 objective function  $F$  of each solution (i.e., each row). Although the objective function  $F$ , is not a variable of the analytical  
 265 model, it is treated in Section 5 as a variable from the point of view of the multivariate data analysis. The solutions are  
 266 sorted in descending order based on their objective function value  $F$ . This means that the best overall solution among the  
 267 1024, lies in the last row of the data table. The analysis of the data table indicates that most variables exhibit multimodal  
 268 histograms, with two or more peaks. However, only the variables  $V_b$  and  $\alpha_b$  are characterized by a unimodal histogram.  
 269 Since the aim of this document is to conduct a multivariate data analysis (MDA), the variables of the data table are split



270 into primary and secondary variables. Primary variables participate in the analysis of multivariate data (i.e., AHC + K-  
 271 Means and PCA), as opposed to secondary variables, which have no role in the calculation. However, secondary variables  
 272 can indeed assist in the interpretation of the data table. In the present study,  $V_b$ ,  $\alpha_b$  and  $\alpha_t$  are considered as secondary  
 273 variables. This choice is primarily driven by the observation that  $V_b$ , and  $\alpha_b$  exhibit unimodal histograms, suggesting that  
 274 they may not significantly contribute to distinguishing different cluster solutions. However, the choice of  $\alpha_t$  as a  
 275 secondary variable is purely practical, since it makes it possible to carry out a multivariate statistical analysis, avoiding  
 276 the problem of circular statistics and, hence, simplifying the calculation.

277 Let's define the data table that contains only primary variables by a matrix  $\mathbb{X}$ . Each row  $i$  of the matrix represents a  
 278 solution vector  $\mathbf{X}_i$ , encompassing the values associated with the nine primary variables. Therefore the solution vector can  
 279 be expressed as  $\mathbf{X}_i = (X_{CO_2}, Y_{CO_2}, R_i, \rho_i, V_{r,max_i}, T_{max_i}, T_{f_i}, V_{t_i}, F_i)^T$  with  $i$  ranging from 1 to  $I$ , where  $I$  represents the  
 280 total number of solutions, in this case  $I = 1024$ . Since the solution vector  $\mathbf{X}_i$  contains  $K = 9$  primary variables, the  
 281 resulting data matrix  $\mathbb{X}$  is an  $I$ -by- $K$  matrix with 1024 rows and 9 columns. For the sake of simplicity, in order to shorten  
 282 the notation, let  $X_{ik}$  be the value of the  $k$ -th primary variable in the  $i$ -th solution. Henceforth, the term "variable" will  
 283 refer to primary variables, unless explicitly specified. Consequently, the dataset within the matrix  $\mathbb{X}$  can be regarded either  
 284 as a collection of rows representing solutions to the optimization problem or as a collection of columns representing  
 285 variables of the analytical model. The focus of the MDA lies in examining the data matrix from both the solution and  
 286 variable perspectives, aiming to identify similarities among solutions based on their variables. In essence, the goal is to  
 287 establish a typology of solutions by identifying groups that exhibit homogeneity in terms of variable similarity. This  
 288 analysis allows for a comprehensive understanding of the relationships and patterns among the solutions, facilitating the  
 289 identification of distinct solution clusters based on their shared variable characteristics. Since a generic solution  $\mathbf{X}$ , is a  
 290 set of  $K = 9$  numerical values,  $\mathbf{X}$  evolves within a space  $\mathbb{R}^K$  (a space with 9 dimensions), called "the solution's space".  
 291 Defining in the solution's space the usual Euclidean metric (i.e., the  $L_2$  norm  $\|\cdot\|_2$ ), then, the squared distance between  
 292 two solutions  $\mathbf{X}_i$  and  $\mathbf{X}_l$  can be expressed by the Euclidean distance  $d_{il}$ :

$$d_{il}^2 = d^2(\mathbf{X}_i, \mathbf{X}_l) = \|\mathbf{X}_i - \mathbf{X}_l\|_2^2 = \sum_{k=1}^K (X_{ik} - X_{lk})^2 \quad (1)$$

293 The distance  $d$  possesses the following metric properties:

$$294 \begin{cases} d(\mathbf{X}_i, \mathbf{X}_l) = 0 \iff i = l \\ d(\mathbf{X}_i, \mathbf{X}_l) = d(\mathbf{X}_l, \mathbf{X}_i) \quad (\text{simmetry}) \\ d(\mathbf{X}_i, \mathbf{X}_l) \leq d(\mathbf{X}_i, \mathbf{X}_j) + d(\mathbf{X}_j, \mathbf{X}_l) \quad (\text{tirangle inequality}) \end{cases}$$

295 The Euclidean distance not only enables distance calculations but also allows for the definition of angles and,  
 296 consequently, orthogonal projections. This concept is fundamental in principal component analysis since PCA relies on  
 297 the Euclidean distance as a key component of its methodology. Conducting the analysis directly on the data matrix  $\mathbb{X}$   
 298 could be misleading without any kind of standardization or normalization. Standardization of the data is essential,  
 299 particularly when variables are expressed in different units (refer to Table 1), as it ensures comparability and removes the  
 300 influence of scale variations. Additionally, even when variables share the same units, disparities in the range of variability  
 301 can skew the analysis. Therefore, normalization operations become crucial to provide equal weight and significance to  
 302 each variable, which ultimately leads to a more comprehensive understanding of the data's structure and relationships.  
 303 Therefore, in the present work the variables are standardized according to the following equation:



$$\hat{X}_{ik} = \frac{X_{ik} - \bar{X}_k}{S_k}, \quad \forall i = 1, \dots, I = 1024 \text{ and } \forall k = 1, \dots, K = 9 \quad (2)$$

304 where  $\bar{X}_k$  denotes the sample mean of the  $k$ -th variable calculated over all  $I$  solutions:  $\bar{X}_k = \frac{1}{I} \sum_{i=1}^I X_{ik}$  and  $S_k$  is the  
 305 sample standard deviation of  $k$ -th variable:  $S_k = \sqrt{\frac{1}{(I-1)} \sum_{i=1}^I (X_{ik} - \bar{X}_k)^2}$ .

306 From a geometric standpoint, the standardization operation holds meaningful interpretations within the solution's space  
 307  $\mathbb{R}^K$ . The centring operation  $X_{ik} - \bar{X}_k$  is equivalent to relocating the origin of the reference system to the centre of mass  
 308 of the point cloud. The centre of mass coordinates,  $\bar{X}_k$  (for  $k = 1, \dots, K$ ), represent the new origin. The standardization  
 309 operation, which consists of considering  $\hat{X}_{ik}$  rather than  $X_{ik}$ , modifies the cloud's shape harmonizing its variability across  
 310 all directions. Finally, the normalized data matrix  $\hat{\mathbb{X}}$  containing the set of vectors  $\hat{\mathbf{X}}_i, i = 1, \dots, I$ , has been used in the  
 311 MDA for the identification of different typology of solutions provided by the TLBO algorithm for the  
 312 simulation/reconstruction of the Sănnicolau Mare downburst. Figure 8 showcase a summary statistic in the form of a box  
 313 plot, illustrating the distribution of the standardized variables. Variables such  $\hat{R}_{max}$  and  $\hat{T}_{max}$  have a large number of  
 314 outliers which indicates extreme values within the dataset. Therefore, even in the context of standardized data, outliers  
 315 can still be informative and may hold important information for distinguishing distinct solution clusters.

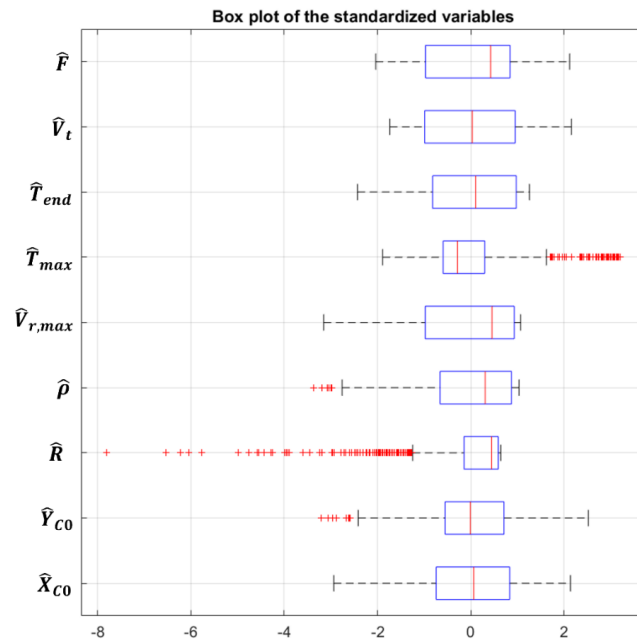


Figure 8. Box plot of the distributions of the standardized variables. Outliers in the data are plotted individually using the red marker symbol + .

## 316 5 Results

317 In the following section the results of multivariate data analysis (MDA) including cluster analysis and principal  
 318 component analysis applied to the data matrix  $\hat{\mathbb{X}}$  is presented. After the clusters have been established a comprehensive



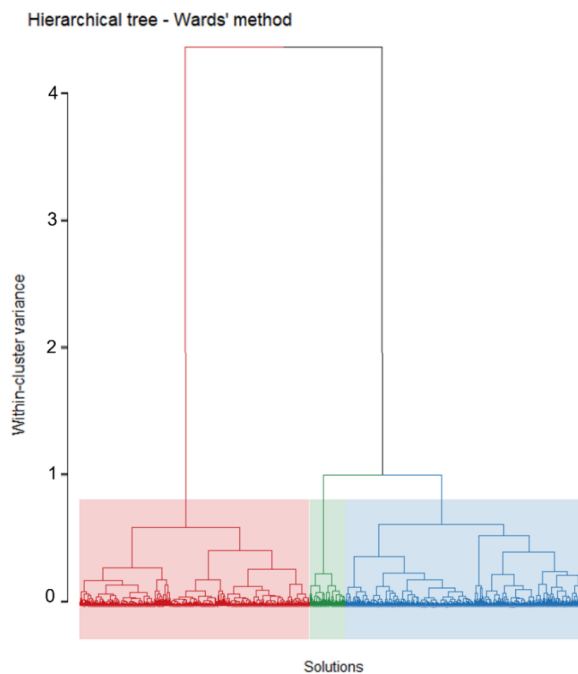
319 description of each of them is provided. This involves examining the variables that contribute to each cluster's composition  
320 as well as identifying specific representative solutions within each cluster. Such an analysis allows for a deeper  
321 understanding of the cluster characteristics and facilitates the interpretation of meaningful patterns and insights within the  
322 data. Sections 5.1 to 5.3 provide an in-depth analysis of data matrix  $\tilde{\mathbf{X}}$  from the variable's perspective, employing  
323 agglomerative hierarchical K – means clustering and principal component analysis. In Section 5.4 the clusters are  
324 analyzed from the point of view of the specific solutions which are the most representative of the clusters. Finally, these  
325 representative solutions are compared with the best overall solution founded from the TLBO algorithm. The comparisons  
326 of the representative solution for each cluster and the best overall solution with the full-scale data is therefore enriched  
327 considering the data from the damage campaign that was carried out after the Sănnicolau Mare downburst event.

### 328 **5.1 Identification of the most meaningful clusters**

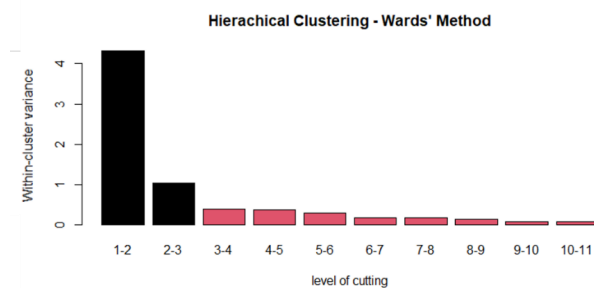
329 In order to identify the appropriate number of clusters for grouping the solutions, the agglomerative hierarchical clustering  
330 (AHC) is firstly employed (Hartigan, (1975), Kaufman and Rousseuw (1990)). In AHC, each individual solution is  
331 initially treated as an independent cluster (leaf). Through a series of iterative steps, the most similar clusters are  
332 progressively merged, forming a hierarchical tree structure known as a dendrogram. This merging process continues until  
333 all the individual clusters are combined into a single cluster (root). Subsequently, the hierarchical tree is analysed, and a  
334 suitable level is chosen to cut the tree, leading to distinct and meaningful clusters. The number of clusters obtained from  
335 the AHC forms a partition of the data set. To refine and optimize this partition, a partitioning clustering algorithm called  
336 K-means (MacQueen, 1967, Hartigan and Wong, 1979) is subsequently applied. Partitioning algorithms, like K-Means,  
337 subdivides the data sets into distinct clusters, ensuring that solutions within each cluster are similar to one another while  
338 exhibiting noticeable differences between clusters. Hence the two steps clustering procedure is called agglomerative  
339 hierarchical K – means clustering (AHK-MC) and is employed to analyse the standardized data matrix  $\tilde{\mathbf{X}}$ . By combining  
340 the strengths of both algorithms AHC and K-means, AHK-MC aims to provide a comprehensive and improved clustering  
341 algorithm of the data, enabling a more accurate identification of distinct solution groups. The hierarchical tree (i.e.,  
342 dendrogram) is constructed following the Wards' method (Ward, (1963)). Figure 9 shows the structure of the dendrogram  
343 obtained according to the Wards' algorithm. Since the total solutions of the optimization problem are  $I = 1024$ , the  
344 dendrogram is very dense at the bottom level (i.e., at the leaf level, where each solution is considered as a cluster by  
345 itself). The hierarchical tree is composed therefore by  $I - 1 = 1023$  nodes, the points were two clusters (solutions or set  
346 of solutions) are merged. The level (height) of each node in the tree is described by the within-cluster variance. The level  
347 of a node in the agglomeration process, when examined from top to bottom, indicates the reduction in within-cluster  
348 variance achieved by merging two connected clusters. This reduction in variance can be visualized using a bar graph, as  
349 depicted in Figure 10. From Figure 10 it is possible to establish the level where to cut the dendrogram and consequently  
350 to establish the number of clusters for partitioning the data set. The choice of the number of clusters is important because  
351 partitioning with too few clusters risk leaving groups which are not at all homogeneous. On the other hand, partitioning  
352 with too many clusters' risks creating classes that are not very different from each other. Being  $\sum_{s=1}^{I-1} \Delta_s = K = 9$  (the  
353 total variance contained in the standardized data), the separation into two groups is able to describe  $\Delta(1,2)/K =$   
354  $4.314/9 = 0.4793$  (47.93 %) of the total variance. Considering the partitioning into three groups, the explained variance  
355 by the three clusters is equal to  $[\Delta(1,2) + \Delta(2,3)]/K = [4.314 + 1.044]/9 = 0.5954$  (59.54 %) of the total variance,  
356 while for four clusters the "explained variance" is equal to  $[\Delta(1,2) + \Delta(2,3) + \Delta(3,4)]/K =$   
357  $[4.314 + 1.044 + 0.406]/9 = 0.6404$  (64.05 %) of the total variance. Therefore, considering more than three clusters  
358 (refer to Figure 10) is going to have a very little impact on the explained variance since very little information is gained



359 and is no longer useful to group together any more classes. For this reason, the dendrogram in this work is partitioned in  
 360 3 clusters (refer to Figure 10) and therefore they can explain approximately 60% of the total variance present in the data.



361 **Figure 9. Hierarchical tree (dendrogram) constructed according to the Wards' Method relative to the solutions of the**  
 362 **optimization problem for the Sănnicolau Mare downburst.**



363 **Figure 10. Bar graph of the relation between the number of merged clusters and the within-cluster variance.**

364 **5.2 Interpretation of the clusters through pca and optimization using k-means**

365 The three clusters of solutions are analyzed using principal component analysis (PCA) to identify the key variables that  
 366 drive the system's behavior. By extracting the principal components, which captures the most significant variation in the  
 367 data, the complexity of the system can be reduced. In particular, the eigenvalues of the correlation matrix  $\mathbb{S} = \frac{1}{(I-1)} \mathbb{X}^T \mathbb{X}$   
 368 quantify the amount of variance accounted by each principal component (Kassambara, 2017). The eigenvalues shows that





369 the first components have larger values, indicating that they capture the most significant variation in the data set. In  
370 contrast, the subsequent components have lower eigenvalues, representing a diminishing level of variation.

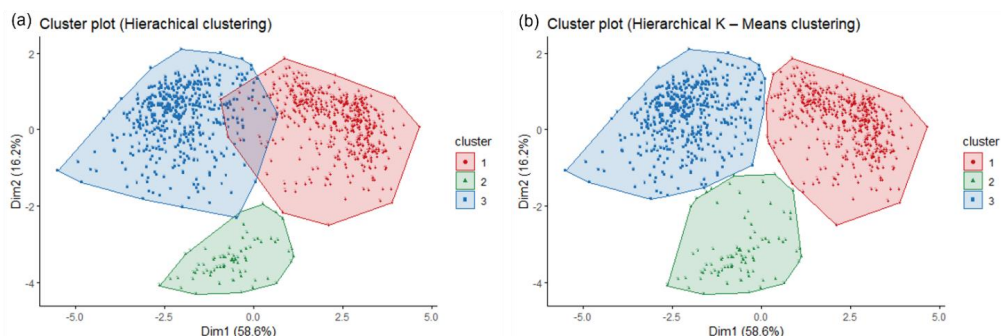
371 Table 3 presents displays the eigenvalues, the percentage of variance explained by each component, and the cumulative  
372 percentage of variance.

373 **Table 3. PCA results in term of the eigenvalues, percentage of variance and cumulative percentage of variance.**

	Dim-1	Dim-2	Dim-3	Dim-4	Dim-5	Dim-6	Dim-7	Dim-8	Dim-9
Eigenvalues ( $\lambda_s$ ) or variance	5.278	1.458	0.884	0.499	0.378	0.195	0.167	0.093	0.048
Percentage of variance	58.645	16.204	9.825	5.542	4.204	2.170	1.852	1.028	0.530
Cumulative perc. of variance	58.645	74.849	84.674	90.216	94.420	96.589	98.441	99.470	100.000

374

375 The first two principal components capture 74.85% of the total variance in the dataset. These components define a plane  
376 that provides significant insights into the underlying patterns and structure of the data. Eigenvalues greater than 1 (Table  
377 3) signify that the respective principal components explain more variance in the data compared to any single standardized  
378 variable. These principal components capture significant patterns and structures in the data, contributing more to the  
379 overall variability. In contrast, eigenvalues less than 1, starting from the third principal component (Table 3) indicate that  
380 the associated principal components explain less variance than individual standardized variables, suggesting they have  
381 relatively less influence on the overall variability in the data. Therefore, it is probably not useful to interpret the next  
382 dimensions and better focusing on the first two principal dimensions for a more meaningful analysis. It is worth  
383 mentioning that the percentage of variance explained by the first principal component (58.65 %) is very close to the  
384 variance explained by the hierarchical tree when is partitioned into three clusters (59.54 %). The three clusters, founded  
385 using the Wards' method only, are represented in terms of solutions in the principal component map (Figure 11a). This  
386 figure shows how solutions are grouped together into three clusters when the overall cloud of solutions is projected into  
387 the first two principal components. Here, cluster 1 is not very well separated from cluster 3, which means that both clusters  
388 share similar solutions. In order to have a better partitioning, the partition is improved (or "consolidated") by applying  
389 the K-Means algorithm to the initial partition (Figure 11a) founded by the Wards' method. Figure 11b shows the principal  
390 component map of the final partitioning of the solutions as result of the application of the K-means algorithm. The  
391 application of the K-Means algorithm improves the partitioning quality since the three clusters this time are very well  
392 separated from each other and are more compact. This final partitioning is therefore maintained and used for the next  
393 analysis of this paper.



394 **Figure 11. (a). Solutions' clusters partitioning on the principal component map, using the Ward's method only. (b) Solutions'**  
 395 **clusters partitioning using the hierarchical K-means method.**

396

397 **5.3 Further considerations on the model's parameters**

398 Table 4 provides insights into the interpretation of the results from the perspective of the variables, focusing on the first  
 399 two principal components. The table displays the correlations between the variables and these components (column 1 and  
 400 4). Additionally, it includes the quality of the representation (i.e., projection) of each variable on the first two principal  
 401 components (columns 2 and 5), as well as the weight of each variable to the construction of these components (columns  
 402 3 and 6) (Husson and Pagès (2017)). It is important to note that the variables in the Table 4 are vectors which represent  
 403 the values observed across the 1024 solutions.

404 Table 4. Principal component analysis results for variables in terms of correlations, quality of representation and contribution to the  
 405 construction relative the first two principal components.

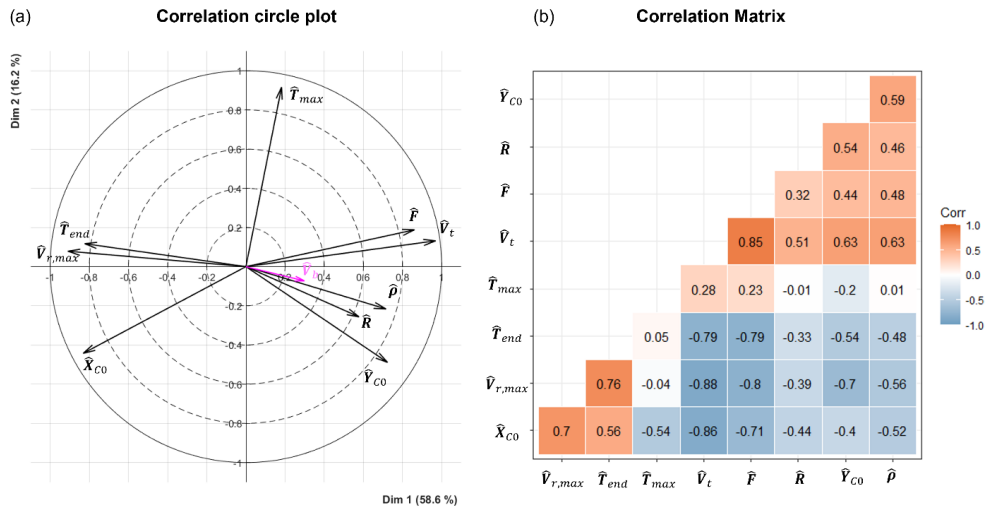
Variables $\hat{V}_k$	Dim-1 $r(\hat{V}_k, p_1)$	Dim-1 $qlt_1(\hat{V}_k)$	Dim-1 $qtr_1(\hat{V}_k)$	Dim-2 $r(\hat{V}_k, p_2)$	Dim-2 $qlt_2(\hat{V}_k)$	Dim-2 $qtr_2(\hat{V}_k)$
$\hat{X}_{CO}$	-0.831	0.691	13.094	-0.443	0.196	13.441
$\hat{Y}_{CO}$	0.723	0.523	9.912	-0.489	0.239	16.377
$\hat{R}$	0.578	0.334	6.326	-0.256	0.066	4.504
$\hat{\rho}$	0.715	0.512	9.699	-0.216	0.047	3.200
$\hat{V}_{r,max}$	-0.909	0.827	15.664	0.079	0.006	0.424
$\hat{T}_{max}$	0.182	0.033	0.628	0.916	0.839	57.502
$\hat{T}_{end}$	-0.823	0.678	12.847	0.117	0.014	0.942
$\hat{V}_t$	0.969	0.939	17.789	0.132	0.017	1.189
$\hat{F}$	0.861	0.741	14.042	0.188	0.035	2.421
<i>Secondary variable</i>						
$\hat{V}_b$	0.299	0.089	-	-0.073	0.005	-

406

407 In Table 4, is also present the secondary variable  $\hat{V}_b$ . The other two variables  $\hat{a}_t$  and  $\hat{a}_b$  are non-considered since their  
 408 interpretation is not consistent with the principal component analysis approach. Despite  $\hat{V}_b$  not being involved in the  
 409 construction of the principal components, it is still possible to evaluate the correlation and the quality of the representation



410 of this variable using the two principal components. To facilitate the interpretation of Table 4, a correlation circle plot  
 411 (Abdi and Williams, 2010) can be used to visually represent the variables. This plot represents each variable as a point in  
 412 a two-dimensional space, where the coordinates of each point correspond to the correlation coefficients between the  
 413 variable and the two principal components (i.e.,  $r(\hat{V}_k, \mathbf{p}_1), r(\hat{V}_k, \mathbf{p}_2)$ ). Figure 12a illustrates the correlation circle plot.

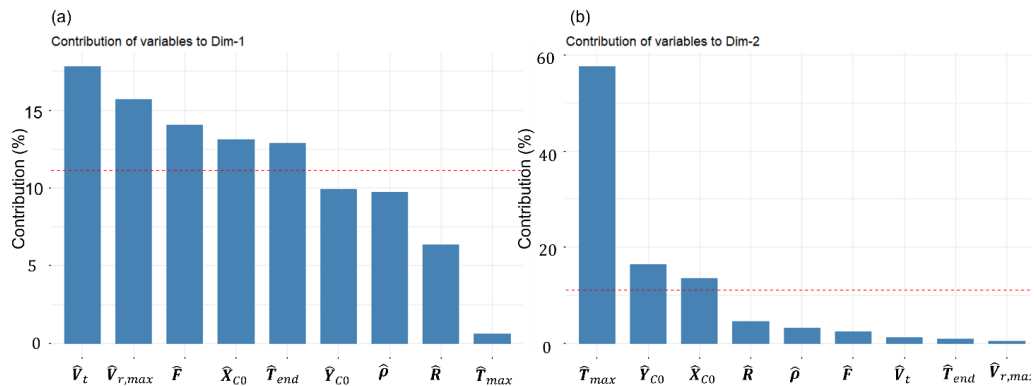


414 **Figure 12. (a) Correlation circle plot. The variables in black are considered as primary variables whereas the variable in**  
 415 **magenta is a secondary variable (b) Correlation matrix plot.**

416 This plot offers a geometric representation of the correlations among all variable pairs, making it easier to observe the  
 417 grouping of positively correlated variables and the positioning of negatively correlated variables on opposite sides relative  
 418 to the origin. The total contribution of each variable's representation across all principal components is always equal to  
 419 1 (i.e.,  $\sum_{s=1}^K qlt_s(\mathbf{V}_k) = 1$ ). If a variable is explained by the first two components, then the sum of its contribution  
 420  $\sum_{s=1}^{K=2} qlt_s(\mathbf{V}_k)$  will be equal to 1. This implies that the variable's location on the correlation circle will exactly lie on the  
 421 circumference of radius 1. Hence, a low-quality variable, which is not very well represented by the first two principal  
 422 components will be positioned close to the center of the circle. Therefore, only well represented variables can be  
 423 interpreted from the correlation circle. Except for the variables  $\hat{V}_b$  and  $\hat{R}$  which are not very well represented by the 2  
 424 principal components, the remaining variables are very well represented since their tip is close to the circle of radius 1.  
 425 The set of variables  $\{\hat{V}_t, \hat{F}, \hat{Y}_{CO}, \hat{\rho}\}$  are positively correlated with each other; this means that an increase in one variable  
 426 is followed by an increase in the other variable. The same is true for the variables  $\{\hat{V}_{r,max}, \hat{X}_{CO}, \hat{T}_{end}\}$ . The variable  $\hat{V}_t$   
 427 is highly correlated with the first dimension (correlation of 0.97). This variable could therefore summarize itself the first  
 428 principal component axis. From Figure 12a is possible to show that the variable  $\hat{V}_t$  is strongly negatively correlated with  
 429 the variables  $\{\hat{V}_{r,max}, \hat{X}_{CO}, \hat{T}_{end}\}$ . This means for example that solutions which are characterized by high value of storm  
 430 motion  $V_t$  will systematically be characterized by low values in the maximum radial velocity  $V_{r,max}$ , “low values” of the  
 431 touch down component  $X_{CO}$  with respect to the station (which means that for lower positive values of  $X_{CO}$ ,  $X_{CO}$  will be  
 432 close to the station, while for lower negative values of  $X_{CO}$ ,  $X_{CO}$  will be far from the station) and low values of the total  
 433 duration of the downburst event  $T_{end}$ . Since  $\hat{V}_t$  is positively correlated with the variables  $\{\hat{F}, \hat{Y}_{CO}, \hat{\rho}\}$ , what is true for  $\hat{V}_t$



434 with respect to the group of variables  $\{\hat{V}_{r,max}, \hat{X}_{CO}, \hat{T}_{end}\}$ , will also remain true for the variables  $\{\hat{F}, \hat{Y}_{CO}, \hat{\rho}\}$ . Finally, from  
 435 the correlation circle plot, it seems that the variable  $\hat{T}_{max}$  is not very well “linearly” correlated with the groups of variables  
 436  $\{\hat{V}_{r,max}, \hat{T}_{end}, \hat{\rho}\}$  since it is nearly orthogonal with these variables. From a quantitative point of view the values of the  
 437 correlation coefficients between all the pairs of variables are plotted in Figure 12b. Table 4 shows also the values of the  
 438 variable’s contribution for the construction of the two principal components (columns 3 and 6 respectively). Also in this  
 439 case, it is possible to plot these values to understand which variable contribute the most for building the first two principal  
 440 axes. Figure 13a and Figure 13b show respectively the contribution of the variables expressed in percentage for the  
 441 reconstruction of the first two principal components.



442 **Figure 13. (a) Contribution of the variables in the reconstruction of the first principal component (Dim-1). (b) Contribution**  
 443 **of the variables in the reconstruction of the second principal component (Dim-2). Variables are sorted from the strongest to**  
 444 **the weakest. The red dashed line indicates the expected average contribution.**

445 The graph displays a red dashed line representing the expected average contribution. If the contribution of variables were  
 446 evenly distributed, the expected value would be calculated as 1 divided by the number of variables  $K$ , which in this case  
 447 is 9. This would result in an expected average contribution of 11.11%. For a given component, a variable with a  
 448 contribution larger than this cutoff could be considered as important in contributing to the construction of the component.  
 449 Therefore, the set of variables  $\{\hat{V}_t, \hat{V}_{r,max}, \hat{F}, \hat{X}_{CO}, \hat{T}_{end}\}$  contribute the most to the construction of the first principal  
 450 component (Dim-1), while the set of variables  $\{\hat{T}_{max}, \hat{Y}_{CO}, \hat{X}_{CO}\}$  contributes the most for the construction of the second  
 451 principal component (Dim-2). Since the contribution can be added, the set of variables that contributes the most for the  
 452 construction of Dim-1 and 2 are given by the set of variables  $\{\hat{V}_t, \hat{X}_{CO}, \hat{T}_{max}, \hat{V}_{r,max}, \hat{F}, \hat{Y}_{CO}\}$  which are ordered from the  
 453 strongest to the weakest. The remaining variables  $\{\hat{T}_{end}, \hat{\rho}, \hat{R}\}$  have a contribution which is lower than the threshold  
 454 11.11 %. It is important to observe the partitioning in strongest variables and weakest ones does not represent necessarily  
 455 a general case, since the partition might depend on the downburst case under study.

#### 456 5.4 Physical description of the solutions corresponding to clusters 1-3

457 Once the partitioning of the solutions of the optimization problems in three cluster is completed, it is important to have a  
 458 closer look at them and describe common features of solutions which belong to the same cluster. From the partition  
 459 analysis, it is found that cluster 1 is made up of 481 solutions, cluster 2 of 85 and cluster 3 of 458 solutions. Table 5  
 460 summarizes a few key statistics related to the three clusters. This table includes primary and secondary (i.e., not used for  
 461 clustering) variables, which are no longer standardized to investigate their physical meaning.

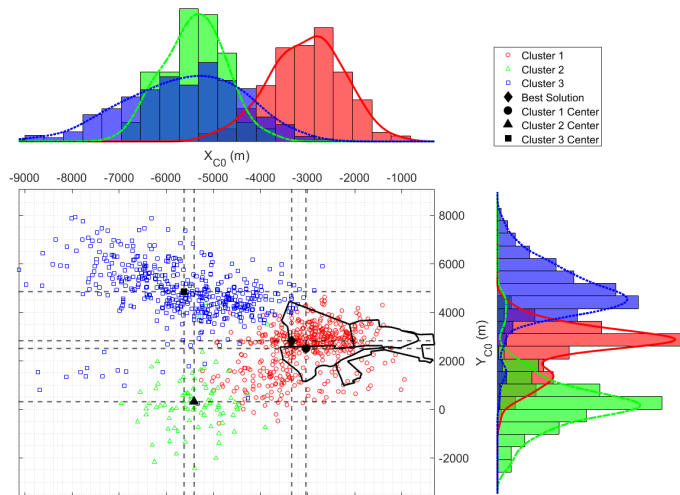


462 **Table 5. Description of the partition by the mean and standard deviation of all the variables.**

Variables $V_k$	Overall Mean	Overall Std	Cluster 1 Mean	Cluster 1 Std	Cluster 2 Mean	Cluster 2 Std	Cluster 3 Mean	Cluster 3 Std
$V_t$ (m/s)	6.025	3.371	2.811	1.042	6.527	1.407	9.307	1.492
$X_{C0}$ (m)	-4386.350	1613.337	-3034.079	789.682	-5410.461	629.282	-5616.465	1209.346
$T_{max}$ (min)	6.954	2.517	5.860	1.172	13.336	1.910	6.919	1.797
$V_{r,max}$ (m/s)	24.293	5.356	28.639	1.465	28.182	1.793	19.006	3.266
$F$ (-)	0.893	0.080	0.823	0.058	0.919	0.043	0.961	0.021
$Y_{C0}$ (m)	3363.669	1809.316	2499.896	975.450	313.553	1257.946	4836.890	1160.613
$T_{end}$ (min)	26.035	3.167	28.269	1.895	27.622	2.295	23.394	2.235
$\rho$ (-)	2.189	0.108	2.126	0.104	2.134	0.100	2.265	0.050
$R$ (m)	1334.478	102.519	1289.518	124.661	1301.969	90.475	1387.728	23.115
Secondary variables								
$\alpha_t$ (deg)	290.383	0.480	276.439	0.416	253.518	0.217	310.868	0.229
$V_b$ (m/s)	6.811	0.670	6.648	0.774	6.705	0.768	7.002	0.449
$\alpha_b$ (deg)	268.218	0.118	264.854	0.138	273.055	0.074	270.827	0.055

463

464 In columns 2-3, the overall mean, and the overall standard deviation (std) are calculated with respect to each variable  
 465 (primary and secondary). In the other columns, the same calculation was repeated taking into consideration the three  
 466 clusters. Mean and the std of the secondary variables  $\alpha_t$  and  $\alpha_b$  have been calculated using circular statistics (Rao and  
 467 Sengupta, 2001). To start clarifying the characteristics of the different clusters, Figure 14 shows the scatter plot and  
 468 distribution of the touchdown components ( $X_{C0}, Y_{C0}$ ) for all the solutions, partitioned into three clusters. In this figure it  
 469 is shown the center (namely the mean) of each cluster and the location of the touchdown position of the best overall  
 470 solution. The figure shows also with a black line the position of the city of Sănnicolau Mare. Also, on the left and on the  
 471 top of this figure is possible to show the histograms of the variable ( $X_{C0}, Y_{C0}$ ) relative to each cluster.



472 **Figure 14. Scatter plot and histogram density distribution for the variables  $(X_{C0}, Y_{C0})$ . The dark black line shows the**  
 473 **contours of the city Sănnicolau Mare.**

474

475 The three clusters appear well separated in terms of touchdown position  $(X_{C0}, Y_{C0})$ . Since it is very unlikely that the  
 476 cluster means coincide with one of the solutions present in the data set, let's define as "cluster solution", the solution  
 477 which is the closest to the mean of the cluster. Accordingly, the cluster solutions, reported in Table 6, will be used to  
 478 interpret the average features of each cluster. The first row of this table is dedicated to the best solution founded by the  
 479 optimization algorithm (i.e., the one that have the lowest objective function  $F$  among all the solutions); the best solution  
 480 belongs to Cluster 1.

481 **Table 6. Overall best solution and clusters representative solutions.**

482

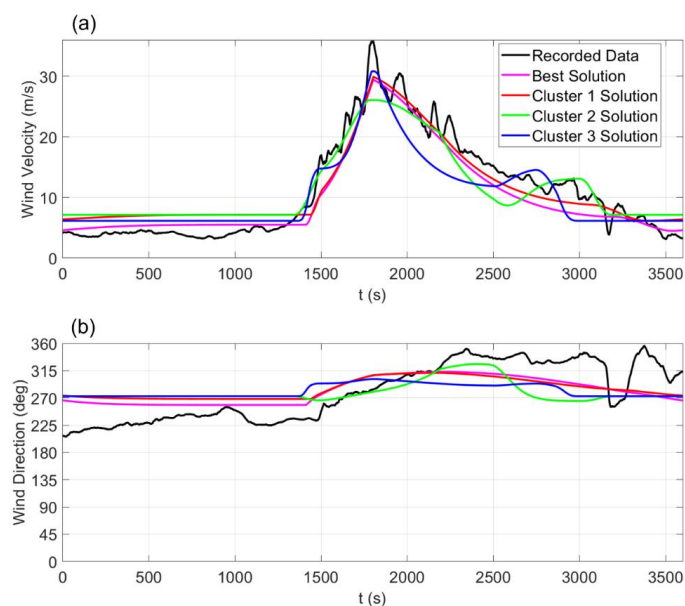
Solutions	$V_t$ (m/s)	$X_{C0}$ (m)	$T_{max}$ (min)	$V_{r,max}$ (m/s)	$F$ (-)	$Y_{C0}$ (m)	$T_{end}$ (min)	$\rho$ (-)	$R$ (m)	$\alpha_t$ (deg)	$V_b$ (m/s)	$\alpha_b$ (deg)
Best solution	2.76	-3339.53	6.50	29.80	0.73	2826.55	29.89	2.15	1381.38	271.74	5.49	58.35
Cluster 1	2.51	-2944.15	6.05	29.54	0.81	2769.36	27.23	2.09	1287.53	278.25	7.15	268.19
Cluster 2	6.14	-5105.66	14.05	27.07	0.91	383.39	28.18	2.14	1295.33	255.36	7.13	272.82
Cluster 3	9.25	-5930.81	7.15	17.36	0.97	4575.50	22.95	2.27	1392.86	307.61	6.15	272.71

483

484 Figure 15 shows the time histories produced by the best solution and the three cluster solutions, in terms of wind velocity  
 485 (Figure 15a) and direction (Figure 15b), compared with the moving averaged recorded data. The figure provides a  
 486 qualitative representation of the goodness of fit between the simulations and the recorded data. The goodness of fit is  
 487 quantitatively measured by the objective function  $F$ . The simulations produced from the best solution and the Cluster 1  
 488 solution fit the data better than Cluster 2 and 3. This is quite obvious since the best solution have the lowest objective



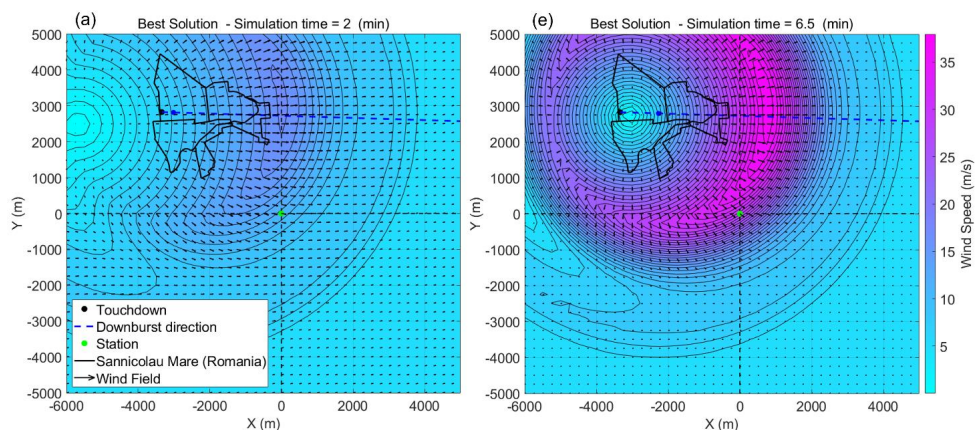
489 function  $F$  and belongs to Cluster 1, whereas Cluster 2 and Cluster 3 solutions have slightly higher objective function  
490 values  $F$  (refer to column 5 in Table 6).



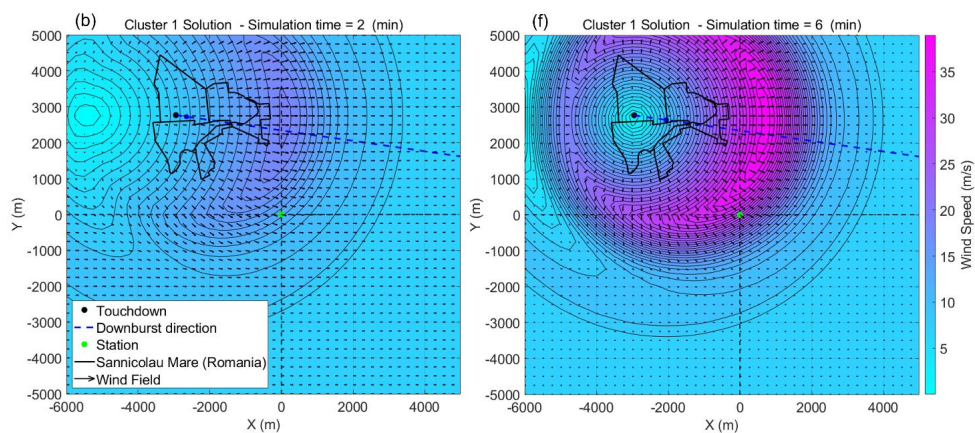
491

492 **Figure 15. Comparison among the moving averaged wind speed (a) and direction (b) obtained from the measurements of the**  
493 **Sânnicolau Mare downburst, along with the best solution and the three cluster solutions.**

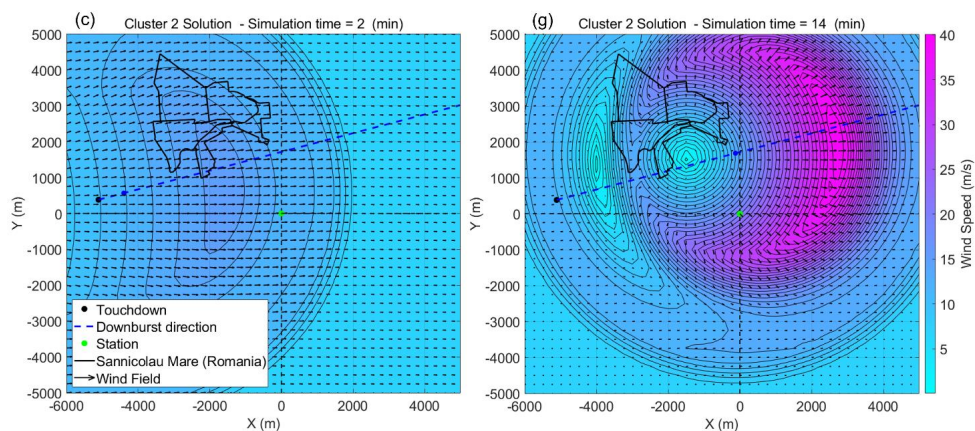
494 In order to better understand the nature of the different solutions relative to each cluster, for each solution present in Table  
495 6, the downburst 2D wind velocity is evaluated at the same height of the anemometric station (i.e., at 50 m AGL). The  
496 left panels of Figure 16 (from (a) to (d)) show for each of the 4 solutions the wind filed reconstruction during the  
497 intensification stage of the downburst, while the right panels (from (e) to (h)) describes the stage of maximum intensity.  
498 Note that the time of maximum intensity is different for each cluster according to the corresponding value of  $T_{max}$   
499 reported in column 3 of Table 6.



500

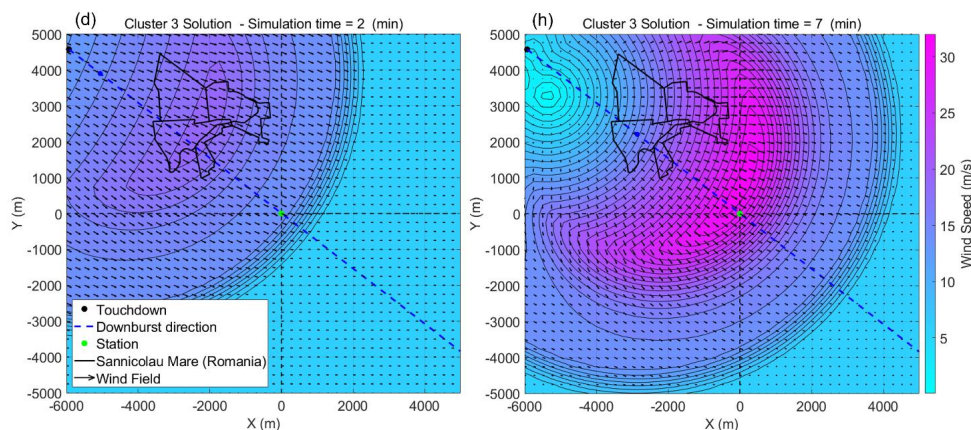


501



502





503

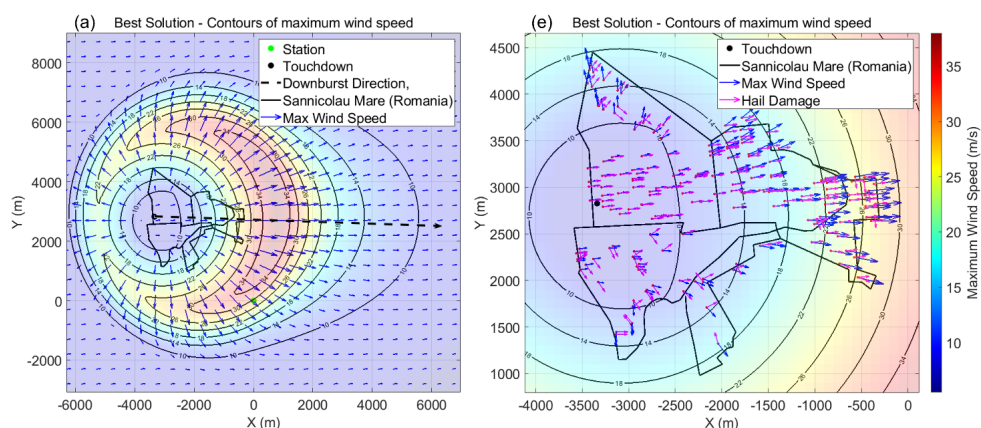
504 **Figure 16. 2D wind field reconstruction at 50 m AGL. From (a) to (d), the best solution, Cluster 1, 2 and 3 solutions are depicted**  
505 **at the intensification stage of the downburst. From (e) to (h), the best solution, Cluster 1, 2 and 3 solutions are depicted at the**  
506 **maximum intensification stage of the downburst.**

507 Cluster 1 touches down very close to the city center and moves slowly eastward, it is characterized by a low value of the  
508 downburst translation velocity  $V_t$ , with mean value 2.8 m/s against the overall mean among all clusters which is 6.0 m/s.  
509 In addition, it has maximum radial velocity  $V_{r,max}$  higher and overall duration of the downburst event  $T_{end}$  longer with  
510 respect to the mean values of the other two clusters. The solutions belonging to Cluster 2 touch down around 2 km  
511 southwest of the city, they propagate northeastward with higher translation velocities compared to Cluster 1 and the  
512 longest intensification periods  $T_{max}$  overall. The solutions in Cluster 3 touch down about 3 km northwest of the city, they  
513 move southeastward with the highest values of downburst translation velocity  $V_t$  but they are the lowest-lasting as the  
514 duration of the downburst event  $T_{end}$  is on average 23.4 min while the overall mean is 26.0 min. They also have the  
515 lowest values of maximum radial velocity  $V_{r,max}$  which compensate the high translation velocities. According to these  
516 descriptions, it is clear that in the solution's space of the model three different solutions exist that can describe similarly  
517 the time-series measured at TM\_424. The existence of different plausible solutions means that the problem of finding the  
518 downburst wind field time-space evolution using a single time-series is an underdetermined problem.

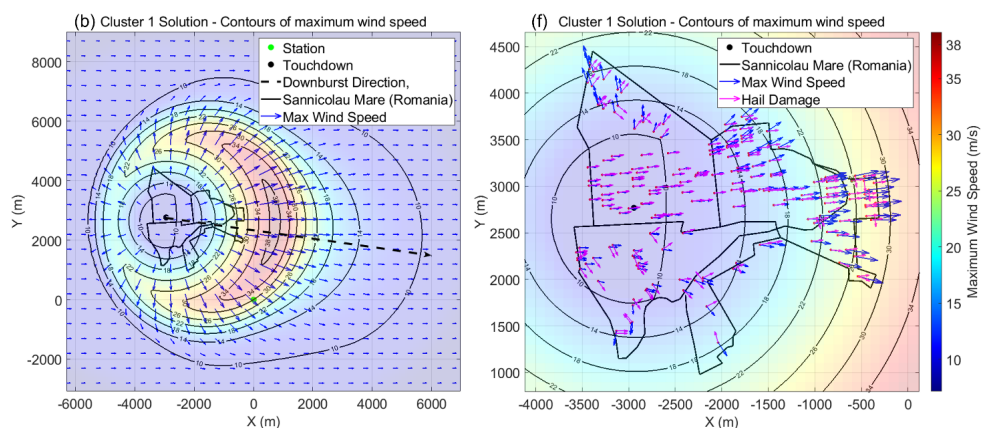
519 The Sannicolau Mare downburst had a strong impact, causing hail damage to numerous buildings in the town. A damage  
520 survey was conducted to assess the affected areas and identify buildings that experienced hail damage during the event.  
521 To estimate the extent of the damage, the simulated wind field generated by the analytical model was utilized. By  
522 analyzing the wind speeds at various locations, the "footprint" of the simulated damage was determined. This footprint  
523 represents the maximum wind speed recorded at different places during the downburst, providing valuable information  
524 on the areas most affected by the event. The left panels of Figure 17, labeled from (a) to (d), depict the complete footprint  
525 of the downburst potential damage area for the best solution and the three cluster solutions. In contrast, the right panels,  
526 labeled from (e) to (h), provide a closer view of the footprints overlaying the simulated maximum wind velocity vectors  
527 (indicated by blue arrows) onto the locations of hail damage. The hail damage is represented by vectors pointing  
528 orthogonally to the damaged facades (represented by pink arrows). The comparison between the facades damage, which  
529 is related to the trajectory of hail transported by the strong downburst-related outflow, and the simulated maximum



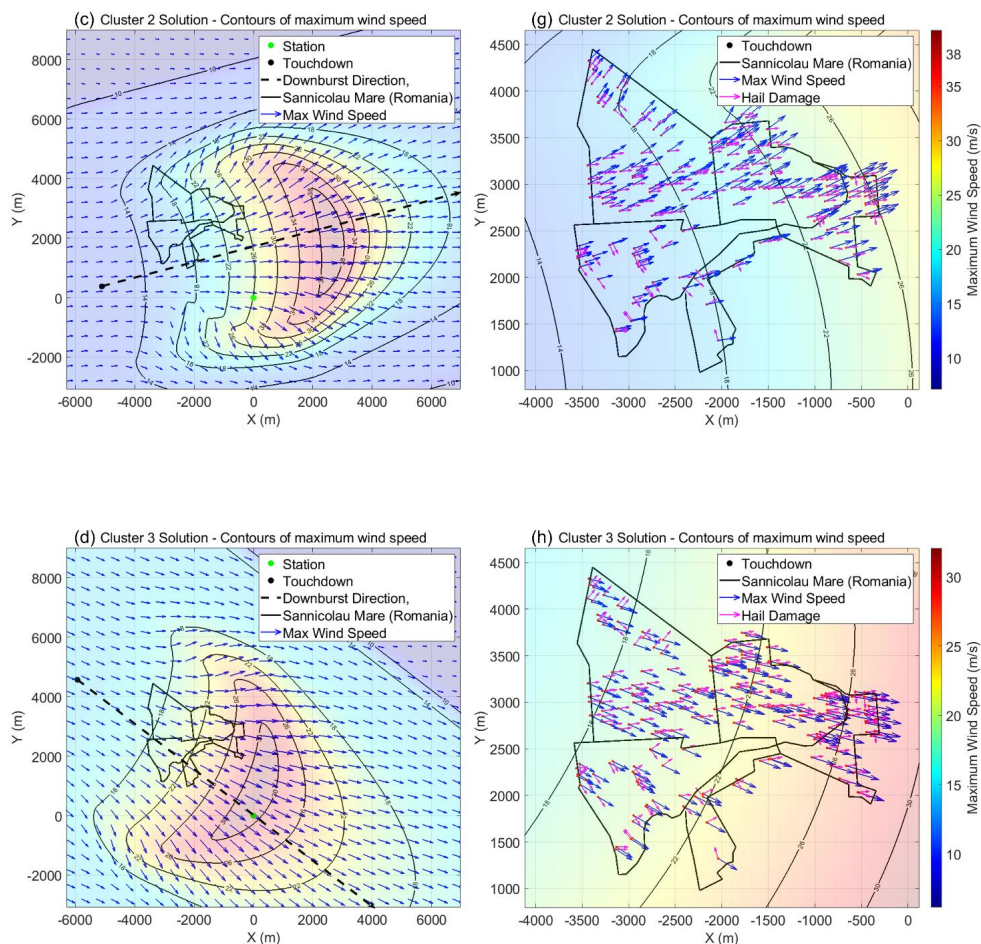
530 velocity reveals interesting findings. Specifically, the best solution and Cluster 1 solutions exhibit the strongest alignment  
531 between the maximum wind velocity vectors and hail damage vectors, particularly in the central part of the city and along  
532 the path of the downburst. In contrast, Cluster 2 and Cluster 3 demonstrate a consistent deviation of the maximum velocity,  
533 with Cluster 2 deviating northward and Cluster 3 deviating southward, relative to the hail trajectories. This observation  
534 suggests that the actual downburst event likely followed a pattern more closely resembling Cluster 1 rather than the other  
535 two potential solutions.



536



537



538

539

540 **Figure 17. Simulated footprints of the downburst that occurred in Sannicolau Mare. Panels from (a) to (d) representing the**  
541 **footprints for the best solution, Cluster 1, Cluster 2, and Cluster 3 respectively. Panels from (e) to (h), representing**  
542 **comparison between hail damage and maximum simulated wind speed for the best solution, Cluster 1, Cluster 2, and Cluster**  
543 **3 respectively.**

544 These observations lead to the conclusion that the optimal (best) solution, which minimizes the objective function  $F$ , is  
545 the most reliable among all possible solutions. This implies that the reconstruction of the downburst wind field should be  
546 based on a large number of simulations to ensure that the optimal solution is obtained. By conducting numerous  
547 simulations, the likelihood of obtaining the most accurate representation of the downburst event is maximized leading to  
548 a more accurate reconstruction of the event.

549

550



## 551 7 Conclusions

552 This study focuses on the analysis of solutions obtained by combining an analytical model (Xhelaj et al., 2020) with a  
553 global metaheuristic optimization algorithm for the reconstruction of the wind field generated during the Sănnicolau Mare  
554 downburst event in Romania on June 25, 2021. The analytical model and optimization algorithm are coupled using the  
555 Teaching Learning Optimization Algorithm (TLBO) to estimate the kinematic parameters of the downburst outflow. The  
556 procedure for this coupling and parameter estimation is described in detail in the study by Xhelaj et al. (2022). Therefore,  
557 the objective was to analyse the differences among the solutions provided by the optimization algorithm and to assess  
558 their physical validity as alternatives to the optimal solution. In the presence of multiple physically sounding solutions, it  
559 has been demonstrated that additional data describing the downburst thunderstorm event is necessary to determine which  
560 solution best represents reality. To support the analysis a comprehensive damage survey was conducted in collaboration  
561 with the University of Genoa (Italy) and the University of Bucharest (Romania) to assess the extent and location of hail  
562 damage on buildings in the affected area. This survey, along with the wind speed and direction signals recorded during  
563 the downburst event by a telecommunication tower located approximately 1 km from the city, significantly enhances the  
564 information available for the reconstruction and simulation of the downburst using the optimization procedure. The  
565 analysis of the solutions generated by the optimization algorithm involves multivariate data analysis (MDA) techniques,  
566 specifically agglomerative hierarchical clustering coupled with the K-means algorithm (AHK-MC) and principal  
567 component analysis (PCA). The AHK-MC is used for classifying the solutions into different clusters based on their  
568 features, while PCA is employed to determine the importance of the variables in the analytical model for the downburst  
569 event reconstruction.

570 The application of AHK-MC resulted in the identification of three main clusters, each with distinct characteristics, among  
571 the 1024 solutions.

- 572 • Solutions belonging to Cluster 1 are characterized by a slow storm motion, small touch down distance from the  
573 city of Sănnicolau Mare and by long duration of the downburst event. The best overall solution belongs to Cluster  
574 1.
- 575 • Solutions belonging to Cluster 2 are characterized by a moderate storm motion and moderate distance of the  
576 touch-down from the town of Sănnicolau Mare. These solutions are also characterized by high duration of the  
577 intensification period of the downburst event.
- 578 • Solutions belonging to Cluster 3 are characterized by a high storm motion and high distance of the touch-down  
579 from Sănnicolau Mare. They are also characterized by low duration of the downburst event and low values of  
580 the maximum radial velocity.

581 The result of the MDA allows also to establish at least for the case under consideration that the set of variables  
582  $\{V_t, X_{CO}, T_{max}, V_{rmax}, F, Y_{CO}\}$  which are ordered from the strongest to the weakest are the more important for the  
583 reconstruction/simulation of the downburst event. The remaining variables  $\{T_{end}, \rho, R\}$  have a lower contribution. It is  
584 important to observe the partitioning in strongest variables and weakest ones does not represent a general case, since the  
585 partition depends on the downburst case under study.

586 Finally, the comparison between the facades damage, which are related to the trajectory of hails transported by the strong  
587 downburst-related outflow and the simulated maximum velocity shows that the best solution and Cluster 1 solutions seem  
588 to have a “good” overlapping between maximum wind velocity vectors and hail damage vectors. Considering the solutions  
589 of Cluster 2 and 3, it seems that the match between maximum wind velocity vectors gradually decreases, with the worst



590 case represented by Cluster 3 solutions. These observations allow to conclude that the optimal solution, that is, the one  
591 that minimizes the objective function  $F$ , is the best with respect to the other three cluster solutions, also from the point of  
592 view of the damage analysis. As a result, for the specific case being examined, relying on the best overall solution provided  
593 by the optimization algorithm appears to yield promising results for reconstructing the downburst wind field. Obviously,  
594 an analysis of this type, conducted on several downburst events, will be able to better confirm this statement.

#### 595 **Author contributions**

596 This paper is based on the Ph.D. thesis of Andi Xhelaj, under the guidance of Prof. Giovanni Solari and Prof. Massimiliano  
597 Burlando. Andi Xhelaj played a crucial role in conceptualizing the study, developing the methodology, organizing the  
598 data, conducting data analysis, and preparing the manuscripts and figures. The study was supervised by Prof.  
599 Massimiliano Burlando, who provided guidance and conducted internal review process.

#### 600 **Declaration of competing interest**

601 The authors affirm that they have no known financial conflicts of interest or personal relationships that could have  
602 influenced the findings presented in this paper.

#### 603 **Acknowledgments**

604 The authors would like to acknowledge the valuable contributions of I. Calotescu, X. Li, M.T. Mengistu, and M.P. Repetto  
605 for providing the time histories of the recorded data and the hail damage map from the survey of the Thunderstorm event  
606 in Sănnicolau Mare, Romania, on 25 June 2021. The monitoring system and damage survey were carried out as part of  
607 a research collaboration between the University of Genoa (UniGe) and the Technical University of Civil Engineering  
608 Bucharest (UTCB), funded by the European Research Council (ERC) under the European Union's Horizon 2020 research  
609 and innovation program (Grant Agreement No. 741273) for the Project THUNDERR - Detection, simulation, modeling,  
610 and loading of thunderstorm outflows to design wind-safer and cost-efficient structures, supported by an Advanced Grant  
611 (AdG) 2016. As the corresponding author is not a native English speaker, he utilized the OpenAI's GPT-4 model as an  
612 editing tool during the creation of this paper to review and amend grammatical and spelling mistakes and to ensure  
613 linguistic consistency and coherence throughout this paper.

614

#### 615 **References**

616

- 617 1. Abdi, H., and Williams, L., J.: Principal component analysis. Wiley Interdisciplinary Reviews: Computational  
618 Statistics 2(4):433 – 459, 2010.
- 619 2. Amato, F., Guignard, F., Robert, S. Kanevski, M.: A novel framework for spatio-temporal prediction of  
620 environmental data using deep learning. Sci Rep 10, 22243, doi:10.1038/s41598-020-79148-7, 2020.  
621
- 622 3. Bjerknes, J., Solberg, H.: Life cycle of cyclones and polar front theory of atmospheric circulation. Geophysiks  
623 Publikationer 3, 3-18, 1922.
- 624 4. Burlando, M., Antonelli M., Ratto, C., F.: Mesoscale wind climate analysis: identification of anemological regions  
625 and wind regimes. Int. J. Climatol. 28(5), 629–641. DOI: 10.1002/joc.1561 – ISSN 1097-0088, 2008.



- 626 5. Burlando, M.: The synoptic-scale surface wind climate regimes of the Mediterranean Sea according to the cluster  
627 analysis of ERA-40 wind fields. *Theor. Appl. Climatol.* 96(1), 69-83. DOI: 10.1007/s00704-008-0033-5 – ISSN  
628 0177-798X – eISSN 1434-4483, 2009.
- 629 6. Burlando, M., Romanic, D., Solari, G., Hangan, H., Zhang, S.: Field data analysis and weather scenario of a  
630 downburst event in Livorno, Italy on 1 October 2012. *Mon. Weather Rev.* 145, 3507-3527, 2017.
- 631 7. Calotescu, I., Bîtcă D., Repetto, M., P.: Full-scale behaviour of a telecommunication lattice tower under wind  
632 loading. *Lightweight Structures in Civil Engineering*, XXVII LSCE Łódź, 2nd – 3rd of December 2021, Szafran J.  
633 and Kaminski M, Eds. p. 15-18, 2021.
- 634 8. Calotescu, I., Repetto, M., P.: Wind and structural monitoring system for a Telecommunication lattice tower. 14th  
635 Americas Conference on Wind Engineering, May 17-19, 2022, Lubbock, TX, 2022.
- 636 9. Calotescu, I., Bîtcă, D., Repetto, M. P.: Full-scale monitoring of a telecommunication lattice tower under synoptic  
637 and thunderstorm winds. *J Wind Eng. Ind. Aerodyn.* (submitted), 2023.
- 638 10. Calotescu, I., Li, X., Mengistu, M.T., Repetto, M., P.: Post-event Survey and Damage Analysis of An Intense  
639 Thunderstorm in Sannicolau Mare, Romania. 14th Americas Conference on Wind Engineering, May 17-19, 2022,  
640 Lubbock, TX, 2022.
- 641 11. Calotescu, I., Li, X., Mengistu, M.T., Repetto, M., P.: Thunderstorm – induced damage to built environment: A field  
642 measurement and post-event survey. *J Wind Eng. Ind. Aerodyn.* (submitted), 2023.
- 643 12. Davenport, A., G.: The application of statistical concepts to the wind loading of structures. *P. I. Civ. Eng.* 19, 449-  
644 472., 1961.
- 645 13. Fujita, T., T.: Manual of downburst identification for project Nimrod. *Satellite and Mesometeorology Research Paper*  
646 156, Dept. of Geophysical Sciences, University of Chicago, 104 pp, 1978.
- 647 14. Fujita, T., T.: Downburst: Microburst and Macrobust. *Univ. Chic. Press II*, p. 122, 1985.
- 648 15. Hartigan, J., A.: *Clustering Algorithms*. Wiley, New York, 1975.
- 649 16. Hartigan, J., A. and Wong, M., A.: A K-means clustering algorithm. *Applied Statistics* 28, 100–108, 1979.
- 650 17. Hjelmfelt, M.R.: Microburst and Macrobust: windstorms and blowdowns. In: Johnson, E.A., Miyaniishi, K., (Eds.),  
651 *Plant Disturbance Ecology*. Academic Press, Amsterdam, pp. 59-102, 2007.
- 652 18. Husson, F., Lê, S., & Pagès, J.: *Exploratory Multivariate Analysis by Example Using R* (2nd edition). CRC Press,  
653 2017.
- 654 19. Jiang, Y., Cooley, D., Wehner, M., F.: Principal Component Analysis for Extremes and Application to U.S.  
655 Precipitation. *Journal of Climate*, Volume 33, 15, 2020.
- 656 20. Kassambara, A.: *Practical Guide to Principal Component Methods in R (Multivariate Analysis Book 2)* (1st edition).  
657 STHDA. ASIN: B0754LHRMV, 2017.
- 658 21. Kaufman, L. and Rousseuw, P.: *Finding Groups in Data. An Introduction to Cluster Analysis*, Wiley & Sons, New  
659 York, 1990.
- 660 22. Letchford, C.W., Mans, C., Chay, M.T.: Thunderstorms – their importance in wind engineering (a case for the next  
661 generation wind tunnel). *J. Wind Eng. Ind. Aerodyn.*, 90, 1415-1433, 2002.
- 662 23. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth*  
663 *Berkeley Symposium on Mathematical Statistics and Probability*, eds L. M. Le Cam & J. Neyman, 1, pp. 281–297,  
664 1967.
- 665 24. Rao, S., J., and Sengupta, A.: *Topics in Circular Statistics*. World Scientific, 2001.



- 666 25. Rao, R.V., Savsani, V.J. and Vakharia, D.P.: Teaching–learning-based optimization: a novel method for constrained  
667 mechanical design optimization problems, *Comput. Aided Des.*, 43(3), pp. 303–315, 2011.
- 668 26. Solari, G.: *Wind science and engineering*. Springer, Switzerland, 2019.
- 669 27. Solari, G.: Emerging issues and new frameworks for wind loading on structures in mixed climates. *Wind Struct.* 19,  
670 295-320, 2014.
- 671 28. Solari, G., Burlando, M., De Gaetano, P., Repetto, M., P.: Characteristics of thunderstorms relevant to the wind  
672 loading of structures. *Wind Struct.* 20, 763-791, 2015.
- 673 29. Solari, G., Burlando, M., Repetto, M., P.: Detection, simulation, modelling and loading of thunderstorm outflows to  
674 design wind-safer and cost-efficient structures. *J. Wind Eng. Ind. Aerodyn.*, 200, 2020.
- 675 30. Ward, J., H., Jr.: Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical*  
676 *Association*, Volume 58, Issue 301, 236-244, 1963.
- 677 31. Weisman, M. L.: Bow echoes: A tribute to T.T. Fujita. *Bulletin of the American Meteorological Society*, 82(1), 97-  
678 116, 2001.
- 679 32. Xhelaj, A., Burlando, M., Solari, G.: A general-purpose analytical model for reconstructing the thunderstorm  
680 outflows of travelling downbursts immersed in ABL flows. *J Wind Eng. Ind. Aerodyn.* 207 104373, 2020.
- 681 33. Xhelaj, A., Burlando, M.: Application of metaheuristic optimization algorithms to evaluate the geometric and  
682 kinematic parameters of downbursts. *Advances in Engineering Software*. Volume 173, November 2022, 103203.,  
683 2022.
- 684
- 685
- 686