

PRECIPITATION CLASSIFICATION FUNCTIONS
FOR NORTHWEST BULGARIA:
GNSS IWV AND INSTABILITY INDICES

Martin Slavchev^{1,2}, Tsvetelina Dimitrova³, Guergana Guerova¹

Received on April 11, 2025

Presented by K. Stoykova, Corresponding Member of BAS, on May 27, 2025

Abstract

Precipitation monitoring during the convective storm season May–September is conducted routinely by the Bulgarian Hail Suppression Agency (HSA). Since 2020, in Northwest Bulgaria data from four stations of the Global Navigation Satellite Systems (GNSS) have been processed in near-real time mode and vertically Integrated Water Vapour (IWV) has been computed in operational manner. As a part of a storm nowcasting demonstrator GNSS IWV and Instability Indices (InI) thresholds are implemented for Sofia and Central Bulgaria. In this work site-specific classification functions are computed for Northwest Bulgaria. A GNSS derived monthly IWV threshold separates well precipitation (P) and no precipitation (nP) groups in July, August and September. Probability of detection is between 77–100% for July and 78–85% for August. For July the false alarm ratio scores are high in the range 30–66%, which limits the use of IWV. Classification functions based on InI and IWV have the best performance with an increase of probability of detection score by 16% in July, 23% in August, and 20% in September and decrease of false alarm ratio score by 23% in July, 22% in August, and 8% in September.

Key words: GNSS, IWV, precipitation, InI, nowcasting

Introduction. Use of GNSS tropospheric products to derive water vapour is a well-established technique for atmospheric monitoring in Europe (GNSS meteorology). GNSS meteorology establishment across Europe was achieved within

This study is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project No BG-RRP-2.004-0008-C01.
<https://doi.org/10.7546/CRABS.2025.08.11>

COST Action ES1206 “Advanced Global Navigation Satellite Systems tropospheric products for monitoring severe weather events and climate” (GNSS4SWEC) [1]. GNSS4SWEC facilitated application of GNSS tropospheric products for severe weather forecasting and nowcasting. A review by [2] concluded that the most frequently studied extreme events with GNSS meteorology are convective storms and heavy rainfall. However, they report that GNSS IWV findings are restricted to isolated severe weather cases, which limits their nowcasting potential. BONAFONI et al. [2] conclude that additional studies are required to define GNSS based thresholds, indicators and/or indices leading to heavy rainfalls and convective storms. SHOJI [3] is the first to propose two indices for water vapour inhomogeneity in Japan using the GNSS carrier phase. A study by GUEROVA et al. [4] uses multiple discriminant analysis to compute calcification functions using the InI, IWV and InI&IWV for thunderstorm days in Sofia, Bulgaria. They report that best scores for probability of detection and false alarm ratio are obtained by a classification function combining InI and IWV.

An active hail suppression campaign coordinated by HSA takes place from May to September over dense agricultural regions in Northwest and South Central parts of Bulgaria. HSA conducts operational forecasting and nowcasting [5] of vigorous thunderstorms producing heavy rain and hail by using radar images, surface and satellite observations in combination with Numerical Weather Prediction model. The challenge remains to resolve the high temporal and spatial variability of severe weather events. From 2017, HSA and Sofia University “St. Kliment Ohridski” (SU) have been partners in the project “Balkan-Mediterranean Real-time Severe Weather Service” (BeRTISS, 2017–2020). As part of the BeRTISS project, the HSA operational tools were augmented by installation of ground based 9 GNSS stations. Since February 2020, the Sofia University GNSS Analysis Centre has provided operational Near-Real Time (NRT) GNSS products to HSA. Using statistical regression analysis, classification functions are obtained that contribute to the thunderstorm forecasting skill.

In 2021, an operational Storm Demonstrator (Storm Demo) was initiated as a prototype for the real-time provision of GNSS products for storm nowcasting in Bulgaria [6]. Storm Demo targets development of service centered at GNSS products for two hail suppression regions in Northwestern and Central Bulgaria and is accessible via a public geoportal. As reported in [6] the real-time IWV and InI are used to derive site-specific thresholds for GNSS stations in Central Bulgaria. The aim of this manuscript is to support Storm Demo operations by evaluation and implementation of site-specific classification functions for precipitation using GNSS IWV and InI for Northwest Bulgaria in the period May–September.

Data and method. Since 2019, the ground based GNSS stations in Northwest Bulgaria have been installed at Gramada (GRAM), Bardarski Geran (BGER), Dolno Cerovene (DCER) and Dolni Dubnik (DDUB) (Fig. 1). The stations are operated by HSA. The stations deliver hourly Rinex files to the operational server

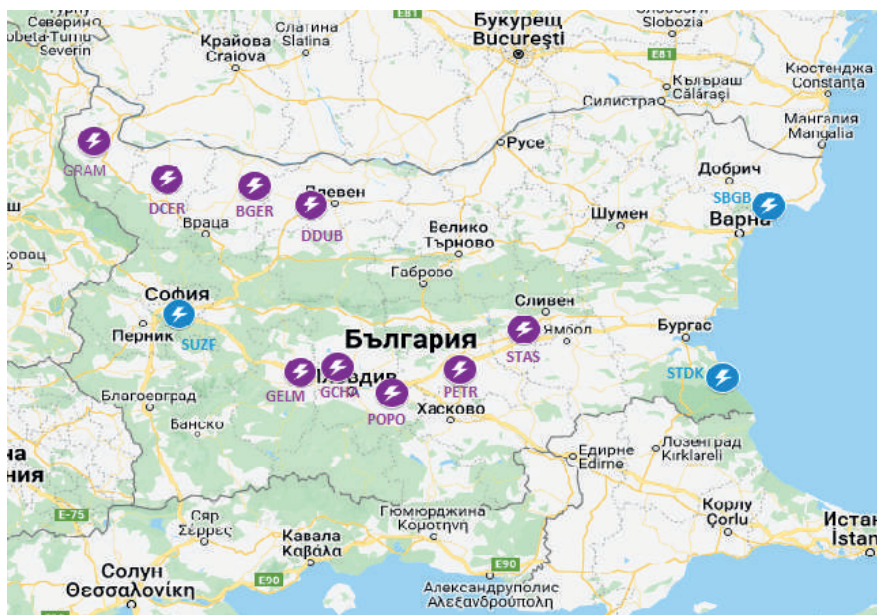


Fig. 1. Map of Bulgaria with location of GNSS stations operated by Hail Suppression Agency (purple circles) and Sofia University (blue circles)

of Sofia University GNSS Analysis Centre (SUGAC). The SUGAC GNSS data processing is conducted in Near Real Time (NRT) mode for ten Bulgarian and nine European IGS stations. The GNSS NRT processing is conducted with a double difference network (DD) approach integrated in the TropNET system, which runs the Bernese 5.2 Engine [7].

In order to minimize the processing redundancy and provide a highly robust solution short data batches are implemented in the TropNET. The robustness is also supported for estimating consistently coordinates and tropospheric parameters within a single process. The following features are characteristics for the TropNET: 1) 4-hour data processing window, 2) adaptable network clustering, 3) dynamic parallelization of all processing steps, 4) all parameters estimated using optimal stacking of previous solutions (NEQs) in temporal and spatial domains, 5) various integrated control points supporting elimination of specific problems, 6) flexible configuration and input orbit product utilization, 7) self-supporting strategy based on GNSS data and precise global products only, 8) self-initialization and automatic update of site coordinates, and 9) tropospheric product quality control and filtering.

The processing outputs are tropospheric files with observable Zenith Total Delay (ZTD). They are saved at a dedicated MySQL database at Sofia University Atmospheric Data Archive designed for storage of GNSS and meteorological information. IWV is computed from ZTD following [8,9].

The proposed in [4] classification function has been applied for separation of Precipitation (P) and no Precipitation (nP) days in the period May–September 2020–2022. Multiple discriminant analysis is performed with StatSoft v13.5.0.17 to classify the days as P or nP using six InI and IWV for the GNSS stations in Northwest Bulgaria, namely GRAM, DCER, BGER and DDUB.

The InI are as follows: Total Totals Index (TT, [12]), K Index (KI, [13]), Severe Weather Threat Index (SWEAT, [12]), Convective Available Potential Energy (CAPE, [14]), Lifted index (LI, [15]), and BRN – Bulk Richardson Number (BRN, [16]). To compute the InI archived vertical profiles data (temperature, relative humidity, pressure and wind speed and direction) at 1200 UTC are used. The profiles are from the Global Data Assimilation System (GDAS) Archive operated by the National Oceanic and Atmospheric Administration (NOAA). GDAS data archive is with resolution 1 deg and can be interpolated to specific location via a web interface (<https://www.ready.noaa.gov/gdas1.php>).

Classification functions are determined separately for each month and station. Two skill scores were calculated for each classification function namely, a) Probability of detection (POD) [10] computed in %, and b) False alarm ratio (FAR) [10] in %.

Results. As a first step monthly IWV threshold is computed for each station for the period May–September 2020–2022. Figure 2 presents: 1) the median, 2) 10–90 percentile, and 3) 25–75 percentile. The monthly IWV(P) values are higher than IWV(nP) for all months and stations. In particular, a very good separation between monthly IWV medians for P and nP groups is seen for all stations for July, August and September. For example in July the IWV(nP) has a median of about 30 kg/m² while the IWV(P) is about 10 kg/m² higher which is about 30%. Furthermore in the period July–September there is no overlap between 25–75 percentiles for IWV(P) and IWV(nP) groups (Fig. 2).

Monthly IWV thresholds are calculated for four GNSS stations separately (data not shown) [11]. A classification function is computed using IWV for each month with the POD and FAR values for each month (data not shown) [11]. The highest number of correctly classified P/nP days was in July and August, with POD values in the range 77–100% and 78–85%, respectively. However, in July the FAR values were also high, 30% for station DCER and 66% for DDUB. For May, June, and September, the FAR values are above 40% for most of the stations and the highest FAR of 61% in DDUB in May. Thus, it can be concluded that due to the high FAR values classification function based on IWV alone is with limited application.

As next iteration, a stepwise discriminant analysis is carried out to compute two classification functions based on 1) InI (F1), and 2) InI in combination with IWV (F2). Table 1 presents the monthly classification functions F1 and F2 that provide the best separation between P and nP groups for station BGER [11]. Stations DCER, DDUB and GRAM are not shown. Interestingly, the InI selected

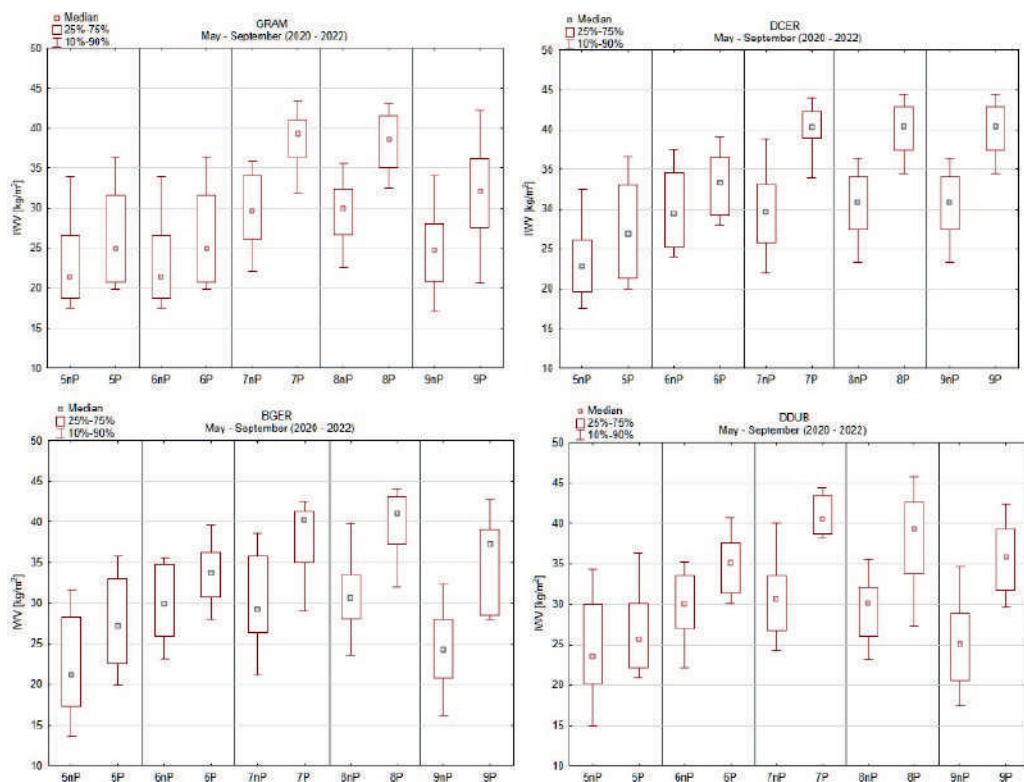


Fig. 2. Box and whisker plot of IWV median (black square) and 10, 25, 75, and 90 percentiles for P and nP days, from May to September 2020–2022 at a) GRAM, b) DCER, c) BGER, and d) DDUB

for the monthly F1 and F2 differ. For example, F1 for station GRAM and May includes five InI (SWEAT, TT, KI, CAPE, and BRN), while F2 is with three InI (KI, BRN, and SWEAT) and IWV. On the other hand, for station BGER in May both F1 and F2 include all six InI. For station DCER the F1 for May includes three InI (LI, TT and KI), while F2 has five InI (SWEAT, LI, TT, CAPE and BRN). For station DCER the F1 for May has two InI (KI and BRN), while F2 has four InI (SWEAT, TT, KI, and CAPE). This comparison also shows that the InI for the same month but different stations are also different. The same combination of InI is not repeated for the stations located in the same climatic region in Northwest Bulgaria. This fact points out the importance of local factors, which drive the severe weather and require separate classification functions for each GNSS station.

The skill scores POD and FAR of the F1 and F2 classification functions are presented in Fig. 3. The dark colour pallet is for F1 and the light colour is for F2. As can be seen, at four stations the POD values for F2 classification function are equal or higher than the respective POD values computed with F1 function for the period June–September. Only for the month of May and stations GRAM and

T a b l e 1

Regression Equations for GRAM and BGER Models

Month	Model	Function	Equation
May	GRAM	F1	$-SWEAT \times 0.0111 + TT \times 0.0245 + KI \times 0.1227 + CAPE \times 0.0009 + BRN \times 0.0091 - 3.0813$
		F2	$KI \times 0.1058 + BRN \times 0.0164 + IWV \times 0.0985 - SWEAT \times 0.0096 - 4.0442$
June	GRAM	F1	$KI \times 0.1957 - CAPE \times 0.0010 - 4.9313$
		F2	$IWV \times 0.1100 - SWEAT \times 0.0087 - LI \times 0.4209 - TT \times 0.0416 + KI \times 0.1473 - CAPE \times 0.0023 - 3.6158$
July	GRAM	F1	$-LI \times 0.5814 - 0.1090$
		F2	$IWV \times 0.3936 + CAPE \times 0.0046 - SWEAT \times 0.0186 - 12.4538$
August	GRAM	F1	$SWEAT \times 0.0113 + LI \times 0.1359 + TT \times 0.3226 - 15.5683$
		F2	$IWV \times 0.4658 + TT \times 0.5670 + LI \times 0.9661 + CAPE \times 0.0031 - 42.1633$
Sep.	GRAM	F1	$SWEAT \times 0.0166 + LI \times 0.4570 + TT \times 0.2271 + KI \times 0.0412 - 14.2820$
		F2	$IWV \times 0.1821 + SWEAT \times 0.0087 + LI \times 0.6273 + TT \times 0.3683 - 24.1305$
May	BGER	F1	$SWEAT \times 0.0007 + LI \times 0.3425 + TT \times 0.2985 + K \times 0.1276 - CAPE \times 0.0004 - BRN \times 0.0028 - 17.1507$
		F2	$IWV \times 0.0972 + SWEAT \times 0.0001 + LI \times 0.4003 + TT \times 0.3646 + K \times 0.0680 - CAPE \times 0.0006 - BRN \times 0.0007 - 21.2268$
June	BGER	F1	$-SWEAT \times 0.0059 - LI \times 0.0508 - TT \times 0.0755 + CAPE \times 0.0000 + K \times 0.1303 - BRN \times 0.0003 + 0.6951$
		F2	$IWV \times 0.1698 - SWEAT \times 0.0041 - LI \times 0.5233 - CAPE \times 0.0025 - BRN \times 0.0012 - 3.8495$
July	BGER	F1	$CAPE \times 0.0045 + K \times 0.0936 - TT \times 0.0570 - 1.2142$
		F2	$IWV \times 0.1203 + SWEAT \times 0.0086 + LI \times 0.2997 - TT \times 0.0860 + KI \times 0.0730 + CAPE \times 0.0060 - BRN \times 0.0057 - 5.0894$
August	BGER	F1	$SWEAT \times 0.0149 + LI \times 0.1343 - TT \times 0.0520 - KI \times 0.0726 + CAPE \times 0.0089 - BRN \times 0.0122 - 3.4237$
		F2	$IWV \times 0.3850 - SWEAT \times 0.0064 + LI \times 0.4313 + TT \times 0.0871 - KI \times 0.0351 + CAPE \times 0.0080 - BRN \times 0.0149 - 17.4159$
Sep.	BGER	F1	$SWEAT \times 0.0203 + LI \times 0.2547 + TT \times 0.0387 + KI \times 0.1070 + CAPE \times 0.0186 - BRN \times 0.0409 - 8.1559$
		F2	$IWV \times 0.3307 - SWEAT \times 0.0054 - LI \times 0.5130 + TT \times 0.2415 + KI \times 0.0161 + CAPE \times 0.0262 - BRN \times 0.1403 - 22.8156$

BGER F1 POD values are 11 and 9% higher than the respective F2 values. The added value of classification function F2, combining IWV and InI, is in improved POD score by: 1) 20% in September for GRAM, 2) 23% in August for BGER, 3) by 11% in August for DCER, and 4) by 16% in July for DDUB. This improvement

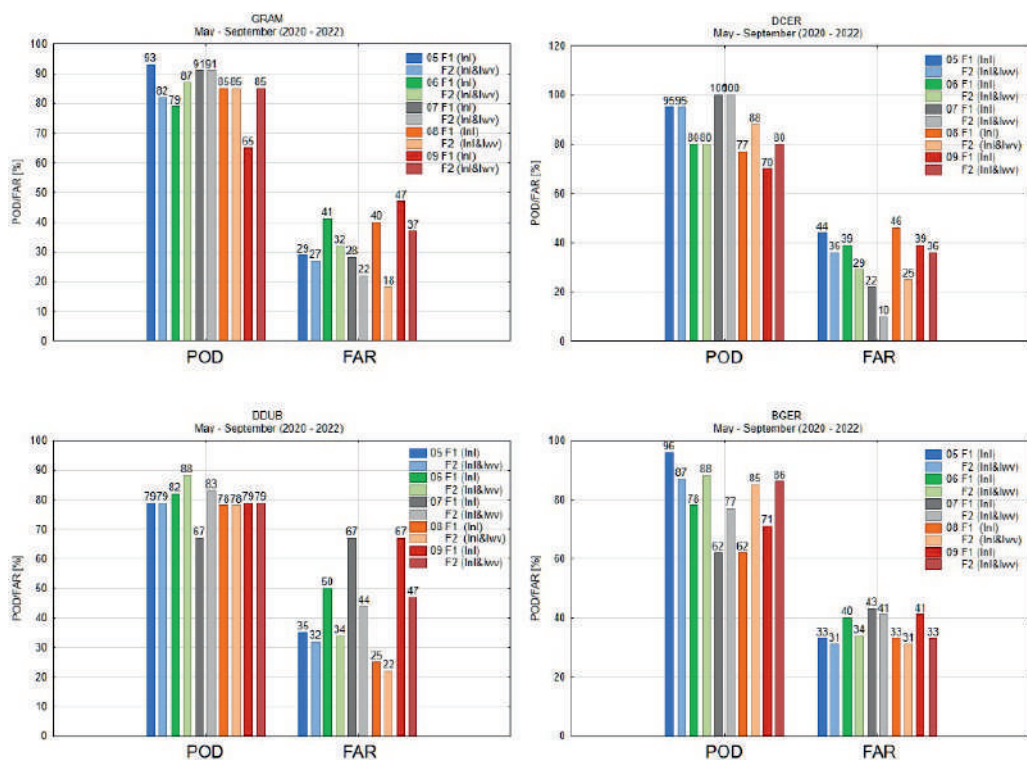


Fig. 3. Monthly Probability of Detection (POD) and False Alarm Ratio (FAR) skill score for classification function with 1) InI (F1, dark colours), and 2) InI and IWV (F2, light colours) for stations: a) Gramada (GRAM), b) Dolno Cerovene (DCER), c) Bardarski Geran (BGER), and d) Dolni Dubnik (DDUB). May (blue bars), June (green bars), July (grey bars), August (orange bars) and September (red bars)

likely reflects the best separation between IWV(P) and IWV(nP) groups for the months July, August and September. A comparison with the results reported in [4] for Sofia the monthly POD scores in the period 2010–2015 improved by 10% in September and 8% in June and August.

With respect to FAR values the best score is obtained for classification function F2 which combines InI and IWV. As seen from Fig. 3a for GRAM station F2 FAR values for all months are lower than the F1 FAR. The largest improvement was in August with the F1 FAR score of 40% (orange bar) reduced to 18% for F2 FAR score (light orange bar). For station BGER the largest FAR reduction between F1 and F2 is by 8% in September (Fig. 3b). As seen from Fig. 3c, d for station DCER the FAR reduction is by 21% in August and 23% in July for station DDUB, respectively. For Sofia station, the largest FAR reduction is 29% in September [4].

Conclusions. In this manuscript site-specific classification functions are computed for precipitation nowcasting during the convective storm period from

May to September for Northwest Bulgaria. A GNSS derived monthly IWV threshold was found to have a very good separation between precipitation (P) and no precipitation (nP) groups in July, August and September. A classification function computed with IWV has the best POD score in July and August in the range 77–100% and 78–85%, respectively. However, in July the FAR scores are also high in the range 30–66%. Due to the high FAR values classification function based on IWV alone has limited application.

Classification functions based on InI and IWV (F2) were found to have higher POD score than InI only function (F1). The POD score increased by 16% in July (station DDUB), 23% in August (station BGER), 11% in August (station DCER) and 20% in September (station GRAM). This is likely a result of good separation between IWV(P) and IWV(nP) values for the months July, August and September. The best FAR score is obtained for F2 as well. The largest FAR improvement is 23% in July (station DDUB), 21 and 22% in August (station DCER and GRAM), and 8% in September (station BGER). As a next step a machine learning model will be developed based on GNSS IWV and gradients.

REFERENCES

- [1] JONES J., G. GUEROVA, J. DOUSA et al. (2020) Advanced GNSS Tropospheric Products for Monitoring Severe Weather Events and Climate. COST Action ES1206 Final Action Dissemination Report, <https://doi.org/10.1007/978-3-030-13901-8>.
- [2] BONAFONI S., R. BIONDI, H. BRENOT, R. ANTHES (2019) Radio occultation and ground-based GNSS products for observing, understanding and predicting extreme events: A review, *Atmospheric Research*, **230**, 104624.
- [3] SHOJI Y. (2013) Retrieval of water vapor inhomogeneity using the Japanese nationwide GPS array and its potential for prediction of convective precipitation, *J. Meteorol. Soc. Jpn.*, Ser. II, **91**, 43–62.
- [4] GUEROVA G., T. DIMITROVA, S. GEORGIEV (2019) Thunderstorm classification functions based on instability indices and GNSS IWV for the Sofia plain, *Remote Sens.*, **11**(24), 2988.
- [5] WANG Y., E. DE CONING, A. HAROU et al. (2017) Guidelines for nowcasting techniques, WMO publication, ISBN 978-92-63-11198-2, published online: <https://library.wmo.int/idurl/4/55666>.
- [6] GUEROVA G., J. DOUSA, T. DIMITROVA et al. (2022) GNSS storm nowcasting demonstrator for Bulgaria, *Remote Sens.* **14**(15), 3746.
- [7] DACH R., P. WALSER, D. ARNOLD et al. (2015) Bernese GNSS Software Version 5.2, <https://doi.org/10.7892/boris.72297>.
- [8] DAVIS J. L., T. A. HERRING, I. I. SHAPIRO et al. (1985) Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length, *Radio Sci.*, **20**(6), 1593–1607.
- [9] BEVIS M., S. BUSINGER, T. A. HERRING et al. (1992) GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System, *J. Geophys. Res.: Atmospheres*, **97**(D14), 15787–15801.

- [10] DONALDSON R., R. M. DYER, M. J. KRAUS (1975) An objective evaluation of techniques for predicting severe weather events. In: Preprints, Ninth Conf. on Severe Local Storms, Norman, OK, Amer. Meteor. Soc.
- [11] SLAVCHEV M. (2022) Diagnosis and Forecast of Convective Processes in Bulgaria Using the GNSS Meteorology and Numerical Modeling Method, PhD Thesis, Sofia University “St. Kliment Ohridski”.
- [12] MILLER R. C. (1972) Notes on analysis and severe-storm forecasting procedures of the air force global weather central (severe weather air mass analysis and forecasting severe weather) [final report].
- [13] GEORGE J. J. (1960) Weather Forecasting for Aeronautics, Academic Press, Waltham.
- [14] GLICKMAN T. S. (2000) Glossary of Meteorology, American Meteorological Society.
- [15] GALWAY J. G. (1956) The lifted index as a predictor of latent instability, Bull. Am. Meteorol. Soc., **37**, 528–529.
- [16] WEISMAN M. L., J. B. KLEMP (1982) The dependence of numerically simulated convective storms on vertical wind shear and buoyancy, Monthly Weather Review, **110**(6), 504–520.

¹*Department of Meteorology and Geophysics, Faculty of Physics, Sofia University “St. Kliment Ohridski”, 5 James Bourchier Blvd, 1164 Sofia, Bulgaria*
e-mails: xmslavchevx@gmail.com, guerova@phys.uni-sofia.bg

²*National Institute of Meteorology and Hydrology, 66 Tsarigradsko Shosse Blvd, 1784 Sofia, Bulgaria*

³*Hail Suppression Agency, 17 Hristo Botev Blvd, 1606 Sofia, Bulgaria*
e-mail: dimitrova_tsvetelina@abv.bg