Verification of forecasts of intense convective phenomena

Andrzej Mazur, Joanna Linkowska

Institute of Meteorology and Water Management – National Research Institute

Report on sub-task 3.1 of COSMO Priority Project AWARE (Appraisal of Challenging WeAther FoREcasts)

Contents

i. Introduction

As it has already been said (cf. sub-task 2.1 report), every weather has its impact. In this part of the work carried out in the frame of AWARE Priority Project, all the activities focused mainly on the verification of the frequency of lightning discharges, predicted by means of various parameterizations. However, since every weather has its impact, each weather element can be treated as an impact source. It's just a question of scale and intensity. Therefore, the general results of the verification of convection indices - determining the possibility of hazardous meteorological situations - in relation to the measurements and calculations performed at aerological stations in Poland are additionally presented.

The verification method may be/could be/should be adapted (and specific) for each element. This report presents once again the basic assumptions of continuous (Mean Error, Root Mean Square Error) and discrete verification (FSS – Fraction Skill Score, SAL – Structure-Amplitude-Location, contingency tables) along with the idea of the VOD method, together with results of both the discrete verification and the continuous method, with the use of the VOD technique (cross-correlation/lagged correlation based on Vector Of Displacement).

ii. Methods

Survey on (basic) methods applicable to the problem (bold marks jobs done/partially done) :

- **SAL (Structure/Amplitude/Location) Verification**[1]
- **FSS (Fraction Skill Score) verification**[2]
- **Categorical analysis (Contingency tables and predictands)**

all the above further on called as "discrete" analysis

- **Standard evaluation at the grid scale** ("continuos" analysis)
- **Cross- (space-lag) correlation approach and verification**

Categorical analysis based on contingency tables

Forecast given	Event observed	
	Yes	No
Yes	Hit (a)	False alarm (c)
No	Miss (c)	Correct non event (d)

Using values *a, b, c* and *d* from the table above, predictands may be constructed as follows:

Table 1. Basic predictands derived from contingency tables

Basic predictands used:	**def. n==a+b+c+d**	**range**	**perfect**
Frequency Bias Index	$\frac{a+b}{a+c}$	-∞ to +∞	1
False Alarm Ratio	$\frac{b}{a+b}$	0 to 1	0
Probability Of Detection	$\frac{a}{a+c}$	0 to 1	1
Probability Of False Detection	$\frac{b}{b+d}$	0 to 1	0
Threat Score	$\frac{a}{a+b+c}$	0 to 1	1
True Skill Statistics	$\frac{a \cdot d - b \cdot c}{(a+c) \cdot (b+d)}$	-1 to 1	1
Equitable Skill Score	$\frac{a-a_r}{(a+b+c-a_r)}$ $a_r = \frac{(a+b) \cdot (a+c)}{n}$	-1/3 to 1	1
Proportion Correct	$\frac{a+d}{(a+b+c+d)}$	0 to 1	1
Success Ratio	$\frac{a}{(a+b)}$	0 to 1	1

[1] Wernli et al., 2008, SAL – a Novel Quality Measure for the Verification of Quantitative Precipitation Forecasts, Mon.Wea.Rev.136(11):4470–4487,https://doi.org/10.1175/2008MWR2415.1

[2] Blaylock and Horel, 2020: Comparison of Lightning Forecasts from the High-Resolution Rapid Refresh Model to Geostationary Lightning Mapper Observations, Wea. Forecasting 35, 402-416

Space lag (cross-) correlation approach

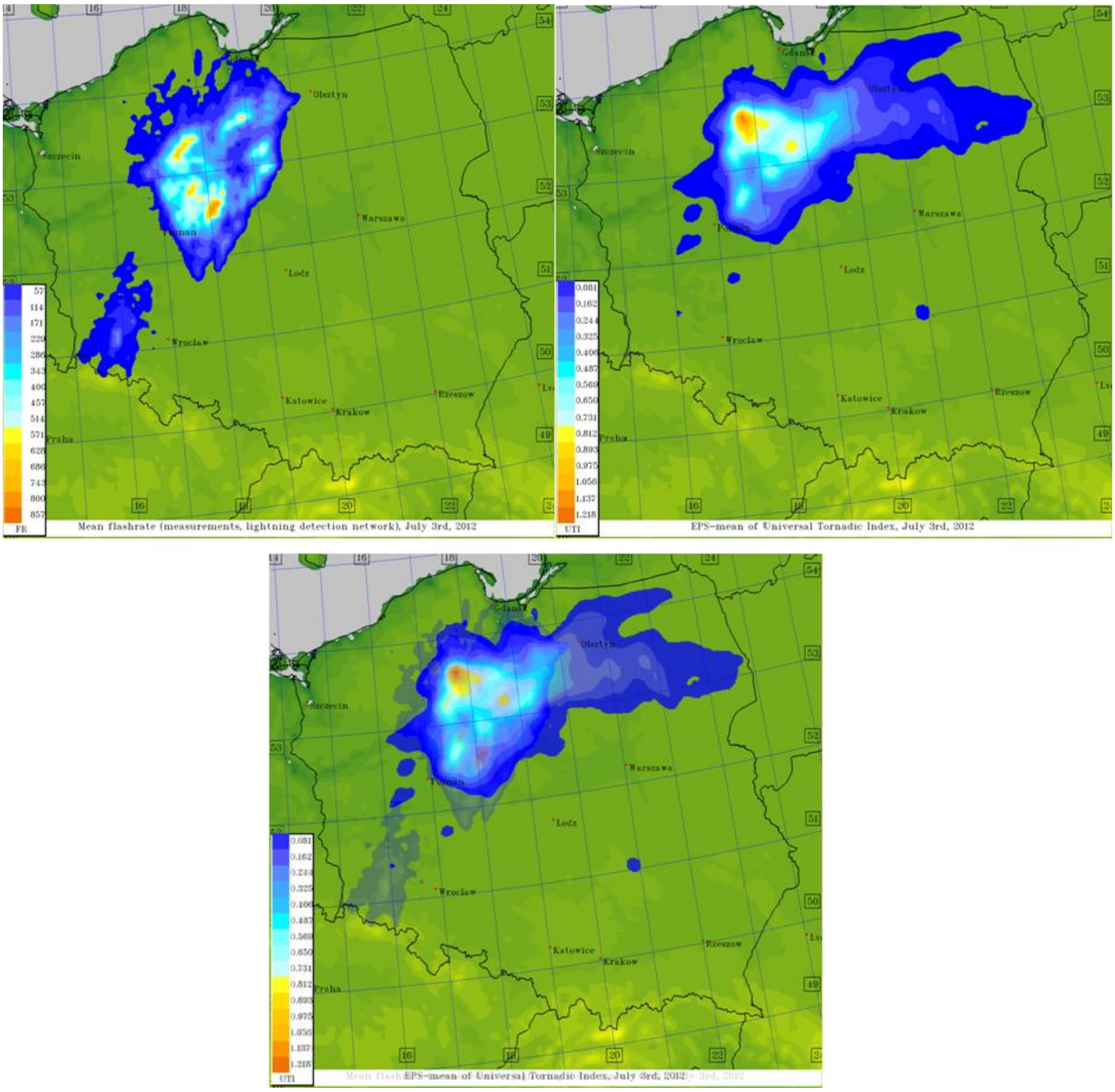

Fig.1 Basic idea of cross-correlation (lagged-correlation) approach

When overlap the upper left (observations field) and the upper right (forecasts) panels, in most cases they do not match. It is possible to improve the forecast by using the cross-correlation (or space lag correlation) method. To do this (using the example from the figure above) one should:

- Calculate coordinates of ”centres of mass” for both distribution patterns (observations vs. forecasts).
- Compute vector of displacement (VOD) of forecasts to observations as a difference of the two above.
- Displace linearly every value of forecasts field by the vector of displacement.

In operational work, VOD is calculated from previous model runs (as compared to observations). It is then assumed to remain constant throughout the next run.

SAL and/or FSS and/or categorical verification for the above period has been applied (both for direct and VOD approach) to the observed and forecasted Flash Rate *FR* as follows:

$$FR = \left(\frac{W}{14.66}\right)^{4.54}$$

with *W* being updraft velocity, calculated as

$$W = 0.3 \cdot \sqrt{2 \cdot CAPE}$$

FR is to be limited with the temperatures of top/bottom cloud temperatures, *CTT* and *CBT*, respectively

$$if CTT > -15°C FR = FR \cdot \left[max\left(0.01, \frac{-CTT}{15.0}\right)\right]$$

and

$$if CBT \leftarrow 5°C FR = FR \cdot \left[max\left(0.01, \frac{15.0 + CBT}{10.0}\right)\right]$$

Another limitation is due to lack of convective clouds – if (forecasted) cloud cover is below 25%, *FR* is set equal to zero. Moreover, case was selected to verification if (for both observations and forecasts) maximum value over the entire domain was greater than 20 strikes/hour , and the duration of the storm was greater than 6 hours.

All the verification (both “continuous” and “discrete”) was done for archive sets of observations (2011-2017).

Results are presented in the following tables and figures.

Basic analysis of the results showed that VOD improved virtually all categorical predictands (like FBI, POD, THS…) from 10 up to 45%.

iii. Examples and detailed results

Table 2. Categorical analysis based on contingency tables

	EQS		FAR		FBI		PFD	
	Direct	*VOD*	*Direct*	*VOD*	*Direct*	*VOD*	*Direct*	*VOD*
2012	0.0302	0.0842	0.8832	0.8240	2.7196	2.3366	0.1736	0.1611
2013	0.0773	0.1140	0.8254	0.7920	2.4679	2.1431	0.1483	0.1232
2014	0.0299	0.0671	0.9060	0.8632	3.4946	2.6446	0.1550	0.1258
2015	0.0263	0.1022	0.8785	0.7970	2.1706	1.8439	0.1311	0.1120
2016	0.0555	0.0751	0.8532	0.8370	2.7295	2.4354	0.1592	0.1344
2017	0.0505	0.0954	0.8296	0.7976	1.9107	1.6072	0.1180	0.0978
Mean	0.0420	0.0867	0.8676	0.8221	2.3164	1.9426	0.1499	0.1283
	POD		**SUC**		**THS**		**TRS**	
	Direct	*VOD*	*Direct*	*VOD*	*Direct*	*VOD*	*Direct*	*VOD*
2012	0.2366	0.4287	0.1169	0.1760	0.0826	0.1398	0.0754	0.2551
2013	0.3245	0.4685	0.1747	0.2081	0.1249	0.1667	0.2012	0.3202
2014	0.2193	0.3863	0.0940	0.1368	0.0681	0.1096	0.0935	0.2313
2015	0.1659	0.3890	0.1215	0.2030	0.0704	0.1543	0.0538	0.2579
2016	0.2644	0.3750	0.1469	0.1630	0.1030	0.1274	0.1299	0.2157
2017	0.1981	0.3433	0.1704	0.2025	0.0925	0.1452	0.1002	0.2253
Mean	0.2349	0.3987	0.1324	0.1779	0.0898	0.1390	0.1066	0.2489

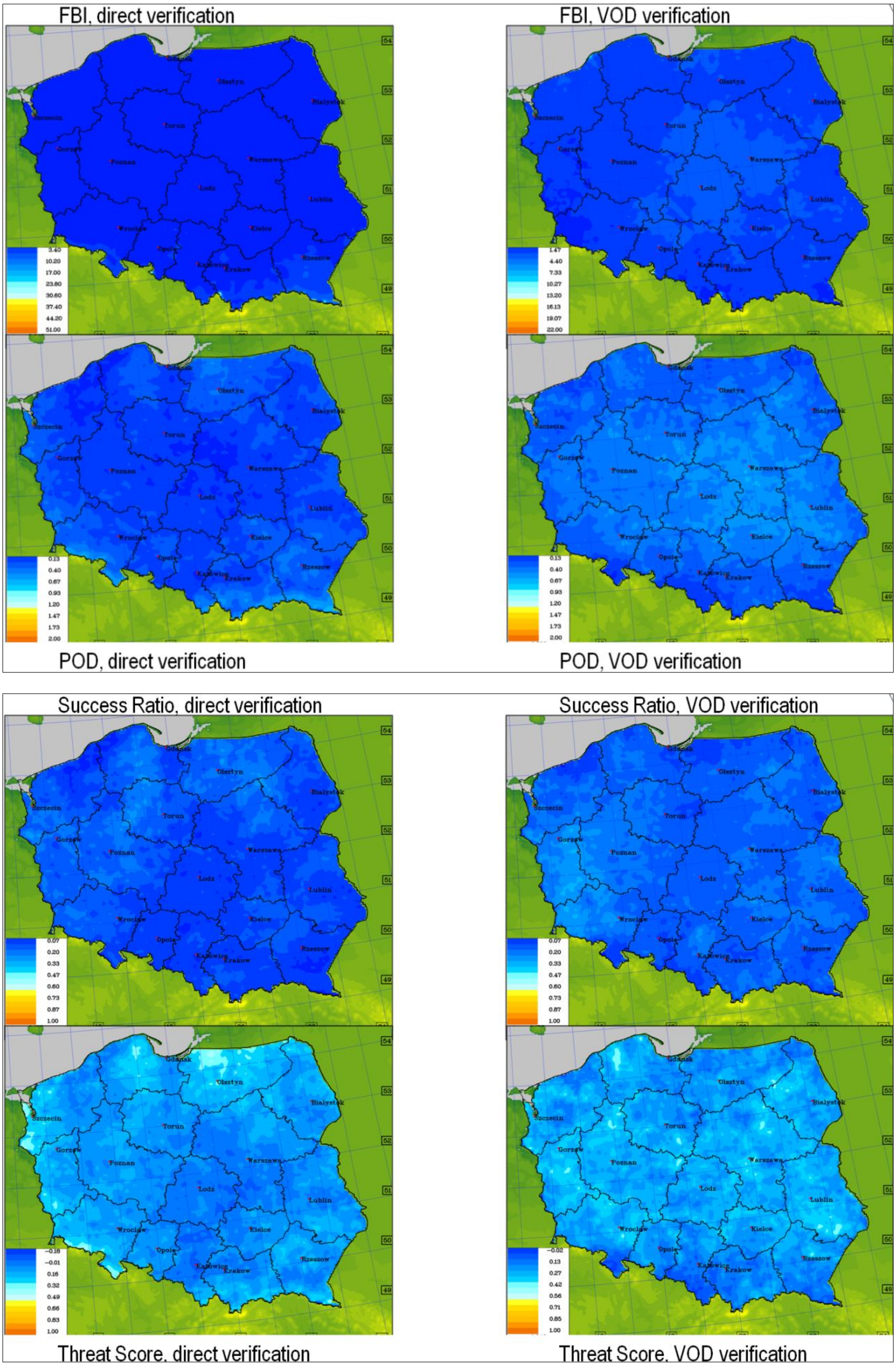

Fig. 2 Examples of results of contingency tables-based verification. Left – DMO, right – verification with VOD applied

Structure-Amplitude-Location (SAL) analysis.
S – structure – compares the volume of the normalized objects.

The structure component S analyses the size and shape of event objects. The values of S are within [-2,2]. The negative values of S correspond to too small and/or too peaked objects, while positive values indicate too large and/or too flat simulated objects. S=0 indicates a perfect structure.

A – amplitude – corresponds to the normalized difference of the domain-averaged values
The amplitude component A evaluates the total amount of event occurrence in a predefined region. The values of A are within [-2,2]. Negative values of A correspond to too little and positive values to too much predicted event occurrence, respectively. A=0 denotes perfect forecasts in terms of amplitude.

L – location – Combinations of a difference of mass centers of fields and averaged distance between the total mass center and individual objects
The location component L quantifies the displacement of observed and simulated precipitation objects, relative to their overall centers of mass. The values of L are within [0,2]. L=0 denotes the perfect value.

The perfect forecast is expected for S = A = L = 0
The examples of input data for SAL analysis and results are shown in the following figures.
It can be noticed that VOD forces some improvement in L-component and (to some extent) in A-component. S-component to a large extent remains unchanged. Forecasts, despite of applying VOD, are evidently overestimated. Choosing smaller domain (when SAL is to be more effective) and selection of more cases resulted, however, in no significant improvement.

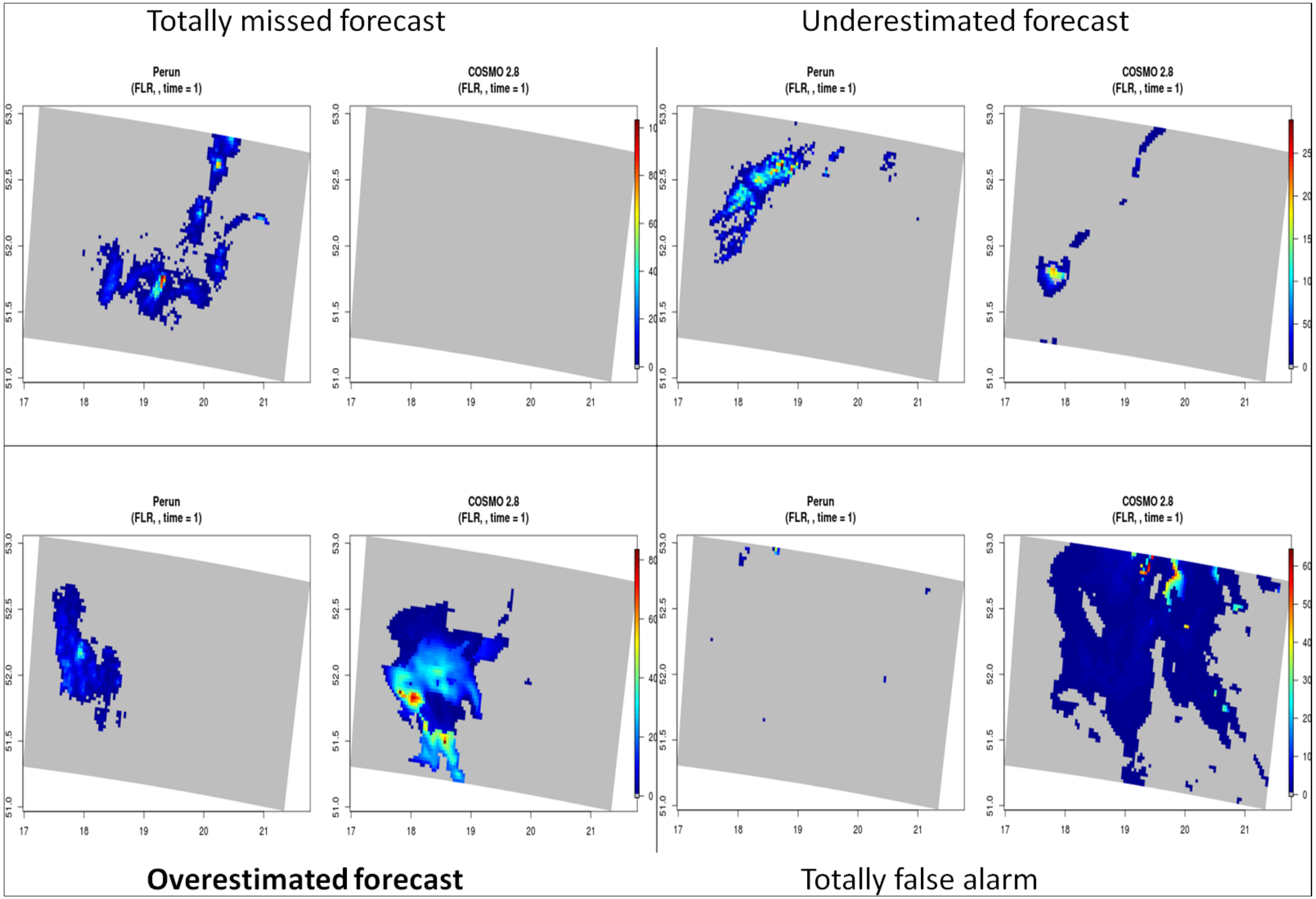

Fig. 3. The examples of input data for SAL analysis and results.

The most common case is marked with bold. The parameterization of Flash Rate based on the CAPE generally overestimates FR compared to the observations. Taking into account all cases from the selected period (2011-2017), the following analysis results were obtained.

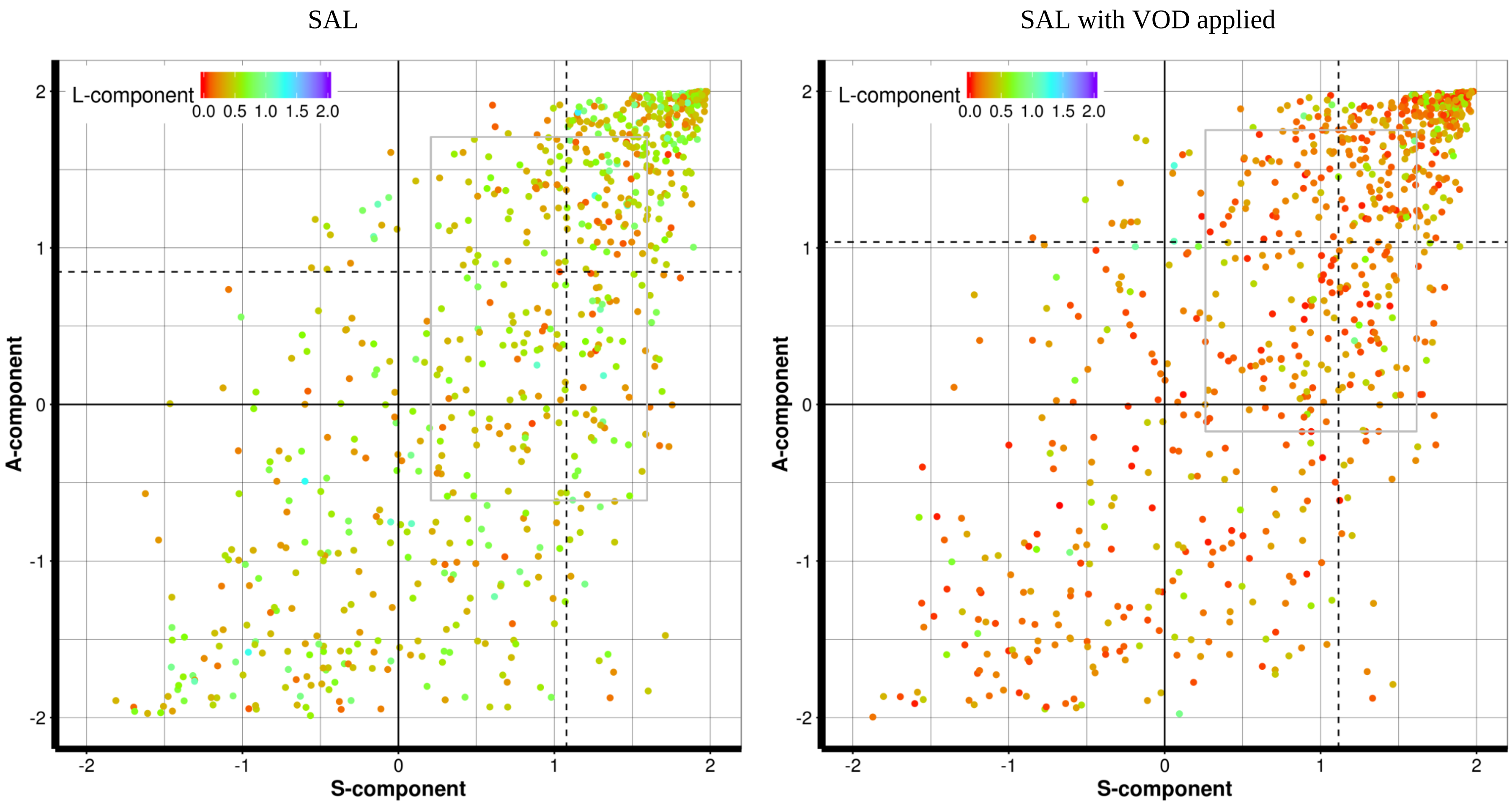

Fig. 4. SAL charts for flashrate, average (2011-2017). Left diagram – direct model output results, right diagram – corrected VOD procedure.

It can be noticed that VOD forces some improvement in L-component and (to some extent) in A-component. S-component to a large extent remains unchanged. Forecasts, despite of applying VOD, are evidently overestimated. Choosing smaller domain (when SAL is to be more effective) and selection of more cases resulted, however, in no significant improvement.

Fraction Skill Scores (FSS) assessment

This method allows for direct comparison of the forecast and of observed fractional coverage of grid-box events in spatial windows of increasing size. It is supposed to be most sensitive to rare events.

Assuming probability of the occurrence of the phenomenon (in the sense of observation) as p_o, and the forecast – p_f , can be defined by the FSS according to the formula below.

$$FSS = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\left(p_f - p_o\right)^2}{\frac{1}{N}\sum_{i=1}^{N}p_f^2 + \frac{1}{N}\sum_{i=1}^{N}p_o^2}$$

with N being number of sub-domains (or windows in overall domain).

When FSS = 0, there is no correspondence between observations and forecasts. If FSS is equal to 1, it describes a perfect match.

Again, results are shown in the following figures.

Results based on the DMO are not very good. VOD, however, significantly improves it, even up to 75%.

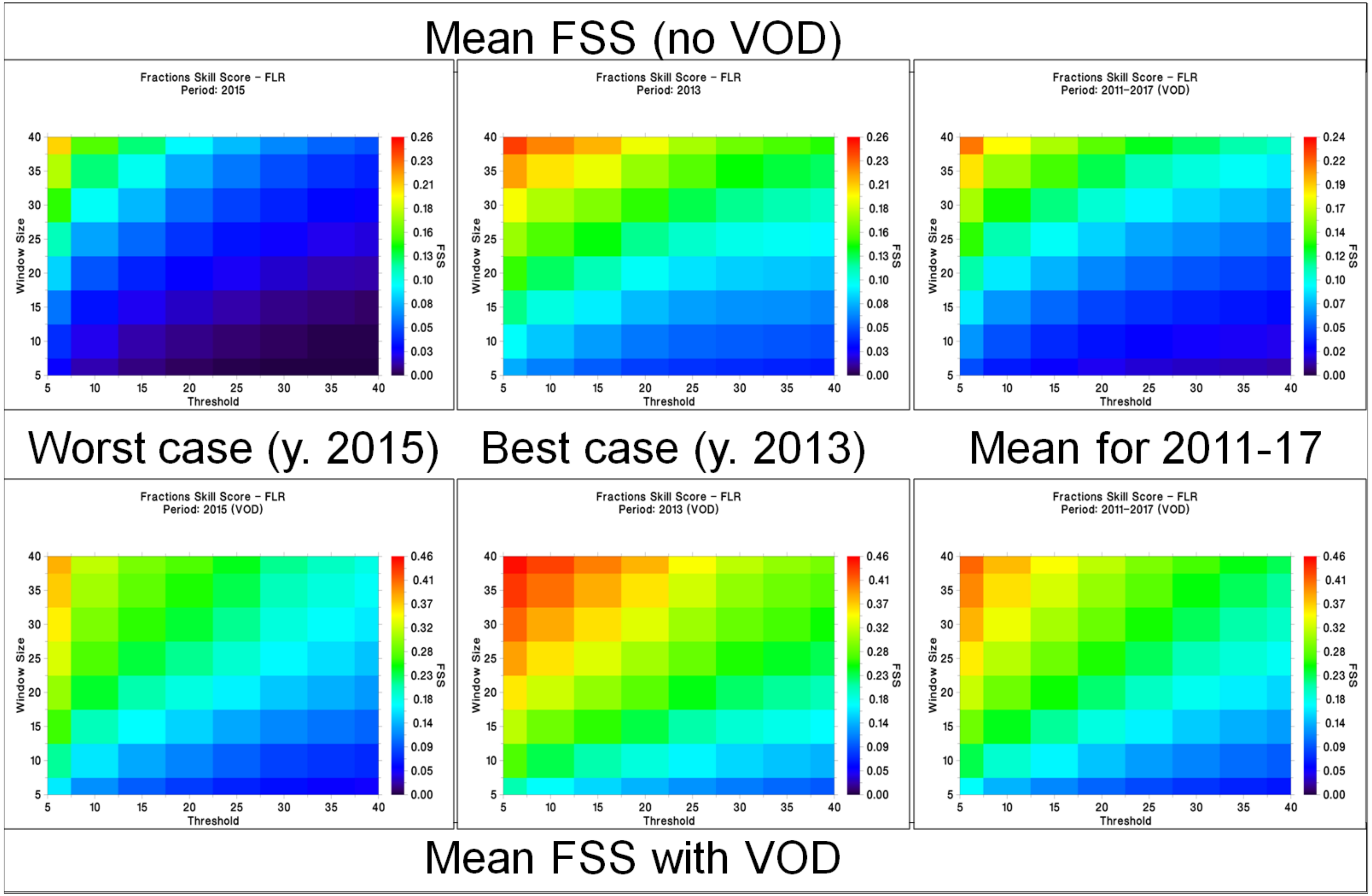

Fig. 5 Values of FSS for flashrate, worst/best/average (2015, 2013, 2011-2017). Upper charts – direct model output results, lower charts – corrected VOD procedure.

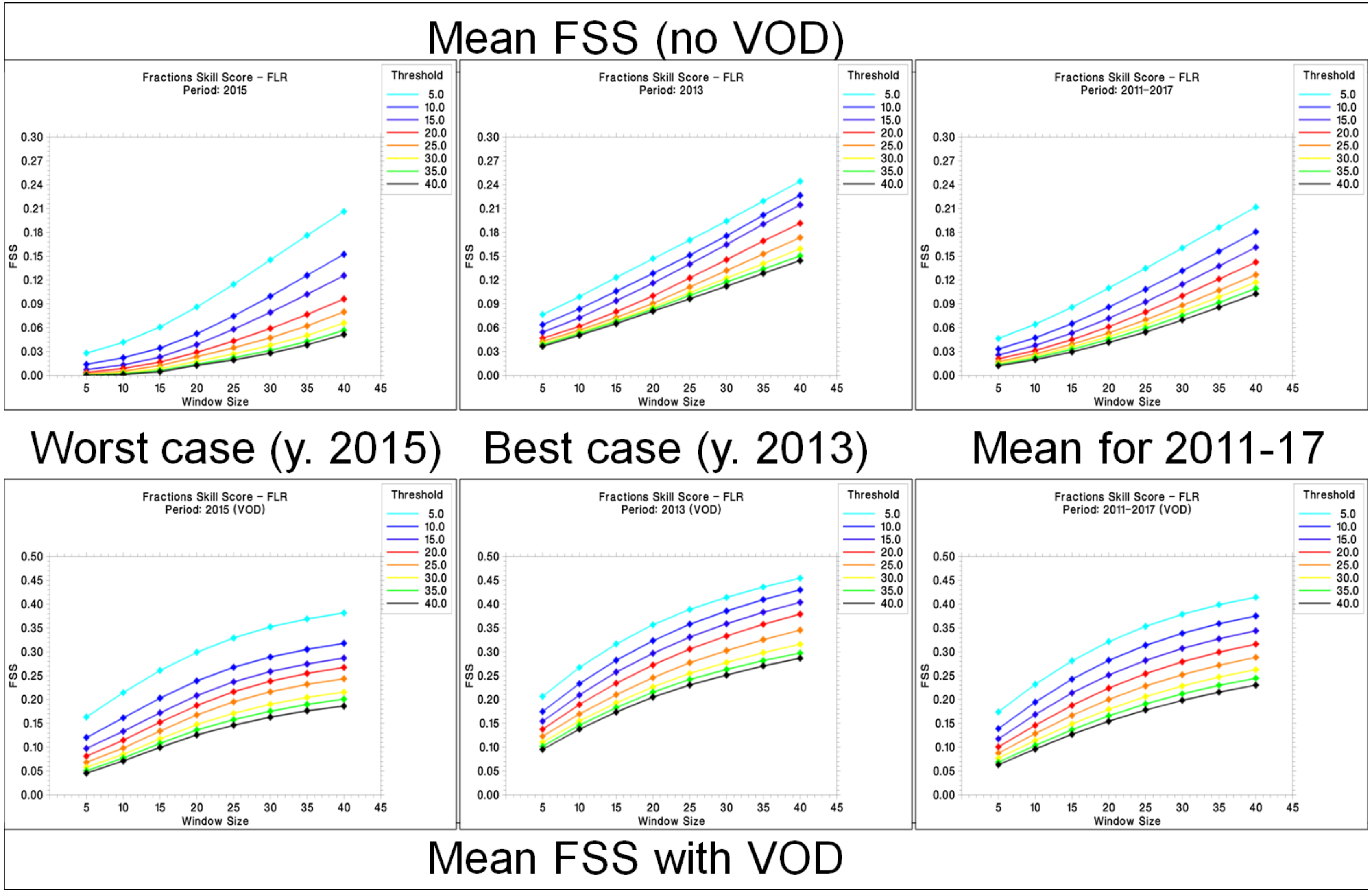

Fig. 6. Values of FSS for flashrate, worst/best/average (2015, 2013, 2011-2017). Upper charts – direct model output results, lower charts – corrected VOD procedure.

Finally, "continuous" analysis requires – in general – the calculation of Mean Error (ME), Mean Absolute Error (MAE) and/or Root Mean Square Error (RMSE). Then, the basic question is - which metric is better?

RMSE has the benefit of penalizing large errors more so can be more appropriate in some cases. However, it does not describe average error alone as MAE does. Yet, distinct advantage of RMSE over MAE is that RMSE doesn't use the absolute value – which is good in many mathematical calculations. Results of calculations – both for DMO and for VOD-applied results – are presented in following tables/figures

Table 3 Values of ME/MAE/RMSE for consecutive years and mean values for 2011-2017 both for "raw" (direct) values and corrected with VOD procedure.

	Direct			VOD		
Year	**ME**	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**
2011	2.128	**4.712**	**18.904**	1.887	**4.213**	**18.051**
2012	-2.811	**5.913**	**18.866**	-3.681	**5.027**	**17.482**
2013	-3.674	**2.184**	**10.556**	1.078	**1.949**	**9.970**
2014	-3.712	**1.516**	**9.186**	-2.192	**1.374**	**8.960**
2015	-2.023	**2.025**	**11.871**	-3.722	**1.819**	**11.391**
2016	-2.291	**3.360**	**14.695**	-0.699	**2.950**	**13.904**
2017	-1.286	**2.817**	**12.761**	-0.176	**2.015**	**11.879**
2011-2017	-1.953	**3.218**	**13.834**	-1.071	**2.764**	**13.091**

Examples of results for year 2013, 2017 (worse, best) and means for the period are presented in following figures.

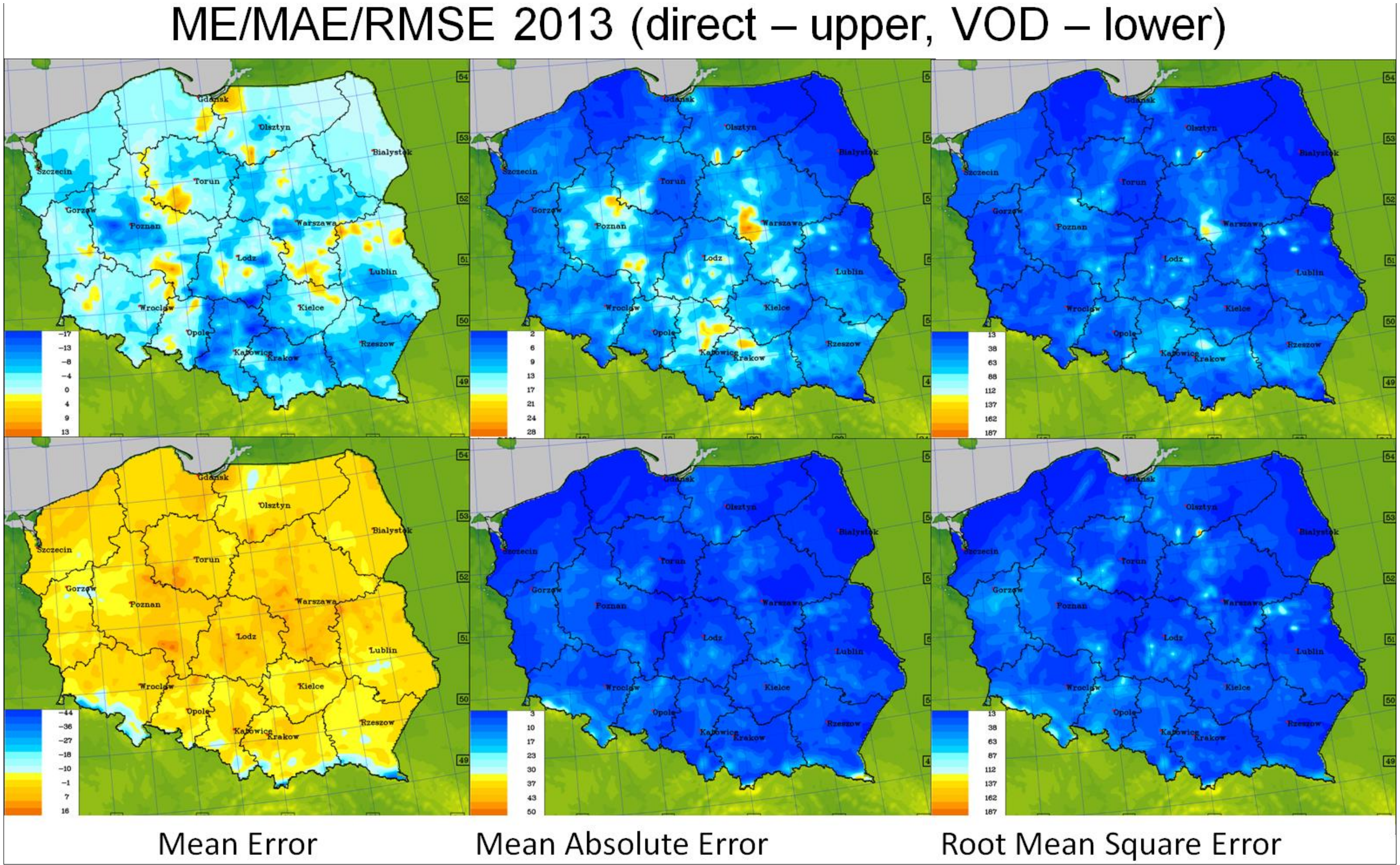

Fig. 7. Values of ME/MAE/RMSE for flashrate, worst avg. year (2013). Upper charts – direct model output results, lower charts – corrected VOD procedure.

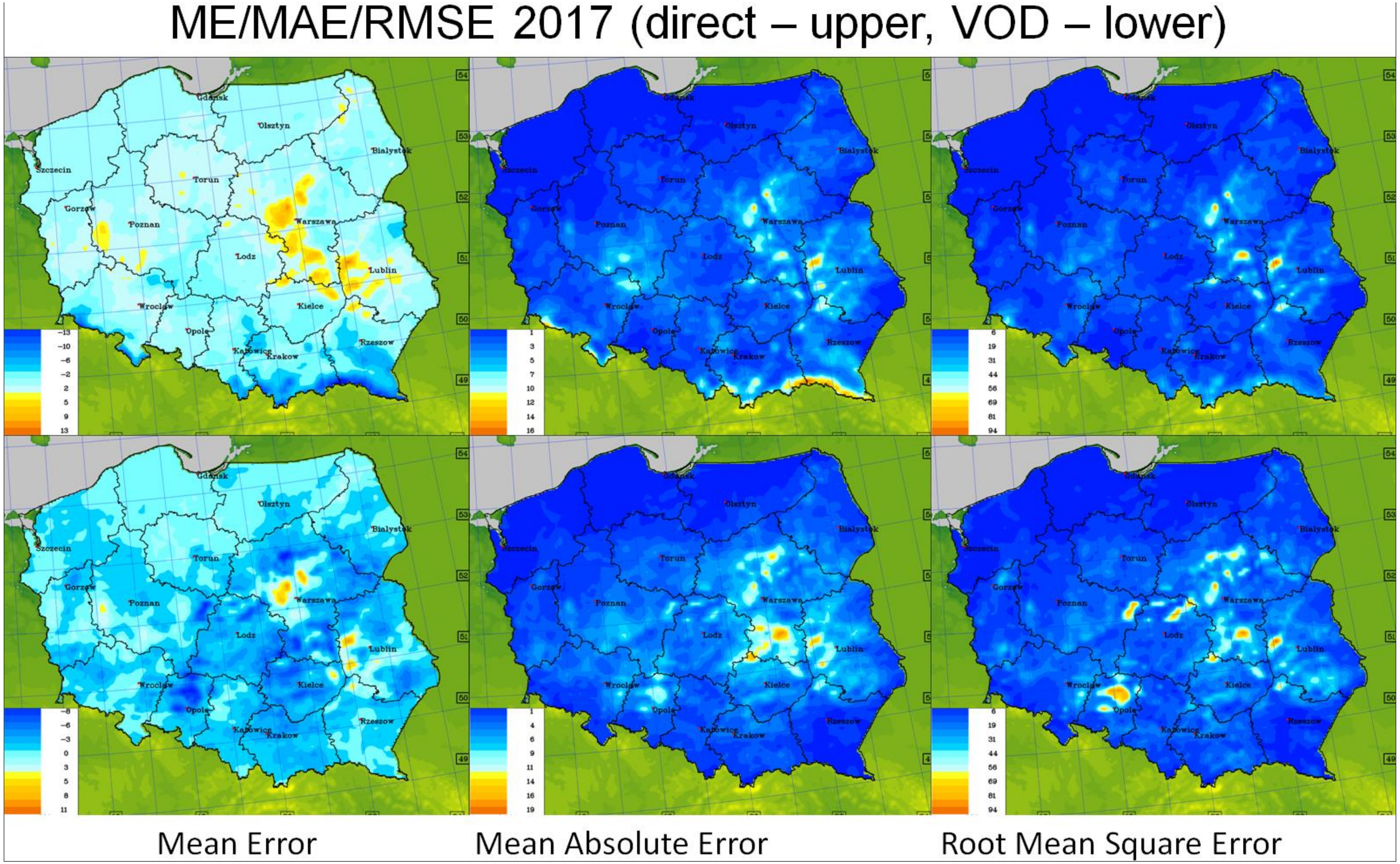

Fig. 8. Values ME/MAE/RMSE for flashrate, best avg. year (2017). Upper charts – direct model output results, lower charts – corrected VOD procedure.

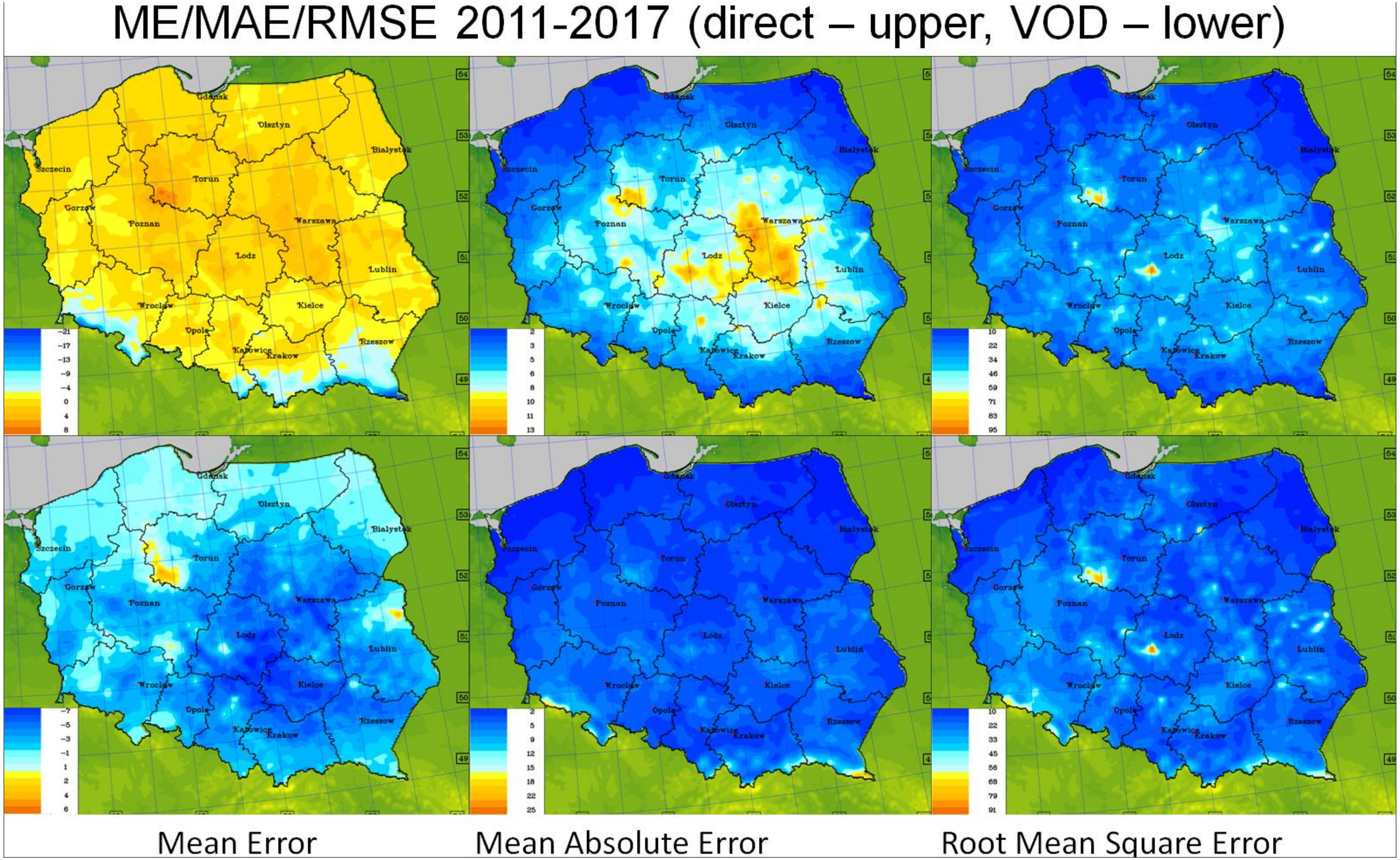

Fig. 9 Average (2011-2017) values ME/MAE/RMSE for flashrate. Upper charts – direct model output results, lower charts – corrected VOD procedure.

When consider MAE/RMSE calculated from DMO it can be seen that the worst values are apparently in mountainous regions. Maybe it is related to the fact, that it's hard(er) to predict thunderstorms in elevated terrain? When VOD procedure is applied to MAE/RMSE, slight improvement can be seen in comparison to direct verification, with a maxima of MAE/RMSE shifted towards domain centre.

In general VOD improves the results in all analyzes - continuous and discrete. This statement can be applied to all the cases presented.

With test period for direct- and VOD-verification extended to 2011-2017 SAL and/or FSS and/or categorical verification for the above period has been applied, both for direct and VOD approach, to four parameterizations of lightning intensity (cf. Fig. 10):

1. CAPE-based with cloud top/bottom temperatures correction (as described before)
2. Lightning Potential Index (LPI) (cf. U. Blahak, X.Lapillonne, D. Cattani)
3. Combination of the two above (cf. P. Lopez, D. Cattani)
4. Graupel flux at -15ºC level/total ice mass (cf. J. Wilkinson, McCaul *et al*. 2019).

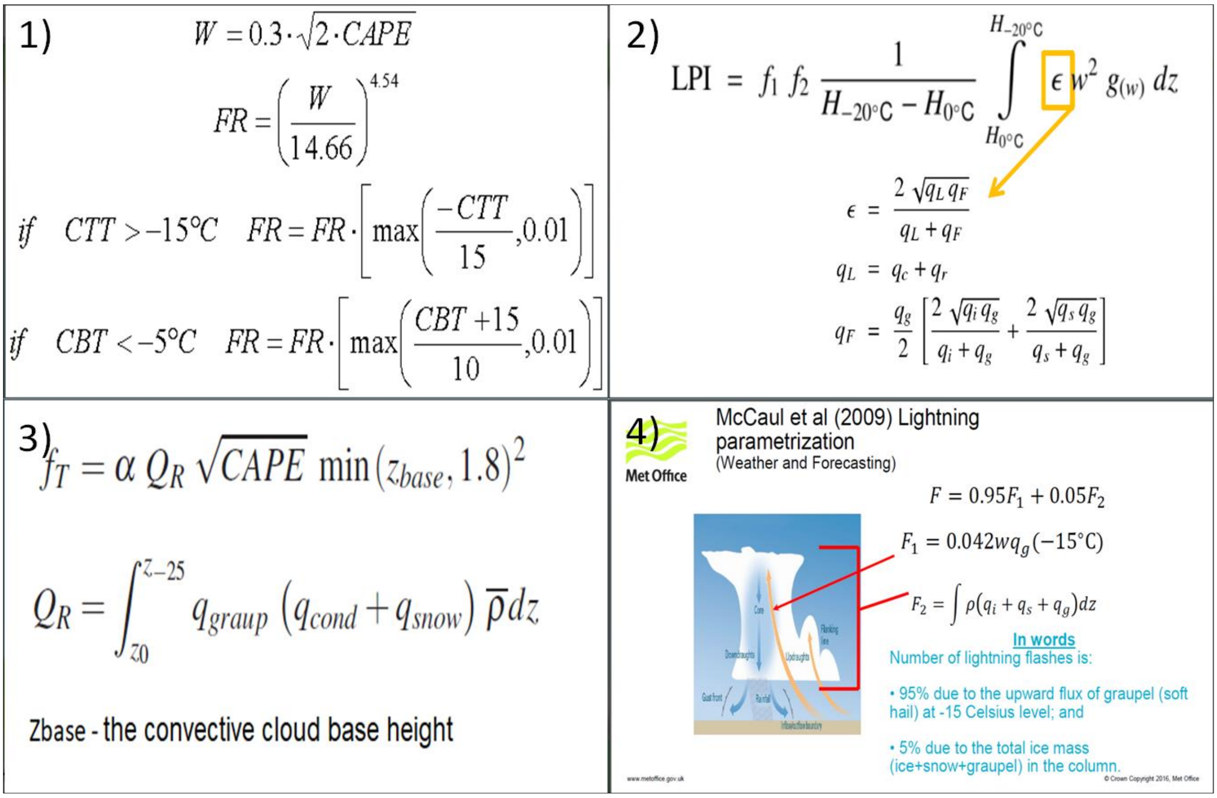

Fig. 10 Basic assumptions of different parameterization of lightning intensity.

These four parameterizations were tested and verified against observations for two periods:

- Case study – August 11th, 2017
- Longer period verification (June-August 2020; 7- and 2.8km only)

Results of the studies are shown in following figures/tables.

Case study August 11th, 2021

Table 3 Continuous verification results. ME – Mean Error, STD – Standard Deviation, MAE – Mean Absolute Error

Resolution	**7.0**				**2.8**				**0.7**			
#Parameterization	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
ME	3.1	3.1	2.0	3.7	1.2	0.4	-0.4	0.6	-0.6	0.2	-1.5	-0.9
STD	15.7	17.7	19.9	18.0	6.9	7.4	10.0	8.2	4.0	3.4	7.2	5.7
MAE	5.7	6.3	7.0	6.4	2.3	2.6	3.4	2.8	1.3	1.0	2.3	1.8

Discrete verification (Fraction Skill Score)

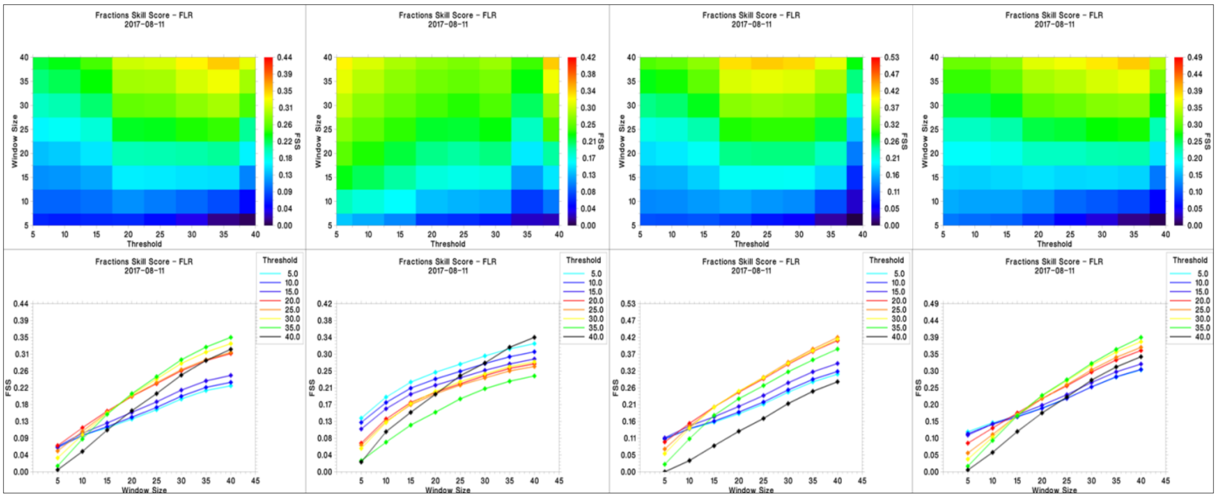

Fig. 11a FSS, 7km. DMO, left to right: parameterization #1-4 (2017.08.11)

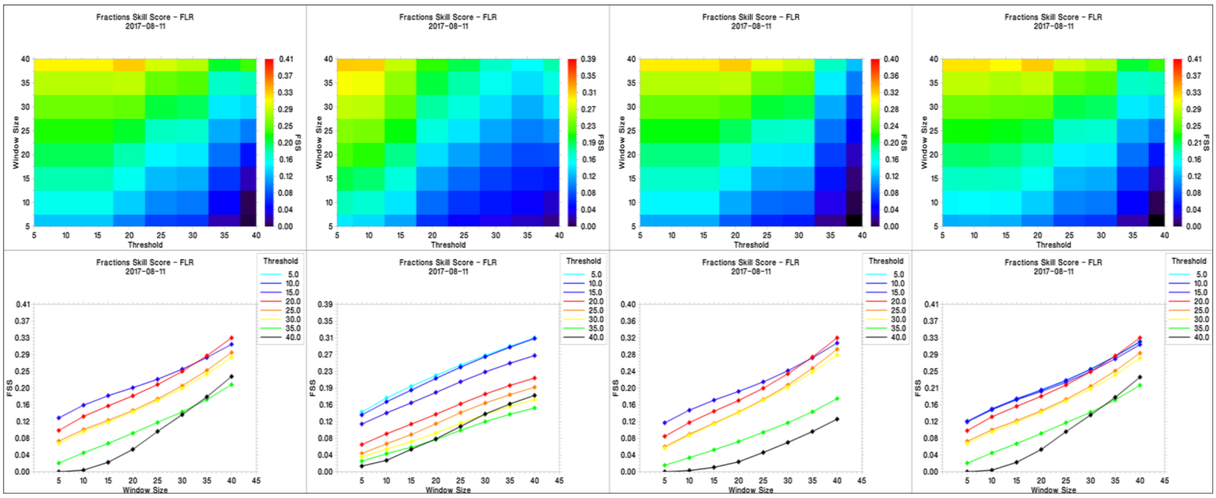

Fig. 11b FSS, 2.8km. DMO, left to right: parameterization #1-4 (2017.08.11)

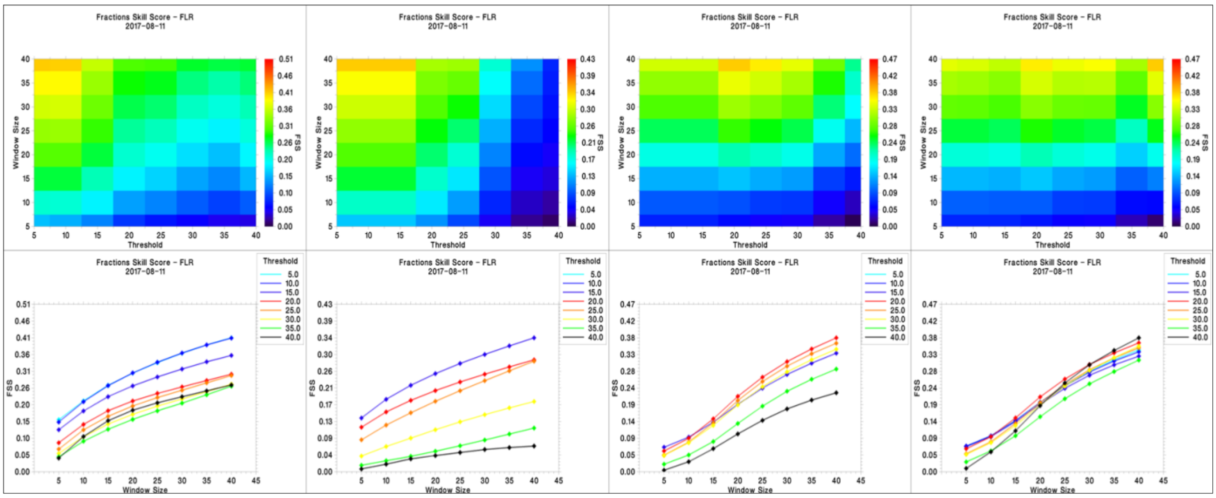

Fig. 11c FSS, 0.7km. DMO, left to right: parameterization #1-4 (2017.08.11)

Summer (June-August) 2020 verification

Table 4a. Verification based on contingency tables, 7km resolution

Parameterization	EQS	FAR	FBI	PFD	POD	SUC	THS
#1	0.051	0.830	1.911	0.118	0.198	0.170	0.093
#2	0.056	0.853	2.730	0.159	0.264	0.147	0.103
#3	0.030	0.906	3.495	0.155	0.219	0.094	0.068
#4	0.030	0.883	2.720	0.174	0.237	0.117	0.083

Table 4b. Verification based on contingency tables, 2.8km resolution

Parameterization	EQS	FAR	FBI	PFD	POD	SUC	THS
#1	0.084	0.823	2.337	0.126	0.386	0.176	0.140
#2	0.095	0.798	1.607	0.098	0.343	0.203	0.145
#3	0.075	0.837	2. 435	0.161	0.429	0.163	0.127
#4	0.067	0.863	2.645	0.134	0.375	0.137	0.110

Table 5. Continuous verification results. ME – Mean Error, MAE – Mean Absolute Error, STD – Standard Deviation.

Resolution	7km			2.8km		
Parameterization	**ME**	**MAE**	**STD**	**ME**	**MAE**	**STD**
#1	-9.32	3.61	23.51	-1.92	5.12	25.75
#2	-5.61	5.39	27.98	-5.66	3.49	21.41
#3	5.36	13.83	49.71	-9.23	12.16	45.61
#4	-7.28	11.87	48.63	4.18	10.71	46.83

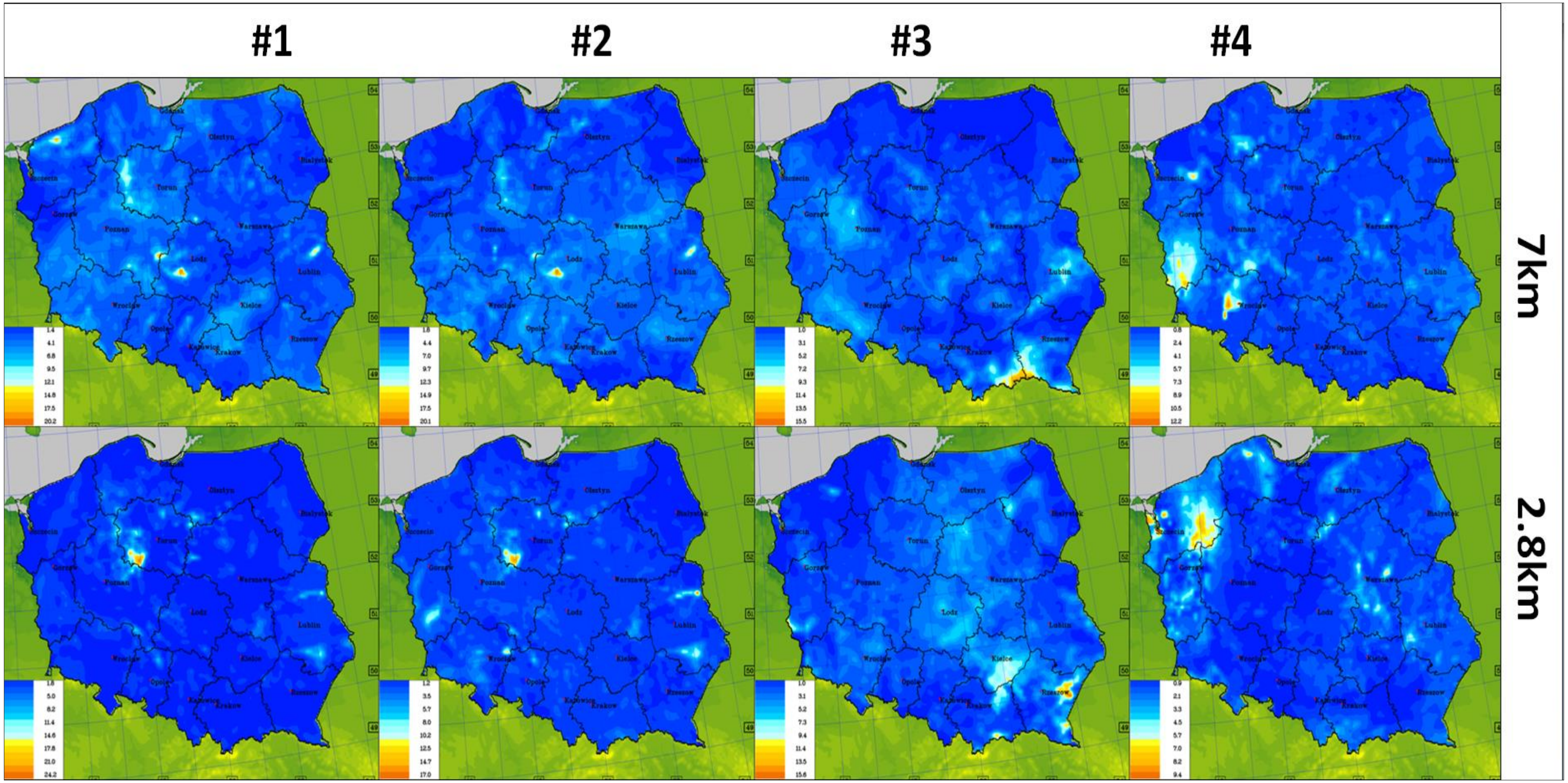

Fig. 12 Flashrate continuous verification – parameterizations (summer 2020). Average values of Standard Deviation.

Considering the four compared parameterizations for continuous verification, the first two seemed to work better than others. Namely, CAPE-based parameterization worked better in coarse resolution while LPI-based – in high resolution.

As far as the discrete verification is concerned (FSS, contingency tables analysis), in low resolution, results of 3rd parameterization seemed to be slightly better than the other two. In high resolution first parameterization worked best.

For a longer period, CAPE-based parameterization again worked better than others, while in low resolution – the one based on LPI.

Stability Indices

As far as these variables are concerned it has to be remembered that compared to the standard predicted values in the models (e.g. temperature, wind or precipitation), the possibilities of verification are significantly limited to data from atmospheric soundings and – to some extension – from satellite scans.

Atmospheric sounding (aerological) stations are located over Europe in much more scattered manner than – for example – SYNOP ones. Figure 11 presents basic available aerological stations in Europe.

Fig. 13 Aerological stations in Europe used to verify forecasts of stability indices. (from http://weather.uwyo.edu/upperair/sounding.html, access: September/October, 2021)

The following table presents an exemplary output from sounding at Wrocław aerological and SYNOP station, July 1st , 2021, 1200 UTC

Table 6. Output results from sounding at Wrocław (#12425), July 1st , 2021, 1200 UTC. http://weather.uwyo.edu/upperair/sounding.html, access: September/October, 2021

Station information and sounding indices

```
                             Station number: 12425
                           Observation time: 210107/1200
                           Station latitude: 51.13
                          Station longitude: 16.98
                          Station elevation: 116.0
                            Showalter index: 8.79
                               Lifted index: 8.76
    LIFT computed using virtual temperature: 8.79
                                SWEAT index: 50.51
                                    K index: 13.10
                         Cross totals index: 23.10
                      Vertical totals index: 23.60
                        Totals totals index: 46.70
      Convective Available Potential Energy: 4.69
             CAPE using virtual temperature: 5.21
                      Convective Inhibition: 0.00
             CINS using virtual temperature: 0.00
                           Equilibrum Level: 864.74
 Equilibrum Level using virtual temperature: 864.36
                   Level of Free Convection: 942.45
             LFCT using virtual temperature: 943.83
                     Bulk Richardson Number: 1.78
          Bulk Richardson Number using CAPV: 1.98
  Temp [K] of the Lifted Condensation Level: 270.88
Pres [hPa] of the Lifted Condensation Level: 952.54
   Equivalent potential temp [K] of the LCL: 284.16
     Mean mixed layer potential temperature: 274.69
              Mean mixed layer mixing ratio: 3.42
              1000 hPa to 500 hPa thickness: 5235.00
Precipitable water [mm] for entire sounding: 9.65
```

On the other hand, some stability indices can be assessed using satellite images. For instance, Showalter index is a measure of thunderstorm potential and severity. In other words, it gives a good indication where the atmosphere is unstable and where convective development may be expected. Fields of Showalter index (obtained via model forecasts, esp. in high resolution) may be compared with Meteosat 8 IR 10.8 satellite images. In some cases the discrepancy between the values of stability indices and the real situation (satellite image) can be noticed. Similarly, CAPE (Convective Available Potential Energy) – as a measure of the amount of energy available for convection – may be compared with Meteosat 8 IR 10.8 satellite images. It should be remembered that CAPE represents potential energy, and will only be used should a parcel be lifted to the level of free convection. The derived stability indices such as convective available potential energy (CAPE), lifted index (LI), total totals (TT), Showalter index (SI), and the K-index (KI) are computed from the retrieved atmospheric moisture and temperature profiles. These indices aid forecasters in nowcasting severe weather by providing them with a plan view of these atmospheric stability parameters. Forecasters use this information to monitor rapid changes in atmospheric stability over time at various geographic locations, thus improving their situational awareness in pre-convective environments for potential watch/warning scenarios.

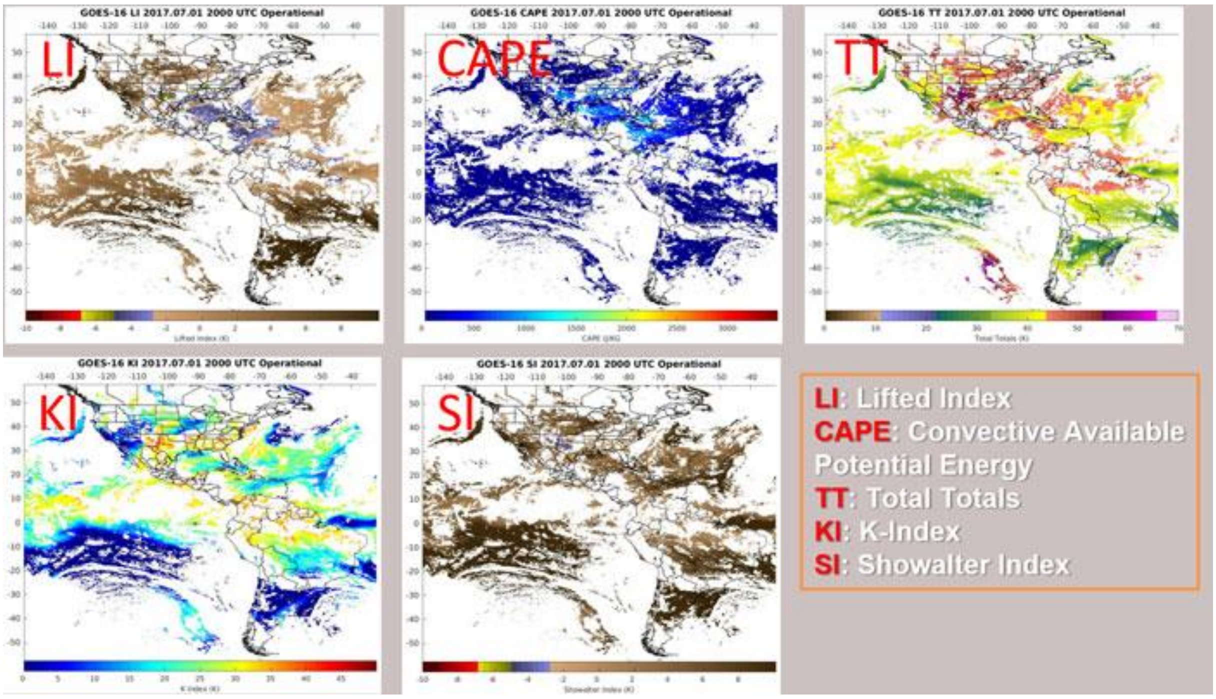

Fig. 14 GOES-16 (Geostationary Operational Environmental Satellites—R Series) derived stability indices product from July 1, 2017, including lifted index (upper left), convective available potential energy (upper middle), total totals (upper right), K-index (lower left) and Showalter index (lower middle). Source: https://www.goes-r.gov/products/baseline-derived-stability-indices.html, access: October 1st , 2021.

Of course, the limitations of satellite soundings (e.g. problems with scanning in cloudy conditions, space resolution etc.) set the limits for possibility of verification of indices. Hence, it is sometimes difficult to satisfactorily define the quality of the forecast of indicators – and the possibility of the severe weather phenomenon occurring – over a large area and / or in high spatial resolution.

In this part of the report authors decided to focus on the soundings-derived values of indices for summer period (June-August) of 2020. In order to maintain a consistent image for 2.8 and

7 km resolution, eight aerological stations, located in the domain for high resolution, were selected, as listed in Table 7.

Table 7. Aerological stations selected for verification of stability indices

Name	WMO Number	Country	Longitude	Latitude
Łeba	1210	Poland	17.50	54.75
Wrocław	12425	Poland	16.98	51.13
Legionowo	12374	Poland	20.93	52.38
Praha	11520	Czech Republic	14.46	50.00
Prostejov	11747	Czech Republic	17.09	49.46
Poprad	11952	Slovakia	20.26	49.05
Greifswald	10184	Germany	13.39	54.09
Lindenberg	10393	Germany	9.89	47.61

Due to small amount of points, lagged correlation procedure(s) has not been carried out. For the same reason, only continuous verification has been performed. For this verification, following indices have been selected: Showalter Index (SI), Lifted Index (LI), SWEAT index, K Index (KI), Totals Totals Index (TTI), Convective Available Potential Energy (CAPE) and Convective INhibition (CIN). Results are presented in the following tables.

Table 8. Mean error (ME) of stability indices forecasts' as compared to values at stations

Name	SI	LI	SWEAT	KI	TTI	CAPE	CIN
Łeba	-1.4	1.1	25	-8	-10.	-49	15
Wrocław	-2.0	1.3	29	6	17	38	-18
Legionowo	1.2	-0.9	19	9	11	35	-14
Praha	2.0	1.2	28	11	-18	50	22
Prostejov	2.5	1.1	-19	-10	-12	68	31
Poprad	2.1	-1.4	21	12	22	-54	-19
Greifswald	1.9	2.1	-23	12	-17	-51	21
Lindenberg	2.9	-2.5	-25	10	21	40	30
Average val.	1.2	0.3	7	5	2	10	9

Table 9. Mean absolute error (ME) of stability indices' forecasts as compared to values at stations

Name	SI	LI	SWEAT	KI	TTI	CAPE	CIN
Łeba	4	2	35	12	18	73	19
Wrocław	5	2	45	11	22	81	21
Legionowo	4	1.5	39	17	24	59	25
Praha	4	2	51	15	32	75	38
Prostejov	6	2	29	19	28	80	45
Poprad	4	2	40	21	35	67	32
Greifswald	5	4	39	22	29	65	34
Lindenberg	6	5	42	17	40	81	51
Average val.	5	3	40	17	29	73	33

Table 10. Root mean square error (RMSE) of stability indices' forecasts as compared to values at stations.

Name	SI	LI	SWEAT	KI	TTI	CAPE	CIN
Łeba	8	4	67	23	34	139	36
Wrocław	10	4	86	21	42	155	40
Legionowo	8	3	74	32	46	113	48
Praha	8	4	97	29	61	143	72
Prostejov	11	3	55	36	53	153	86
Poprad	8	4	76	40	67	128	61
Greifswald	10	8	74	42	55	124	65
Lindenberg	11	10	80	32	76	155	97
Average val.	9	5	76	32	54	139	63

iv. Conclusions

In every parameterizations, taking into account MAE/RMSE calculated from DMO it can be seen that the worst values are apparently in mountainous regions. Authors suggest that this effect may be related to the fact that it's hard(er) to predict thunderstorms in elevated terrain. Similar correlation is hard to find considering stability indices and measurements at aerological stations. This may be, in turn, caused by the small amount of verification point and their space locations.

Comparing ME/MAE/RMSE with the boundary values of individual stability indices that determine the change in the convection situation, it should be stated that – perhaps – only in the case of CAPE the compliance of the forecast with the measurements does not substantially affect the determination of this situation. In other cases, a forecast error may result in incorrect determination of the possibility (or lack thereof) of high-impact weather. An open question remains about the compatibility of measurements (and stability indices values, which, as it should be remembered, are not DMO) on aerological stations with reality.

When VOD procedure is applied to MAE/RMSE, slight improvement can be seen in comparison to direct verification, with a maxima of MAE/RMSE shifted towards domain centre. In general VOD improves the results in all analyzes - continuous and discrete. This statement can be applied to all the cases presented.

Further works are planned to improve the Flash Rate parameterization and verify the results obtained in this way, accordingly.

And last but not least important conclusion that could be drawn from all the above results is that if there is a possibility it is strongly suggested do both discrete and continuous verification.