OVERSHOOTING TOPS IN MSG 2.5-MINUTE RAPID SCAN DATA

Michaela Radová [1, 2], Jakub Seidl [3], Martin Setvák [1], Jindřich Šťástka [1, 2]

(1) Czech Hydrometeorological Institute, Satellite Department
(2) Department of Atmospheric Physics, Faculty of Mathematics and Physics, Charles University in Prague
(3) Institute of Plasma Physics AS CR

with contribution of **Kristopher Bedka [4]**
(4) NASA Langley Research Center

Prague, 3 November 2015

email: michaela.radova@chmi.cz

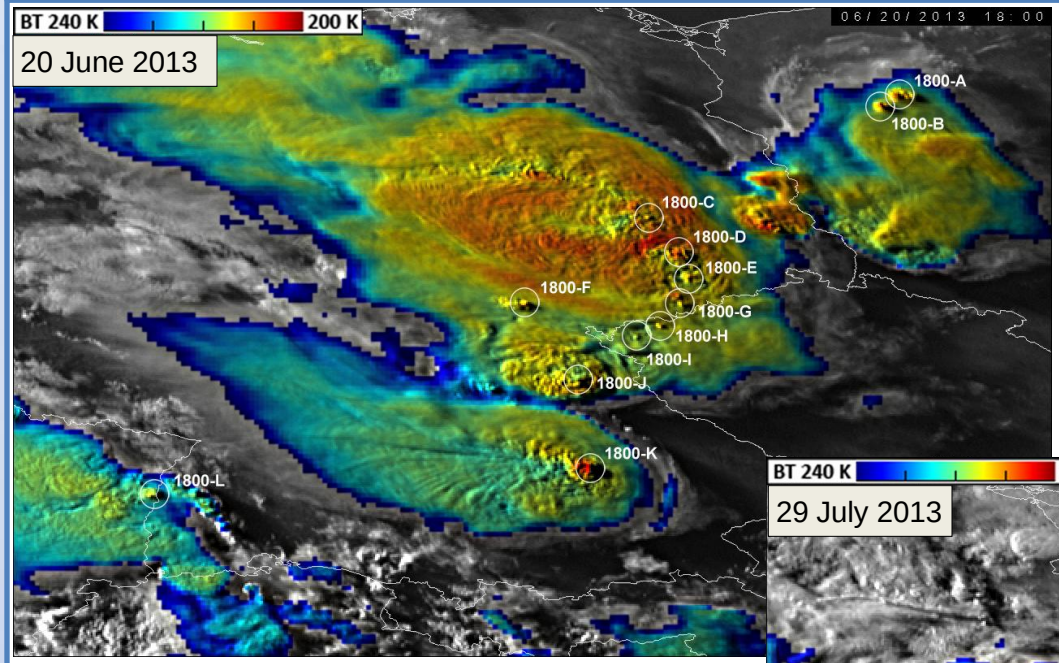
Outline

- **Subjective database of overshooting tops (OTs)**
- **Previously published algorithms of automatic OT detection**
- **Automatic OT detection using AdaBoost method**

Outline

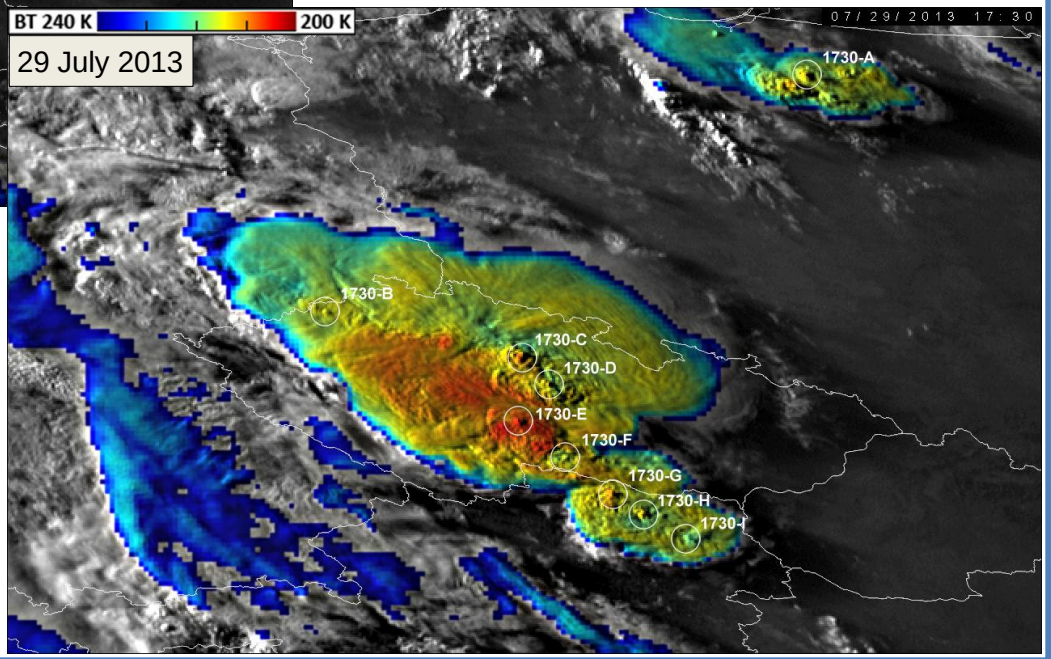
- **Subjective database of overshooting tops (OTs)**
- **Previously published algorithms of automatic OT detection**
- **Automatic OT detection using AdaBoost method**

Subjective database of overshooting tops (OTs)



used data:

- 20 June 2013 9:00 – 19:30 UTC
- 29 July 2013 13:00 – 19:00 UTC
- Meteosat-8 satellite: experimental **2.5 min RSS** data (see M. Setvák's presentations)



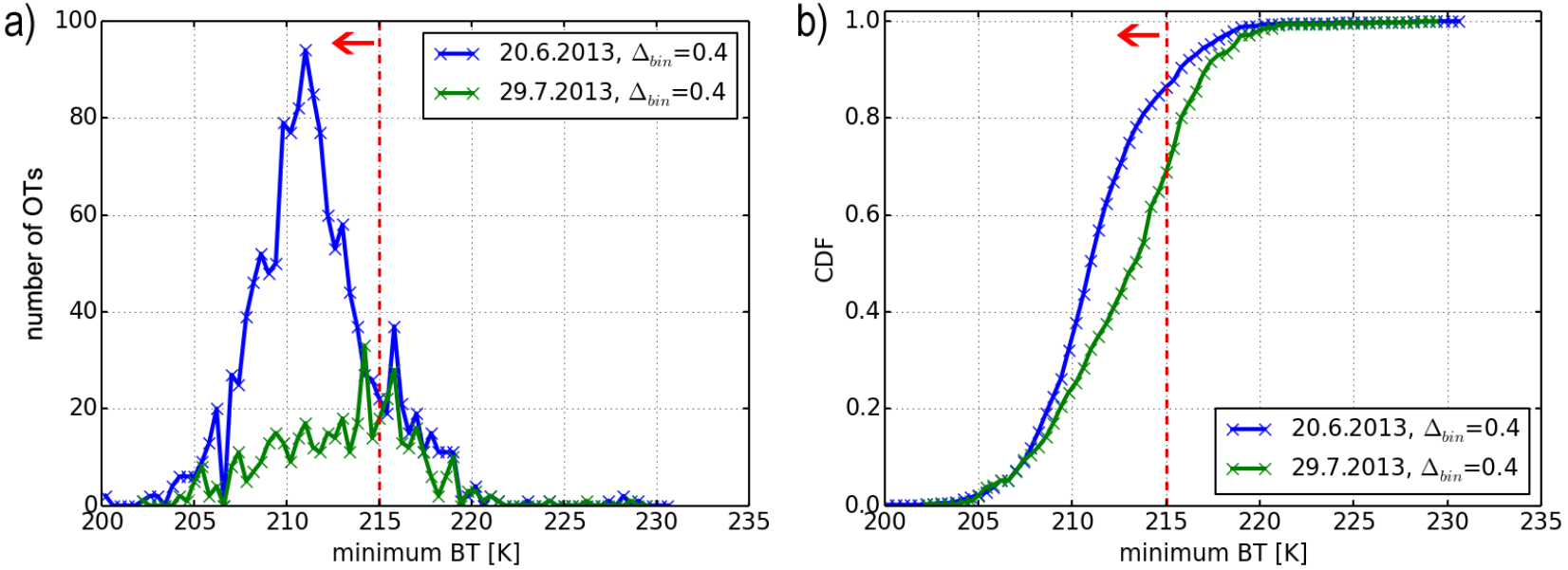
database of OTs:

- subjective detection based mainly on HRV imagery
- 20 June 2013 – **1365 OTs**
- 29 July 2013 – **446 OTs**

(see M. Setvák's presentation)

OT characteristics
- algorithms of automatic OT detection

Minimal brightness temperature (BT, IR10.8) for OTs in the database:



- **215 K** – upper threshold for BT in IR window (IRW) band (10-13 μm) often used as a criterion for OT detection (e.g. Bedka et al., 2010; Mikuš and Strelec Mahović, 2013)
- **30%** of OTs from 29 July 2013 and around **14%** from 20 June 2013 **do not meet** this condition and would not be detected by the automatic algorithm with such parameterization

Outline

- Subjective database of overshooting tops (OTs)
- **Previously published algorithms of automatic OT detection**
- Automatic OT detection using AdaBoost method

Previously published OT detection methods

Methods:

BTD (Brightness Temperature Difference) methods (Mikuš and Strelec Mahović, 2013):

method	criteria	
WV-IRW	IR10.8 < **215 K**	WV6.2 - IR10.8 > **4 K**
CO2-IRW		IR3.4 - IR10.8 > **3.5 K**
O3-IRW		IR9.7 - IR0.8 > **13 K**
COMB		WV6.2 - IR10.8 > **4 K** & IR9.7 - IR10.8 > **13 K**

modified BTD methods (adjusted for the OT database):

method	criteria	
CO2-IRW_4,5	IR10.8 < **215 K**	IR13.4 - IR10.8 > **4.5 K**
CO2-IRW_5		IR13.4 - IR10.8 > **5 K**
CO2-IRW_5,5		IR13.4 - IR10.8 > **5.5 K**
O3-IRW_16		IR9.7 - IR10.8 > **16 K**
COMB_3_16		WV6.2 - IR10.8 > **3 K** & IR9.7 - IR10.8 > **16 K**

based on thresholding

IRW (InfraRed Window) texture method (Bedka et al., 2010):

- detection of clusters of cold pixels with a size matching that of commonly observed OTs
- IRW BT ≤ 215 K & IRW BT ≤ $T_{tropopause}$
- IRW BT - IRW BT_{anvil} ≥ 6,5 K
- output of the method was provided by K. Bedka

based on spatial characteristics of IRW BT field

Comparison of previously published OT detection methods against the subjective database

Comparison:

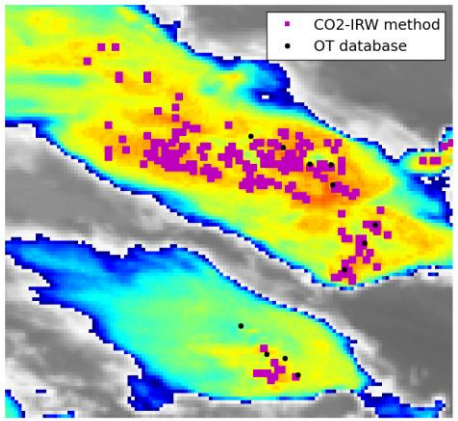
method	POD [%]	FAR [%]
IRW texture	20.6	62.5
WV-IRW	7.7	93.4
COMB	7	94
CO2-IRW_4,5	20.6	93.8
CO2-IRW_5	11	88.2
CO2-IRW_5,5	4.5	82.9
O3-IRW_16	10.6	92
COMB_3_16	7.5	91.5

- **POD (probability of detection)** and **FAR (false alarm ratio)** were computed for ***d* = 10 km**
- ***d*** is the maximal distance between detected OT and the closest OT from the database, in which we still consider both OTs as the same one

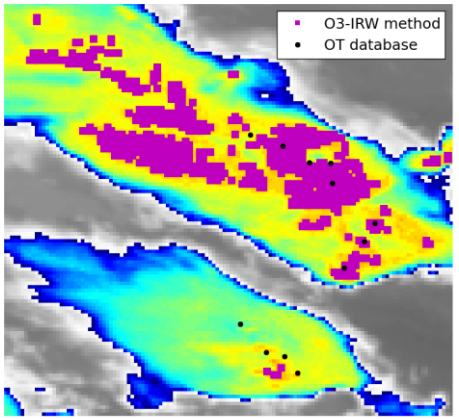
- best results:
  - IRW texture method
  - POD ~ 20.6 %, FAR ~ 62.5%

Areas detected by **CO2-IRW, O3-IRW** methods were typically too large to represent an OT.

CO2-IRW



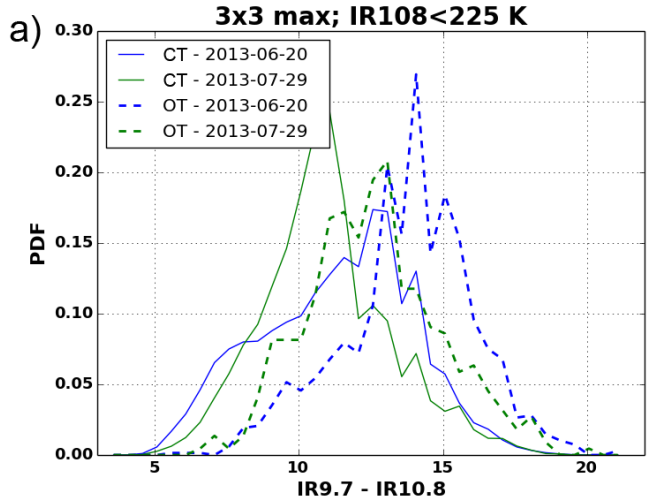
O3-IRW

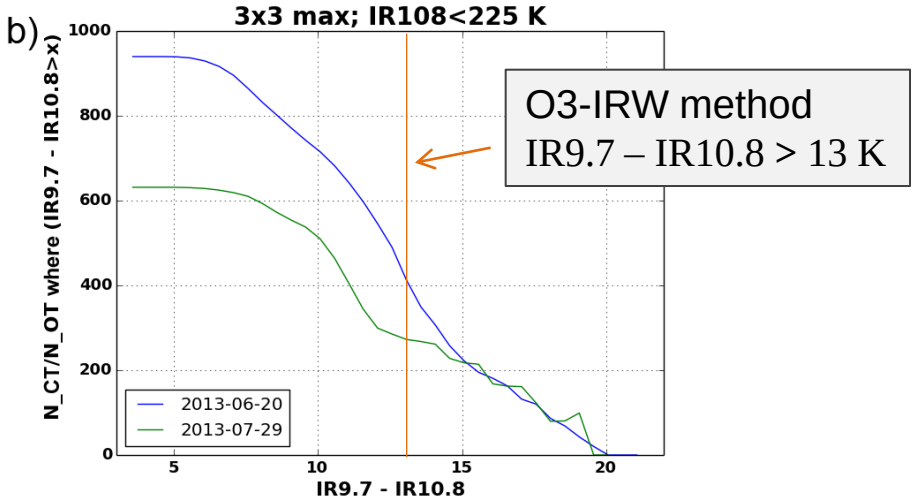


Comparison of previously published OT detection methods against the subjective database

BTD methods

- values of used BTD **typical for OTs** occur also **in non-negligible parts of surrounding cloud top (CT), N_CT >> N_OT**





Simple thresholding of one band or BTD of two bands can not sufficiently distinguish OTs from the surrounding cloud top.

Could a combination of more thresholds be more successful?

Comparison of previously published OT detection methods against the subjective database

IRW texture method

- assumes **sufficiently strong gradient in IRW BT** field in the vicinity of OT
- comparison with OTs from the database that were **accompanied by distinct local BT minimum** (6 % of OTs) → **POD ~ 81 %**
- **successful** for detection of **OTs** that are accompanied by **distinct local BT minimum.**

Spatial distribution of values measured in individual spectral bands may be important for OT detection.

Could a combination of both principles, detection based on thresholding and detection based on spatial characteristics, increase success rate of OT detection?

Outline

- Subjective database of overshooting tops (OTs)
- Previously published algorithms of automatic OT detection
- **Automatic OT detection using AdaBoost method**

Automatic OT detection using AdaBoost method

Adaboost method:

- **supervised machine learning method**
- applicable when a **larger amount of weak classifiers** can be constructed, each possibly with success rate just slightly above the level of random choice
- aim of the method: to create **new, more robust and more successful classifier** as a linear combination of these weak classifiers
- construction of classifier: iteratively – N steps – in each step, a new weak classifier is constructed and added to the emerging strong classifier in such a way that it minimizes its error over a training data set

OT detection:

- **binary classification** of individual spatial points – the output for each spatial point is logical value **YES/NO – the point corresponds to an OT or not**
- **weak classifier** – **binary decision tree**

- 3 phases - **training** + **testing** (construction of classifier)
  - **application** (OT detection)

Automatic OT detection using AdaBoost method

OT detection - training, testing, application:

- **training and testing phase** – construction of classifier

  - input - **2 disjunctive data sets** (training data set + testing data set)

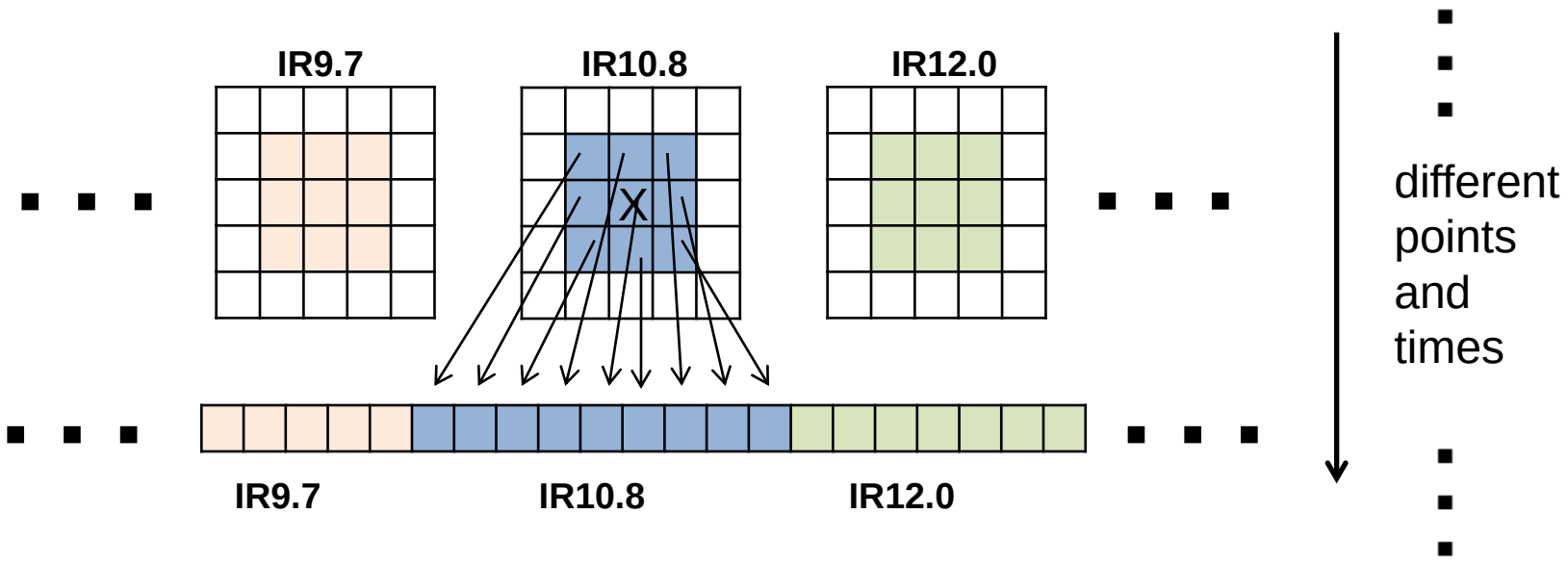
- purpose of **training** – **capture characteristics** of OTs that distinguish them from the surrounding cloud top
- purpose of **testing** – **prevent from overfitting** (classifier should reflect general rules in the data, but should not adapt to the training data set too firmly)
- **application** – OT detection

Automatic OT detection using AdaBoost method

- input – **values measured in individual bands of SEVIRI instrument** in the times corresponding to the appropriate **data set** (training, testing, application)

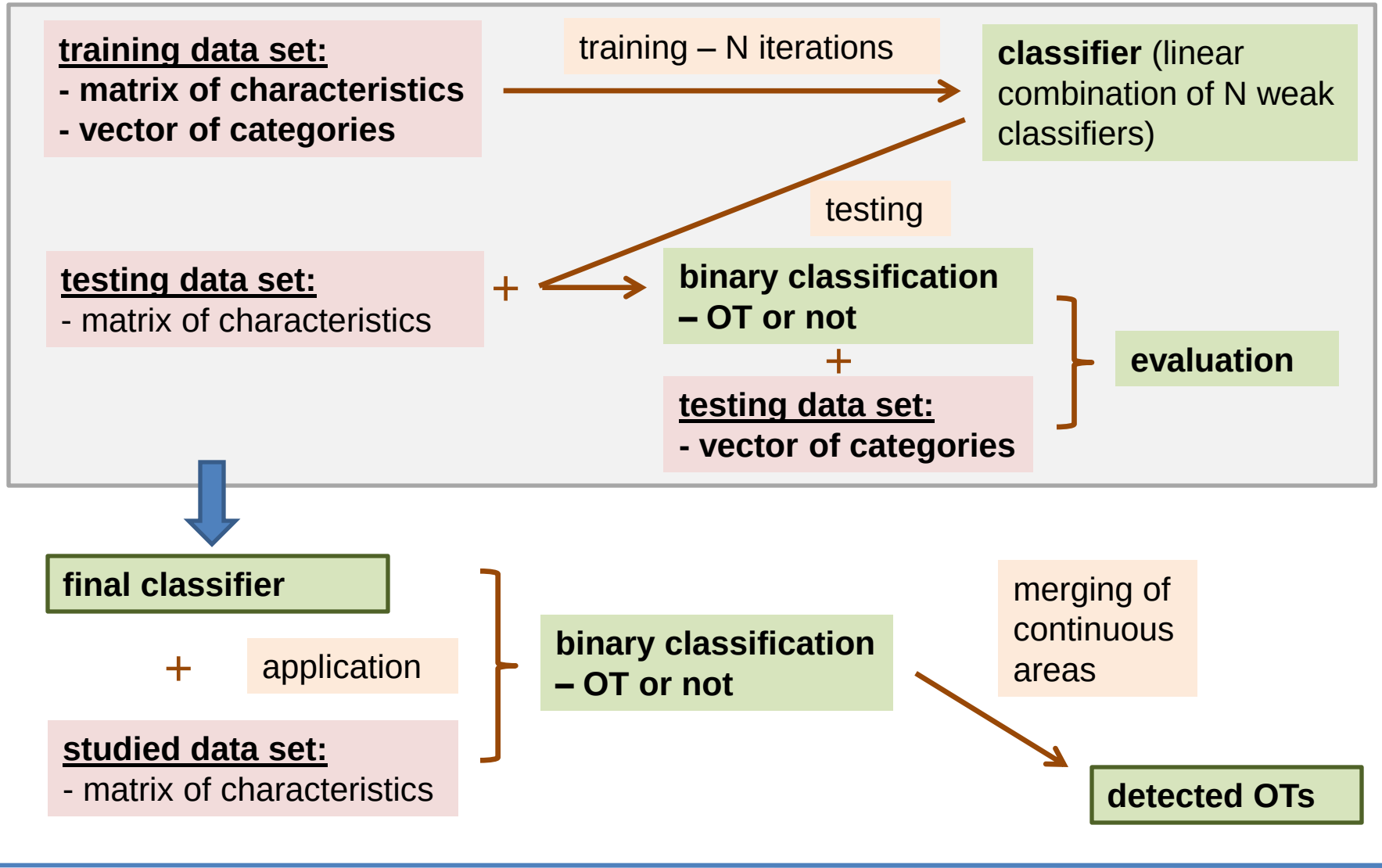
  – **information whether given spatial point corresponds to an OT or not – „vector of categories“** (training, testing)

- for each point and time of the data set we construct a **vector of characteristics** – values from all bands at the point + values from surrounding pixels (3 x 3 or 5 x 5 pixels; in pre-defined ordering)



→ **matrix of characteristics *V* – size: (# characteristics) x (#points ∗ # times)**

Automatic OT detection using AdaBoost method

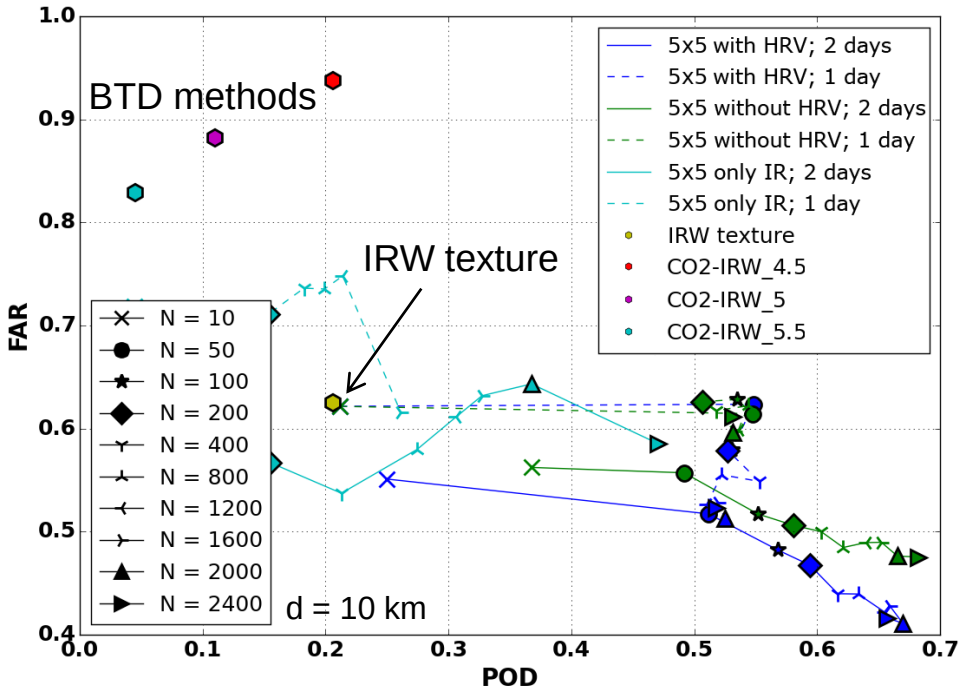


Automatic OT detection using AdaBoost method

different classifiers were tested:

- N = 10, 50, 100, 200, 400, 800, 1200, 1600, 2000 and 2400
- 2 variants of training and testing data sets:
  - training data set 20.6. + testing data set 29.7.
  - both data sets include half of the data from each day
- 3 variants of used bands of SEVIRI instrument:
  - using all bands (“with HRV“)
  - using all bands without HRV band (“without HRV“)
  - using only thermal bands (“only IR“)

Automatic OT detection using AdaBoost method



- IR bands only + training data set from 1 day – success rate comparable to the IRW texture method
- **including solar bands increases success rate** – POD ~ 55 %
- **training data set based on both days – increases success rate – POD ~ 66 %, FAR ~ 42 %**

„1 day“ = 20.6. as training data set + 29.7. as testing data set, „2 days“ = each set covers 50 % of each day

Proposed algorithm is more successful for studied days and region in comparison with BTD methods and IRW texture method.

Automatic OT detection using AdaBoost method

Importance of individual SEVIRI bands

Importance [%]			
Method / Band	5x5 with HRV N=2400 (2 days)	5x5 without HRV N=2400 (2 days)	5x5 only IR N=2400 (2 dny)
HRV	**46.4**	-	-
VIS0.6	8.5	**15.8**	-
VIS0.8	7.3	**16**	-
NIR1.6	9.1	**17.5**	-
IR3.9	3.6	6.3	-
WV6.2	5.9	9.7	**18.7**
WV7.3	3.2	6.3	14.0
IR8.7	3.7	5.7	12.8
IR9.7	4.1	5.5	12
IR10.8	1.7	4.0	11.4
IR12.0	3.1	6.2	15.6
IR13.4	3.4	7.0	15.5

But HRV band cannot be used at night!

Optimal algorithm – **band importance changing with daytime** or two different algorithms for day / night detection.

Future plans

current state of the algorithm = **proof of concept**

Optimization:

- **enlarge the input dataset** (database of OTs) to cover more complex set of synoptic situations → **add further input characteristics** – time of scanning, tropopause temperature and height
- test (non)linear combinations of different bands as input characteristics – e.g. use **principal component analysis**
- test **different weak classifiers**

Summary

- **Success rate of previously published OT detection methods was evaluated** using database of 1811 subjectively detected OTs.
- **BTD and IRW texture methods are not sufficiently reliable for OT detection at least for the studied days and geographic region** – the IRW texture method was the most successful: POD ~ 20.6 %, FAR ~ 62.5 %.
- It is **not possible** to base sufficiently reliable OT detection on **simple threholding** of 1 band or BTD of 2 bands.
- We propose OT detection algorithm based on the AdaBoost method - uses **complex system of thresholding conditions** and takes **spatial characteristics** into account – **POD ~ 66 %, FAR ~ 42 %. Broadening of training data set may lead to further increase of success rate**.
- **Bands in visible and near infrared spectrum bring the highest benefit for the OT detection.** Optimal: **algorithm takes into account importance of bands according to the daytime** or there are two different algorithms for day and night.

Summary

- It is important to be aware **what is the purpose of each method** and **how is it proposed**.

- **IRW texture method** - proposed to detect OTs with distinct gradient in IRW BT field and its POD for detection od such OTs is high.

- **detection based on AdaBoost method** – training phase is based on subjective database – some of OTs may have escaped their detection, while some of the detected ones may appear somewhat problematic

Thank you for the attention.

References:

BEDKA K., BRUNNER J., DWORAK R., FELTZ W., OTKIN J., GREENWALD T. (2010). Objective Satellite-Based Detection of Overshooting Tops Using Infrared Window Channel Brightness Temperature Gradients. Journal of applied meteorology and climatology, **49**, 2, 181-202.

MIKUŠ P., STRELEC MAHOVIĆ N. (2013). Satellite-based overshooting top detection methods and an analysis of correlated weather conditions. Atmospheric Research, **123**, 268-280.

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated.