Detection of convective overshooting tops using Himawari-8 AHI, CloudSat CPR, and CALIPSO data

Miae Kim¹, Jungho Im¹, Seonyoung Park¹

¹Ulsan National Institute of Science and Technology (UNIST), South Korea
Detection of convective overshooting tops using Himawari-8 AHI, CloudSat CPR, and CALIPSO data

Contents

01. Introduction
02. Research methods
03. Research results
04. Summary and future studies
Introduction

Overshooting Tops (OT): “a domelike protrusion above a cumulonimbus anvil, representing the intrusion of an updraft through its equilibrium level”

[American Meteorological Society’s *Glossary of Meteorology*]

**Importance of research on overshooting top**

- Cumulonimbus clouds with OT can cause severe weather conditions such as ground lightning, large hail, strong winds, and heavy rainfall, significantly influencing in-flight and ground aviation operations.
- The accurate detection of OT is important for inclement weather, lightning, and aircraft turbulence.
**IRW-Texture method:** as it uses gradients (i.e. texture) in brightness temperature, it is called “IRW-texture”. The method identifies a group of pixels with about 15 km in diameter and brightness temperatures significantly colder than the surrounding anvil cloud.

**WV-IRW BTD method:** it uses the difference of brightness temperatures between water vapor and infrared channel.
Research methods

Flow diagram of detecting overshooting top algorithm

Input data (Infrared channel brightness temperature)

OT & non-OT

- CloudSat Cloud profile
- CALIPSO Cloud lidar

Pixel-based Input variables

Object-based input variables

- Decision Trees
- Random Forest
- Support Vector Machines

Overshooting top result

Input variables

- CloudSat Cloud profile
- CALIPSO Cloud lidar

Decision Trees
Random Forest
Support Vector Machines
Data used in OT detection algorithm

Satellite data

<table>
<thead>
<tr>
<th>Satellite/Sensor</th>
<th>Channel information</th>
<th>Period</th>
<th>Spatial res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Himawari-8 Advanced Himawari Imager (AHI)</td>
<td>Infrared 10.4 μm (Band 13)</td>
<td>June 2015</td>
<td>2 km</td>
</tr>
</tbody>
</table>

Ancillary data (reference data)

<table>
<thead>
<tr>
<th>Satellite/Sensor</th>
<th>Used data</th>
<th>Period</th>
<th>Spatial resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>CloudSat Cloud Profiling Radar (CPR)</td>
<td>Cloud Geometrical Profile (2B-GEOPROF)</td>
<td>June 2015</td>
<td>Vertical res.: 480 m Swath: 1.3 km</td>
</tr>
<tr>
<td>CALIPSO lidar</td>
<td>Cloud lidar profile</td>
<td>June 2015</td>
<td>Vertical res.: 60 m Horizontal res.: 5 km</td>
</tr>
</tbody>
</table>
## Data used in OT detection algorithm

### Used input data

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Analysis method</th>
<th>Used variables</th>
<th>Period</th>
<th>Spatial res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Himawari-8 AHI</td>
<td>Pixel-based</td>
<td>● 10.4 μm channel brightness temperature&lt;br&gt;● 10 min. before 10.4 μm channel brightness temperature&lt;br&gt;● 10.4 μm channel average and standard deviation (Moving window size (MWS) = 5, 7, 11)&lt;br&gt;● Difference of 10.4 μm channel brightness temperature and 10 min. before one (MWS = 1, 3, 5)</td>
<td>June 2015</td>
<td>2 km</td>
</tr>
<tr>
<td></td>
<td>Object-based</td>
<td>● Object-based variables (a total of 13 variables)&lt;br&gt; - Area, asymmetric, compactness, mean, 10 min. before mean, radius of the major/minor axis, roundness, skewness, 10 min. before skewness, STD(standard deviation), 10 min. before STD, width</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Research methods

- Machine learning methods used for detection of OT

### Decision Trees (DT)
- Feature vector: \( \mathbf{v} \in \mathbb{R}^N \)
- Split functions: \( f_i : \mathbb{R}^N \rightarrow \mathbb{R} \)
- Thresholds: \( t_i \in \mathbb{R} \)
- Classifications: \( P_i(v) \)

### Random Forest (RF)
(Ensemble of several decision trees)

- Average prediction:
  \[
  (0.23 + 0.19 + 0.34 + 0.22 + 0.26 + \ldots + 0.31) / \text{# Trees} = 24
  \]

### Support Vector Machines (SVM)

- Decision Trees: A flowchart showing decision points based on feature vectors.
- Random Forest: A visualization of an ensemble of decision trees.
- SVM: A 3D representation of classification boundaries in predictor space.
Research results

◆ Construction of OT cases using CloudSat and CALIPSO data & Sampling for OT and non-OT region
  ❖ Construction of OT cases using CloudSat CPR data

At about 03:50
June 12, 2015
Research results

◆ Construction of OT cases using CloudSat and CALIPSO data & Sampling for OT and non-OT region
  ❖ Construction of OT cases using CloudSat CPR data

Himawari-8 image
at about 3:50 June 12, 2015
Research results

◆ Construction of OT cases using CloudSat and CALIPSO data & Sampling for OT and non-OT region

❖ Construction of OT cases using CALIPSO data

At about 18:20
June 13, 2015
Research results

- Inter-comparison of machine learning results (decision trees, random forest, support vector machines)
- Pixel-based OT detection result – Variable importance of DT & RF

### Attribute Usage (DT)

- 10 min. before 10.4μm TB
- 10.4μm TB mean (MWS 11)
- 10.4μm TB
- 10.4μm TB STD (MWS 11)
- 10.4μm TB - 10 min. before 10.4μm TB (MWS 5)

### Mean Decrease Accuracy (RF)

- 10.4μm TB mean (MWS 11)
- 10.4μm TB
- 10.4μm TB - 10 min. before 10.4μm TB (MWS 5)
- 10.4μm TB mean (MWS 7)
- 10.4μm TB STD (MWS 5)
- 10.4μm TB - 10 min. before 10.4μm TB (MWS 1)
- 10.4μm TB STD (MWS 7)
- 10.4μm TB - 10 min. before 10.4μm TB (MWS 3)
Research results

◆ Inter-comparison of machine learning results (decision trees, random forest, support vector machines)

❖ Pixel-based OT detection result – Qualitative validation using Himawari-8 image for DT & RF model result

Himawari-8 image at about 3:50 June 12, 2015

<table>
<thead>
<tr>
<th>DT model result</th>
<th>RF model result</th>
</tr>
</thead>
</table>

Yellow line: a track of CloudSat passing through OT occurrence region

- □ OT occurrence region identified by CloudSat
- ◈ Location of OT delineated by visual interpretation
Research results

- Inter-comparison of machine learning results (decision trees, random forest, support vector machines)
  - Pixel-based OT detection result – Qualitative validation using Himawari-8 image for SVM model result

**SVM model result**

Himawari-8 image at about 3:50 June 12, 2015

Yellow line: a track of CloudSat passing through OT occurrence region

- OT occurrence region identified by CloudSat
- Location of OT delineated by visual interpretation
Research results

◆ Detection of object-based OT detection
  ❖ Construction of OT cases based on visual interpretation

At about 3:10 June 12, 2015

At about 3:30 June 12, 2015
Research results

◆ Detection of object-based OT detection
  ❖ Result of segmentation for input variables using e-Cognition software

Himawari-8 image at about 3:50 June 12, 2015
Detection of object-based OT detection

- SVM result

Himawari-8 image at about 3:50 June 12, 2015

- Yellow line: a track of CloudSat passing through OT occurrence region
- OT occurrence region identified by CloudSat
- Location of OT delineated by visual interpretation
Research results

- Detection of object-based OT detection
  - Animation of RF result

Himawari-8 image at about 3:50 June 12, 2015

OT occurrence region identified by CloudSat

Unit: K

2015.06.12 2:50
Summary and future studies

- The result of OT detection using machine learning methods (decision trees, random forest, support vector machines) presented the best performance in SVM model based on qualitative validation for both pixel and object-based analysis.

- According to the information of variable importance from DT and RF model, average and standard deviation of brightness temperature (WMS 11), brightness temperature, difference of 10.4 μm channel brightness temperature and 10 min. before one (WMS 5), 10 min. before brightness temperature were identified as important variables for detection of OT in common.

- SVM model showed similar results for object and pixel-based OT detection results.

Further research

- To find OT reference data with satellite image and CloudSat and CALIPSO data so as to add more OT cases in training dataset and perform qualitative/quantitative validation with more OT cases for reliable OT algorithm.

- To develop day/night OT detection algorithm.

- To compare between machine learning approach and existing algorithm qualitatively and quantitatively.
Reference

- RDCA ATBD, JMA 2012
- ICCV Tutorial, Boosting and Random Forest for Visual Recognition, [http://www.iis.ee.ic.ac.uk/icvl/iccv09_tutorial.html](http://www.iis.ee.ic.ac.uk/icvl/iccv09_tutorial.html)
- NOAA NESDIS center for satellite applications and research, ATBD, Overshooting Top and Enhanced-V Detection, Version 1.0
- NOAA NESDIS center for satellite applications and research, ATBD, Flight Icing Threat, Version 1.0
Thank you

Intelligent Remote sensing and geospatial Information Systems (IRIS)

School of Urban and Environmental Engineering
Ulsan National Institute of Science and Technology, Ulsan, S. Korea

UNIST-gil 50, Ulsan 689-798, Republic of Korea
Tel : +82 52 217 2887
E-mail : iris-lab@unist.ac.kr